False Discovery Rate and Asymptotics

I n a u g u r a l - D i s s e r t a t i o n

zur
Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Thorsten-Ingo Dickhaus
aus Berlin-Kreuzberg

Januar 2008

Aus dem Institut für Biometrie und Epidemiologie des
Deutschen Diabetes-Zentrums, Leibniz-Institut an der
Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Arnold Janssen

Korreferent: PD Dr. Helmut Finner

Tag der mündlichen Prüfung: 15. Januar 2008

Contents

List of Abbreviations and Symbols

AORC	Asymptotically Optimal Rejection Curve
$B(p, q)$	Beta function, $B(p, q) = \Gamma(p)\Gamma(q)/\Gamma(p + q)$
BP	Boundary Point
$\lceil x \rceil$	Smallest integer larger than or equal to x
χ^2_ν	Chi-square distribution with ν degrees of freedom
$\complement M$	Complement of the set M
CP	Crossing Point
cdf.	Cumulative distribution function
$\delta_{i,j}$	Kronecker symbol
ecdf.	Empirical cumulative distribution function
ε_a	Dirac measure in point a
$=^d$	Equality in distribution
EER	Expected Error Rate
F_X	Cumulative distribution function of a real-valued random variable X
FDR	False Discovery Rate
FWER	Family Wise Error Rate
$\lfloor x \rfloor$	Largest integer lower than or equal to x
$\Gamma(\cdot)$	Gamma function, $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt,\ x > 0$
$\mathrm{im}(X)$	Image of the random entity X

i.i.d.	independent and identically distributed
$\mathbf{1}_M$	Indicator function of set M
LCP	Largest Crossing Point
$\mathcal{L}(X)$	Law of distribution of random variable X
LFC	Least Favorable Configuration
$\lambda\!\!\lambda$	Lebesgue measure
MTP_2	Multivariate total positivity of order 2
$\mathbb{N}_n$	$\{1, \ldots, n\}$
$\mathcal{N}(\mu, \sigma^2)$	Normal distribution with parameters μ and σ^2
Φ	Cumulative distribution function of the $\mathcal{N}(0,1)$ distribution
$\varphi(\cdot)$	Probability density function of the $\mathcal{N}(0,1)$ distribution
PRD	Positive regression dependency
PRDS	Positive regression dependency on subsets
pdf.	Probability density function
pmf.	Probability mass function
SD	Step-down
SU	Step-up
SUD	Step-up-down
UNI$[a, b]$	Uniform distribution on the interval $[a, b]$

Overview

The False Discovery Rate (FDR) is a rather young paradigm in controlling errors of a multiple test procedure. Especially in the context of genetics and microarray analyses, the FDR has become a very popular error control criterion over the last decade, because it is less restrictive than the classical Family Wise Error Rate (FWER). This is especially important since in several of today's application fields like genome-wide association (GWA) studies, sometimes ten thousands or even some hundred thousands of hypotheses have to be tested simultaneously and the analyses (at least at a first stage) have mainly explorative character so that in this stage of the analysis one is often more interested in getting some significances than in avoiding a few false ones. Instead of controlling the probability of making at least one false rejection, the FDR controls the *expected proportion* of falsely rejected (true) null hypotheses among all rejections. Due to the massive multiplicity of some of the current applications, asymptotic considerations become more and more relevant. Therefore, in this work special focus will be laid on the asymptotic behaviour of the False Discovery Rate with the number n of hypotheses tending to infinity. Other applications include astronomy (cf., e. g., [176]) and proteomics, cf. Application 2.4.

The remainder of this work is organized as follows. In Chapter 1, some theoretical foundations will be presented, including a formal definition of the FDR. Most of the results in that chapter are already known so that it has a repetitious character. Furthermore, some notational aspects are covered.

Chapter 2 then deals with a popular FDR controlling multiple test procedure, namely the linear step-up procedure based on Simes' critical values introduced in the pioneering article by Benjamini and Hochberg from 1995, see [13]. Since it is well known that this method controls the FDR for positively dependent test statistics being at hand, we study its asymptotic conservativeness in some special distributional situations.

In Chapter 3 we present and investigate a new rejection curve designed to asymptotically exhaust the whole FDR level α under some extreme parameter configurations.

Besides these theoretical considerations, we will apply some of the test procedures presented in Chapters 2 and 3 to real life data and investigate FDR "at work".

Chapter 4 contains a systematic (numerical) comparison of some recently developed test procedures which aim at improving the linear step-up procedure. Under various distributional settings, we investigate their behaviour with respect to type I error and power. This allows us to discuss assets and drawbacks of each of the considered procedures.

In Chapter 5, finally, our results will be summarized and we give an outlook on some pursuing issues.

Some numerical computations and computer simulations referring to the theoretical results in Chapters 2 and 3 are presented in the Appendix. Moreover, we briefly discuss some notions of positive dependency there.

The research that has lead to this work has been part of the first period of a research project sponsored by the Deutsche Forschungsgemeinschaft (DFG), grant No. FI 524/3-1, under the responsibility of my advisor Helmut Finner and of Prof. Guido Giani. In the application to this grant, the aims of Chapters 2 and 3 have already been formulated and parts of the elaborations in these chapters are joint work with Helmut Finner and Markus Roters as well. Main results of Chapter 2 are pre-published in [86] and [88]. An article containing the main results of Chapter 3 has been accepted for publication, see [87]. I am grateful to the DFG for financing my tenure at the German Diabetes Center from July 2005 to April 2007 and to Helmut Finner for providing me with the interesting topics and for some valuable preliminary notes from his treasure chest.

Chapter 1

Introduction

1.1 Multiple testing and False Discovery Rate

The goal of multiple testing consists of testing $n > 1$ hypotheses simultaneously and controlling some kind of overall error rate. The most conservative and highly intuitive method is controlling the Family Wise Error Rate (FWER) in the strong sense. The Family Wise Error is defined as the event that at least one false rejection among the n individual tests is performed and the FWER (in the strong sense) for a multiple test procedure $\varphi = (\varphi_1, \ldots, \varphi_n)$ is the probability for the latter event and it can therefore loosely by defined as

$$\text{FWER}_n(\varphi) = \mathbb{P}\left(\exists\, 1 \leq i \leq n : \{\varphi_i = 1 \text{ and } H_i \text{ is true}\}\right). \tag{1.1}$$

There also exists a definition of the FWER in the *weak sense* aiming at error control under the global hypothesis that all n null hypotheses are true. However, we only consider the FWER in the strong sense here. A rather simple and naive method for controlling the FWER is the Bonferroni procedure, where each individual test φ_i is carried out at level $\alpha_i = \alpha/n$. Due to subadditivity, we immediately get the FWER-controlling property of the Bonferroni method, because of

$$\text{FWER}_n(\varphi) \leq \sum_{k=1}^{n} \alpha_k,$$

with α_k denoting the individual level for φ_k. The disadvantage of the Bonferroni method is that these individual levels become extremely small for a large number of hypotheses n at hand which results in a very low power of the Bonferroni method for large n. Therefore, many improvements of the Bonferroni method have been developed. The maybe most advanced method towards constructing a multiple level α-test consists in the so-called *partitioning principle* developed by Finner and Straßburger, see [94].

It shall be mentioned here that a multiple test procedure $\varphi = (\varphi_1, \ldots, \varphi_n)$ which controls the

FWER at a pre-specified level α can also be used to perform a level α-test for the global intersection hypothesis $H_0 = \bigcap_{i=1}^n H_i$ (assuming that H_0 is not empty). We simply reject H_0 iff there exists an index $1 \leq k \leq n$ with $\varphi_k = 1$. The type I error controlling property of this test method is immediate if we keep in mind that φ has the property that the right-hand side of (1.1) is bounded by α. If the test φ is constructed according to the Bonferroni method, the corresponding intersection hypothesis test ψ (say) simply becomes $\psi = \mathbf{1}_{\{p_{1:n} \leq \alpha/n\}}$, where $p_{1:n}$ denotes the smallest p-value, cf. Section 1.2. One improvement with respect to power has been developed by Simes, cf. [264], for independent p-values. We mention it here because its critical values will be used in a different context later. Simes' method is described in Algorithm 2.1 at the beginning of Chapter 2.

A more radical approach towards gaining of power in a multiple testing problem is relaxation of the underlying error measure. Especially for large values of n, controlling the FWER may be a much too conservative goal, especially if we consider a screening experiment where it is more important to get some significances than to avoid a few false ones. A more liberal and nowadays widely used error measure in the latter situation is the False Discovery Rate (FDR). In contrast to the FWER, not the probability of performing at least one false rejection is controlled, but the *expected proportion* of falsely rejected hypotheses with regard to all rejected hypotheses. In order to formalize this task, we need some notation.

Definition 1.1
Let $(\Omega, \mathcal{A}, \{\mathbb{P}_\vartheta : \vartheta \in \Theta\})$ denote a statistical experiment and $\mathbb{N}_n = \{1, \ldots, n\} \subset \mathbb{N}$. Let $\varphi = (\varphi_1, \ldots, \varphi_n)$ be a multiple test procedure for the family $(H_1, \ldots, H_n)$ of non-empty hypotheses with $H_i \subset \Theta$ for all $i \in \mathbb{N}_n$. A hypothesis $H_k, k \in \mathbb{N}_n$, is called true if $\vartheta \in H_k$ and false otherwise. Then we define

$$R_n(\varphi) = |\{i \in \mathbb{N}_n : \varphi_i = 1\}|, \tag{1.2}$$

$$V_n(\varphi) = |\{i \in \mathbb{N}_n : \varphi_i = 1 \text{ and } H_i \text{ is true}\}|, \tag{1.3}$$

$$FDR_n(\varphi) = \mathbb{E}_\vartheta \left(\frac{V_n(\varphi)}{R_n(\varphi) \vee 1} \right), \tag{1.4}$$

and say that φ controls the FDR at a pre-chosen level of significance $\alpha \in (0,1)$ iff

$$\sup_{\vartheta \in \Theta} FDR_n(\varphi) \leq \alpha.$$

The ratio $V_n(\varphi)/[R_n(\varphi) \vee 1]$ is called the false discovery proportion (FDP).

If it is clear which procedure φ is investigated, the argument φ is often dropped and we simply write $V_n = V_n(\varphi)$ and $R_n = R_n(\varphi)$. The meaning of the quantities V_n and R_n is illustrated in the following table.

	Test decision		
Hypothesis	0	1	
true	U_n	V_n	n_0
false	T_n	S_n	n_1
	$n - R_n$	R_n	n

Table 1.1: Quantities in a multiple test procedure

It is important to notice that V_n is unobservable, because it contains information about the validity of the hypotheses in the family. Another measurement of type I errors is the *expected error rate* (EER), which will be considered in Chapter 2 together with the FDR. It is sometimes also called *per comparison error rate* (PCER) and is defined as the expected proportion of type I errors with regard to the size of the family of hypotheses, as formalized in the following definition.

Definition 1.2 (Expected Error Rate)
For given $\vartheta \in \Theta$, we define the expected error rate (EER) of a multiple test procedure φ by

$$EER_n(\varphi) = \mathbb{E}_\vartheta \left[\frac{V_n(\varphi)}{n} \right].$$

Moreover, in a multiple test problem type I errors and type II errors (counted by T_n in the nomenclature introduced in Table 1.1) can occur simultaneously. Subject to type I error rate control (measured by the FWER or the FDR, for example), it may also be worthy to introduce a measurement for *multiple power* in order to compare different (e. g., FDR-controlling) test procedures with each other. In the literature, there is no common agreement on how to define multiple power. One possibility that is frequently encountered in the literature (cf., e. g., [43]) is to define the power of a multiple test φ as the expected proportion of false hypotheses that can be rejected with φ, formally expressed in the following definition.

Definition 1.3 (Multiple power)
For given $\vartheta \in \Theta$, we define the quantity

$$\bar{\beta}_\vartheta(\varphi) = \mathbb{E}_\vartheta \left[\frac{S_n(\varphi)}{n_1 \vee 1} \right]$$

as power of a multiple test procedure φ.

We will use this power definition throughout this work, although we are aware of the fact that it has weaknesses. Especially, Definition 1.3 does not take into account *how false* the rejected false hypotheses are, i.e., how far the corresponding alternative parameters are away from the null parameters. Establishing a power definition based on the latter considerations would offer the possibility to measure the "statistical resolution" of φ which is interesting as well.

The following two short remarks establish a well known connection between the FDR and the FWER.

Remark 1.4
Definition (1.4) can equivalently be expressed as

$$\mathrm{FDR}_n(\varphi) = \mathbb{E}_\vartheta\left(\frac{V_n(\varphi)}{R_n(\varphi)} \mid R_n(\varphi) > 0\right) \cdot \mathbb{P}_\vartheta(R_n(\varphi) > 0).$$

Remark 1.5
If all hypotheses are true, it holds:

$$\mathrm{FDR}_n(\varphi) = \mathbb{P}_\vartheta(R_n(\varphi) > 0) = \mathbb{P}_\vartheta(V_n(\varphi) > 0) = \mathrm{FWER}_n(\varphi).$$

In general, we have that $V_n(\varphi)/[R_n(\varphi) \vee 1] \leq \mathbf{1}_{\{V_n(\varphi)>0\}}$ and therefore it holds

$$\mathrm{FDR}_n(\varphi) = \mathbb{E}_\vartheta\left(\frac{V_n(\varphi)}{R_n(\varphi) \vee 1}\right) \leq \mathbb{E}_\vartheta\left(\mathbf{1}_{\{V_n(\varphi)>0\}}\right) = \mathrm{FWER}_n(\varphi).$$

Starting in 1995 with the famous article by Benjamini and Hochberg ([13]), over the last decade a variety of FDR-controlling procedures has been developed, although the general idea for applying this error measure is older. Since the defining equation (1.4) is a highly complicated construct, proofs of FDR-control of a certain multiple test procedure often are technically cumbersome and partly need strong assumptions about the dependency structure of the underlying test statistics. In the next chapter, we focus on the well known and widely spread linear step-up procedure φ^{LSU} based on Simes' critical values originally proposed in [13] and study some its properties in detail.

Before doing so, we will finish our introductory comments with a short survey of some recent developments in the rapidly growing field of FDR-research which has become one of the leading research topics in (bio-)statistics and (bio-)informatics in the past few years. As said before, due to the rising complexity and massive multiplicity encountered especially in genomics and cosmologic applications, asymptotic FDR-considerations with the number of tests growing to infinity have attracted special attendance. Early valuable convergence investigations with respect to the number of type I errors can be found in [91] and [92] for independent test statistics. Initiated by the work of John D. Storey (cf. [271], [272], [273], [275]), procedures relying on estimation of the proportion of true null hypotheses have recently received particular attention. Although not much discussed yet, such test procedures can behave very conservatively in certain situations like in discrete models or in case of composite null hypotheses. We will explain and study this further in Section 3.8 and in Chapter 4.
Another data-adaptive approach consists of multi-stage testing. On the one hand, such a strategy can be utilized to use the number of rejections in the first step of the procedure as an estimate for the number of true null hypotheses in the following stages (cf, e. g., [15]), on the other hand, it is

possible to reduce the complexity of the test problem iteratively in order to increase the power for each remaining individual test in each step (cf. [323]). Moreover, *empirical Bayes* approaches are discussed in order to estimate the posterior probability for the validity of a specific null hypothesis given its p-value (for a definition of the p-value, see Section 1.2 below). These estimation techniques are known as *local fdr*-theory (cf., e.g., [23], [69], [72], [73], [75]). Bayesian criterions for comparing the quality of FDR-controlling test procedures can be found in [102] and [103], for example.

For dependent test statistics, only few results concerning FDR-control are available yet. Decisive for models with positive dependency (see Appendix B) assumptions were the works of S. K. Sarkar (especially [237]) and Benjamini and Yekutieli (see [17]), which proved the conservativity of φ^{LSU} under certain assumptions independently from each other. The FDR-behavior of the linear step-up procedure for certain kinds of exchangeable test statistics will be the topic of Chapter 2. If positive dependency cannot be assumed and / or if there is insufficient knowledge about the distribution of the test statistics, up to now resampling techniques are often used to simulate this distribution (cf. [67], [68], [297]) under the null hypotheses.

1.2 The concept of p-values

Instead of explicitly carrying out a particular statistical test, statistical software systems often report so-called p-values, because they do not depend on a pre-defined significance level. These p-values are sometimes referred to as *observed* levels of significance. To formalize how we understand a p-value, consider again a statistical experiment $(\Omega, \mathcal{A}, \{\mathbb{P}_\vartheta : \vartheta \in \Theta\})$ and assume we have a test φ for the pair of hypotheses H_0 versus H_1 concerning the parameter $\vartheta \in \Theta$ relying on a test statistic $T = T(X_1, \ldots, X_k)$, where the $X_i, i = 1, \ldots, k$, are i.i.d. $\mathbb{P}_\vartheta$-distributed random variables mapping Ω onto $\mathbb{R}$, representing the experiment. For given realizations $(x_1, \ldots, x_k)$ in a sample of size k, the corresponding p-value denotes the smallest level of significance $\alpha_{\text{min}}(x_1, \ldots, x_k)$, for which the null hypothesis is rejected given the actual observed data. If we denote the rejection region of φ for a given level α with Γ_α, then the p-value for the realizations $x = (x_1, \ldots, x_k)$ computes as

$$p(\varphi, x) = \inf_{\{\Gamma_\alpha:\, T(x) \in \Gamma_\alpha\}} \mathbb{P}^*(T \in \Gamma_\alpha),$$

where $\mathbb{P}^*$ is chosen such that $\mathbb{P}^*(T \in \Gamma_\alpha) = \sup_{\vartheta \in H_0} \mathbb{P}_\vartheta(T \in \Gamma_\alpha)$ if H_0 consists of more than one element. Often, a unique measure $\mathbb{P}^*$ yielding the aforementioned supremum exists; e. g., if ϑ is a location parameter and the test problem is of the structure $H_0 : \{\vartheta \leq \vartheta_0\}$ versus $H_1 : \{\vartheta > \vartheta_0\}$, we typically have $\mathbb{P}^* = \mathbb{P}_{\vartheta_0}$. If especially the test statistic T tends to larger values under the alternative and the test φ is of the form $\varphi(X_1, \ldots, X_k) = \mathbf{1}_{[c,\infty)}(T(X_1, \ldots, X_k))$, the definition

of the p-value given above simplifies to

$$p(\varphi, x) = \mathbb{P}^*(T \geq T(x_1, \ldots, x_k)).$$

Obviously, this is the probability under the null hypothesis of the event that the test statistic T has a value that is not more likely for H_0 than the value $T(x_1, \ldots, x_k)$ for the actually observed realizations and therefore indeed equals the smallest level of significance which leads to rejection of H_0 in case of having observed $(x_1, \ldots, x_k)$.

It is also possible to comprehend the p-values themselves as random entities (cf., e. g., [234]). If we again assume $\varphi(X_1, \ldots, X_k) = \mathbf{1}_{[c,\infty)}(T(X_1, \ldots, X_k))$, the tuple of data $x = (x_1, \ldots, x_k)$ in this interpretation is assigned to the probability of the test statistic T exceeding the fixed deterministic value $T(x_1, \ldots, x_k)$. The formal description

$$x = (x_1, \ldots, x_k) \to p(\varphi, x) = \mathbb{P}^*(T \geq T(x_1, \ldots, x_k)) = 1 - F_{\mathbb{P}^*}(T(x_1, \ldots, x_k))$$

together with the principle of quantile transformation yields immediately, that the random entity $p(\varphi, x)$ in this consideration is uniformly distributed on the interval $[0, 1]$ under the null hypothesis, if $\mathbb{P}^*$ is continuous and H_0 consists of only one element. This statement means reworded, that the number of rejections of a level α-test averages $\alpha \cdot 100\%$ of the performations, if always the circumstances of the null hypothesis are at hand.
The latter consideration becomes a deeper meaning in a multiple testing procedure with many (say n) hypotheses to be tested. If we have drawn samples for each individual test problem, it is possible to compute the corresponding p-values and their ecdf. $F_n(\cdot)$ (say). If the graph of F_n then significantly deviates from the bisecting line on the unit interval, this can give information about how many of the n hypotheses are wrong. Moreover, many multiple test procedures are defined in terms of p-values.
In case that $\mathbb{P}^*$ is a discrete probability measure or H_0 is a composite null hypothesis, we obtain that $p(\varphi, x)$ under H_0 is stochastically larger than a UNI$[0, 1]$-distributed random variable.

Remark 1.6 (Multiple test procedures in terms of p-values)
Consider a multiple test procedure φ consisting of n one-sided tests $\varphi_i(T_i) = \mathbf{1}_{[c_i,\infty)}(T_i)$ for $i = 1, \ldots, n$ with test statistics $T_i, i = 1, \ldots, n$ which are i.i.d. with continuous cdf. F_T under the null hypotheses. Then φ can equivalently be expressed by n one-sided tests $\varphi_i'(P_i) = \mathbf{1}_{(0,\alpha_i]}(P_i)$ for $i = 1, \ldots, n$ in terms of the corresponding p-values P_i (regarded as random entities with cdf. F_P and therefore written with capital letter here) and critical values $\alpha_1, \ldots, \alpha_n$, iff the following condition holds:

$$\mathbb{P}(T_1 \geq c_i) = 1 - F_T(c_i) = F_P(\alpha_i) = \mathbb{P}(P_1 \leq \alpha_i) \text{ for all } i = 1, \ldots, n. \tag{1.5}$$

Therefore, the modeling of the underlying distributional situation can be done either by F_T or by F_P. Since many multiple test procedures are defined in terms of p-values, it is sometimes useful

not to model the distribution of the test statistics, but the distribution of the p-values. Moreover, expressing the test problem in terms of p-values has the advantage that this makes it independent of the scale of the original test statistics since p-values always have support $[0, 1]$.

1.2.1 p-value adjustment for multiplicity

In a multiple testing context, it may be considered as appropriate not only to report an observed level of significance for each particular hypothesis separately without taking the multiplicity into account, but to provide a per-hypothesis observed significance level with regard to a multiple error measure. This can be done by finding a suitable *adjustment* to apply to the original (sometimes called raw) p-values in that way that the i-th adjusted p-value $p_i^{(\text{adj})}$ has the property that for a given overall level of significance α it holds

$$p_i^{(\text{adj})} < \alpha \text{ is equivalent to } i\text{-th hypothesis can be rejected}$$

while keeping an underlying overall error rate. For example, the Bonferroni-adjusted p-value $p_i^{(\text{Bonf.-adj.})}$ (say) for hypothesis H_i is simply given by $p_i^{(\text{Bonf.-adj.})} = n \cdot p_i$, where p_i denotes the i-th raw p-value. If then $p_i^{(\text{Bonf.-adj.})} < \alpha$, hypothesis H_i can be rejected while keeping the FWER. The dual problem consists of finding adjusted critical values for the underlying test statistics.
In the quite popular article by Ge, Dudoit and Speed [100], a nice overview of multiple testing concepts, various error rates and corresponding p-value adjustments is given. In Section 2.2, the authors present various error rates and in Section 2.3, the defining equations for the corresponding adjusted p-values are given. Sections 3 and 4 then especially deal with several FWER controlling procedures and resulting p-value adjustments. In Section 5, finally, the same investigations are done for the FDR as underlying error measure.
In our work, we will not further consider this technique of p-value adjustment. We describe multiple test procedures controlling the FDR or the FWER, respectively, by suitable critical values for the raw p-values.

Chapter 2

FDR control with Simes' critical values

The first article dealing systematically with the FDR has been published in 1995 by Benjamini and Hochberg, see [13]. The authors give some motivation, the formal definition of the FDR reported in Definition 1.1 and present a multiple test procedure designed to control the FDR in case that the test statistics are independent under the null hypotheses. This procedure, described in Algorithm 2.2 below, employs critical values for the ordered p-values which were originally introduced in a different context. More specifically, in 1986, R. J. Simes proposed the following test algorithm for the global intersection hypothesis H_0:

Algorithm 2.1 (Simes' test for the intersection hypothesis $H_0 = \bigcap_{i=1}^n H_i$)

1. Compute the p-values $p_1, \ldots, p_n$ for each individual test.
2. Denote the ordered p-values by $p_{1:n} \leq \ldots \leq p_{n:n}$.
3. Reject H_0 if there exists an index $1 \leq k \leq n$, such that $p_{k:n} \leq \alpha_k = k\alpha/n$.

For the remainder of this work, we will confer to the critical values $\alpha_k = k\alpha/n$ for the p-values used in this algorithm as *Simes' critical values*. Simes (1986) proved that his Algorithm 2.1 controls the type I error with respect to the global hypothesis H_0 at level α if the underlying test statistics (and, consequently, the corresponding p-values) are i.i.d. He furthermore conjectured that this property is preserved for positively correlated test statistics. This conjecture was proven by S. K. Sarkar in 1998, see [236].

Benjamini and Hochberg employed Simes' critical values in the context of FDR control. They developed the linear step-up test procedure φ^{LSU} which works as follows.

Algorithm 2.2 (The linear step-up test procedure φ^{LSU})

1. Compute the p-values $p_1, \ldots, p_n$ for each individual test.
2. Denote the ordered p-values by $p_{1:n} \leq \ldots \leq p_{n:n}$.

3. Determine $k = \max\{i : p_{i:n} \leq \alpha_i\}$.

4. If such a k exists, reject the hypotheses $H_{1:n}, \ldots, H_{k:n}$ corresponding to $p_{1:n}, \ldots, p_{k:n}$. Otherwise, reject no hypotheses.

In [13], the authors proved that, assuming that n_0 hypotheses are true and the other $n-n_0$ hypotheses are false, φ^{LSU} controls the FDR at level $n_0\alpha/n \leq \alpha$ in case of independence of the vector of p-values corresponding to true null hypotheses from the vector of p-values corresponding to false null hypotheses and i.i.d. UNI$[0,1]$-distributed p-values under the n_0 true null hypotheses. Later investigations even revealed that

$$\text{FDR}_n = \frac{n_0}{n}\alpha \ \ \forall n > 1, \alpha \in (0,1),$$

i.e., the so-called *Benjamini-Hochberg bound* $n_0\alpha/n$ for the FDR is exactly obtained for any size of the family of hypotheses under the aforementioned assumptions. Different proofs of this fact can be found in [91], [237], [275] or [17]. In [17], the FDR controlling property of φ^{LSU} was extended to the case of PRDS test statistics (cf. Definition B.7).
In the following, we are interested in the asymptotic sharpness of the Benjamini-Hochberg bound in the latter situation. We investigate some examples of multivariate PRDS distributions and study the FDR behaviour of φ^{LSU} for n tending to infinity. First, we present a general theoretical framework, the Dirac-exchangeable setup, and then apply the resulting formulas to some concrete distributional examples. For this reason, a slight re-formulation of Algorithm 2.2 in terms of the ecdf. of the p-values given in the following remark will be helpful.

Remark 2.3
Algorithm 2.2 can equivalently be expressed as

1. Compute the p-values $p_1, \ldots, p_n$ for each individual test.

2. Let F_n denote the ecdf. of the p-values, that is,

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{[0,t]}(p_i), t \in [0,1].$$

3. Compute $t^* = \sup\{t \in [0,\alpha] : \ F_n(t) \geq t/\alpha\}$.

4. Reject all H_i with corresponding $p_i \leq t^*$.

We will call t^* the *largest crossing point* and denote the function $t \to t/\alpha$ for $t \in [0,\alpha]$ by *Simes' line*, the continuous version of the set of Simes' critical values.

This type of connection between critical values of a multiple test procedure and the ecdf. of p-values is indicated in [250]. The aforementioned algorithm can be carried out in practice by

drawing F_n and Simes' line together in one graph and determining t^*. The following figure shows an example with $\alpha = 0.05$.

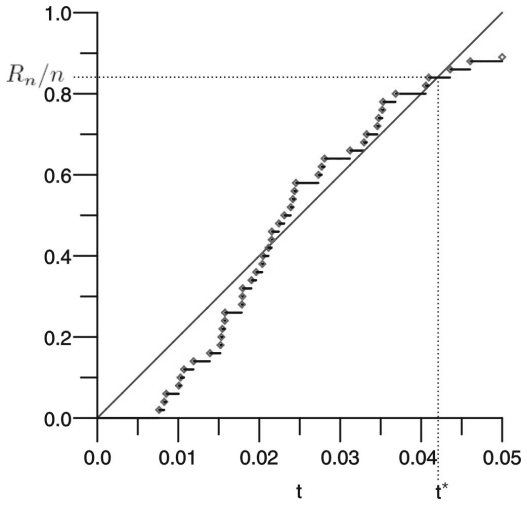

Figure 2.1: $F_{50}(t)$ and Simes'line on $[0, \alpha]$.

The abscissa t^* of the largest crossing point determines the threshold for the p-values and the value R_n/n of its ordinate reflects the proportion of hypotheses that are rejected by the linear step-up procedure. The simplicity and intuitivity of this algorithm has lead to a wide spread of φ^{LSU} over the last decade. How it works in practice shall be demonstrated with two real life applications. Of course, it has to be conceded that the dependency assumptions guaranteeing FDR control of the linear step-up procedure are difficult to verify in practical applications. Especially in quantitative trait analyses (e.g., when processing gene expression data), often negative dependency of some kind is likely to occur. For example, overexpression at one gene locus can lead to underexpression at another, linked locus. Therefore, it seems possible that φ^{LSU} does not control the FDR strictly in the following application examples. However, in both cases we investigate situations where our statistical analyses can be viewed as screening instruments at a first stage of an analysis which includes more stringent error control methods at later stages.

Application 2.4 (Evaluation of a proteomics experiment)
In a proteomics experiment carried out in the biochemical department of our research institute, 1330 protein spots from two groups A and B were detected and matched by a spot detection software. The protein material consisted of pooled tissue from two different mice stems under investigation in a diabetes-specific context. Tissue differences with regard to different spot intensities in the two groups should be found out. Group A was processed on four independent sheets and group B was processed on three independent sheets (the fourth sheet for group B was defect). In some data cleaning and preparation steps, we filtered out only spots with a minimal measurement number of three per group, i.e., all measurements for group B had to be successful and there

was only one missing measurement allowed in group A. Furthermore, intensities below 0.5 were excluded because of lacking courtesy and relevance. In such quantitative trait analyses, often a log-normal distribution for the intensity ratios is assumed. Therefore, the remaining intensities were transformed by applying the natural logarithm. After these steps, $n = 393$ spots remained. After some diagnostic plots, it turned out that the normal distribution assumption for the remaining log-intensities was justified and therefore, we carried out two-sided two-sample t-tests for the logarithmic intensity differences per spot and collected the corresponding p-values. This resulted in a multiple testing problem of dimension $n = 393$. As significance levels for the FDR, we chose $\alpha_1 = 0.05$ and $\alpha_2 = 0.1$.

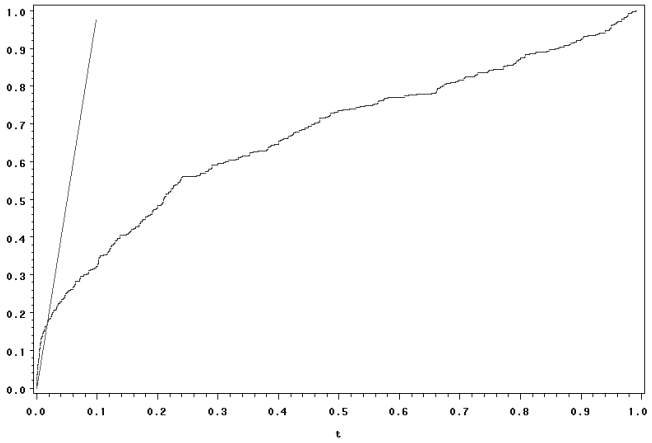

Figure 2.2: Proteomics example: Simes' line and ecdf. of 393 p-values

Figure 2.2 shows the case $\alpha_2 = 0.1$. Simes' line and the ecdf. of the obtained p-values are displayed. Obviously, we have a unique crossing point of the two objects on $(0, \alpha)$ and with the Benjamini-Hochberg procedure we got 47 rejections in case of $\alpha_1 = 0.05$ and 64 rejections in case of $\alpha_2 = 0.1$.
A discussion with the head of the proteomics department showed a good accordance of our "detected" spots with the ones found by a commercial proteomics analysis software and with the spots that were identified by experts in the department.

Application 2.5 (Adenocarcinoma data Notterman et al.)
Our second application concerns a data set taken over from the literature. In the article [203] from 2001, Notterman et al. published data from a cancer research project. The aim was detecting differentially expressed gene and R(D)NA profiles in tumor tissue in comparison with normal (healthy) tissue. To this end, a case-control study was carried out. In group A, there were 18 adenocarcinomic cancer patients and group B consisted of 18 (paired) healthy patients. From these 36 individuals, expression data for 7457 different RNA, DNA and gene entities was collected. The

complete data is available as supplementary material to [203].
After some Affymetrix preprocessing (cf. the "Materials and Methods" section in [203]), the comparison between the two groups was performed by applying t-tests to the log-transformed data. This lead to $n = 7457$ p-values. Again, we analysed this multiple testing problem utilizing φ^{LSU} at FDR level $\alpha = 0.1$. Figure 2.3 illustrates Simes' line with parameter $\alpha = 0.1$ and the ecdf. of the obtained p-values.

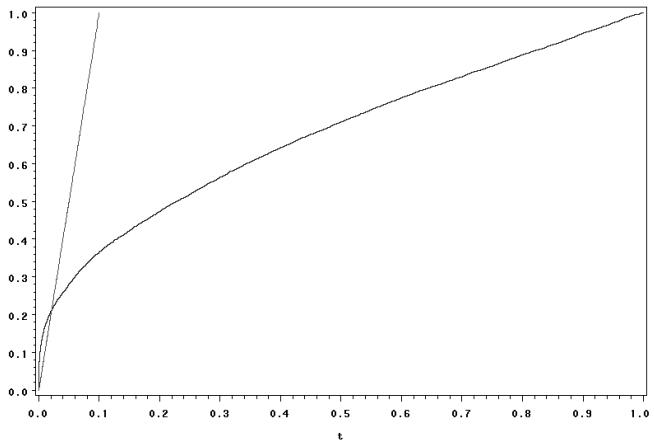

Figure 2.3: Notterman example: Simes' line and ecdf. of 7457 p-values

Again, the concave shape of F_{7457} leads to a unique crossing point on $(0, \alpha)$. With the Benjamini-Hochberg procedure, the hypotheses corresponding to the 1582 smallest p-values get rejected with a thresholding value of 0.0212.

Before we start our main theoretical investigations, we motivate our goal of investigating the sharpness of the Benjamini-Hochberg bound and present one (maybe surprising) example, how slight modifications of the p-value distribution can have an enormous effect on the resulting FDR behaviour.

Example 2.6
Assume that independent p-values for a multiple test procedure of family size n are not uniformly distributed on the unit interval under the null hypotheses, but their support shall be bounded by some value $b_n > 0$ on its left side, i.e., $P_i \sim \text{UNI}[b_n, 1]$ if H_i is true. Then, clearly, $\mathbb{P}(V_n = j) = 0$ for all $0 < j < \left\lceil \frac{nb_n}{\alpha} \right\rceil$ for the linear step-up procedure. We will show that this has a large impact on the FDR behaviour of φ^{LSU}. Noting that

$$\left\{ P_i \leq \frac{k\alpha}{n} \right\} = \left\{ \frac{P_i - b_n}{1 - b_n} \leq \frac{k\alpha - nb_n}{n(1 - b_n)} \right\} = \left\{ U \leq \frac{k\alpha - nb_n}{n(1 - b_n)} \right\},$$

where U denotes a UNI$[0, 1]$-distributed random variable for such k with $k\alpha/n \geq b_n$ and 0 for

smaller values of k, we obtain the transformed critical values

$$\tilde{\alpha}_{k:n} = \frac{k\alpha - nb_n}{n(1-b_n)}$$

corresponding to UNI$[0,1]$-distributed, transformed random p-values for $k \geq \lceil \frac{nb_n}{\alpha} \rceil$. In other words, this test problem can equivalently be regarded as one with UNI$[0,1]$-distributed p-values under the null hypotheses which have to be compared with the $\tilde{\alpha}_{k:n}$'s in step-up manner. For the sake of simplicity, we only treat the case $n_0 = n$ and note that

$$\mathbb{P}(V_n > 0) = \sum_{j=\lceil \frac{nb_n}{\alpha} \rceil}^{n} \mathbb{P}(V_n = j).$$

In [92], the exact distribution of V_n in case of φ^{LSU} and UNI$[0,1]$-distributed p-values is given as

$$\mathbb{P}(V_n = j) = \binom{n}{j}(1-\beta)(1-\beta+(n-j)\tau)^{n-j-1}(\beta-(n-j)\tau)^j,$$

if the critical values for the $P_{k:n}$ are of the structure $\gamma_{k:n} = \beta - (n-k)\tau$. For our $\tilde{\alpha}_{k:n}$'s from above we have in this nomenclature

$$\beta = \frac{\alpha - b_n}{1-b_n} \text{ and } \tau = \frac{\alpha}{n(1-b_n)}.$$

The authors also derived the limiting distribution of V_n with n tending to infinity, which is exponentially decreasing and has most of its distributional mass in the small outcomes. Choosing $b_n = 1/n$ has the effect that $V_n \in \{0\} \cup \{\lceil 1/\alpha \rceil, \ldots, n\}$ almost surely for every $n > 1$ and results in very small values for the probability $\mathbb{P}(V_n > 0)$, even for very large n and in the limiting case. This is due to of the limiting distribution properties stated before, because small outcomes for V_n with large mass are almost surely not realized. This effect is rather surprising, since b_n tends to 0 and therefore the distribution of the p-values tends to UNI$[0,1]$ in this situation and one should assume that the FDR should tend to its bound α for large n. This, however, is not the case.

As we will see, the fractional structure of the FDP leads to some more surprising results. Our first major goal consists of computation of the FDR of φ^{LSU} under positive dependence or, more precisely, exchangeability under the null hypotheses. This will be done quite generally in the following Section 2.1 before we investigate specific distributional settings in Sections 2.2 (exponential distributions), 2.3 (normal distributions) and 2.4 (t-distributions).

2.1 General theoretical framework in the exchangeable setup

In this section, we present our basic statistical model with exchangeable test statistics. It will be the basis for the concrete applications carried out in the following sections.
Let therefore $X_i,\ i = 1, \ldots, n$, be real-valued independent random variables with support $\mathcal{X}$. Moreover, let Z be a further real valued random variable, independent of the X_i's, with support $\mathcal{Z}$

whose cdf. will be denoted denoted by W_Z. Denote the cdf. of X_i by W_i. Suppose the cdf. W_i depends on a parameter $\vartheta_i \in [\vartheta_0, \infty)$, where ϑ_0 is known. Without loss of generality it will be assumed that $\vartheta_0 = 0$. Consider the multiple testing problem

$$H_i : \vartheta_i = 0 \quad \text{versus } K_i : \vartheta_i > 0,\ i = 1, \ldots, n.$$

Suppose that $T_i = g(X_i, Z)$ (with support $\mathcal{T}$) is a suitable real-valued test statistic for testing H_i, that is, it will be assumed that T_i tends to larger values if ϑ_i increases. The sets $\mathcal{X}, \mathcal{Z}$ and $\mathcal{T}$ are assumed to be intervals. Suppose that g is continuous, strictly increasing in the first argument and strictly monotone or constant in the second argument.
Examples which will play a role in the remainder are $T_i = g(X_i, Z) = X_i - Z$ (Sections 2.2 and Section 2.3) and $T_i = g(X_i, Z) = X_i/Z$ (Section 2.4).
In case that H_i is true, the cdf. of X_i will be denoted by W_X and the cdf. of T_i will be denoted by W_T. For $Z = z$, we define p-values $p_i = p_i(z)$ as a function of z by

$$p_i(z) = 1 - W_T(g(x_i, z)),\ i = 1, \ldots, n.$$

The ecdf. of these p-values is denoted by $F_n(\cdot|z)$. Clearly, the Glivenko-Cantelli lemma applies. The ordered p-values $p_{1:n} \leq \cdots \leq p_{n:n}$ are given by $p_{i:n} = 1 - W_T(g(x_{n-i+1:n}, z)), i = 1, \ldots, n$.

2.1.1 Two models with exchangeable test statistics

Assuming that all hypotheses are true, the limiting ecdf. of the p-values will be denoted by $F_\infty(\cdot|z)$. For the sake of simplicity it will be assumed that the model implies that $F_\infty(x|z)$ is continuous in $x \in [0, 1]$ and differentiable from the right at $x = 0$ with $F_\infty(0|z) = 0$ for all $z \in \mathcal{Z}$. Finally, let $g_1 : \mathcal{T} \times \mathcal{Z} \to \mathcal{X}$ be such that for all $x \in \mathcal{X}, z \in \mathcal{Z}, w \in \mathcal{T}$

$$g(x, z) = w \text{ iff } x = g_1(w, z),$$

and let $g_2 : \mathcal{X} \times \mathcal{T} \to \mathcal{Z}$ be such that for all $x \in \mathcal{X}, z \in \mathcal{Z}, w \in \mathcal{T}$

$$g(x, z) = w \text{ iff } z = g_2(x, w).$$

We refer to this setup as the *EX(1) model.* In practical examples, the conditions concerning the functions g, g_1 and g_2 need to be fulfilled only for arguments belonging to sets of measure 1, for example, if $g(x, z) = x/z$. The following lemma provides a formula for computing the limiting ecdf. F_∞ and the defining equation for a point of intersection with Simes' line in this model.

Lemma 2.7
Given EX(1)*, it holds for* $\mathbb{P}^Z$*-almost all* $z \in \mathcal{Z}$

$$F_\infty(t|z) \quad = \quad 1 - W_X(g_1(W_T^{-1}(1 - t), z)),\ t \in (0, 1).$$

Moreover, the limiting empirical cdf crosses (or contacts) the Simes line, that is, $F_\infty(t|z) = t/\alpha$ *for some* $t \in (0, \alpha)$*, if*

$$W_X^{-1}(1 - t/\alpha) = g_1(W_T^{-1}(1 - t), z),$$

or equivalently,

$$z = z(t) = g_2(W_X^{-1}(1 - t/\alpha), W_T^{-1}(1 - t)).$$

Proof: We make use of the representation

$$F_n(t|z) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{[0,t]}(p_k(z)), t \in [0, 1].$$

This representation implies that $\lim_{n\to\infty} F_n(t|z) = \mathbb{P}_{H_k}(p_k(z) \le t) = 1 - \mathbb{P}_{H_k}(W_T(g(X_i, z)) < 1 - t) = 1 - \mathbb{P}_{H_k}(g(X_i, z) < W_T^{-1}(1 - t)) = 1 - \mathbb{P}_{H_k}(X_i < g_1(W_T^{-1}(1 - t), z)) = 1 - W_X(g_1(W_T^{-1}(1 - t), z))$, since W_X is assumed to be continuous. Furthermore, the Glivenko-Cantelli lemma guarantees that this convergence is almost surely. ■

Remark 2.8
Given $Z = z$, the p-values $p_i(z)$, $i = 1, \ldots, n$, may under $H_0 = \bigcap_{i=1}^n H_i$ be interpreted as realizations of conditionally i.i.d. random variables with common cdf. $F_\infty(\cdot|z)$.

In case that a proportion $\zeta_n = n_0/n$ of hypotheses is true and the rest is false, that is, n_0 hypotheses are true and $n_1 = n - n_0$ hypotheses are false, we make the following additional assumption in order to avoid additional limiting considerations. It will be assumed that under an alternative $K_i : \vartheta_i > 0$ the parameter value $\vartheta_i = \infty$ is possible. Moreover, for $\vartheta_i = \infty$ it will be assumed that the p-value p_i has a Dirac distribution with point mass in 0. In this case, the ecdf. of the p-values will be denoted by $F_n(\cdot|z, \zeta_n)$. We refer to this situation as the *D-EX*(ζ_n) model.

Lemma 2.9
Given D-EX(ζ_n) *with* $\lim_{n\to\infty} \zeta_n = \zeta \in (0, 1]$*, the limiting cdf of the p-values is given by*

$$F_\infty(t|z, \zeta) = (1 - \zeta) + \zeta(1 - W_X(g_1(W_T^{-1}(1 - t), z))),\ t \in (0, 1),\ z \in \mathcal{Z}.$$

Moreover, F_∞ *crosses (or contacts) the Simes line, that is,* $F_\infty(t|z, \zeta) = t/\alpha$ *for some* $t \in (\alpha(1 - \zeta), \alpha)$*, if*

$$W_X^{-1}((1 - t/\alpha)/\zeta) = g_1(W_T^{-1}(1 - t), z),$$

or equivalently,

$$z = z(t|\zeta) = g_2(W_X^{-1}((1 - t/\alpha)/\zeta), W_T^{-1}(1 - t)).$$

Note that $F_\infty(t|z) = F_\infty(t|z, 1)$.

Proof: Denote the set of indices corresponding to true hypotheses with I_0. In analogy to the EX(1) model, we notice that

$$\begin{aligned} F_n(t|z,\zeta_n) &= \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{[0,t]}(p_k(z)), t \in [0,1] \\ &= \frac{1}{n}\left(\sum_{k\in I_0} \mathbf{1}_{[0,t]}(p_k(z)) + \sum_{k\in I_0^c} \mathbf{1}_{[0,t]}(p_k(z))\right) \\ &= \frac{n_0}{n}\left(\frac{1}{n_0}\sum_{k\in I_0} \mathbf{1}_{[0,t]}(p_k(z))\right) + \frac{n_1}{n}. \end{aligned}$$

This representation (together with the assertion of Lemma 2.7) implies that $\lim_{n\to\infty} F_n(t|z,\zeta_n) = (1-\zeta) + \zeta\mathbb{P}_{H_k}(p_k(z) \le t) = (1-\zeta) + \zeta(1 - W_X(g_1(W_T^{-1}(1-t), z)))$. ■

Remark 2.10

Under the assumptions of Lemma 2.9, the Glivenko-Cantelli lemma again yields

$$\lim_{n\to\infty} \sup_{t\in[0,1]} |F_n(t|z,\zeta_n) - F_\infty(t|z,\zeta)| = 0 \text{ almost surely for all } z \in \mathcal{Z}.$$

Moreover,

$$\mathbb{E}[F_\infty(t|Z,\zeta)] = \int F_\infty(t|z,\zeta) w_Z(z) d\lambda^1(z) = 1 - \zeta + \zeta t \text{ for all } t \in [0,1].$$

2.1.2 Largest crossing points and computation of EER and FDR

In order to characterize the asymptotic behavior of the linear step-up algorithm in a D-EX(ζ_n) model, the largest crossing point of the limiting ecdf. of the conditional p-values and Simes' line is of crucial importance. For $\lim_{n\to\infty} \zeta_n = \zeta \in (0,1]$, we therefore define

$$t(z|\zeta) = \sup\{t \in [\alpha(1-\zeta), \alpha] : F_\infty(t|z,\zeta) = t/\alpha\}. \tag{2.1}$$

If there exists an $\epsilon > 0$ such that $F_\infty(t|z,\zeta) > t/\alpha$ for all $t \in [t(z|\zeta)-\epsilon, t(z|\zeta))$ and $F_\infty(t|z,\zeta) < t/\alpha$ for all $t \in (t(z|\zeta), t(z|\zeta)+\epsilon]$, then $t(z|\zeta)$ will be called the largest crossing point (LCP) of $F_\infty(\cdot|z,\zeta)$ and Simes' line. The set of LCPs will be denoted by C_ζ. Moreover, set $D_\zeta = \{z \in \mathcal{Z} : t(z|\zeta) \in C_\zeta\}$. Note that there may be some boundary points (BPs) $t(z|\zeta)$ satisfying (2.1). However, it will be assumed that $\mathbb{P}^Z(D_\zeta) = 1$. In practical examples, C_ζ is a finite union of intervals.

Obviously, for $\zeta \in (0,1)$ we always have a well-defined LCP or BP $t(z|\zeta) \ge \alpha(1-\zeta) > 0$. For $\zeta = 1$ the LCP may be 0 for a large set of z-values which makes the calculation of the limiting EER and limiting FDR much subtler.

For the remainder of this chapter, we make use of the notation

$$\begin{aligned} \text{FDR}_n(\zeta_n|z) &= \mathbb{E}[\frac{V_n}{R_n \vee 1}|Z=z], \ \text{FDR}_n(\zeta_n) = \mathbb{E}[\frac{V_n}{R_n \vee 1}] \\ \text{FDR}_\infty(\zeta|z) &= \lim_{n\to\infty} \text{FDR}_n(\zeta_n|z), \ \text{FDR}_\infty(\zeta) = \lim_{n\to\infty} \text{FDR}_n(\zeta_n), \end{aligned}$$

and the corresponding expressions for EER. Moreover, the notation $V_n(z)$, $R_n(z)$ will be used if $Z = z$ is given.

The further considerations heavily depend on an assumption about the proportion $\zeta_n = n_0/n$ of true hypotheses and its limit ζ for n tending to infinity. As we will point out, it makes a crucial difference for the FDR computation if we assume all hypotheses to be true ($\zeta_n \equiv 1$ or $\zeta = 1$) or if we have $\zeta \in (0,1)$. We therefore subdivide this section according to these two cases.

2.1.3 All LCPs greater than zero

We first consider the case $\zeta \in (0,1)$. As the following theorem and its proof point out, here the asymptotic FDR- and EER-behavior for a given $z \in \mathcal{Z}$ can directly be deduced from the LCP $t(z|\zeta)$.

Theorem 2.11

Given D-EX(ζ_n) *with* $\lim_{n\to\infty} \zeta_n = \zeta \in (0,1)$*, it holds for all* $z \in D_\zeta$

$$\lim_{n\to\infty} \frac{V_n(z)}{n} = \frac{t(z|\zeta)}{\alpha} - (1-\zeta) \text{ a. s.}, \tag{2.2}$$

$$\lim_{n\to\infty} \frac{V_n(z)}{R_n(z) \vee 1} = 1 - \frac{\alpha(1-\zeta)}{t(z|\zeta)} \text{ a. s.} \tag{2.3}$$

Proof: We will show that the proportion of rejected hypotheses $R_n(z)/n$ converges almost surely to $t(z|\zeta)/\alpha$. This then immediately implies (2.2) and (2.3).

Therefore, note that $R_n(z) = \sup\{k \in \{1,\dots,n\} : k/n \le F_n(k\alpha/n|z,\zeta_n)\}$ and that for any $n \in \mathbb{N}, z \in \mathcal{Z}$ it holds $R_n(z)/n \in [0,1]$. The latter statement implies that any subsequence $(R_{n_k}(z)/n_k)_{k\in\mathbb{N}}$ of $R_n(z)/n$ has a convergent subsequence. With a subsequence technique similar to the proof of Lemma A.2 in [91], it can now be shown that the limits of all these convergent subsequences have the same value $\eta(z)$ (say) and that $\eta(z)$ has to fulfill the defining equation $F_\infty(\eta(z)\alpha|z,\zeta) = \eta(z)$. Recalling our definition of $t(z|\zeta)$, the assertion is proven. ■

Remark 2.12

Under the assumptions of Theorem 2.11 it holds

$$\text{EER}_\infty(\zeta|z) = \mathbb{E}\left[\lim_{n\to\infty} \frac{V_n(z)}{n}\right] = \frac{t(z|\zeta)}{\alpha} - (1-\zeta), \tag{2.4}$$

$$\text{FDR}_\infty(\zeta|z) = \mathbb{E}\left[\lim_{n\to\infty} \frac{V_n(z)}{R_n(z) \vee 1}\right] = 1 - \frac{\alpha(1-\zeta)}{t(z|\zeta)}. \tag{2.5}$$

It remains to calculate $\mathrm{EER}_\infty(\zeta)$ and $\mathrm{FDR}_\infty(\zeta)$. This may be done in two ways. The first is to integrate (2.2) and (2.3) with respect to $Z = zd\mathbb{P}^Z$. In this case the main problem is the computation of $t(z)$. In general, $t(z)$ cannot be determined explicitly and, furthermore, its numerical calculation can be very cumbersome. The second possibility seems more convenient and is summarized in the following theorem.

Theorem 2.13

Under the assumptions of Theorem 2.11 suppose that $F_\infty(t|z)$ is strictly decreasing in z for $t \in (0, \alpha]$. Let $C_{\zeta,1} = \{t/\alpha - 1 + \zeta : t \in C_\zeta\}$ and $C_{\zeta,2} = \{1 - \alpha(1-\zeta)/t : t \in C_\zeta\}$. Define

$$\begin{aligned} G_{\zeta,1}(u) &= 1 - W_Z(z(\alpha(u+1-\zeta)|\zeta)) \text{ for } u \in C_{\zeta,1}, \\ G_{\zeta,2}(u) &= 1 - W_Z(z(\frac{\alpha(1-\zeta)}{1-u}|\zeta)) \text{ for } u \in C_{\zeta,2}, \end{aligned}$$

and continue these functions on $[0, \zeta]$ by linear interpolation.
Then

$$\mathrm{EER}_\infty(\zeta) = \int_{C_{\zeta,1}} u dG_{\zeta,1}(u), \tag{2.6}$$

$$\mathrm{FDR}_\infty(\zeta) = \int_{C_{\zeta,2}} u dG_{\zeta,2}(u). \tag{2.7}$$

Proof: For $\zeta \in (0,1)$ and $t \in C_\zeta$ we get from (2.2) in Theorem 2.11 and from the antitonicity of $F_\infty(t|z)$ in $z \in \mathcal{Z}$ that

$$\{z \in D_\zeta : \lim_{n\to\infty} \frac{V_n(z)}{n} > \frac{t}{\alpha} - (1-\zeta) \text{ a. s. } \} = \{z \in D_\zeta : z < z(t|\zeta)\}.$$

Therefore, the substitution $u = t/\alpha - (1-\zeta)$ yields

$$\begin{aligned} W_Z(z(\alpha(u+1-\zeta)|\zeta)) &= \mathbb{P}^Z(\{z \in D_\zeta : \lim_{n\to\infty} \frac{V_n(z)}{n} > \frac{t}{\alpha} - (1-\zeta) \text{ a. s. } \}) \\ &= 1 - G_{\zeta,1}(u) \end{aligned}$$

for all $u \in C_{\zeta,1}$. Moreover,

$$\mathbb{P}^Z\left(\{z \in D_\zeta : \lim_{n\to\infty} \frac{V_n(z)}{n} \in C_{\zeta,1} \text{ a. s. } \}\right) = 1.$$

Hence, $G_{\zeta,1}$ as defined in the theorem is the cdf. of $\lim_{n\to\infty} V_n(Z)/n$ which implies (2.6). Similarly, we obtain from (2.3) in Theorem 2.11 that

$$\lim_{n\to\infty} \frac{V_n(z)}{R_n(z) \vee 1} > 1 - \frac{\alpha(1-\zeta)}{t} \text{ a. s.} \quad \text{iff} \quad z < z(t|\zeta).$$

Therefore, a similar argumentation as before yields that $G_{\zeta,2}$ as defined in the theorem is the cdf. of $\lim_{n\to\infty} V_n(Z)/(R_n(Z) \vee 1)$, and (2.7) follows. ■

Theorem 2.13 is a key step towards computation of $\mathrm{EER}_\infty(\zeta)$ and $\mathrm{FDR}_\infty(\zeta)$ in D-EX(ζ_n) models with $\zeta_n \to \zeta \in (0,1)$. In practical examples, it remains to determine the sets $C_{\zeta,1}$ and $C_{\zeta,2}$ and to evaluate the corresponding integrals.

2.1.4 Some LCPs equal to zero

If an LCP is equal to zero, the behavior of the FDR heavily depends on the gradient in zero of the cdf. of the p-value distribution. The next lemma covers the finite case. It has mainly preparing character, but will be also be applied directly in Section 2.2.

Lemma 2.14

Let $\alpha \in (0,1)$, $0 < \gamma < 1/\alpha$, $n_0, n \in \mathbb{N}$, $n_0 \leq n$ and let $\xi_1, \ldots, \xi_{n_0}$ be i.i.d. random variables with values in $[0,1]$ with cdf. F_ξ satisfying $F_\xi(t) = \gamma t$ for all $t \in [0,\alpha]$. Furthermore, let $\xi_{n_0+1}, \ldots, \xi_n$ be random variables with values in $[0,1]$, independent of $(\xi_j : 1 \leq j \leq n_0)$. For $c_i = i\alpha/n$, $i = 1, \ldots, n$, define $R'_n = \max\{k \leq n : \xi_{k:n} \leq c_k\}$ and $V'_n = |\{i \in \{1, \ldots, n_0\} : \xi_i \leq c_{R'_n}\}|$ (with $c_{R'_n} = -\infty$ for $R'_n = -\infty$). Then

$$\mathbb{E}\left(\frac{V'_n}{R'_n \vee 1}\right) = \frac{n_0}{n}\gamma\alpha. \tag{2.8}$$

Proof: For $1 \leq i \leq n_0$, denote the $(n-1)$-dimensional random vector $(\xi_1, \ldots, \xi_{i-1}, \xi_{i+1}, \ldots, \xi_n)$ by $\xi^{(i)}$, define for $1 \leq k < n$ the sets $D_k^{(i)}(\alpha) = \{\xi^{(i)}_{k:n-1} > c_{k+1}, \ldots, \xi^{(i)}_{n-1:n-1} > c_n\}$ and set $D_0^{(i)}(\alpha) = \emptyset$, $D_n^{(i)}(\alpha) = \Omega$. Then the left hand side of (2.8) (cf., e.g., Lemma 3.2 and formula (4.4) in [237]) is equal to

$$\frac{1}{n}\sum_{i=1}^{n_0}\mathbb{P}(\xi_i \leq c_n) \; + \sum_{i=1}^{n_0}\sum_{j=2}^{n}\left[\frac{\mathbb{P}(\xi_i \leq c_{j-1})}{j-1} - \frac{\mathbb{P}(\xi_i \leq c_j)}{j}\right]\mathbb{P}(D_{j-1}^{(i)}(\alpha)).$$

Noting that $\mathbb{P}(\xi_i \leq c_n) = \gamma\alpha$ for all $1 \leq i \leq n_0$ and $\mathbb{P}(\xi_i \leq c_j)/j = \gamma\alpha/n$ for all $1 \leq j \leq n$, the assertion follows immediately. ■

The following result extends Lemma 2.14 and is a helpful tool in case that LCPs are in 0.

Lemma 2.15

Under the assumptions of Lemma 2.14, but only supposing that $F_\xi(t) = \gamma t$ for all $t \in [0,t^*]$ for some $t^* \in (0,\alpha)$, let $A_n(t^*) = \{F_n(t) < t/\alpha \; \forall t \in (t^*,\alpha]\}$, where F_n denotes the ecdf. of $\xi_1, \ldots, \xi_n$. Then, setting $r = \max\{i \in \mathbb{N}_0 : i\alpha/n \leq t^*\}$,

$$\mathbb{E}\left(\frac{V'_n}{R'_n \vee 1}\mathbf{1}_{A_n(t^*)}\right) = \frac{n_0}{n}\gamma\alpha\mathbb{P}(D_r^{(1)}(\alpha)). \tag{2.9}$$

Proof: It is clear that $A_n(t^*) = \{R'_n \leq r\}$, hence, for $r > 0$, the left-hand-side of (2.9) is now equal to

$$\frac{1}{r}\sum_{i=1}^{n_0}\mathbb{P}(\xi_i \leq c_r)\mathbb{P}(D_r^{(i)}(\alpha)) + \sum_{i=1}^{n_0}\sum_{j=2}^{r}\left[\frac{\mathbb{P}(\xi_i \leq c_{j-1})}{j-1} - \frac{\mathbb{P}(\xi_i \leq c_j)}{j}\right]\mathbb{P}(D_{j-1}^{(i)}(\alpha)).$$

The assertion follows in similarity to the proof of Lemma 2.14. ■

The next theorem is an important step for the understanding of the asymptotic behavior of both EER and FDR in D-EX-(ζ_n) models given a fixed value $Z = z$ such that the LCP is in 0.

Theorem 2.16

Given D-EX(ζ_n) *with* $\lim_{n\to\infty} \zeta_n = 1$*, let* $z \in \mathcal{Z}$ *such that* $F_\infty(t|z) < t/\alpha$ *for all* $t \in (0, \alpha]$*. Then, setting*

$$\gamma(z) = \lim_{t\to 0^+} \frac{F_\infty(t|z)}{t},$$

it holds

$$\mathit{EER}_\infty(\zeta|z) = 0, \tag{2.10}$$

$$\mathit{FDR}_\infty(\zeta|z) = \alpha\gamma(z). \tag{2.11}$$

Proof: The assumptions concerning F_∞ imply that $\lim_{n\to\infty} R_n(z)/n = 0$ almost surely. Noting that $V_n(z)/n \le R_n(z)/n$ for all $n \in \mathbb{N}$, (2.10) is obvious.
In order to prove (2.11), we nest F_∞ between two cdf's being linear in a neighborhood of zero. To this end, let $t^* \in (0, \alpha]$ be fixed, $B = [0, t^*)$, $m_\ell(t^*) = \inf_{t\in B\setminus\{0\}} F_\infty(t|z)/t$, $m_u(t^*) = \sup_{t\in B\setminus\{0\}} F_\infty(t|z)/t$, and

$$\begin{aligned} F_\ell(t) &= m_\ell(t^*)t \cdot \mathbf{1}_B(t) + F_\infty(t|z) \cdot \mathbf{1}_{B^c}(t), \\ F_u(t) &= m_u(t^*)t \cdot \mathbf{1}_B(t) + \max\{m_u(t^*)t^*, F_\infty(t|z)\} \cdot \mathbf{1}_{B^c}(t). \end{aligned}$$

This results in $F_\ell(t) \le F_\infty(t|z) \le F_u(t)$ for all $t \in [0, 1]$.
For $n \in \mathbb{N}$, let the event $A_n(t^*)$ be defined as in Lemma 2.15.
Then

$$\begin{aligned} \mathrm{FDR}_n(\zeta_n|z) &= \mathbb{E}\left(\frac{V_n(z)}{R_n(z)\vee 1}\mathbf{1}_{A_n(t^*)}\right) + \mathbb{E}\left(\frac{V_n(z)}{R_n(z)\vee 1}\mathbf{1}_{A_n^c(t^*)}\right) \\ &= \Lambda_n + \lambda_n \ \text{(say)}. \end{aligned}$$

With $r_n = \max\{i \in \mathbb{N}_0 : i\alpha/n \le t^*\}$ we obtain similarly to the argumentation in the proof of Lemma 2.14 that

$$\begin{aligned} \Lambda_n &= \mathbb{E}\left(\frac{V_n(z)}{R_n(z)\vee 1}\mathbf{1}_{\{R_n(z)\le r_n\}}\right) \\ &= n_0 \sum_{j=1}^{r_n} \frac{\mathbb{P}(p_1(z) \le \alpha_j)}{j}\left[\mathbb{P}(D_j^{(1)}(\alpha)) - \mathbb{P}(D_{j-1}^{(1)}(\alpha))\right]. \end{aligned}$$

Due to the pointwise order of F_ℓ, F_∞ and F_u, we get

$$\begin{aligned} \zeta_n m_\ell(t^*)\alpha\mathbb{P}(D_{r_n}^{(1)}(\alpha)) &\le \Lambda_n &&\le \zeta_n m_u(t^*)\alpha\mathbb{P}(D_{r_n}^{(1)}(\alpha)), \\ \zeta_n m_\ell(t^*)\alpha\mathbb{P}(D_{r_n}^{(1)}(\alpha)) + \lambda_n &\le \mathrm{FDR}_n(\zeta_n|z) &&\le \zeta_n m_u(t^*)\alpha\mathbb{P}(D_{r_n}^{(1)}(\alpha)) + \lambda_n. \end{aligned}$$

Since $\zeta_n \to 1$, $\mathbb{P}(D_{r_n}^{(1)}(\alpha)) \to 1$ and $\mathbb{P}(A_n(t^*)) \to 1$ for $n \to \infty$, we obtain $\lambda_n \to 0$ and $m_\ell(t^*)\alpha \leq \liminf_{n\to\infty} \mathrm{FDR}_n(\zeta_n|z) \leq \limsup_{n\to\infty} \mathrm{FDR}_n(\zeta_n|z) \leq m_u(t^*)\alpha$. The assertion now follows by noticing that $\lim_{t^*\to 0^+} m_\ell(t^*) = \lim_{t^*\to 0^+} m_u(t^*) = \gamma(z)$. ■

Remark 2.17

In [92], the distribution and expectation of V_n has been computed for uniform p-values under the assumption that all hypotheses are true. Assuming $\zeta_n = 1$ for all $n \in \mathbb{N}$, the nesting method described before together with the technique in [92] may be used to prove

$$\lim_{n\to\infty} \mathbb{E}[V_n(z)] = \frac{\alpha\gamma(z)}{(1-\alpha\gamma(z))^2}.$$

It is important to note that the latter formula is only valid for $\zeta_n = 1$. If in contrast n_1 tends to infinity while $\lim_{n\to\infty} n_1/n = 0$ and $\gamma(z) > 0$, we get that $\mathbb{E}[V_n(z)]$ diverges to infinity in this case even if $\gamma(z) < 1/\alpha$. To see this, we utilize the assertion of Lemma 4.3 in [91], which is

$$\mathbb{E}[V_n] \geq \frac{n_1\alpha}{\zeta_n^{-1} - \alpha}$$

for uniformly on $[0,1]$ distributed p-values. Obvious modifications lead to the assertion for $\mathbb{E}[V_n(z)]$.

In order to get a complete picture for $\zeta = 1$, the following theorem puts things together.

Theorem 2.18

Given D-EX(ζ_n) with $\lim_{n\to\infty} \zeta_n = 1$, suppose that $F_\infty(t|z)$ is strictly decreasing in z for $t \in [0,\alpha]$. Moreover, let $G_{1,1}$ be defined according to $G_{\zeta,1}$ in Theorem 2.13 and let $E_0 = \{z \in \mathcal{Z} : t_1(z) = 0\}$ and $E_1 = \mathcal{Z} \setminus E_0$. Then

$$\mathit{EER}_\infty(1) = \int_{C_1\setminus\{0\}} u dG_{1,1}(u), \tag{2.12}$$

$$\mathit{FDR}_\infty(1) = \mathbb{P}^Z(E_1) + \alpha \int_{E_0} \gamma(z) d\mathbb{P}^Z(z). \tag{2.13}$$

Proof: Using the disjoint decomposition $\mathcal{Z} = E_0 + E_1$, we obtain

$$\begin{aligned}
\mathrm{EER}_\infty(\zeta) &= \lim_{n\to\infty} \int_{\mathcal{Z}} \frac{V_n(z)}{n} d\mathbb{P}^Z(z) \\
&= \int_{E_0} \lim_{n\to\infty} \frac{V_n(z)}{n} d\mathbb{P}^Z(z) + \int_{E_1} \lim_{n\to\infty} \frac{V_n(z)}{n} d\mathbb{P}^Z(z) \\
&= A_1 + A_2 \text{ (say).}
\end{aligned}$$

Now, Theorem 2.16 immediately yields $A_1 = 0$ and in analogy to the argumentation in the proof of Theorem 2.13 we get that $A_2 = \int_{C_{1,1}\setminus\{0\}} u dG_{1,1}(u)$. Therefore, (2.12) is proven. Applying

the same decomposition (together with the considerations in Theorem 2.16) to $\mathrm{FDR}_\infty(\zeta)$ and observing that $\lim_{n\to\infty} V_n(z)/(R_n(z) \vee 1) = 1$ if $z \in E_1$ (similar to (2.3) with $\zeta = 1$) finally proves (2.13). ■

In the remaining sections of this chapter, we will apply our general results to some concrete well-known and often used distributional settings.

2.2 Exchangeable exponentially distributed variables

The exponential distribution often arises in reliability and life time analysis. For example, consider the situation that we have n (technical) systems consisting of several independent components each and we are interested in testing the reliability of these systems with respect to a reference system. In order to describe this task formally, we denote the times until failure of each individual component with $(X_{i,j})_{i=1,\ldots,n;j=1,\ldots,m(i)}$ and assume that the $X_{i,j}$'s are independent random variables, each underlying an exponential distribution with parameter $\lambda_{i,j}$. If we then define the reliability of the i-th entire system, denoted by Y_i, by the minimum failure time of all components belonging to the i-th system, we obtain that

$$Y_i \sim \mathrm{Exp}\left(\lambda_i =: \sum_{j=1}^{m(i)} \lambda_{i,j}\right).$$

Furthermore, we denote the minimum time until failure of the components of the reference system with Y_0 and assume that $Y_0 \sim \mathrm{Exp}(\lambda_0)$, independent of the Y_i's. Consequently, we get that

$$\mathbb{E}(Y_i) = \frac{1}{\lambda_i} \quad \text{mean expected survival time of system } i \text{ and}$$

$$\mathbb{E}(Y_0) = \frac{1}{\lambda_0} \quad \text{mean expected survival time of the reference system.}$$

A well known and often arising multiple test problem is now given by

$$H_i = \{\frac{1}{\lambda_i} = \frac{1}{\lambda_0}\} \text{ vs. } K_i = \{\frac{1}{\lambda_i} > \frac{1}{\lambda_0}\} \text{ for } i = 1, \ldots, n.$$

A parametric approach towards this test problem consists of applying Cox's F-test, cf. e. g. [166], pp. 236-237. However, non-parametric techniques are more commonly used in this setting.
In order to fit in our general setup, we investigate a slightly different test problem, involving the two-parametric exponential distribution. Under the general framework given in Section 2.1, we assume that the X_i follow an exponential distribution with scale parameter λ and location parameter ϑ_i and Z is exponentially distributed with scale parameter λ and location parameter 0. The underlying test problems shall be $H_i : \{\vartheta_i = 0\}$ vs. $K_i : \{\vartheta_i > 0\}$ and the test statistics shall be given by $T_i = X_i - Z$. Noting that large values of T_i favour the alternative K_i, the corresponding p-values for a given realization t^* of T_i are given by

$$p_i(t^*) = \mathbb{P}_{H_i}(T_i \geq t^*) = 1 - W_T(t^*),$$

with W_T denoting the cdf. of the difference $X_i - Z$ of two independent exponentially distributed random variables. We will now study the behaviour of the FDR in this setup, denoted as the D-EX-EXP model. The next two auxiliary results prepare the computation of the limiting ecdf. of the p-values in such a D-EX-EXP model.

Lemma 2.19 (Distribution of the difference of two exponentially distributed random variables) *Let $X \sim Exp(0, \lambda_1)$ and $Z \sim Exp(0, \lambda_2)$ be two independent exponentially distributed random variables. Then the cdf. of the difference is given by*

$$\mathbb{P}(X - Z \leq t) = \begin{cases} \frac{\lambda_1}{\lambda_1+\lambda_2} \exp(\lambda_2 t) & \text{for } t \leq 0 \\ \frac{\lambda_1}{\lambda_1+\lambda_2} + \frac{\lambda_2}{\lambda_1+\lambda_2} (1 - \exp(-\lambda_1 t)) & \text{for } t > 0. \end{cases}$$

Proof: Computation of the cross-correlation function of X and Z. ■

Lemma 2.20 (Conditional probability for the difference in the exponential case)
Assume that $X \sim Exp(0, \lambda_1)$ and $Z \sim Exp(0, \lambda_2)$, independent of X. Denoted the difference of X and Z by $T = X - Z$.
Then it holds:

$$\mathbb{P}(T > W_T^{-1}(1-t)|Z = z) = \begin{cases} \exp(-\lambda_1 z) \cdot \frac{\lambda_1+\lambda_2}{\lambda_2} t & \text{for } 0 \leq t \leq \frac{\lambda_2}{\lambda_1+\lambda_2}, \\ \exp(-\lambda_1 z) \cdot \left[\frac{\lambda_1+\lambda_2}{\lambda_1} \cdot (1-t)\right]^{-\frac{\lambda_1}{\lambda_2}} & \text{for } \frac{\lambda_2}{\lambda_1+\lambda_2} < t \leq u(z), \\ 1 & \text{for } u(z) < t \leq 1, \end{cases}$$

with $u(z) = 1 - \frac{\lambda_1}{\lambda_1+\lambda_2} \exp(-\lambda_2 z)$.

Proof: Analogously to the notation in Section 2.1, denote the (unconditional) cdf. of T by W_T and the cdf. of X by W_X. Then we obtain (due to the fact that X and Z are independent) that

$$\begin{aligned} \mathbb{P}(X - Z > W_T^{-1}(1-t)|Z = z) &= \mathbb{P}(X - z > W_T^{-1}(1-t)) = \mathbb{P}(X > W_T^{-1}(1-t) + z) \\ &= 1 - W_X(W_T^{-1}(1-t) + z). \end{aligned}$$

Noticing that $W_X(x) = 0$ for $x < 0$, we have $\mathbb{P}(X - Z > W_T^{-1}(1-t)|Z = z) = 1$ if $W_T^{-1}(1-t) + z < 0$. Deducing W_T^{-1} from Lemma 2.19, we therefore obtain the assertion for the case $u(z) < t \leq 1$. Assuming $W_T^{-1}(1-t) + z \geq 0$, we have

$$1 - W_X(W_T^{-1}(1-t) + z) = \exp(-\lambda_1 z) \exp(-\lambda_1 W_T^{-1}(1-t)).$$

Obtaining W_T^{-1} from Lemma 2.19 and plugging in the resulting expressions for the different cases, we obtain the assertion in the remaining two cases. ■

If we now return to our test problem in the D-EX-EXP model, we obtain by combining the argu-

ments of Lemmas 2.19 and 2.20 that

$$\begin{aligned} p_i(t) = \mathbb{P}_{H_i}(T_i \geq t) = 1 - W_T(t) &= \begin{cases} 1 - \frac{1}{2}\ \exp(\lambda t) & \text{for } t \leq 0, \\ \frac{1}{2}\ \exp(-\lambda t) & \text{for } t > 0, \text{ and} \end{cases} \\ F_\infty(t|z) &= 2\exp(-\lambda z)t \text{ for } 0 \leq t \leq 1/2. \end{aligned}$$

This reveals, that $F_\infty(t|z)$ for given $z \in \mathcal{Z}$ is a linear function in t on $[0, 1/2]$. For the computation of the False Discovery Rate in such a situation, a simple consequence of Lemma 2.14 is helpful.

Corollary 2.21 (False Discovery Rate for linear conditional limiting ecdf.'s)
Consider our general D-EX(ζ_n) model. Assume that $F_\infty(t|z) = m(z)t \ \ \forall t \leq \alpha$, where $m(z)$ is the slope of a straight line depending only on z, and $m(z) < 1/\alpha$ for all $z \in \mathcal{Z}$. Then it holds:

(i) $FDR_n(\zeta_n|z) = \alpha\ \zeta_n\ m(z)\ \ \forall n \geq 1.$

(ii) $FDR_n(\zeta_n) = \alpha\ \zeta_n\ \int m(z)\ d\mathbb{P}^Z(z)\ dz.$

Proof:
ad (i): We will apply Lemma 2.14. Therefore, we set $F_\xi(t) = F_\infty(t|z)$ and note that the $p_i(z)$'s are conditionally i.i.d. with $p_1(z) \sim F_\infty(\cdot|z)$ if H_i is true.
ad (ii): Follows immediately from (i) via integrating with respect to Z. ■
Applying the latter corollary, we finally get the FDR results in the D-EX-EXP model by plugging in $m(z) = 2\exp(-\lambda z)$ as follows.

Corollary 2.22
In the D-EX-EXP(ζ_n) model, the FDR computes as

$$\begin{aligned} FDR_n(\zeta_n|z) &= 2\alpha\ \zeta_n\ \exp(-\lambda z), \\ FDR_n(\zeta_n) &= \alpha\zeta_n \int_0^\infty 2\lambda \exp^2(-\lambda z) dz \\ &= \alpha\zeta_n \quad \textit{for any } \ n > 1, \alpha < 1/2, \lambda > 0. \end{aligned}$$

This has the interpretation that in this special case, the FDR of the linear step-up procedure based on Simes' critical values exactly equals the Benjamini-Hochberg bound for any size n of the family of hypotheses if $\alpha < 1/2$, although the underlying test statistics (and therefore the corresponding p-values) are not independent.

Remark 2.23
It is remarkable that the MTP_2 property holds in this setting so that the Benjamini-Hochberg bound for the FDR applies. This is an immediate consequence of Proposition 3.7. in [148], because the the pdf. of the Exp(λ) distribution is PF_2 for any $\lambda > 0$, cf. [147].

The following Figure 2.4 displays $F_\infty(t|z)$ in case of $\lambda = 1$ for different values of z together with Simes' line for $\alpha = 0.1$. It is remarkable that the angle between $F_\infty(t|z)$ and Simes' line determines the limiting FDR.

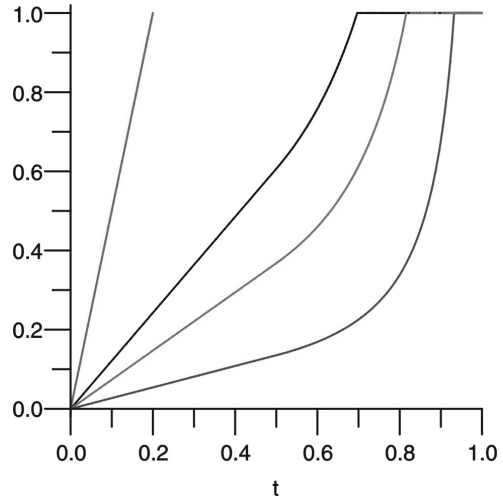

Figure 2.4: $F_\infty(t|z)$ for $z = 0.5$, $z = 1.0$ and $z = 2.0$ in case of $\lambda = 1$ together with Simes' line for $\alpha = 0.1$

Remark 2.24

From Lemma 2.19, it follows that the one-dimensional marginal cumulative distribution function of an individual T_i under the corresponding null hypothesis H_i in case of $\lambda = 1$ can be expressed by

$$W_T(t) = \begin{cases} 1/2 \ \exp(t) & \text{for } \ t \leq 0, \\ 1 - 1/2 \ \exp(-t) & \text{for } \ t > 0. \end{cases}$$

The corresponding distribution is the well-known double-exponential or Laplace distribution. It may be interesting to have a brief look on the multivariate Laplace distribution with the same correlation structure as present for our T_i. In [164], the density function w_n for the multivariate Laplace distribution in dimension n was given in the context of copulas as

$$w_n(z) = \frac{\Gamma(n/2)}{\Gamma(n)} \frac{1}{2\pi^{n/2}\sqrt{\det(\Sigma_n)}} \exp\left(-\sqrt{z^t \Sigma_n^{-1} z}\right), z \in \mathbb{R}^n$$

with Σ_n denoting the $n \times n$-dimensional correlation matrix under the assumption that the n components are standardized. For our T_i we have $\text{Var}(T_i) = 2$ for all $i = 1, \ldots, n$ as well as $\rho(T_i, T_j) = 1/2$ for all i not equal to j and, therefore, in the corresponding multivariate Laplace case the MTP_2 property does hold as well. The latter result is due to the fact that the multivariate Laplace distribution belongs to the class of spherical (elliptical) distributions (cf. [152]) for which the MTP_2 condition is equivalent to the property that the correlation matrix is invertible and its

off-diagonal elements are non-negative (this fact can e. g. be deduced from Theorem 3 in [230] in connection with the generalization methods derived in [152]). Since in our case it holds

$$\Sigma_n = (\sigma_{i,j})_{i,j=1,\ldots,n} \quad \text{with} \quad \sigma_{i,j} = \frac{1}{2}\left(1 + \delta_{i,j}\right),$$

the assertion is obvious.

2.3 Exchangeable normally distributed variables

Our next concrete example treats the case of exchangeable normally distributed variables and has very high practical relevance. Our notation will be as follows.

Notation 2.25 (D-EX-N(ζ_n) model)
Let $X_i \sim \mathcal{N}(0,1)$, $i = 0, \ldots, n$, be independent standard normal random variables and let $T_i = \vartheta_i + \sqrt{\bar{\rho}} X_i - \sqrt{\rho} X_0$ with $\vartheta_i \geq 0$, $i = 0, \ldots, n$, where $\rho \in (0,1)$ is assumed to be known and $\bar{\rho} = 1 - \rho$. Then $T = (T_1, \ldots, T_n)$ is multivariate normally distributed with mean vector $\vartheta = (\vartheta_1, \ldots, \vartheta_n)$, Var$[T_i] = 1$ for $i = 1, \ldots, n$, and Cov$(T_i, T_j) = \rho$ for $1 \leq i, j \leq n$ with i not equal to j. Consider the multiple testing problem $H_i : \vartheta_i = 0$ versus $K_i : \vartheta_i > 0$, $i = 1, \ldots, n$. For $\rho \in (0,1)$ the distribution of T is MTP$_2$ so that the Benjamini-Hochberg bound applies, cf. [17] or [237].
In the following we use the notation introduced in the Section 2.1. Note that Z is replaced by X_0 and $W_X = W_{X_0} = W_T = \Phi$. Suitable p-values for testing the H_i's are given by $p_i = p_i(x) = 1 - \Phi(\vartheta_i + \sqrt{\bar{\rho}} x_i - \sqrt{\rho} x_0)$, $i = 1, \ldots, n$. Again we add $\vartheta_i = \infty$ to the model such that $p_i = 0$ a. s. if $\vartheta_i = \infty$, $i = 1, \ldots, n$. We denote the corresponding D-EX(ζ_n) model by D-EX-N(ζ_n).

Remark 2.26
This setup includes the well-known many-one multiple comparisons problem which usually reads as follows. Let $\overline{Y}_{i\cdot} \sim N(\nu_i, \sigma^2/m_i)$, $i = 0, \ldots, n$, denote independently normally distributed sample means with $\sigma^2 > 0$ (known), $m_1 = \cdots = m_n$ and $\nu_i \geq \nu_0$ for $i = 1, \ldots, n$. Suppose one is interested in testing $\tilde{H}_i : \nu_i = \nu_0$ versus $\tilde{K}_i : \nu_i > \nu_0$ for $i = 1, \ldots, n$ by using the test statistics $W_i = (1/m_0 + 1/m_1)^{-1/2}(\overline{Y}_{i\cdot} - \overline{Y}_{0\cdot})/\sigma$, $i = 1, \ldots, n$. Then $\mathbb{E}[W_i] = (1/m_0 + 1/m_1)^{-1/2}(\nu_i - \nu_0)/\sigma = \vartheta_i$ (say), Var$[W_i] = 1$ and Cov$(W_i, W_j) = \sqrt{m/(m+m_0)} = \rho$ (say).

Our policy in the remainder of this chapter will be to express the EER and the FDR in the D-EX-N(ζ_n) models with respect to the correlation ρ. First of all, we now determine the EER-values and FDR-values for the extreme cases $\rho = 0$ (independence) and full dependency, i.e. $\rho = 1$.

Theorem 2.27 (FDR and EER in case of $\rho = 1$ in the D-EX-N(ζ_n)-model)
In the case of full dependency in the D-EX-N(ζ_n)-model, i.e. $\rho = 1$, we obtain

$$\mathit{FDR}_n(\zeta_n) = \alpha\, \zeta_n = \mathit{EER}_n(\zeta_n) \ \text{for any} \ n \in \mathbb{N}.$$

Proof: The test statistics T_i simplify in case of $\rho = 1$ to $T_i = \vartheta_i - X_0$ and therefore the corresponding conditional p-values are $p_i(x_0) = 1 - \Phi(\vartheta_i - x_0)$. Since in our D-EX-N($\zeta_n$)-model it holds $\vartheta_i = 0$ for all $i \in I_0$ and $\vartheta_i = \infty$ for all $i \in \complement I_0$, where I_0 denotes the set of indices corresponding to true null hypotheses, it follows

$$\begin{aligned} F_n(t|x_0, \zeta_n) &= \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{[0,t]}(p_i(x_0)) \\ &= (1 - \zeta_n) + \zeta_n \mathbf{1}_{[0,t]}(\Phi(x_0)). \end{aligned}$$

This representation of F_n implies, that the conditional ecdf. of the p-values is a step function with exactly one step at $t = \Phi(x_0)$ for every $n \in \mathbb{N}$ in case of $\rho = 1$. Consequently, it follows for the LCP $t^*(x_0)$ of F_n and Simes' line in this setting, that

$$t^*(x_0) = \begin{cases} \alpha\,(1 - \zeta_n) & \text{for } x_0 > \Phi^{-1}(\alpha) \\ \alpha & \text{for } x_0 \leq \Phi^{-1}(\alpha) \end{cases}$$

and the number of falsely rejected hypotheses in our model is given as

$$V_n(x_0) = \begin{cases} 0 & \text{for } x_0 > \Phi^{-1}(\alpha) \\ n_0 & \text{for } x_0 \leq \Phi^{-1}(\alpha). \end{cases}$$

Noting that $V_n(x_0)$ can only realize these two values in this setting, $\mathrm{EER}_n(\zeta_n)$ can immediately be expressed via a discrete expectation formula, i.e.

$$\mathrm{EER}_n(\zeta_n) = \zeta_n\, \mathbb{P}(V_n = n_0) = \zeta_n\, \mathbb{P}(X_0 \leq \Phi^{-1}(\alpha)) = \alpha\, \zeta_n.$$

Since $R_n(x_0) = V_n(x_0) + n_1$, it holds

$$\frac{V_n(x_0)}{R_n(x_0) \vee 1} = \begin{cases} 0 & \text{for } x_0 > \Phi^{-1}(\alpha), \\ \zeta_n & \text{for } x_0 \leq \Phi^{-1}(\alpha), \end{cases}$$

and the assertion for $\mathrm{FDR}_n(\zeta_n)$ follows analogously. ■

Remark 2.28 (FDR and EER for $\rho = 0$)

(i) In case of $\rho = 0$, we are in the i.i.d. situation originally investigated by Benjamini and Hochberg. As stated before, in this situation the False Discovery Rate has the value $\mathrm{FDR}_n(\zeta_n) = \alpha\, \zeta_n$ for all $n \in \mathbb{N}$.

(ii) As pointed out in [91], it holds for the limiting EER in the independent case, i.e. $\rho = 0$, that $\mathrm{EER}_\infty(\zeta) = \alpha(1 - \zeta)/(1 - \alpha\zeta)$.

In case of $\rho \in (0, 1)$, however, the computation of $\mathrm{FDR}_\infty(\zeta)$ and $\mathrm{EER}_\infty(\zeta)$ becomes substantially more difficult. We first focus on the case $\zeta = 1$ (the proportion of true hypotheses tends to one).

2.3.1 The special case $\zeta = 1$

Utilizing Lemma 2.9, we obtain that the conditional ecdf. of the p-values given $X_0 = x_0$ has the limit $F_\infty(\cdot|x_0)$ with formal representation

$$F_\infty(t|x_0) = 1 - \Phi\left(\frac{\Phi^{-1}(1-t)}{\sqrt{1-\rho}} + \sqrt{\frac{\rho}{1-\rho}}x_0\right). \tag{2.14}$$

Some important properties of this limiting conditional ecdf. are listed in the following lemma, which can be verified by elementary analytic calculations.

Lemma 2.29

The function $F_\infty(\cdot|x_0)$ defined in (2.14) has the following properties.

(i) For any fixed $t \in [0,1]$, $F_\infty(t|x_0)$ is non-increasing in x_0.

(ii) $\lim_{t\downarrow 0}(\partial/\partial t)F_\infty(t|x_0) = 0$.

(iii) $F_\infty(t|x_0)$ is convex for $t \in [0, \Phi(x_0/\sqrt{\rho})]$ and concave for $t \in [\Phi(x_0/\sqrt{\rho}), 1]$.

From these considerations, it can be concluded that F_∞ has (depending on x_0) either zero or exactly two points of intersection or exactly one boundary point with Simes' line on the interval $[0, \alpha]$. These three possible situations shall be demonstrated in the following figure with $\rho = 0.90$.

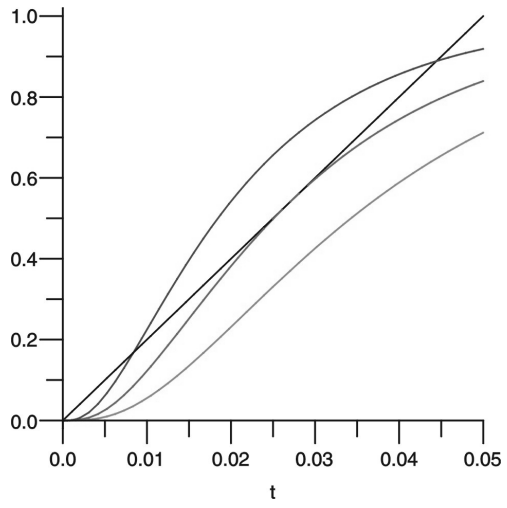

Figure 2.5: $F_\infty(t|x_0)$ for $x_0 = -1.92$, $x_0 \approx -2.06453$ and $x_0 = -2.2$ together with Simes' line on $[0, \alpha]$

The constellation corresponding to the curve in the middle, i.e. the existence of exactly one boundary point of the two objects resulting from the special outcome $X_0 = \bar{x}_0$ (say) is of particular importance for the computation of the False Discovery Rate, because it holds:

Theorem 2.30 (The FDR in case of $\zeta = 1$ and $\rho \in (0,1)$)
Let $\bar{x}_0$ be the value of X_0 for which the limiting conditional ecdf. $F_\infty(\cdot|x_0)$ of the p-values has exactly one boundary point with Simes' line on the interval $[0,\alpha]$ and set $t_2 = t(\bar{x}_0)$. Then the limiting Expected Error Rate and False Discovery Rate compute as follows:

$$\begin{aligned} \mathit{EER}_\infty(1) &= t_2\Phi(\bar{x}_0)/\alpha + \int_{t_2/\alpha}^{1} \Phi\left(x_0(\alpha t|1)\right) dt, && (2.15)\\ \mathit{FDR}_\infty(1) &= \mathbb{P}(X_0 \le \bar{x}_0) = \Phi(\bar{x}_0). && (2.16) \end{aligned}$$

Proof: We will apply Theorem 2.18. Therefore, we have to determine the set C_1 of largest crossing points and the corresponding sets E_0 and E_1 for the x_0-values. We note that for every given $t \in (0,\alpha)$, it is possible to determine the corresponding $x_0 = x_0(t)$, so that $F_\infty(\cdot|x_0)$ intersects Simes' line in the point $(t, t/\alpha)$. Regarding this value $x_0(t)$ as a function of t, we obtain after equating $F_\infty(t|x_0) = t/\alpha$ and solving for x_0:

$$x_0(t) = \sqrt{\frac{1-\rho}{\rho}}\Phi^{-1}(1-\frac{t}{\alpha}) - \frac{\Phi^{-1}(1-t)}{\sqrt{\rho}}, t \in (0,\alpha).$$

Studying the analytic properties of the function $x_0(t)$ yields $\lim_{t\downarrow 0} x_0(t) = \lim_{t\uparrow\alpha} x_0(t) = -\infty$ and since $x_0(\cdot)$ is a continuous concave function, it exists a unique maximum $\bar{x}_0$ of $x_0(t)$ on $(0,\alpha)$. It is worth noting that this implies that there is exactly one abscissa t_2 for which $x_0(t)$ equals $\bar{x}_0$. This can geometrically be interpreted as the boundary point situation. Furthermore, as stated before, for a fixed $t \in (0,\alpha)$, the limiting conditional ecdf. $F_\infty(t|x_0)$ (regarded as a function of x_0) is non-increasing in x_0. Therefore, no intersection of $F_\infty(\cdot|x_0)$ and Simes' line occurs for values $X_0 > \bar{x}_0$ and for values $X_0 < \bar{x}_0$, we obtain two points of intersection, the larger of which is larger than t_2. Consequently, it holds

$$C_1 = \{0\} \cup (t_2, \alpha), E_1 = (-\infty, \bar{x}_0] \text{ and } E_0 = (\bar{x}_0, \infty).$$

Theorem 2.18 then immediately yields the formula for $\mathrm{EER}_\infty(1)$ by plugging in the actual cdf. $G_{1,1} = \Phi$. For the computation of $\mathrm{FDR}_\infty(1)$, we recall that $\lim_{t\downarrow 0}(\partial/\partial t)F_\infty(t|x_0) = 0$ for any $x_0 \in \mathbb{R}$. Therefore, the second summand in (2.13) vanishes and we finally obtain (2.16). ■

After this preparing considerations, computation of the FDR in case of $\zeta = 1$ is equivalent to the task of computing the boundary point $(t^*, t^*/\alpha)$ of the limiting conditional ecdf. $F_\infty(\cdot|x_0)$ and Simes' line on $(0,\alpha)$ as well as the corresponding $\bar{x}_0 = x_0(t^*)$ for given values α and ρ. Necessary and sufficient conditions for having a boundary point of $F_\infty(\cdot|x_0)$ and Simes line at the point t are, that functional and derivative value of the two objects at the point t are equal. This means, formally expressed,

$$\begin{aligned} F_\infty(t|x_0) &= \frac{t}{\alpha}, && (2.17)\\ \frac{d}{dt}F_\infty(t|x_0) &= \frac{1}{\alpha}. && (2.18) \end{aligned}$$

If we now define the distance function $d(\cdot|x_0)$ by $d(t|x_0) := F_\infty(t|x_0) - t/\alpha$ and furthermore substitute $u := \Phi^{-1}(1-t) \Leftrightarrow t = 1 - \Phi(u) = \Phi(-u)$, we obtain

$$d(u|x_0) = \Phi\left(-\frac{u}{\sqrt{1-\rho}} - \sqrt{\frac{\rho}{1-\rho}}x_0\right) - \frac{\Phi(-u)}{\alpha} \quad \text{as well as} \tag{2.19}$$

$$\frac{d}{du}d(u|x_0) = \frac{\varphi(u)}{\alpha} - \frac{1}{\sqrt{1-\rho}}\varphi\left(\frac{u}{\sqrt{1-\rho}} + \sqrt{\frac{\rho}{1-\rho}}x_0\right) \tag{2.20}$$

and the conditions (2.17) and (2.18) from above read re-formulated

$$d(u|x_0) = 0, \tag{2.21}$$

$$\frac{d}{du}\ d(u|x_0) = 0. \tag{2.22}$$

This is a system of two equations in the two variables u and x_0. Equation (2.22) corresponds to a quadratic form and can be solved explicitly for u. The solutions are given by

$$u_{1,2} = \frac{-x_0}{\sqrt{\rho}} \pm \sqrt{\frac{1-\rho}{\rho}}\sqrt{x_0^2 - 2\ln\left(\frac{\sqrt{1-\rho}}{\alpha}\right)}.$$

Since for a fixed x_0 the largest crossing or boundary point (or, more exactly, the one with the largest abscissa in t) decides over the retention and rejection of hypotheses in the linear step-up procedure, the smaller value

$$u_2 = \frac{-x_0}{\sqrt{\rho}} - \sqrt{\frac{1-\rho}{\rho}}\sqrt{x_0^2 - 2\ln(\frac{\sqrt{1-\rho}}{\alpha})} \tag{2.23}$$

is the demanded solution, because the transformation from t to u was a strictly decreasing one. Plugging u_2 into (2.21), we get the following defining equation for $\bar{x}_0$:

$$\frac{\Phi\left(\frac{1}{\sqrt{\rho}}\left[\bar{x}_0 + \sqrt{1-\rho}\,\sqrt{\bar{x}_0^2 - 2\ln(\frac{\sqrt{1-\rho}}{\alpha})}\right]\right)}{\Phi\left(\frac{1}{\sqrt{\rho}}\left[\sqrt{1-\rho}\,\bar{x}_0 + \sqrt{\bar{x}_0^2 - 2\ln(\frac{\sqrt{1-\rho}}{\alpha})}\right]\right)} = \alpha. \tag{2.24}$$

Unfortunately, it is not possible to solve (2.24) analytically for $\bar{x}_0$. Therefore, a numerical algorithm has to be employed to determine the value $\bar{x}_0$ approximately. This can be done e.g. by the well-known Newton-Raphson iteration method, which computes $\bar{x}_0$ up to an arbitrary precision. For its application, we can define the distance between $F_\infty(u_2|x_0)$ and Simes' line as a function $\tilde{d}$ (say) of x_0, given by

$$\begin{aligned}\tilde{d}(x_0) = {} & \Phi\left(\frac{1}{\sqrt{\rho}}\left[\sqrt{1-\rho}\,x_0 + \sqrt{x_0^2 - 2\ln(\frac{\sqrt{1-\rho}}{\alpha})}\right]\right) \\ & - \frac{1}{\alpha}\,\Phi\left(\frac{1}{\sqrt{\rho}}\left[x_0 + \sqrt{1-\rho}\,\sqrt{x_0^2 - 2\ln(\frac{\sqrt{1-\rho}}{\alpha})}\right]\right),\end{aligned}$$

and additionally use its derivative

$$\begin{aligned}\frac{d}{dx_0}\,\tilde{d}(x_0) &= \varphi\left(\sqrt{\frac{1-\rho}{\rho}}x_0+\frac{1}{\sqrt{\rho}}\sqrt{x_0^2-2\ln(\frac{\sqrt{1-\rho}}{\alpha})}\right)\\ &\quad\times\left(\sqrt{\frac{1-\rho}{\rho}}+\frac{1}{\sqrt{\rho}}\frac{x_0}{\sqrt{x_0^2-2\ln(\frac{\sqrt{1-\rho}}{\alpha})}}\right)\\ &- \frac{1}{\alpha}\,\varphi\left(\frac{x_0}{\sqrt{\rho}}+\sqrt{\frac{1-\rho}{\rho}}\sqrt{x_0^2-2\ln(\frac{\sqrt{1-\rho}}{\alpha})}\right)\\ &\quad\times\left(\sqrt{\frac{1}{\rho}}+\sqrt{\frac{1-\rho}{\rho}}\frac{x_0}{\sqrt{x_0^2-2\ln(\frac{\sqrt{1-\rho}}{\alpha})}}\right)\end{aligned}$$

in each iteration step.

Remark 2.31

It may be asked why the method described above should be preferred over a numerical (grid) search for the maximum of $x_0(t)$ on $(0,\alpha)$. A first answer from the practical point of view is that the numerical computation of $\Phi(\cdot)$ is substantially more feasible than the one of $\Phi^{-1}(\cdot)$, but we also give a theoretical one. That is to say that it is possible to give an upper bound for the FDR in the considered setup by employing the explicit solutions for $u_{1,2}$. In case of u_1 being equal to u_2, i.e. the distance function $d(u|x_0)$ having a saddle point in $u^* := u_1 = u_2$, the corresponding value x_0^*, for which the discriminant $\sqrt{x_0^2-2\ln\left(\frac{\sqrt{1-\rho}}{\alpha}\right)}$ of the quadratic equation (2.22) for determining $u_{1,2}$ vanishes, is larger than the exact solution $\bar{x}_0$. Consequently, it holds FDR $\leq \Phi(x_0^*)$. However, x_0^* can be computed very easily and we obtain $x_0^* = -\sqrt{2\ln(\frac{\sqrt{1-\rho}}{\alpha})}$ for $\rho \leq 1-\alpha^2$.

If we now let $\rho \downarrow 0$, it even reveals that the limiting FDR has the corresponding right side limit $\Phi(-\sqrt{-2\ln(\alpha)})$, at least for $\alpha < 1/2$.

Theorem 2.32 (Limiting value of the FDR for $\rho \downarrow 0$ in the D-EX-N(ζ_n)-model with $\zeta = 1$)
For $\alpha \in (0,1/2)$, it holds in the D-EX-N(ζ_n) model:

$$\lim_{\rho\downarrow 0} \mathit{FDR}_\infty(1) = \Phi(-\sqrt{-2\ln(\alpha)}).$$

Proof: From geometric considerations, we have that for any $\rho \in (0,1)$ there exists a unique solution $(u,x_0) = (u_\rho, x_{0,\rho})$ (say) of (2.21) and (2.22). Moreover, notice that $\alpha \in (0,1/2]$ implies $u_\rho > 0$ because of $t \in (0,\alpha)$ and the substitution $u = \Phi^{-1}(1-t)$ and therefore (see 2.23) $x_{0,\rho} < 0$. Since u_ρ has to be a real number in (2.23), we furthermore obtain that $\limsup_{\rho\to 0^+} x_{0,\rho} \leq -\sqrt{-2\ln(\alpha)}$. We will now additionally show that $\liminf_{\rho\to 0^+} x_{0,\rho} \geq -\sqrt{-2\ln(\alpha)}$.

To this end, for $\delta \in (0, \alpha)$, we consider the ansatz $x_0 = x_0(\delta) = -\sqrt{-2\ln(\delta)} < -\sqrt{-2\ln(\alpha)} = x_0(\alpha)$ covering the entire range of possible values for $x_{0,\rho}$ and define

$$u = u(\rho, \delta) = \frac{-x_0(\delta)}{\sqrt{\rho}} \text{ and } w = w(\rho, \delta) = \frac{u(\rho, \delta)}{\sqrt{\bar{\rho}}} + \sqrt{\frac{\rho}{\bar{\rho}}} x_0(\delta).$$

Then we get from (2.19) that $d(u(\rho, \delta)|x_0(\delta)) = \Phi(-w) - \Phi(-u)/\alpha$. Employing the asymptotic relationship $(x \to \infty)$ $\Phi(-x)/\varphi(-x) \sim 1/x$ for Mills' ratio (cf. [193]), we get

$$\frac{\Phi(-u)}{\Phi(-w)} \sim \frac{w}{u} \frac{\varphi(u)}{\varphi(w)} = \frac{w}{u} \exp((w^2 - u^2)/2).$$

Since $\exp((w(\rho, \delta)^2 - u(\rho, \delta)^2)/2) = \delta < \alpha$ independent of ρ and $\lim_{\rho \to 0^+} w(\rho, \delta)/u(\rho, \delta) = 1$, we obtain that $\lim_{\rho \to 0^+} d(u(\rho, \delta)|x_0(\delta)) > 0$ for all $\delta \in (0, \alpha)$ and consequently conclude that $\liminf_{\rho \to 0^+} x_{0,\rho} \geq -\sqrt{-2\ln(\alpha)}$. Together with the result $\limsup_{\rho \to 0^+} x_{0,\rho} \leq -\sqrt{-2\ln(\alpha)}$ from above, we finally obtain $\lim_{\rho \to 0^+} x_{0,\rho} = -\sqrt{-2\ln(\alpha)}$ and the assertion follows from formula (2.16). ■

Remark 2.33

Note that the latter result implies a discontinuity of the FDR (looked at with respect to its dependence on ρ), because for $\rho = 0$ it holds FDR $= \alpha$ as stated above. In practice, it is often assumed that there may be some kind of weak dependence between test statistics (cf. e. g. [275]) being close to independence in some sense. However, Theorem 2.32 suggests that for large n and small $\rho > 0$ the actual FDR may be much smaller than in the independence model if only a small number of hypotheses is false. For example, for $\alpha = 0.05$ it is $-\sqrt{-2\ln(\alpha)} \approx -2.4477$ and $\Phi(-\sqrt{-2\ln(\alpha)}) \approx 0.0072$. This seems to be quite contradictory to the weak dependence paradigm. A deeper view into this matter however reveals that if we change the order of limits, the results again become what one would expect. More specifically, we have that

$$\begin{aligned} \lim_{\rho \to 0^+} \left(\lim_{n \to \infty} \text{FDR}_n(1) \right) &= \Phi(-\sqrt{-2\ln(\alpha)}) << \alpha, \text{ but} \\ \lim_{n \to \infty} \left(\lim_{\rho \to 0^+} \text{FDR}_n(1) \right) &= \alpha. \end{aligned}$$

Taking this into consideration, one may argue that Theorem 2.32 has mainly academic value since the first order of limit has no practical application. A nice visual illustration of the discrepancy of the two results is given in Appendix A.1.

Remark 2.34

If ζ equals 1, there exists for any value of $\rho \in (0, 1)$ an $\bar{x}_0$, such that the function $F_\infty(t|\bar{x}_0)$ and the Simes line have a boundary point.

As a summarization, the following figure illustrates the graph of the FDR- and EER- curves with regard to ρ in case of $\zeta = 1$.

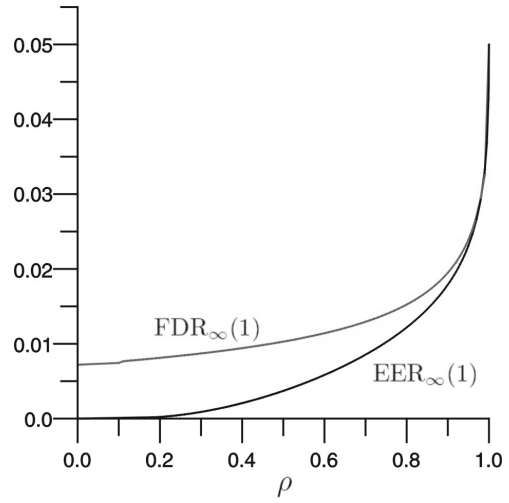

Figure 2.6: FDR$_\infty(1)$ and EER$_\infty(1)$ for varying values of $\rho \in [0,1]$ and $\alpha = 0.05$

It becomes obvious that the EER is always bounded by the FDR. This fact is immediately clear if we consider the representations

$$\begin{aligned} \text{FDR}_n(1) = \mathbb{P}(V_n > 0) &= \sum_{i=1}^{n} \mathbb{P}(V_n = i), \\ \text{EER}_n(1) = \mathbb{E}\left(\frac{V_n}{n}\right) &= \sum_{i=1}^{n} \frac{i}{n} \mathbb{P}(V_n = i) \leq \text{FDR}_n(1), \end{aligned}$$

which hold for all $n \in \mathbb{N}$. Furthermore, it is illustrated that the two entities converge against the same limiting value α for $\rho \to 1$ as proven in Theorem 2.27.

2.3.2 The general case $\zeta < 1$

We will now lay focus on the more general case $\zeta < 1$. Recall that the underlying model assumption is now that a proportion ζ of the hypotheses are true and the remaining proportion $(1-\zeta)$ of the hypotheses are totally wrong with p-values equal to zero almost surely (according to the Dirac-exchangeable model definition). As a consequence, the limiting conditional ecdf. $F_\infty(\cdot|x_0,\zeta)$ of the p-values is now given by

$$F_\infty(t|x_0,\zeta) = (1-\zeta) + \zeta\left(1 - \Phi\left(\frac{\Phi^{-1}(1-t)}{\sqrt{1-\rho}} + \sqrt{\frac{\rho}{1-\rho}}\, x_0\right)\right).$$

Most of the substantial properties of the graph of this function are preserved from the case $\zeta = 1$, but there is one major change: Now, $F_\infty(\cdot|x_0,\zeta)$ always starts above Simes' line, because obviously it holds $F_\infty(0|x_0,\zeta) = 1-\zeta > 0$. Therefore, the two objects always have at least one point of intersection. For some values x_0, together with certain parameter constellations for ζ and ρ, however, we may get two ore three points of intersection, but never more than three. The following figure illustrates the three possible scenarios for the example case $\zeta = 0.9$ and $\rho = 0.95$.

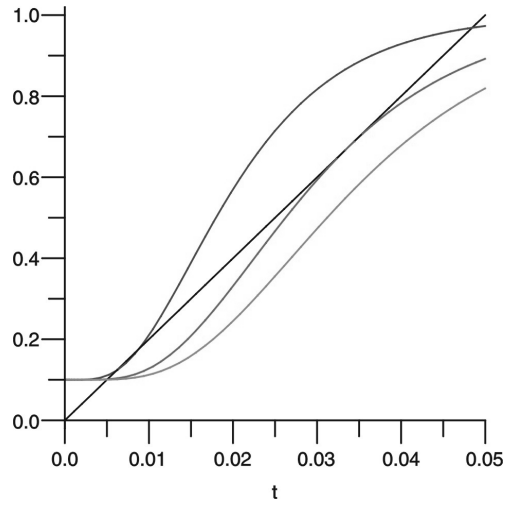

Figure 2.7: $F_\infty(t|x_0,\zeta)$ for $x_0 = -1.88$, $x_0 \approx -1.95738$, and $x_0 = -2.12$ on $[0,\alpha]$ in case of $\zeta = 0.9$ and $\rho = 0.95$.

Again, for a given $t \in (0,\alpha)$, it is possible to determine the value $x_0(t)$, such that the limiting conditional ecdf. and Simes' line cross each other in the point $(t, t/\alpha)$. A straightforward calculation yields

$$x_0(t) = \sqrt{\frac{1-\rho}{\rho}}\ \Phi^{-1}\left(\frac{1-t/\alpha}{\zeta}\right) - \frac{\Phi^{-1}(1-t)}{\sqrt{\rho}}, t \in (0,\alpha).$$

The computation of the FDR, however, is more complicated in this case. The actual FDP, i.e. the proportion of falsely rejected null hypotheses with respect to all rejected hypotheses for a given x_0, has now the limit

$$\lim_{n\to\infty} \frac{V_n(x_0)}{R_n(x_0)\vee 1} = \frac{t(x_0)/\alpha - (1-\zeta)}{t(x_0)/\alpha} = 1 - \frac{\alpha\,(1-\zeta)}{t(x_0)},$$

where $t(x_0)$ denotes the largest point of intersection of $F_\infty(\cdot|x_0,\zeta)$ and Simes' line. The FDR is defined as the expectation of this ratio and therefore it holds

$$\mathrm{FDR}_\infty(\zeta) = \mathbb{E}\left(\lim_{n\to\infty} \frac{V_n(x_0)}{R_n(x_0)\vee 1}\right) = 1 - \alpha\,(1-\zeta)\cdot\mathbb{E}\left(\frac{1}{t(x_0)}\right). \tag{2.25}$$

Again, the boundary point situation is of crucial interest. If it occurs for a certain $\bar{x}_0$, the function $F_\infty(\cdot|\bar{x}_0,\zeta)$ intersects Simes' line at a point $t_1 \in (\alpha(1-\zeta), \Phi(\bar{x}_0/\sqrt{\rho}))$ and touches Simes' line at the boundary point $t_2 \in (\Phi(\bar{x}_0/\sqrt{\rho}), \alpha)$ with $t_1 < t_2$. For $X_0 > \bar{x}_0$, we have exactly one point of intersection (automatically the largest) and its t-coordinate lies in the interval $I_1 := (\alpha(1-\zeta), t_1)$. For $X_0 < \bar{x}_0$, however, three points of intersection occur and the largest crossing point abscissa, which is relevant for the computation of the FDR, comes from the interval $I_2 := (t_2, \alpha)$. Summarized, there are the two disjoint intervals I_1 and I_2 of possible largest crossing point abscissas and therefore t_1 and t_2, respectively, provide integration bounds for the expectation formula (2.25) for the computation of the limiting FDR.

For the determination of the boundary point, we utilize a technique analogue to the case $\zeta = 1$ described above. With our already introduced notation, we obtain here

$$\begin{aligned} d(u|x_0,\zeta) &= (1-\zeta) + \zeta\left(\Phi(-\frac{u}{\sqrt{1-\rho}} - \sqrt{\frac{\rho}{1-\rho}}x_0)\right) - \frac{\Phi(-u)}{\alpha}, \\ \frac{d}{du}\ d(u|x_0,\zeta) &= \frac{\varphi(u)}{\alpha} - \frac{\zeta}{\sqrt{1-\rho}}\varphi\left(\frac{u}{\sqrt{1-\rho}} + \sqrt{\frac{\rho}{1-\rho}}x_0\right) \end{aligned}$$

and

$$u_2 = \frac{-x_0}{\sqrt{\rho}} - \sqrt{\frac{1-\rho}{\rho}}\sqrt{x_0^2 - 2\cdot\ln(\frac{\sqrt{1-\rho}}{\alpha\zeta})}. \tag{2.26}$$

Plugging this u_2 into the equation $d(u|x_0,\zeta) = 0$, which represents the condition for the distance function (in analogy to the case $\zeta = 1$), results in a determining equation for $\bar{x}_0$ given by

$$\frac{\Phi\left(\frac{1}{\sqrt{\rho}}\left[\bar{x}_0 + \sqrt{1-\rho}\,\sqrt{\bar{x}_0^2 - 2\ln\left(\frac{\sqrt{1-\rho}}{\alpha\zeta}\right)}\right]\right)}{(1-\zeta) + \zeta\Phi\left(\frac{1}{\sqrt{\rho}}\left[\sqrt{1-\rho}\,\bar{x}_0 + \sqrt{\bar{x}_0^2 - 2\ln\left(\frac{\sqrt{1-\rho}}{\alpha\zeta}\right)}\right]\right)} = \alpha. \tag{2.27}$$

Also (2.27) cannot be solved analytically for $\bar{x}_0$ and therefore only an approximative numerical solution can be computed via root-finding iteration methods using the functions

$$\begin{aligned} \tilde{d}(x_0|\zeta) &= (1-\zeta) + \zeta\,\Phi\left(\frac{1}{\sqrt{\rho}}\left[\sqrt{1-\rho}\,x_0 + \sqrt{x_0^2 - 2\ln(\frac{\sqrt{1-\rho}}{\alpha\zeta})}\right]\right) \\ &\quad - \frac{1}{\alpha}\,\Phi\left(\frac{1}{\sqrt{\rho}}\left[x_0 + \sqrt{1-\rho}\,\sqrt{x_0^2 - 2\ln(\frac{\sqrt{1-\rho}}{\alpha\zeta})}\right]\right) \end{aligned}$$

and

$$\begin{aligned} \frac{d}{dx_0}\ \tilde{d}(x_0|\zeta) &= \zeta\,\varphi\left(\sqrt{\frac{1-\rho}{\rho}}x_0 + \frac{1}{\sqrt{\rho}}\sqrt{x_0^2 - 2\ln(\frac{\sqrt{1-\rho}}{\alpha\zeta})}\right) \\ &\quad\times\left(\sqrt{\frac{1-\rho}{\rho}} + \frac{1}{\sqrt{\rho}}\frac{x_0}{\sqrt{x_0^2 - 2\ln(\frac{\sqrt{1-\rho}}{\alpha\zeta})}}\right) \\ &\quad - \frac{1}{\alpha}\,\varphi\left(\frac{x_0}{\sqrt{\rho}} + \sqrt{\frac{1-\rho}{\rho}}\sqrt{x_0^2 - 2\ln(\frac{\sqrt{1-\rho}}{\alpha\zeta})}\right) \\ &\quad\times\left(\sqrt{\frac{1}{\rho}} + \sqrt{\frac{1-\rho}{\rho}}\frac{x_0}{\sqrt{x_0^2 - 2\ln(\frac{\sqrt{1-\rho}}{\alpha\zeta})}}\right). \end{aligned}$$

With this (approximate) solution for $\bar{x}_0$, t_2 is given immediately by back substitution, i.e. $t_2 = \Phi(-u_2)$. For the smaller point of intersection and its t-coordinate t_1, we convert the formulas for

the distance function and its derivative with respect to u into the simple form

$$\begin{aligned} d(u|\bar{x}_0,\zeta) &= \Phi\left(\frac{u}{\sqrt{1-\rho}}+\sqrt{\frac{\rho}{1-\rho}}\,\bar{x}_0\right)-\frac{1}{\alpha\zeta}\left(\Phi(u)+\alpha-1\right) \text{ and} \\ \frac{d}{du}\; d(u|\bar{x}_0,\zeta) &= \frac{\varphi\left(\frac{u}{\sqrt{1-\rho}}+\sqrt{\frac{\rho}{1-\rho}}\,\bar{x}_0\right)}{\sqrt{1-\rho}}-\frac{\varphi(u)}{\alpha\zeta} \end{aligned}$$

and apply an iteration method for root-finding in an environment of $\Phi^{-1}(1-\alpha(1-\zeta))$. Having obtained $t_1=\Phi(-u_1)$ in this manner, the limiting EER and FDR finally compute as given in the following theorem.

Theorem 2.35

Given model D-EX-N(ζ_n) with $\lim_{n\to\infty}\zeta_n=\zeta\in(0,1)$, the set of LCP's is given by $C_\zeta=(\alpha(1-\zeta),t_1)\cup(t_2,\alpha)$ and EER$_\infty(\zeta)$ and FDR$_\infty(\zeta)$, respectively, compute as

$$\begin{aligned} \mathit{EER}_\infty(\zeta) &= \frac{t_2-t_1}{\alpha}\cdot\Phi\left(x_0\left(t_1|\zeta\right)\right)+\int_{1-\zeta}^{t_1/\alpha}\Phi\left(x_0\left(\alpha\, t|\zeta\right)\right)dt+\int_{t_2/\alpha}^{1}\Phi\left(x_0\left(\alpha\, t|\zeta\right)\right)dt, \\ \mathit{FDR}_\infty(\zeta) &= (z_2-z_1)\cdot\Phi\left(x_0\left(\frac{\alpha\,(1-\zeta)}{1-z_1}|\zeta\right)\right) \\ &\quad+\int_0^{z_1}\Phi\left(x_0\left(\frac{\alpha\,(1-\zeta)}{1-z}|\zeta\right)\right)dz+\int_{z_2}^{\zeta}\Phi\left(x_0\left(\frac{\alpha\,(1-\zeta)}{1-z}|\zeta\right)\right)dz, \end{aligned}$$

where $z_i=1-\alpha(1-\zeta)/t_i, i=1,2$.

Proof: The assertions follow from the general Theorem 2.13 via integration by parts. Denote the pdf. corresponding to $G_{\zeta,1}$ by $g_{\zeta,1}$ and notice that $C_{\zeta,1}=(0,t_1/\alpha-(1-\zeta))\cup(t_2/\alpha-(1-\zeta),\zeta)$. From Theorem 2.13, we get

$$\mathrm{EER}_\infty(\zeta) = \int_0^{t_1/\alpha-(1-\zeta)} u g_{\zeta,1}(u)du+\int_{t_2/\alpha-(1-\zeta)}^{\zeta} u g_{\zeta,1}(u)du.$$

Since $W_Z=W_{X_0}=\Phi$, $x_0(t_1|\zeta)=x_0(t_2|\zeta)$ and $\lim_{t\uparrow\alpha}x_0(t|\zeta)=-\infty$, we get

$$\begin{aligned} \mathrm{EER}_\infty(\zeta) &= \left[\frac{t_1}{\alpha}-(1-\zeta)\right]\left[1-\Phi(x_0(t_1|\zeta))\right]+\zeta \\ &\quad-\left[\frac{t_2}{\alpha}-(1-\zeta)\right]\left[1-\Phi(x_0(t_1|\zeta))\right]-\left(\frac{t_1}{\alpha}-(1-\zeta)\right)-\zeta \\ &\quad+\frac{t_2}{\alpha}-(1-\zeta)+\int_0^{t_1/\alpha-(1-\zeta)}\Phi(x_0(\alpha(u+1-\zeta)|\zeta))du \\ &\quad+\int_{t_2/\alpha-(1-\zeta)}^{\zeta}\Phi(x_0(\alpha(u+1-\zeta)|\zeta))du \\ &= \frac{t_2-t_1}{\alpha}\Phi(x_0(t_1|\zeta))+\int_{1-\zeta}^{t_1/\alpha}\Phi(x_0(\alpha t|\zeta))dt+\int_{t_2/\alpha}^{1}\Phi(x_0(\alpha t|\zeta))dt. \end{aligned}$$

In order to compute $\text{FDR}_\infty(\zeta)$, note that for $z \in (0, z_1) \cup (z_2, \zeta)$

$$G_{\zeta,2}(z) = 1 - \Phi(x_0(\frac{\alpha(1-\zeta)}{1-z}|\zeta)).$$

In view of $\lim_{t \downarrow \alpha(1-\zeta)} x_0(t|\zeta) = \infty$, it is $G_{\zeta,2}(z_1) = G_{\zeta,2}(z_2)$, $G_{\zeta,2}(0) = 0$ and $G_{\zeta,2}(\zeta) = 1$. Denoting the corresponding pdf of $G_{\zeta,2}$ by $g_{\zeta,2}$, we obtain

$$\begin{aligned}
\text{FDR}_\infty(\zeta) &= \int_0^{z_1} z g_{\zeta,2}(z)dz + \int_{z_2}^{\zeta} z g_{\zeta,2}(z)dz \\
&= z_1 G_{\zeta,2}(z_1) + \zeta G_{\zeta,2}(\zeta) - z_2 G_{\zeta,2}(z_2) - \int_0^{z_1} G_{\zeta,2}(z)dz - \int_{z_2}^{\zeta} G_{\zeta,2}(z)dz \\
&= (z_2 - z_1)\Phi(x_0(\frac{\alpha(1-\zeta)}{1-z_1}|\zeta)) \\
&\quad + \int_0^{z_1} \Phi(x_0(\frac{\alpha(1-\zeta)}{1-z}|\zeta))dz + \int_{z_2}^{\zeta} \Phi(x_0(\frac{\alpha(1-\zeta)}{1-z}|\zeta))dz,
\end{aligned}$$

and the formulas given in the theorem are proven. ■

Remark 2.36

In contrast to the situation regarded in Remark 2.34, there are parameter combinations for $\zeta < 1$ and ρ, so that for every value of x_0 only exactly one point of intersection of $F_\infty(t|x_0)$ and the Simes line is at hand. Such a constellation can be detected by noticing that the function $x_0(t)$ then decreases monotonously on the whole range $[\alpha\ (1-\zeta), \alpha]$ of possible arguments. In this situation, the abscissa of the demanded (largest) crossing point can lie anywhere in this interval and therefore it then holds

$$\begin{aligned}
\text{EER}_\infty(\zeta) &= \int_{1-\zeta}^{1} \Phi\left(x_0\left(\alpha t\right)\right) dt, \\
\text{FDR}_\infty(\zeta) &= \int_0^{\zeta} \Phi\left(x_0\left(\frac{\alpha(1-\zeta)}{1-z}\right)\right) dz.
\end{aligned}$$

This can be formally interpreted as $t_1 = t_2$.

For $\alpha < 1/2$, there is another way of detecting such a situation. First we notice that $\alpha < 1/2$ implies $u_2 > 0$ in (2.26) and consequently $x_0 < 0$. Now, $t_1 = t_2$ (no tangent point possible for (ρ, ζ)) is equivalent to $d(u_2(\underline{x}_0)|\underline{x}_0, \zeta) \geq 0$ for $\underline{x}_0 = -\sqrt{2\ln(\sqrt{\bar{\rho}}/(\alpha\zeta))}$. To see this, we notice that $F_\infty(t|\underline{x}_0, \zeta)$ has a unique abscissa t^* with same derivative as Simes' line, because the discriminant in (2.26) vanishes for $x_0 = \underline{x}_0$, and it holds $t^* = \Phi(\underline{x}_0/\sqrt{\rho})$, i.e., t^* is the abscissa of the point of inflection of $F_\infty(\cdot|\underline{x}_0, \zeta)$. If now $x_0 < \underline{x}_0 < 0$, no tangent point can occur, because $d(u_2(x_0)|x_0, \zeta) > d(u_2(\underline{x}_0)|\underline{x}_0, \zeta) \geq 0$. On the other hand, if $\underline{x}_0 < x_0 < 0$, no tangent point can occur because we obtain a negative discriminant in (2.26) for such an x_0.

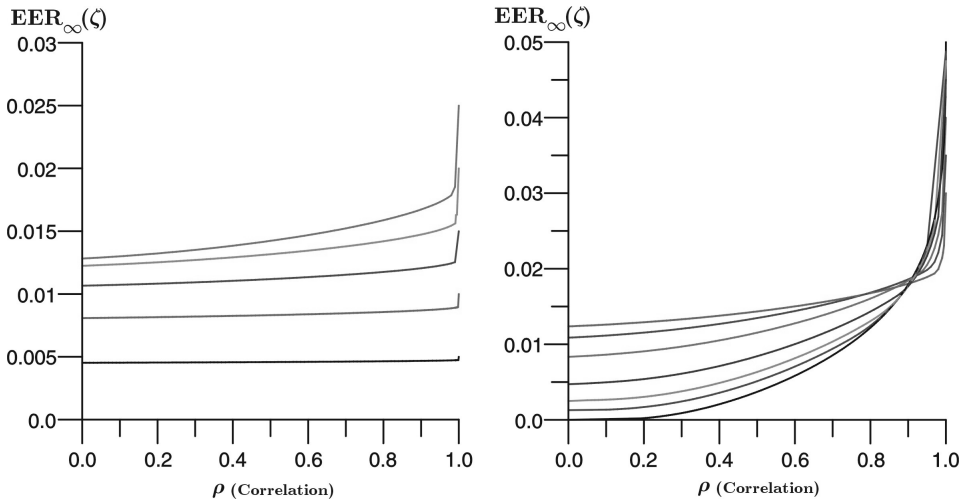

Figure 2.8: $\text{EER}_\infty(\zeta)$ in the D-EX-N(ζ_n) model for $\alpha = 0.05$ and $\zeta = 0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.5$ (left graph) and $\zeta = 0.6,\ 0.7,\ 0.8,\ 0.9,\ 0.95,\ 0.975,\ 1$ (right graph).

We end this section with a depictive representation of $\text{EER}_\infty(\zeta)$ and $\text{FDR}_\infty(\zeta)$ in our D-EX-N(ζ) model. We start with $\text{EER}_\infty(\zeta)$ in Figure 2.8. For $\rho \to 0$, $\text{EER}_\infty(\zeta)$ tends to $\alpha(1-\zeta)/(1-\alpha\zeta)$ as expected, see Remark 2.28 and for $\rho \to 1$, $\text{EER}_\infty(\zeta)$ tends to $\alpha\zeta$ according to Theorem 2.27. Moreover, it seems that $\text{EER}_\infty(\zeta)$ is increasing in ρ with largest values for large ρ and ζ. If ρ is not too large (< 0.9), $\text{EER}_\infty(\zeta)$ is largest for $\zeta \approx 1/2$.

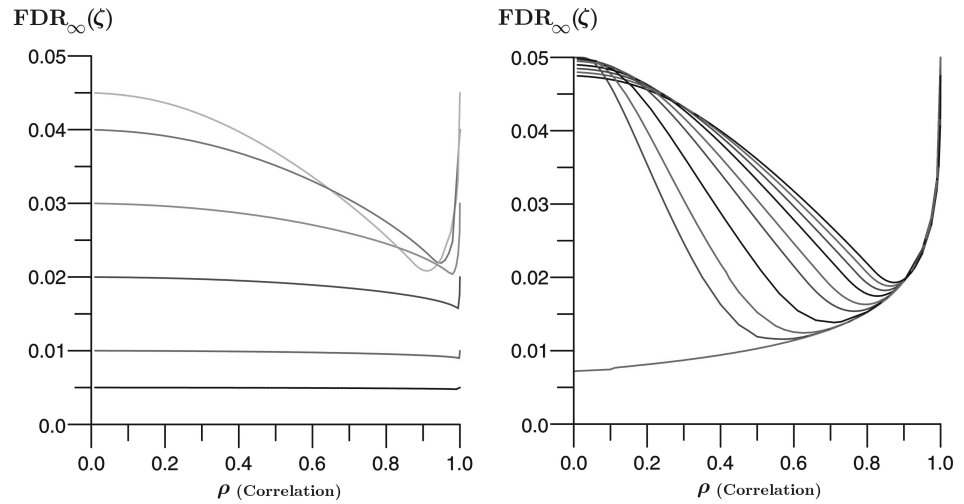

Figure 2.9: $\text{FDR}_\infty(\zeta)$ in the D-EX-N(ζ_n) model for $\alpha = 0.05$ and $\zeta = 0.1,\ 0.2,\ 0.4,\ 0.6,\ 0.8,\ 0.9$ (left graph) and $\zeta = 0.95,\ 0.96,\ 0.97,\ 0.98,\ 0.99,\ 0.995,\ 0.999,\ 0.9999,\ 0.99999,\ 1$ (right graph).

Figure 2.9 displays $\text{FDR}_\infty(\zeta)$ for various values of ζ for $\rho \in [0,1]$. For $\zeta \in (0,1)$, the FDR tends

to the Benjamini-Hochberg bound for $\rho \to 0$ and $\rho \to 1$. For $\rho = 1$ we have total dependence so that $\mathrm{FDR}_n(\zeta_n) = \zeta_n\alpha$ in the D-EX-N(ζ_n) model, again according to Theorem 2.27. For large values of ζ the computation of $\mathrm{FDR}_\infty(\zeta)$ is extremely cumbersome. The main reason is that the BP's are very close to zero so that an enormous numerical accuracy is required. Finally, it is interesting to note that for $\zeta = 1$, $\mathrm{FDR}_\infty(\zeta)$ is the true level of Simes' [264] global test for the intersection hypothesis.

2.4 Exchangeable studentized normal variables

The last distributional setting we investigate deals with equi-correlated studentized normal variables formally introduced in the following definition of the D-EX-t(ζ_n) model.

Notation 2.37 (D-EX-t(ζ_n) model)
Let $X_i \sim \mathcal{N}(\vartheta_i, \sigma^2)$, $i = 1, \ldots, n$, be independent normal random variables and let $\nu S^2/\sigma^2 \sim \chi^2_\nu$ be independent of the X_i's. Without loss of generality we assume $\sigma^2 = 1$ and the cdf. of $\sqrt{\nu}S$ will be denoted by F_{χ_ν}. Again we consider the multiple testing problem $H_i : \vartheta_i = 0$ versus $K_i : \vartheta_i > 0$, $i = 1, \ldots, n$. Let $T_i = X_i/S$, $i = 1, \ldots, n$. Then $(T_1, \ldots, T_n)$ has a multivariate equi-correlated t-distribution. The cdf. (pdf.) of a univariate (central) t-distribution with ν degrees of freedom will be denoted by F_{t_ν} (f_{t_ν}) and a β-quantile of the t_ν-distribution will be denoted by $t_{\nu,\beta}$. With respect to the notation introduced in Section 2.1, Z is replaced by S, $W_X = \Phi$, $W_S(s) = F_{\chi_\nu}(s/\sqrt{\nu})$ and $W_T = F_{t_\nu}$.
Suitable p-values (as a function of s) are defined by $p_i(s) = 1 - F_{t_\nu}(x_i/s)$. Again we add $\vartheta_i = \infty$ to the model such that $p_i = 0$ a. s. if $\vartheta_i = \infty$. We denote the corresponding D-EX(ζ_n) model by D-EX-t(ζ_n). It is outlined in [17] by employing PRDS arguments that the Benjamini-Hochberg bound applies in this model for $\alpha \in (0, 1/2)$.

Computation of the asymptotic False Discovery Rate and the asymptotic Expected Error Rate, respectively, can be done in the D-EX-t(ζ_n) model quite similarly to the description in Chapter 2.3 for the D-EX-N(ζ_n) model. Again, the largest crossing point of the conditional limiting ecdf. $F_\infty(\cdot|s, \zeta)$ of the p-values and Simes' line determines the limiting proportion of rejected hypotheses in the linear step-up procedure. Therefore, we first give the explicit representation of F_∞ in this case, namely

$$F_\infty(t|s,\zeta) = \begin{cases} 1 - \Phi(s \cdot F_{t_\nu}^{-1}(1-t)) & \text{for } \zeta = 1, \\ (1-\zeta) + \zeta\,(1 - \Phi(s \cdot F_{t_\nu}^{-1}(1-t))) & \text{for } \zeta < 1. \end{cases} \tag{2.28}$$

As the following figures show, this function behaves very similarly to its analogue in the normal case treated before. Again, we will first take a closer look at the situation $\zeta = 1$.

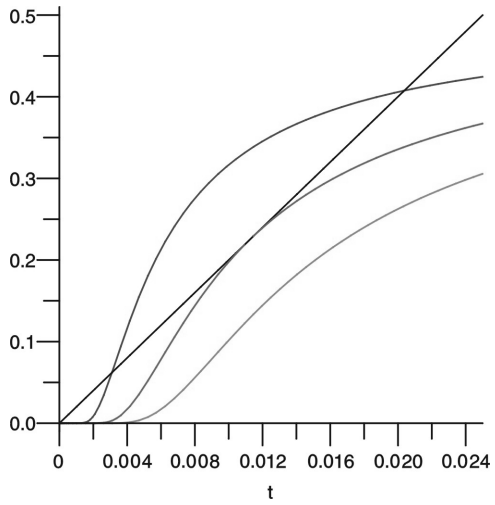

Figure 2.10: $F_\infty(t|s, 1.0)$ for $\nu = 1$ and $s = 0.015$, $s \approx 0.026710$ and $s = 0.04$ on $[0, \alpha/2]$

For $s^2 < (\nu + 1)/\nu$, we again obtain that the limiting ecdf. is first convex and then concave. Due to the same argumentation as in the normal case, we have to investigate the boundary point situation (depicted in Figure 2.10 by the curve in the middle) and determine the corresponding value $\bar{s}$. The limiting FDR then computes as $\text{FDR}_\infty(1) = \mathbb{P}(S \leq \bar{s}) = F_{\chi^2_\nu}(\nu\bar{s}^2)$, because the derivative of $F_\infty(\cdot|s, \zeta)$ in the origin is zero and $F_\infty(t|s, \zeta)$ is non-increasing in s for any fixed t. Formally, we can state the following properties of $F_\infty(t|s) = F_\infty(t|s, 1)$ which will be helpful in what follows.

Lemma 2.38

The function $F_\infty(\cdot|s) = F_\infty(\cdot|s, 1)$ defined in (2.28) has the following properties.

(i) For any fixed $t \in [0, 1]$, $F_\infty(t|s)$ is non-increasing in s.

(ii) $\lim_{t \downarrow 0}(\partial/\partial t)F_\infty(t|s) = 0$.

(iii) Defining $a(s, \nu) = \sqrt{(\nu + 1)/s^2 - \nu}$, $F_\infty(t|s)$ is convex for $t \in [0, F_{t_\nu}(-a(s, \nu))]$ and concave for $t \in [F_{t_\nu}(-a(s, \nu)), 1]$ for $s^2 < (\nu + 1)/\nu$.

In case of $\zeta < 1$, too, most of the essential characteristics of F_∞ are preserved, as the following figure for $\zeta = 0.95$ shows. The crucial difference again consists in the fact that $F_\infty(0|s, \zeta) = 1 - \zeta > 0$ for $\zeta < 1$ such that an LCP larger that zero is guaranteed.

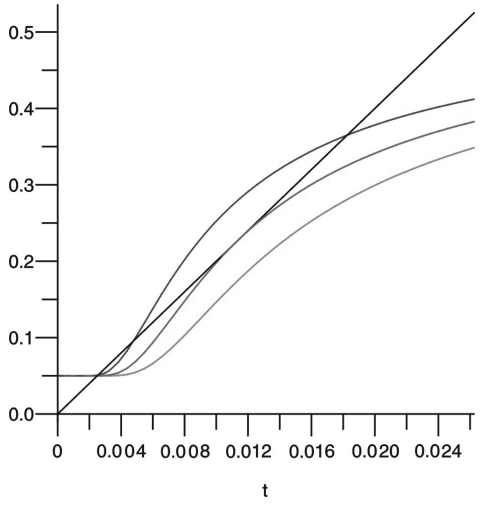

Figure 2.11: $F_\infty(t|s, 0.95)$ for $\nu = 1$ and $s = 0.025$, $s = 0.03180$ and $s = 0.04$ on $[0, \alpha\,(1-\zeta/2)]$

Since the value of s determines if and where F_∞ and Simes' line have a crossing or a boundary point, it is again of interest to give a formal representation of $s(t)$ so that the functional values of the two objects coincide at t. We obtain

$$s(t) = \frac{\Phi^{-1}\left(1/\zeta - t/(\alpha\zeta)\right)}{F_{t_\nu}^{-1}(1-t)} \tag{2.29}$$

and notice that crossing or boundary points (for $\alpha < 1/2$) can consequently only occur in the interval $(\alpha(1-\zeta), \alpha(1-\zeta/2)) \ni t$, because only strictly positive values for s are possible. The limits of $s(t)$ are given in the following lemma.

Lemma 2.39

Let $t_u = \alpha\,(1-\zeta)$ and $t_o = \alpha\,(1-\zeta/2)$. Then $s(t)$ as defined in (2.29) has the following limits:

$$\begin{cases} \lim_{t\downarrow t_u} s(t) = \lim_{t\uparrow t_o} s(t) = 0 & \text{for } \zeta = 1, \\ \lim_{t\downarrow t_u} s(t) = \infty \text{ and } \lim_{t\uparrow t_o} s(t) = 0 & \text{for } \zeta < 1. \end{cases}$$

Furthermore, we give (for the determination of extrema) the derivative of $s(t)$ with respect to t:

$$\frac{d}{dt}\, s(t) = \frac{1}{\left(F_{t_\nu}^{-1}(1-t)\right)^2} \cdot \left(\frac{\Phi^{-1}(1/\zeta - t/(\alpha\zeta))}{f_{t_\nu}(F_{t_\nu}^{-1}(1-t))} - \frac{F_{t_\nu}^{-1}(1-t)}{\alpha\zeta\varphi(\Phi^{-1}(1/\zeta - t/(\alpha\zeta)))} \right).$$

Therefore, for points with horizontal tangent to the curve $s(t)$ the condition

$$\frac{\alpha\zeta\Phi^{-1}(1/\zeta - t/(\alpha\zeta)) \cdot \varphi(\Phi^{-1}(1/\zeta - t/(\alpha\zeta)))}{F_{t_\nu}^{-1}(1-t) \cdot f_{t_\nu}(F_{t_\nu}^{-1}(1-t))} = 1 \tag{2.30}$$

must hold. Condition (2.30) simplifies in the special case $\nu = 1$, because it can in this case entirely be expressed by elementary functions and Φ. Moreover, for $\nu = 1$ the t_1-distribution is the well-known Cauchy distribution so that we devote one subsection to this special case.

2.4.1 The special case $\nu = 1$

It is well known that in case of $\nu = 1$, Student's t-distribution with one degree of freedom coincides with the Cauchy distribution with characteristics as given in the following lemma.

Lemma 2.40 (Cauchy distribution)
Let f denote the pdf. and let F denote the cdf. of the Cauchy distribution. Then it holds:

$$f(y) = \frac{1}{\pi\,(1+y^2)}, \tag{2.31}$$

$$F(y) = \frac{1}{2} + \frac{\arctan(y)}{\pi}, \tag{2.32}$$

$$F^{-1}(y) = \tan(\pi\,(y - \frac{1}{2})) = t_{1,y}. \tag{2.33}$$

Plugging in the thereby obtained quantiles for Student's t-distribution with one degree of freedom into the formal representation of $s(t)$ in (2.29) yields

$$s(t) = \frac{\Phi^{-1}(1/\zeta - t/(\alpha\zeta))}{\tan(\pi(\frac{1}{2} - t))}$$

and, consequently,

$$\frac{d}{dt}\,s(t) = \frac{1}{\tan^2(\pi\,(\frac{1}{2}-t))}\left[\pi(1+\tan^2(\pi\,(\frac{1}{2}-t)))\cdot\Phi^{-1}(\frac{1}{\zeta}-\frac{t}{\alpha\zeta}) - \frac{\tan(\pi\,(\frac{1}{2}-t))}{\alpha\zeta\varphi(\Phi^{-1}(\frac{1}{\zeta}-\frac{t}{\alpha\zeta}))}\right].$$

Therefore, the condition $\frac{d}{dt}\,s(t) = 0$ is equivalent to

$$\pi(1+\tan^2(\pi\,(\frac{1}{2}-t)))\cdot\Phi^{-1}(\frac{1}{\zeta}-\frac{t}{\alpha\zeta}) = \frac{\tan(\pi\,(\frac{1}{2}-t))}{\alpha\zeta\varphi(\Phi^{-1}(1/\zeta - t/(\alpha\zeta)))}. \tag{2.34}$$

Substituting $u := \Phi^{-1}(1/\zeta - t/(\alpha\zeta))$, we get $t = \alpha(1-\zeta\,\Phi(u))$ and (2.34) reads re-formulated

$$\frac{\alpha\zeta\pi u\varphi(u)\cdot\left[1+\tan^2(\pi\,(\frac{1}{2}-\alpha(1-\zeta\,\Phi(u))))\right]}{\tan(\pi(1/2-\alpha(1-\zeta\,\Phi(u))))} = 1. \tag{2.35}$$

Computation of an approximate numerical solution of (2.35) with respect to u and following back substitution enables us immediately to determine an arbitrary precise numerical value of the limiting FDR for $\zeta = 1$ in the Cauchy case of $\nu = 1$. We obtain

$$\begin{aligned} \bar{u} &\approx 0.751350, \\ \bar{t} &\approx 0.011311, \\ \bar{s} &\approx 0.026710 \quad \text{and, finally,} \\ \mathrm{FDR}_\infty(1) &\approx 0.021309. \end{aligned}$$

Remark 2.41

If ν is not equal to 1, the same computational steps have to be carried out using the target equation (2.30). Although conceptionally the same, this is numerically much more cumbersome, because $F_{t_\nu}^{-1}$ for arbitrary ν makes numerical problems.

Remark 2.42

The limiting EER in case of $\zeta = 1$ can (according to Theorem 2.18) be computed as

$$\mathrm{EER}_\infty(1) = \bar{t} \cdot F_{\chi^2_\nu}\left(\nu \bar{s}^2\right) + \int_{\bar{t}/\alpha}^{1/2} F_{\chi^2_\nu}\left(\nu s^2\left(\alpha z\right)\right) dz.$$

Remark 2.43

Due to the limits at the boundaries of the domain of $s(t)$ and its uniform flexional behaviour, the value $\bar{s}$ is a global maximum of $s(t)$ in case of $\zeta = 1$.

Again, this property is no longer preserved in the general case $\zeta < 1$. Then, either two points with horizontal tangent occur (corresponding to a minimum and a maximum of $s(t)$) or there are no roots of the derivative of $s(t)$ at all. Clearly, numerical algorithms can only detect one of the two extrema in the first situation. But since we know the flexion of $s(t)$ in case of $\zeta < 1$, too, a suitable choice of the initial value for the applied algorithm guarantees that the smaller solution in u and consequently the larger solution in t will be found. This is then the maximum of $s(t)$ and is of crucial interest for the further steps towards FDR and EER computation. The remaining computational steps can then be carried out in analogy to Chapter 2.3.

2.4.2 The general case $\nu > 1$ and $\zeta < 1$

In order to determine the local maximum $\bar{s}$ of $s(t)$ (which corresponds to the boundary point situation) in the general case, we apply the substitution $u := F_{t_\nu}^{-1}(1-t)$, equivalent to $t = 1 - F_{t_\nu}(u)$, and achieve the following representations for $s(u)$ and its derivative with respect to the newly introduced variable u:

$$s(u) = \frac{1}{u}\Phi^{-1}\left(\frac{1}{\zeta} - \frac{1 - F_{t_\nu}(u)}{\alpha\zeta}\right) \quad \text{and} \tag{2.36}$$

$$\frac{d}{du}\, s(u) = \frac{1}{u^2}\left(\frac{u f_{t_\nu}(u)}{\alpha\zeta\varphi\left(\Phi^{-1}\left(\frac{1}{\zeta} - \frac{1-F_{t_\nu}(u)}{\alpha\zeta}\right)\right)} - \Phi^{-1}\left(\frac{1}{\zeta} - \frac{1 - F_{t_\nu}(u)}{\alpha\zeta}\right)\right). \tag{2.37}$$

With the transformed bounds $u_u := F_{t_\nu}^{-1}(1-t_o)$ and $u_o := F_{t_\nu}^{-1}(1-t_u)$, the point with horizontal tangent on $s(u)$ with the smaller abscissa u_2 is demanded. Since the roots of (2.37) cannot be determined analytically, a numerical algorithm for maximum searching in a neighborhood of u_u has to be employed. The so obtained numerical value for u_2 provides the value $\bar{s} = s(u_2)$. Via back substitution, we also immediately get the abscissa of the boundary point expressed in t-coordinates as $t_2 = 1 - F_{t_\nu}(u_2)$. Again, this t_2 is one of the two necessary integration bounds

for the expectation formulas expressing $\text{FDR}_\infty(\zeta)$ and $\text{EER}_\infty(\zeta)$. In order to compute the lower bound t_1, we consider the distance $d(u|\bar{s})$ between the conditional limiting ecdf. and Simes' line in the boundary point situation as a function of the transformed variable u and obtain

$$d(u|\bar{s}) = (1-\zeta) + \zeta\ \Phi(-\bar{s}u) - \frac{1 - F_{t_\nu}(u)}{\alpha}.$$

Applying a root finding algorithm to the latter function in a suitable neighborhood of u_o with result u_1 yields $t_1 = 1 - F_{t_\nu}(u_1)$ and with the transformations

$$\begin{aligned} z_i &= 1 - \frac{\alpha\ (1-\zeta)}{t_i} \quad \text{for } i = 1, 2, \text{ and} \\ z_3 &= \frac{\zeta}{2-\zeta}, \end{aligned}$$

the limiting FDR and EER can finally (according to Theorem 2.13) be computed as

$$\begin{aligned} \text{FDR}_\infty(\zeta) &= (z_2 - z_1) \cdot F_{\chi^2_\nu}\left(\nu s^2\left(\frac{\alpha(1-\zeta)}{1-z_1}\right)\right) \\ &\quad + \int_0^{z_1} F_{\chi^2_\nu}\left(\nu s^2\left(\frac{\alpha(1-\zeta)}{1-z}\right)\right) dz + \int_{z_2}^{z_3} F_{\chi^2_\nu}\left(\nu s^2\left(\frac{\alpha(1-\zeta)}{1-z}\right)\right) dz, \\ \text{EER}_\infty(\zeta) &= \frac{t_2 - t_1}{\alpha} \cdot F_{\chi^2_\nu}\left(\nu s^2(t_1)\right) \\ &\quad + \int_{1-\zeta}^{t_1/\alpha} F_{\chi^2_\nu}\left(\nu s^2\left(\alpha t\right)\right) dt + \int_{t_2/\alpha}^{1-\zeta/2} F_{\chi^2_\nu}\left(\nu s^2\left(\alpha t\right)\right) dt. \end{aligned}$$

However, if $s(t)$ decreases monotonously on the entire interval $[t_u, t_o]$ or it holds $\frac{d}{du}\ s(u) > 0$ for all u in the interval $[u_u, u_o]$, respectively, there is no possible choice for s leading to a boundary point situation. Consequently, the largest crossing point of F_∞ and Simes' line can then lie anywhere in the interval $[t_u, t_o]$ and it holds ($t_1 = t_2$)

$$\begin{aligned} \text{FDR}_\infty(\zeta) &= \int_0^{z_3} F_{\chi^2_\nu}\left(\nu s^2\left(\frac{\alpha(1-\zeta)}{1-z}\right)\right) dz, \\ \text{EER}_\infty(\zeta) &= \int_{1-\zeta}^{1-\zeta/2} F_{\chi^2_\nu}\left(\nu s^2\left(\alpha t\right)\right) dt. \end{aligned}$$

Again, an alternative approach towards determining $\bar{s}$ consists of working with the distance function $d(u|s)$ and its derivative with respect to u. In the boundary point situation, both objects must simultaneously equal zero for a tuple $(\bar{u}, \bar{s})$. In case of $\zeta = 1$, this is equivalent to the pair of equations

$$\alpha\Phi(-\bar{s}\bar{u}) = F_{t_\nu}(-\bar{u}) \tag{2.38}$$

$$\bar{s}\alpha\varphi(\bar{s}\bar{u}) = f_{t_\nu}(\bar{u}). \tag{2.39}$$

From (2.38) und (2.39), asymptotic results for $\nu \to \infty$ can be deduced as follows.

Lemma 2.44
Let $\bar{s} = 1 - (-\ln(\alpha))^{1/2}\nu^{-1/2} + o(\nu^{-1/2})$ *and* $\bar{u} = \sqrt{2}\,(-\ln(\alpha))^{\frac{1}{4}}\,\nu^{\frac{1}{4}} + o(\nu^{-1/4})$.
Then it holds:

$$\lim_{\nu\to\infty} \frac{f_{t_\nu}(\bar{u})}{\varphi(\bar{s}\bar{u})} = \alpha.$$

Proof: Noting that $\bar{s}^2 = 1 - 2\,(-\ln(\alpha))^{\frac{1}{2}}\,\nu^{-\frac{1}{2}} + o(\nu^{-1/2})$ and $\bar{u}^2 = 2\,(-\ln(\alpha))^{\frac{1}{2}}\,\nu^{\frac{1}{2}} + o(1)$, we obtain for the product that $\bar{u}^2\bar{s}^2 = 2(-\ln(\alpha))^{\frac{1}{2}}\nu^{\frac{1}{2}} + 4\ln(\alpha) + o(1)$ and, consequently,

$$\varphi(\bar{s}\bar{u}) = \frac{1}{\sqrt{2\pi}}\exp(-(-\ln(\alpha))^{\frac{1}{2}}\nu^{\frac{1}{2}} - 2\ln(\alpha) - o(1)) = \frac{1}{\sqrt{2\pi}\alpha^2\exp(\sqrt{-\ln(\alpha)})^{\sqrt{\nu}}o(1)}.$$

Furthermore, we have

$$f_{t_\nu}(\bar{u}) = \frac{1}{\sqrt{\nu}B(\frac{\nu}{2},\frac{1}{2})}\left(1 + \frac{2(-\ln(\alpha))^{\frac{1}{2}}}{\nu^{\frac{1}{2}}} + o(\nu^{-1})\right)^{-\frac{\nu+1}{2}}$$

and therefore we get for the ratio of both expressions

$$\frac{f_{t_\nu}(\bar{u})}{\varphi(\bar{s}\bar{u})} = \frac{\sqrt{2\pi}}{\sqrt{\nu}B(\frac{\nu}{2},\frac{1}{2})}\alpha^2\exp\left(\sqrt{-\ln(\alpha)} + o(\nu^{-1/2})\right)^{\sqrt{\nu}}\left(1 + \frac{2(-\ln(\alpha))^{\frac{1}{2}}}{\nu^{\frac{1}{2}}} + o(\nu^{-1})\right)^{-\frac{\nu+1}{2}}.$$

Since

$$\lim_{\nu\to\infty}\frac{\sqrt{2\pi}}{\sqrt{\nu}B(\frac{\nu}{2},\frac{1}{2})} = 1,$$

it remains to study the expression

$$g_\alpha(\nu) = \exp\left(\sqrt{-\ln(\alpha)} + o(\nu^{-1/2})\right)^{\sqrt{\nu}}\left(1 + \frac{2(-\ln(\alpha))^{\frac{1}{2}}}{\nu^{\frac{1}{2}}} + o(\nu^{-1})\right)^{-\frac{\nu+1}{2}}. \qquad (2.40)$$

Substituting $n = \sqrt{\nu}$ and $a = \sqrt{-\ln(\alpha)}$ we obtain a new function h_a (say), given by

$$h_a(n) = \left(\exp(a + o(n^{-1}))\right)^n\left(1 + \frac{2a}{n} + o(n^{-2})\right)^{-\frac{n^2+1}{2}},$$

which has the same limiting behaviour for $n \to \infty$ as $g_\alpha(\nu)$ for $\nu \to \infty$. Taking the natural logarithm in the definition of $h_a(n)$ results in

$$\ln(h_a(n)) = n(a + o(n^{-1})) - \frac{n^2+1}{2}\ln\left(1 + \frac{2a}{n} + o(n^{-2})\right).$$

Now we make use of the series expansion

$$\ln(1+\beta) = \sum_{k=0}^{\infty}\frac{(-1)^k\,\beta^{k+1}}{k+1}$$

and obtain the representation

$$\ln(h_a(n)) = na + o(1) - \frac{n^2}{2}\left[\frac{2a}{n} - \frac{2a^2}{n^2} + o(n^{-2})\right] - \frac{1}{2}\ln\left(1 + \frac{2a}{n} + o(n^{-2})\right). \qquad (2.41)$$

Equation 2.41 immediately yields $\lim_{n\to\infty} \ln(h_a(n)) = a^2$ and therefore it holds $\lim_{n\to\infty} h_a(n) = \exp(a^2)$. Back substitution yields

$$\lim_{\nu\to\infty} g_\alpha(\nu) = \frac{1}{\alpha}$$

and this completes the proof. ■

Remark 2.45

An analogous calculation even yields, that for any fixed constant $k \in \mathbb{R}$, it holds:

$$\lim_{\nu\to\infty} \frac{f_{t_\nu}(k\bar{u})}{\varphi(k\bar{s}\bar{u})} = \alpha^{k^2(2-k^2)}. \tag{2.42}$$

Lemma 2.44 shows that the chosen values for $\bar{s}$ and $\bar{u}$ satisfy the condition (2.39) asymptotically, because $\bar{s}$ tends to 1 with ν tending to infinity.
Furthermore, it can be shown that the tuple $(\bar{u}, \bar{s})$ also satisfies (2.38) asymptotically and that it is the unique solution of the pair of equations (2.38) and (2.39). This is the task of the following lemma.

Lemma 2.46

Let $\alpha \in (0, 1/2)$ and define

$$s = s_\nu(x) = 1 - \nu^{-1/2}(-\ln(x))^{1/2} + o(\nu^{-1/2}),\ x \in (0, 1/2).$$

Then, given model D-EX-t(ζ_n) with $\lim_{n\to\infty} \zeta_n = \zeta = 1$, it holds for sufficiently large ν that $F_\infty(\cdot|s_\nu(x))$ has (i) two CP for all $x \in (0, \alpha)$, and, (ii) no CP for all $x \in (\alpha, 1/2)$.

Proof: For $s^2 < (\nu+1)/\nu$, the unique point of inflection of $F_\infty(\cdot|s)$ on $(0, 1/2)$ is given by $t^*(\nu|s) = F_{t_\nu}(-a(s,\nu))$ with $a(s,\nu) = ((\nu+1)/s^2 - \nu)^{1/2}$. Hence, it suffices to show that

$$F_\infty(t^*(\nu|s_\nu(x))|s_\nu(x)) > t^*(\nu|s_\nu(x))/\alpha \text{ for } x \in (0, \alpha)$$

for sufficiently large ν and that the derivative of $F_\infty(\cdot|s_\nu(x))$ in $t = t^*(\nu|s_\nu(x))$ is less than $1/\alpha$ for all $x \in (\alpha, 1/2)$ for sufficiently large ν. Therefore, the assertion follows if

$$\lim_{\nu\to\infty} \frac{F_{t_\nu}(-a(s_\nu(x),\nu))}{\Phi(-s_\nu(x)a(s_\nu(x),\nu))} < \alpha \text{ for } x \in (0, \alpha), \tag{2.43}$$

$$\lim_{\nu\to\infty} \frac{f_{t_\nu}(a(s_\nu(x),\nu))}{s_\nu(x)\varphi(s_\nu(x)a(s_\nu(x),\nu))} > \alpha \text{ for } x \in (\alpha, 1/2). \tag{2.44}$$

For $x_\nu \in (0, \infty)$ with $\lim_{\nu\to\infty} x_\nu^4/\nu = \beta \in [0, \infty]$ it is shown in [88] that

$$\lim_{\nu\to\infty} \frac{f_{t_\nu}(x_\nu)}{\varphi(x_\nu)} = \lim_{\nu\to\infty} \frac{F_{t_\nu}(-x_\nu)}{\Phi(-x_\nu)} = \exp(\beta/4).$$

Note that for $u \to \infty$ and $s \to 1$ it holds (Mills' ratio)

$$\frac{F_{t_\nu}(-u)}{\Phi(-su)} \sim \frac{F_{t_\nu}(-u)}{\Phi(-u)} \frac{\varphi(u)}{\varphi(su)}.$$

It can easily be verified that $\lim_{\nu\to\infty} a(s_\nu(x),\nu)^4/\nu = -4\ln(x)$ and $\lim_{\nu\to\infty} a(s_\nu(x),\nu)^2(1 - s_\nu(x)^2) = -4\ln(x)$. As a consequence, (2.43) follows by noting that

$$\begin{aligned}
&\lim_{\nu\to\infty} \frac{F_{t_\nu}(-a(s_\nu(x),\nu))}{\Phi(-s_\nu(x)a(s_\nu(x),\nu))} \\
&= \lim_{\nu\to\infty} \left[\frac{F_{t_\nu}(-a(s_\nu(x),\nu))}{\Phi(-a(s_\nu(x),\nu))} \frac{\varphi(-a(s_\nu(x),\nu))}{\varphi(-s_\nu(x)a(s_\nu(x),\nu))}\right] \\
&= \exp(-4\ln(x)/4) \lim_{\nu\to\infty} \exp(-\frac{1}{2}a(s_\nu(x),\nu)^2(1 - s_\nu(x)^2)) \\
&= \frac{1}{x}\exp(2\ln(x)) \\
&= x.
\end{aligned}$$

An analogous calculation yields (2.44). Hence, Lemma 2.46 is proved. ■

Noticing that the FDR in case of $\zeta = 1$ can be computed as

$$\mathbb{P}(S \le \bar{s}) = F_{\chi^2_\nu}(\nu\bar{s}^2),$$

we finally obtain that $\text{FDR}_\infty(1)$ for $\nu \to \infty$ in the D-EX-t(ζ_n) model with $\lim_{n\to\infty}\zeta_n = 1$ tends to the same limiting value as the one given in Theorem 2.32 for $\rho \to 0$ in the D-EX-N(ζ_n) model treated in Section 2.3.

Theorem 2.47

In the D-EX-t(ζ_n) model, it holds:

$$\lim_{\nu\to\infty} \mathit{FDR}_\infty(1) = \Phi(-\sqrt{-2\ \ln(\alpha)}).$$

Proof: The χ^2_ν-distribution has expectation $\mu_{\chi^2_\nu} = \nu$ and variance $\sigma^2_{\chi^2_\nu} = 2\nu$. Therefore, the transformed variable

$$\frac{\nu S^2 - \mu_{\chi^2_\nu}}{\sigma_{\chi^2_\nu}} = \frac{\nu S^2 - \nu}{\sqrt{2\nu}}$$

is standardized with expectation 0 und variance 1. Now it holds

$$\frac{\nu\bar{s}^2 - \nu}{\sqrt{2\nu}} = -\sqrt{-2\ln(\alpha)} + o(1)$$

and according to the Central Limit Theorem we finally get

$$\begin{aligned}
\lim_{\nu\to\infty} \mathbb{P}(\frac{\nu S^2 - \nu}{\sqrt{2\nu}} \le \frac{\nu\bar{s}^2 - \nu}{\sqrt{2\nu}}) &= \lim_{\nu\to\infty} \mathbb{P}\left(\frac{\nu S^2 - \nu}{\sqrt{2\nu}} \le -\sqrt{-2\ln(\alpha)} + o(1)\right) \\
&= \Phi(-\sqrt{-2\ln(\alpha)}). \ ■
\end{aligned}$$

Again, we close this section with two figures displaying $\text{EER}_\infty(\zeta)$ and $\text{FDR}_\infty(\zeta)$, respectively, in the D-EX-t(ζ_n) model.

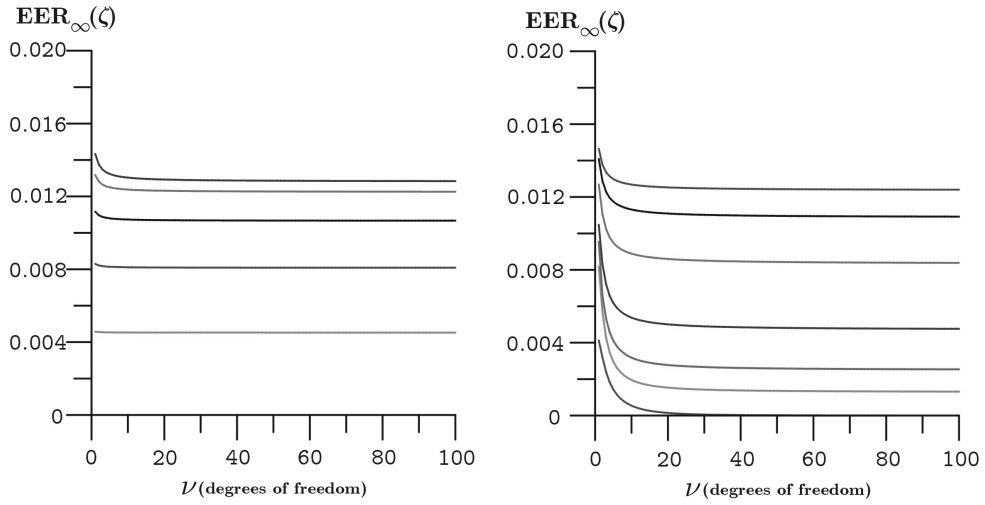

Figure 2.12: $\mathrm{EER}_\infty(\zeta)$ in the D-EX-t(ζ_n) model for $\alpha = 0.05$ and different ζ's.

Figure 2.12 displays $\mathrm{EER}_\infty(\zeta)$ for various values of ζ. It seems that $\mathrm{EER}_\infty(\zeta)$ is decreasing in ν. For $\nu \to \infty$, $\mathrm{EER}_\infty(\zeta)$ again tends to the value $\alpha(1-\zeta)/(1-\alpha\zeta)$ as expected, cf. [91]. Note that $\mathrm{EER}_\infty(\zeta)$ is close to this limit if ν is not too small. As expected, for $\zeta \approx 1/2$ and ν not too small, $\mathrm{EER}_\infty(\zeta)$ is largest.

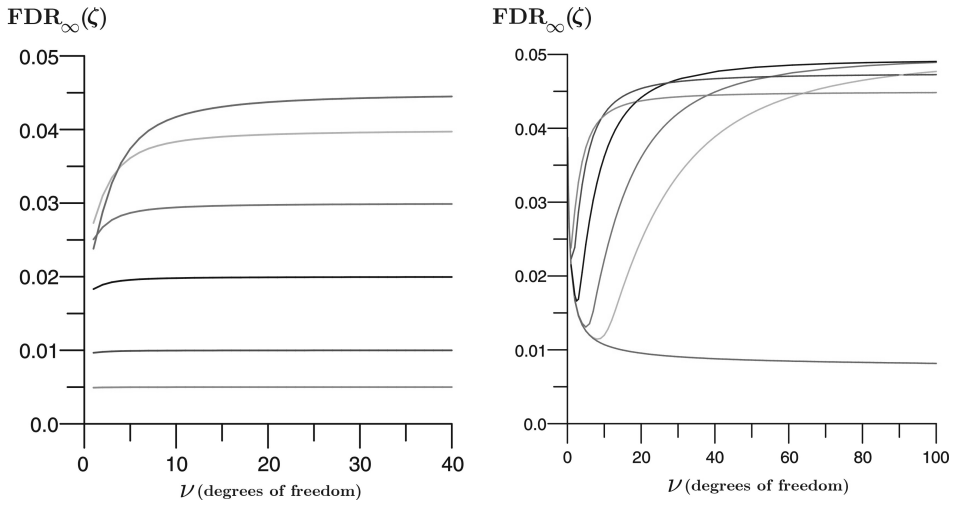

Figure 2.13: $\mathrm{FDR}_\infty(\zeta)$ in the D-EX-t(ζ_n) model for $\alpha = 0.05$ and different ζ's.

In Figure 2.13, $\mathrm{FDR}_\infty(\zeta)$ is displayed for various values of ζ. Except for $\zeta = 1$, the FDR tends to the Benjamini-Hochberg bound $\alpha\zeta$ for increasing degrees of freedom. The limit for $\nu \to 0$ is not clear. In the latter case the density of the t-distribution becomes more and more flat making the computation of $\mathrm{FDR}_\infty(\zeta)$ extremely difficult. But anyhow, looking at $0 < \nu < 1$ has only academic value.

As in the D-EX-N(ζ) model with $\zeta = 1$, $\text{FDR}_\infty(\zeta)$ is the true level of Simes' global test for the intersection hypothesis, cf. [264].

2.5 Conclusions

The investigations in this chapter show that the false discovery proportion FDP$= V_n/[R_n \vee 1]$ of the LSU-procedure can be very volatile in case of dependent p-values, that is, the actual FDP may be much larger (or smaller) than in the independent case. The same is true for V_n, V_n/n, R_n and R_n/n. Under mild assumptions, the ecdf. of the p-values converges to a fixed curve under independence (cf. [91]), which implies convergence of V_n/n and R_n/n to a fixed value. On the other hand, the shape of the ecdf. of the p-values under exchangeability heavily depends on the (realization of the) disturbance variable Z. In the latter case, the limit distribution of V_n/n and R_n/n typically has positive variance. It is often assumed that there may be some kind of "weak dependence" between test statistics (cf., e. g., [275]) being close to independence in some sense. The results in Theorems 2.32 and 2.47 and the numerical calculations reflected in Figures 2.9 and 2.13 suggest that for large n and $\zeta_n \to 1$ small deviations from independence (small ρ or large ν) may result in a substantially smaller FDR than the Benjamini-Hochberg bound. However, simulations for small ρ and large ν show that $\text{FDR}_n(\zeta_n)$ approaches its limit $\text{FDR}_\infty(1)$ only for unrealistically large values of n if $\zeta_n \to 1$ (cf. Appendix A.1). A possible explanation may be that $\lim_{\rho\to 0^+} \text{FDR}_n(1) = \alpha$, $\lim_{\nu\to\infty} \text{FDR}_n(1) = \alpha$, hence the order of limits plays a severe role. Moreover, for small ρ it seems that n has to be very large such that the ecdf. reproduces the shape of F_∞ close to 0. For $\zeta < 1$, the FDR_∞-curves in Figures 2.9 and 2.13 reflect the FDR for realistically large n (e.g. $n = 1000$) very well. The reason is that the shape behavior of F_∞ close to 0 is not that crucial as for $\zeta = 1$.

Section 2.2 shows that the FDR under dependency may also have the same behavior as in the independent case. Therefore, it seems very difficult to predict what happens with EER, FDR and FDP in models with more complicated dependency structure, e.g., in a multivariate normal model with arbitrary covariance matrix. In any case, results of the LSU-procedure, or more general, of any FDR-controlling procedure, should be interpreted with some care under dependency taking into account that the FDR refers to an expectation and that the procedure at hand may lead to much more false discoveries than expected.

Finally, with slight modifications of the methods presented in this chapter one can also treat statistics like $T_i = |X_i - Z|$ or $T_i = |X_i|/Z$. Somewhat more effort will be necessary if the disturbance variable Z is two-dimensional, as for example in $T_i = |X_i - Z_1|/Z_2$.

Chapter 3

A new rejection curve

This chapter deals with a new method for gaining power in a multiple test situation by discussing procedures based on a new and in some sense (cf. Section 3.6) asymptotically optimal rejection curve. The fact that the FDR of the original linear step-up procedure is bounded by $n_0\alpha/n$ if n_0 hypotheses are true and the remaining $n_1 = n-n_0$ hypotheses are false implies that the pre-defined error level is not entirely exhausted for $n_0 < n$ by this method which raises the possibility of improving the Benjamini-Hochberg procedure with regard to power. We will tackle this problem from the perspective of rejection curves, which will be formally introduced in Definition 3.2 below, and no longer use Simes' line for determining the indices of hypotheses to be rejected, but another, more sophisticated function of $t \in [0,1]$, parametrized by the pre-defined FDR-level α.

3.1 Notation and preliminaries

Before we can state the main results of this chapter, we need some additional notation. Especially, we have to distinguish carefully between several probability measures and data models which we formalize in this section.

Notation 3.1 (General setup for Chapter 3)
Let $(\Omega, \mathcal{A}, \{\mathbb{P}_\vartheta : \vartheta \in \Theta\})$ denote a statistical experiment and let $(H_n)_{n\in\mathbb{N}}$ be a sequence of non-empty null hypotheses with $H_n \subset \Theta$. The corresponding alternatives are given by $K_n = \Theta \setminus H_n$. Let $(p_n)_{n\in\mathbb{N}}$ denote a sequence of p-values with $p_n : (\Omega, \mathcal{A}) \longrightarrow ([0,1], \mathcal{B})$, where $\mathcal{B}$ denotes the Borel-σ-field over $[0,1]$. Let $I_0 = I_0(\vartheta) = \{i \in \mathbb{N} : \vartheta \in H_i\}$, $I_1 = I_1(\vartheta) = \mathbb{N} \setminus I_0 = \{i \in \mathbb{N} : \vartheta \in K_i\}$ and $I_{n,j} = I_{n,j}(\vartheta) = I_j \cap \mathbb{N}_n$, $j = 0, 1$. As usual, let a p-value p_i for testing H_i satisfy $0 < \mathbb{P}_\vartheta(p_i \leq x) \leq x$ for all $\vartheta \in H_i$, $i \in \mathbb{N}$ and $x \in (0,1]$. We also assume that for every $\vartheta \in \Theta$ and $i \in I_0(\vartheta)$ there is a probability measure $\mathbb{P}_{\vartheta^i}$ defined on $(\Omega, \mathcal{A})$ for which the sequence $(p_n)_{n\in\mathbb{N}}$ has the same distribution under $\mathbb{P}_{\vartheta^i}$ as the sequence $(p_n^i)_{n\in\mathbb{N}}$ under $\mathbb{P}_\vartheta$, the only difference between these two sequences of p-values being that $p_i^i \equiv 0$. This is a technical assumption which will be used in Section 3.5 for the determination of upper bounds for the FDR.

Notice that the $\mathbb{P}_{\vartheta^i}$'s need not be contained in $\{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$. For a non-empty set $I_0 \subseteq \mathbb{N}$ we denote by I_0' the set $I_0 \setminus \{\min I_0\}$ in the sequel. For notational convenience, we introduce $F_{n,j}$, $j = 0, 1$, as the ecdf.'s of the p-values corresponding to the true ($j = 0$) and the false ($j = 1$) hypotheses, respectively. Finally, let $\mathcal{H}_n = (H_i : i \in \mathbb{N}_n)$ and let $\varphi_{(n)} = (\varphi_i : i \in \mathbb{N}_n)$ denote a non-randomized multiple test procedure for $\mathcal{H}_n$.

The original linear step up-procedure by Benjamini and Hochberg for independent p-values employing Simes' critical values for the family $\mathcal{H}_n = \{H_i : i \in \mathbb{N}_n\}$ of dimension n will consequently be abbreviated by $\varphi_{(n)}^{\text{LSU}}$. As outlined in Remark 2.3, this procedure can be rewritten in terms of the empirical cdf F_n of the p-values $p_1, \ldots, p_n$. Let $t(\varphi_{(n)}^{\text{LSU}}) = \sup\{t \in [0,1] : F_n(t) \geq t/\alpha\}$. Then $\varphi_{(n)}^{\text{LSU}}$ rejects H_i iff $p_i \leq t(\varphi_{(n)}^{\text{LSU}})$. The rejection curve $r_\alpha(t) = t/\alpha$ is Simes' line, cf. Remark 2.3. Notice that $\alpha_{i:n}^{\text{LSU}} = r_\alpha^{-1}(i/n)$. More generally, many multiple test procedures can be described in terms of the ecdf. F_n of the p-values and a rejection curve r defined as below.

Definition 3.2 (Critical value function and rejection curve)
Let $\rho : [0,1] \to [0,1]$ be non-decreasing, continuous with $\rho(0) = 0$ and positive values on $(0,1]$. Define critical values $\alpha_{i:n} = \rho(i/n) \in (0,1]$ for $i = 1, \ldots, n$. We call ρ a critical value function. Moreover, r defined by $r(x) = \inf\{u : \rho(u) = x\}$ for $x \in [0,1]$ ($\inf \emptyset = \infty$), will be called a rejection curve.

For illustrative purposes, a plot of F_n together with the rejection curve r is useful in order to demonstrate the decision procedure. Note that we have (cf. [250]) the following relationship between the ecdf. F_n of distinct p-values $p_1, \ldots, p_n$, the ordered p-values, the critical values $\alpha_{i:n} = \rho(i/n)$ and the rejection curve r:

$$F_n(p_{i:n}) \geq r(p_{i:n}) \text{ if and only if } p_{i:n} \leq \alpha_{i:n}. \tag{3.1}$$

In analogy to the notation in Chapter 2, a point $t = p_{i:n}$ satisfying $F_n(p_{i:n}) \geq r(p_{i:n})$ and $F_n(p_{i+1:n}) < r(p_{i+1:n})$ is called a crossing point (CP) between F_n and r. We consider test procedures which determine one of the CPs as a threshold t^* in order to reject all H_i with $p_i \leq t^*$. It is immediately clear that the proportion $(R_n - V_n)/(n_1 \vee 1)$ of rejected false null hypotheses with respect to all false null hypotheses is non-decreasing in the threshold t^*. Therefore, we look for procedures which maximize t^* for any given set of p-values subject to FDR control, because this leads to maximization of the multiple power defined in Definition 1.3 as the expectation of the aforementioned ratio.
In order to formally express upper bounds and least favorable parameter configurations for the FDR, we finally introduce the Dirac-uniform models as follows.

Definition 3.3 (Dirac-uniform models)
We assume that the measurable space $(\Omega, \mathcal{A})$ is large enough to accommodate probability measures $\mathbb{P}_{I_0}$, $I_0 \subseteq \mathbb{N}$, under which all p-values p_i, $i \in I_0$, are i.i.d. uniformly distributed on $[0,1]$,

and all p_i, $i \in I_1$, follow a Dirac distribution with point mass 1 at 0. We refer to $\mathbb{P}_{I_0}$ as a Dirac-uniform configuration. Under each $\mathbb{P}_{I_0}$, the Extended Glivenko-Cantelli Theorem (cf. [262], p. 105) applies for the ecdf F_n (say) of the p-values, that is,

$$\lim_{n\to\infty} \sup_{t\in[0,1]} |F_n(t) - (\frac{n_1(n)}{n} + \frac{n_0(n)}{n}t)| = 0 \quad [\mathbb{P}_{I_0}], \tag{3.2}$$

where $n_j = n_j(n) = |I_j \cap \mathbb{N}_n|$, $j = 0, 1$. Notice that the $\mathbb{P}_{I_0}$'s need not be contained in $\{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$.

As in Chapter 2, let $\zeta_n = n_0(n)/n$ denote the proportion of true hypotheses among the first n hypotheses. We refer to this situation as the Dirac-uniform finite model, $DU_n(\zeta_n)$.

Now suppose that

$$\lim_{n\to\infty} \zeta_n = \zeta \in [0, 1].$$

Then (3.2) implies that, for n tending to infinity, the empirical distribution function F_n of the observed p-values converges to

$$F_\infty(t|\zeta) = (1 - \zeta) + \zeta t \text{ for all } t \in [0, 1] \ [\mathbb{P}_{I_0}].$$

This situation will be called the Dirac-uniform asymptotic model, $DU_\infty(\zeta)$ for short.

3.2 Motivation and heuristic derivation

Our new rejection curve has the representation

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

Figure 3.1 displays this new rejection curve together with Simes' line and several limiting ecdf.'s $F_\infty(\cdot|\zeta)$ of p-values under Dirac-uniform configurations with limiting proportion ζ of true hypotheses. The motivation for choosing f_α as "asymptotically optimal rejection curve" reveals, if we investigate the points of intersection of each $F_\infty(\cdot|\zeta)$ and f_α on $(0, 1)$ for $\zeta \in (\alpha, 1)$. It holds that

$$F_\infty(t|\zeta) = f_\alpha(t) \text{ iff } t = \frac{\alpha\ (\zeta - 1)}{\zeta\ (\alpha - 1)} = t_\zeta \text{ (say)}.$$

If we now consider a single-step multiple test procedure $\varphi^{\mathrm{ss}(t_\zeta)}$ (say), which rejects all hypotheses with p-values smaller than or equal to t_ζ, we get for the asymptotic FDR of $\varphi^{\mathrm{ss}(t_\zeta)}$ in the $\mathrm{DU}_\infty(\zeta)$ model that

$$\mathrm{FDR}_\zeta^\infty(\varphi^{\mathrm{ss}(t_\zeta)}) = \frac{\zeta t_\zeta}{(1-\zeta)+\zeta t_\zeta} \equiv \alpha, \text{ independent of } \zeta \in (\alpha, 1).$$

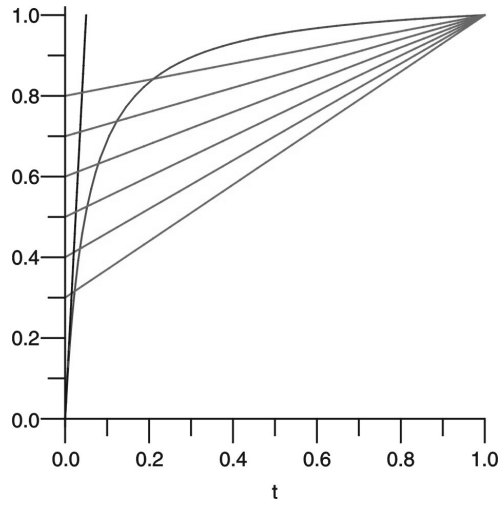

Figure 3.1: Simes-line, f_α and limiting ecdf.'s of Dirac-uniform configurations

For $\zeta \in [0, \alpha)$ one may set $t_\zeta = 1$, which implies that all hypotheses are rejected and $\mathrm{FDR}_\zeta^\infty(\varphi^{\mathrm{ss}(1)}) = \zeta < \alpha$. Since Dirac-uniform configurations can be viewed as least favorable for certain stepwise multiple test procedures (see Section 3.4 below), the latter considerations show that the choice of f_α as rejection curve is asymptotically optimal in the sense that the FDR level α is fully exhausted in the least favorable case. In other words (since $\mathrm{FDR}_\zeta^\infty(\varphi^{\mathrm{ss}(t)})$ and $\bar{\beta}_\vartheta(\varphi^{\mathrm{ss}(t)})$ are both increasing in t), t_ζ is the largest possible threshold for given $\zeta \in (\alpha, 1)$ such that the asymptotic FDR is still controlled by α and it therefore maximizes the asymptotic power under the constraint of asymptotic FDR control. But, since the proportion of true hypotheses is unknown in practice, we have to find suitable test procedures that automatically generate the correct threshold. Typical candidates are stepwise test procedures which choose a CP of F_n and a rejection curve as rejection threshold. Therefore, such stepwise test procedures based on f_α are the topic of the next section.

3.3 Procedures based on the new rejection curve

Before deriving the announced test procedures based on f_α, we start with some properties of this new rejection curve.

Lemma 3.4 (Properties of the new rejection curve)
Comparing the Simes-line $r_\alpha(t) = t/\alpha$ and the new rejection curve $f_\alpha(t)$, we obviously have $r_\alpha(t) > f_\alpha(t)$ for $t > 0$ and the derivative in $t = 0$ is $1/\alpha$ for both curves. Moreover, notice that f_α obeys the symmetry property

$$f_\alpha^{-1}(t) = \frac{\alpha t}{1 - (1-\alpha)t} = 1 - f_\alpha(1-t) \text{ for all } t \in [0,1],$$

where f_α^{-1} denotes the inverse function of f_α. Clearly, f_α^{-1} is a critical value function. Furthermore, f_α is differentiable on $[0,1]$ and it holds

$$\frac{d}{dt}f_\alpha(t) = \frac{\alpha}{(t(1-\alpha)+\alpha)^2}.$$

The question is how to implement the new rejection curve f_α, which will be called the asymptotically optimal rejection curve (AORC), not only in the Dirac-uniform models but also in more general models into a multiple test procedure which controls the FDR-level α strictly or at least asymptotically. The critical values induced by f_α are given by

$$\alpha_{i:n} = f_\alpha^{-1}(\frac{i}{n}) = \frac{\frac{i}{n}\alpha}{1-\frac{i}{n}(1-\alpha)} = \frac{i\alpha}{n-i(1-\alpha)},\ i=1,\ldots,n. \tag{3.3}$$

Remember that

$$F_n(p_{i:n}) \geq f_\alpha(p_{i:n}) \text{ if and only if } p_{i:n} \leq \alpha_{i:n}.$$

It is tempting to use $\alpha_{1:n} \leq \cdots \leq \alpha_{n:n}$ in a step-up procedure for testing n hypotheses. Unfortunately, $\alpha_{n:n} = 1$, so that this procedure always rejects all hypotheses. This pitfall is due to the fact that $f_\alpha(1) = F_n(1)$. Therefore, we need some adjustment with respect to f_α or the step-up procedure. In the remainder of this section, we consider some candidates for asymptotic FDR control avoiding the aforementioned pitfall. They will again be studied in Corollary 3.19 which proves asymptotic FDR control for the procedures proposed in the following three examples.

Example 3.5 (Step-up-down procedures)
An interesting class of procedures based on critical values $0 < \alpha_{1:n} \leq \cdots \leq \alpha_{n:n} \leq 1$ are step-up-down (SUD) procedures introduced in [280] and studied in [237] in terms of FDR control. For $\lambda_n \in \mathbb{N}_n$ a SUD-procedure $\varphi^{\text{SUD}}_{n,\lambda_n} = (\varphi_1,\ldots,\varphi_n)$ of order λ_n is defined as follows. If $p_{\lambda_n:n} \leq \alpha_{\lambda_n:n}$, set $m_n = \max\{j \in \{\lambda_n,\ldots,n\} : p_{i:n} \leq \alpha_{i:n} \text{ for all } i \in \{\lambda_n,\ldots,j\}\}$, whereas for $p_{\lambda_n:n} > \alpha_{\lambda_n:n}$, put $m_n = \sup\{j \in \{1,\ldots,\lambda_n\} : p_{j:n} \leq \alpha_{j:n}\}$ ($\sup\emptyset = -\infty$). Define $\varphi_i = 1$ if $p_i \leq \alpha_{m_n:n}$ and $\varphi_i = 0$ otherwise ($\alpha_{-\infty:n} = -\infty$). Note that $\lambda_n = 1$ yields a step-down (SD) procedure and $\lambda_n = n$ yields a SU-procedure. The order of a SUD-procedure can be defined in terms of a fixed parameter $\lambda \in [0,1]$ by setting $\lambda_n = \inf\{j \in \mathbb{N}_n : \alpha_{j:n} \geq \lambda\}$ ($\inf\emptyset = n$). Then $\lambda = 0$ ($\lambda = 1$) corresponds to a SD- (SU-) procedure. A SUD-procedure of order $\lambda_n = \lambda_n(\lambda),\ \lambda \in [0,1)$, based on f_α resolves the problems around the point $t = 1$ in an elegant way. It is obvious in view of Lemma 3.4 that in case of $\lambda \geq \alpha$ the new step-up-down procedure based on f_α rejects at least all hypotheses rejected by the linear step-up procedure, possibly more. Therefore, one cannot expect that the FDR-level is controlled in the finite case. However, it will be shown that the FDR is controlled asymptotically. Note that $\varphi^{\text{SUD}}_{n,\lambda_n}$ is component-wise non-decreasing in λ. For the computation of the starting index $\lambda_n(\lambda)$, notice that

$$\alpha_{j:n} \geq \lambda \text{ iff } \frac{j\alpha}{n-j(1-\alpha)} \geq \lambda \text{ iff } j \geq \frac{n\lambda}{\alpha+\lambda(1-\alpha)},$$

such that we choose the starting index $\lambda_n(\lambda) = \left\lceil \frac{n\lambda}{\alpha+\lambda(1-\alpha)} \right\rceil$. For example, if we set $\lambda = 1/2$, we obtain $\lambda_n(1/2) = \left\lceil \frac{n}{1+\alpha} \right\rceil$ and the choice $\lambda = \alpha$ leads to $\lambda_n(\alpha) = \left\lceil \frac{n}{2-\alpha} \right\rceil$.

Application 3.6 (Applications 2.4 and 2.5 revisited)
We return to our introductory example applications 2.4 and 2.5. In both cases, the actual ecdf.'s have concave shape such that we obtain a unique crossing point of F_{393} in the proteomics example and F_{7457} in the adenocarcinoma example, respectively, with f_α on $(0, 1)$ (cf. Figures 3.2 and 3.3 below). In such cases, all step-up-down procedures with parameters $\lambda \in (0, 1)$ are equivalent.

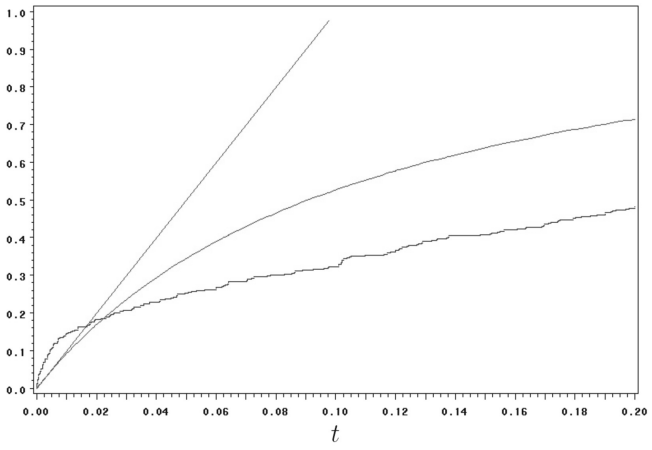

Figure 3.2: Proteomics example: Simes' line, optimal rejection curve and ecdf. of p-values on $[0, 0.2]$

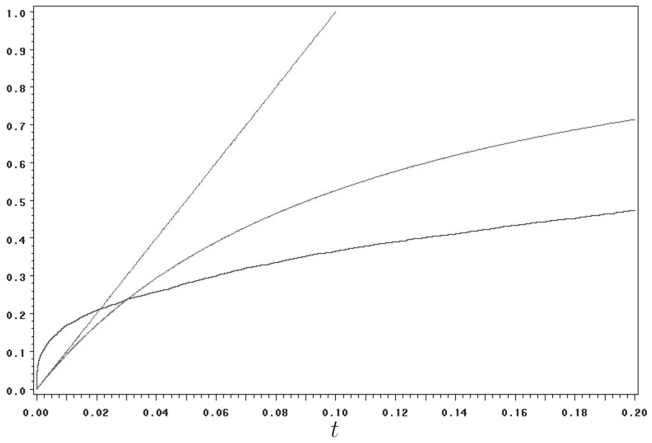

Figure 3.3: Notterman example: Simes' line, optimal rejection curve and ecdf. of p-values on $[0, 0.2]$

In order to demonstrate the differences between the results of the linear step-up procedure and a step-up-down procedure based on f_α, the figures only cover the subinterval $[0, 2\alpha] \ni t$ for a better resolution. It becomes apparent that a step-up-down procedure based on f_α leads to more rejections than the linear step-up procedure.

More specifically, with the Benjamini-Hochberg procedure, we got 47 rejections in case of $\alpha_1 = 0.05$ and 64 rejections in case of $\alpha_2 = 0.1$ in our proteomics example. Using a step-up-down procedure of order $\lambda_n = \left\lfloor \frac{n}{2-\alpha_i} \right\rfloor$, $i = 1, 2$, with critical values based on the AORC, we obtain 53 rejections for α_1 and 74 rejections for α_2.

In the adenocarcinoma example, the hypotheses corresponding to the 1582 smallest p-values got rejected by φ^{LSU}, while a step-up-down procedure of order $\lambda_n = \left\lfloor \frac{n}{2-\alpha} \right\rfloor$ with critical values based on the AORC with $\alpha = 0.1$ rejects 1772 hypotheses. The thresholding values are 0.0212 for the Benjamini-Hochberg procedure and 0.0303 for the step-up-down procedure based on $f_{0.1}$.

Example 3.7 (Adjusted step-up procedures based on f_α)
As noted before, a step-up procedure based on $\alpha_{i:n} = f_\alpha^{-1}(i/n)$ cannot work. Therefore, some adjustment of f_α in a step-up procedure is necessary. We first consider the case where the adjusted rejection curve f_α^{adj} satisfies that $f_\alpha^{\text{adj}}(x)/x$ is non-increasing in x, an important property for the calculation of the FDR. One may specify some $\kappa \in (0, 1)$ and define a new rejection curve

$$f_{\alpha,\kappa}^{\text{adj}}(x) = \begin{cases} f_\alpha(x), & 0 \le x < \kappa, \\ h(x), & \kappa \le x \le 1, \end{cases} \tag{3.4}$$

such that $f_\alpha^{\text{adj}}(x)/x$ is non-increasing in x and $f_\alpha^{\text{adj}}(x^*) = 1$ for some $x^* < 1$. For example, one may choose $h(x) = ax + b$ with $a \le f_\alpha(\kappa)/\kappa$ (which implies $x^* \ge \kappa/f_\alpha(\kappa)$) and $f_\alpha(\kappa) = h(\kappa)$. We consider two possible choices of h (h_1 and h_2 say) and refer to the resulting rejection curves as $f_{\alpha,\kappa}^{(i)},\ i = 1, 2$. Let

$$\begin{aligned} h_1(x) &= f'_\alpha(\kappa)(x - \kappa) + f_\alpha(\kappa) \\ &= \frac{\alpha}{(\alpha + \kappa(1-\alpha))^2} x + \frac{\kappa^2(1-\alpha)}{(\alpha + \kappa(1-\alpha))^2},\ x \in [\kappa, 1). \end{aligned}$$

Then $h_1'(x) = f'_\alpha(\kappa)$, $h_1(\kappa) = f_\alpha(\kappa)$ and $h_1(x^*) = 1$ for $x^* = \kappa(1-\alpha)(2-\kappa) + \alpha$. If we want to pre-define x^*, we have to choose $\kappa = \kappa(x^*) = \arg\left(f_\alpha(x) + f'_\alpha(x)(x^* - x) = 1\right)$, leading to

$$\kappa = 1 - \sqrt{\frac{1 - x^*}{1 - \alpha}}.$$

The resulting modified curve and modified critical values are given by

$$
\begin{aligned}
f^{\text{adj}}_{\alpha,\kappa}(x) &= \begin{cases} f_\alpha(x) & \text{for } x \le \kappa, \\ f_\alpha(\kappa) + (x-\kappa) f'_\alpha(\kappa) & \text{for } x > \kappa, \end{cases} \\
{f^{\text{adj}}_{\alpha,\kappa}}^{-1}(u) &= \begin{cases} \frac{u\alpha}{1-u(1-\alpha)} & \text{for } u \le f_\alpha(\kappa), \\ 1/f'_\alpha(\kappa)\,(u + \kappa f'_\alpha(\kappa) - f_\alpha(\kappa)) & \text{for } u > f_\alpha(\kappa), \end{cases} \\
\alpha^{\text{adj}}_{i:n} = {f^{\text{adj}}_{\alpha,\kappa}}^{-1}(i/n) &= \begin{cases} \frac{i\alpha}{n-i(1-\alpha)} & \text{for } i/n \le f_\alpha(\kappa), \\ 1/f'_\alpha(\kappa)\,(i/n + \kappa f'_\alpha(\kappa) - f_\alpha(\kappa)) & \text{for } i/n > f_\alpha(\kappa). \end{cases}
\end{aligned}
$$

The largest possible slope of h in (3.4) is $a = f_\alpha(\kappa)/\kappa$. This leads to the second choice, that is, $h_2(x) = x f_\alpha(\kappa)/\kappa$. This is close to the truncated step-up procedure in Example 3.8 below. Note that $h_2(x^*) = 1$ if $x^* = \kappa(1-\alpha) + \alpha$. For example, suppose that $\kappa = f_\alpha^{-1}(i/n)$ for some fixed $i \in \mathbb{N}_n$. Then the step-up critical values are given by

$$
\gamma_{j:n} = \begin{cases} f_\alpha^{-1}(j/n), & 1 \le j \le i, \\ \frac{j}{n}\frac{\kappa}{f_\alpha(\kappa)}, & i+1 \le j \le n. \end{cases}
$$

Example 3.8 (Truncated step-up procedures based on f_α)
Let $\kappa \in (0,1)$ be fixed and define

$$
\rho_\alpha(x) = \begin{cases} f_\alpha^{-1}(x), & 0 \le x \le f_\alpha(\kappa), \\ \kappa, & f_\alpha(\kappa) < x \le 1. \end{cases}
$$

With $\gamma_{i:n} = \min\{f_\alpha^{-1}(i/n), \kappa\}$ we have $\gamma_{j:n} = \rho_\alpha(i/n)$ for $j = i, \ldots, n$. Hence, the truncated step-up procedure is well defined in terms of ρ_α. It is worth mentioning that this type of procedure differs substantially from the adjusted procedures discussed in Examples 3.5 and 3.7, because the monotonicity behavior of the ratio $\rho_\alpha(x)/x$ changes at $x = f_\alpha(\kappa)$, which makes FDR calculation much subtler.

In Figure 3.4, the three possible adjustments mentioned in Examples 3.7 and 3.8 are illustrated. The parameters κ_1, κ_2 and κ_3 are chosen such that all three curves have the property that no hypothesis with p-value larger than 0.5 can be rejected (no critical value larger than 0.5 occurs).

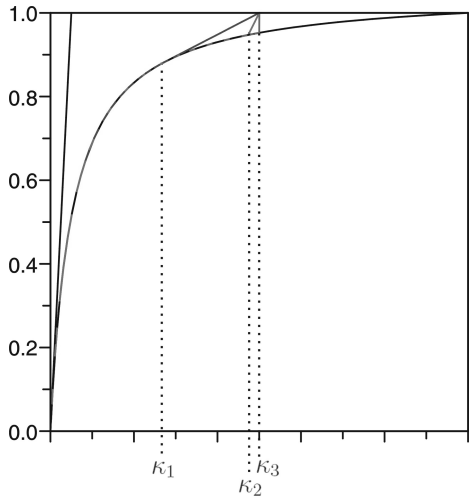

Figure 3.4: Simes-line, $f^{(1)}_{\alpha,\kappa_1}$, $f^{(2)}_{\alpha,\kappa_2}$, the truncated version of f_α with $\kappa_3 = 1/2$ and f_α

3.4 LFC results and upper FDR bounds

Suppose that R_n and $\varphi_{(n)}$, respectively, are defined in terms of p-values $p_1, \ldots, p_n$ and critical values $\alpha_{i:n} = \rho(i/n)$ for some critical value function ρ and consider the following three sets of possible assumptions.
The first two assumptions concern the structure of the test procedure (test assumptions):

(T1) $\forall i \in \mathbb{N}_n : p_i \leq \alpha_{1:n}$ implies $\varphi_i = 1$.

(T2) $\forall j \in \mathbb{N}_n : R_n = j$ implies $\forall i \in \mathbb{N}_n : [\, \varphi_i = 1 \Leftrightarrow p_i \leq \alpha_{j:n} \,]$.

The second set of assumptions concerns properties of distributions of p-values and R_n (distributional assumptions):

(D1) $\forall \vartheta \in \Theta : \forall j \in \mathbb{N}_n : \forall i \in I_{n,0}(\vartheta) : \mathbb{P}_\vartheta(R_n \geq j | p_i \leq t)$ is non-increasing in $t \in (0, \alpha_{j:n}]$.

(D2) $\forall \vartheta \in \Theta : \forall j \in \mathbb{N}_n : \forall i \in I_{n,0}(\vartheta) : \forall t \in (0, \alpha_{j:n}] : \mathbb{P}_\vartheta(R_n \geq j | p_i \leq t) \leq \mathbb{P}_{\vartheta^i}(R_n \geq j)$.

(D3) $\forall \vartheta \in \Theta : \forall i \in I_{n,0}(\vartheta) : p_i \sim U([0,1])$.

Finally we have two possible independence assumptions:

(I1) $\forall \vartheta \in \Theta :$ The p_i's, $i \in I_{n,0}(\vartheta)$, are iid.

(I2) $\forall \vartheta \in \Theta : (p_i : i \in I_{n,0})$ and $(p_i : i \in I_{n,1})$ are independent random vectors.

The simple structure of step-up tests often simplifies derivations concerning properties of these tests. If $\varphi_{(n)}$ is a step-up-down test, the properties of a step-up-test remain valid in the step-up

branch of such a procedure. For example, it is important to note (cf. [237], p. 248) that in case of a step-up-down test of order λ_n and assuming (D3) and (I1)-(I2), we get for all $\vartheta \in \Theta$ and all $i \in I_{n,0}(\vartheta)$

$$\forall j = 1, \ldots, \lambda_n : \ \forall t \in (0, \alpha_{j:n}] : \mathbb{P}_\vartheta(R_n \geq j | p_i \leq t) \quad = \quad \mathbb{P}_{\vartheta^i}(R_n \geq j), \tag{3.5}$$

$$\forall j = 1, \ldots, \lambda_n : \ \forall t \in (0, \alpha_{j:n}] : \mathbb{P}_\vartheta(R_n = j | p_i \leq t) \quad = \quad \mathbb{P}_{\vartheta^i}(R_n = j). \tag{3.6}$$

For $\lambda_n = n$, i.e., for a step-up test, we even get

$$\forall j = 1, \ldots, n : \ \forall t \in (0, \alpha_{j:n}] : \mathbb{P}_\vartheta(R_n \geq j | p_i \leq t) = \mathbb{P}_{\vartheta^i}(R_n \geq j).$$

Assumptions (T1) and (T2) concern possible structures of test procedures. Step-up-down tests satisfy both of these assumptions.

The monotonicity assumption in (D1) is somewhat weaker than the PRDS-assumption (PRDS: positive regression dependency on subsets). More precisely, from the $\mathbb{P}_\vartheta^{p_i}$-almost sure antitonicity of the factorized conditional probability $\mathbb{P}_\vartheta(R_n \geq j | p_i = t)$ in $t \in [0, \alpha_{j:n}]$ we obtain the property formulated in (D1), where the equality in the condition is replaced by an inequality. This type of conclusion is indicated in [171] and can be proved in an easy way by using Wijsman's inequality, cf. [306]. Anyhow, (D1) is the decisive condition for dependent p-values in order to prove FDR-control or to derive upper bounds for the FDR. Examples of distributions being PRDS are extensively studied in [17] and [237]. Important examples are multivariate normal distributions with positive correlations and (absolute valued) multivariate t-distributions, cf. the discussion of the examples in Chapter 2.

Property (D2) will only be used under (I1) and (I2), i.e., if the p-values are independent, and is an important tool for deriving LFC results. In case of dependency, (D2) is often violated.

Assumptions (D3) and (I1) concern the distribution of p-values under the corresponding null hypotheses.

To demonstrate the usefulness and the essentiality of the derived set of assumptions and the equalities (3.5) and (3.6), we present a new proof for FDR control of the classical Benjamini-Hochberg procedure which unifies, simplifies and slightly extends the results and the proofs given in [17] and [237], respectively.

Theorem 3.9

Let $\alpha \in (0, 1)$ and let $\varphi_{(n)}$ be a multiple test procedure for $\mathcal{H}_n$ defined in terms of Simes' critical values $\alpha_{i:n} = i\alpha/n$, $i = 1, \ldots, n$. Let $\vartheta \in \Theta$ such that n_0 hypotheses are true and the remaining ones are false. If (T1), (T2) and (D1) are satisfied, then

$$FDR_\vartheta(\varphi_{(n)}) \leq \frac{n_0}{n}\alpha,$$

with " $=$ " if $\varphi_{(n)}$ is a step-up test and (D3), (I1) and (I2) are satisfied in addition.

Proof: Consider the following chain of (in)equalities:

$$\begin{aligned}
\mathrm{FDR}_\vartheta(\varphi_{(n)}) &= \sum_{i\in I_{n,0}(\vartheta)} \sum_{j=1}^{n} \frac{1}{j}\mathbb{P}_\vartheta(R_n = j, \varphi_i = 1) \\
&= \sum_{i\in I_{n,0}(\vartheta)} \sum_{j=1}^{n} \frac{1}{j}\mathbb{P}_\vartheta(p_i \le \alpha_{j:n})\mathbb{P}_\vartheta(R_n = j | p_i \le \alpha_{j:n}) \qquad (3.7) \\
&\le \sum_{i\in I_{n,0}(\vartheta)} \sum_{j=1}^{n} \frac{\alpha_{j:n}}{j}\mathbb{P}_\vartheta(R_n = j | p_i \le \alpha_{j:n}) \qquad (3.8) \\
&\le \sum_{i\in I_{n,0}(\vartheta)} [\alpha_{1:n}\mathbb{P}_\vartheta(R_n \ge 1 | p_i \le \alpha_{1:n}) \qquad (3.9) \\
&\quad + \sum_{j=2}^{n} \left[\frac{\alpha_{j:n}}{j} - \frac{\alpha_{j-1:n}}{j-1}\right] \mathbb{P}_\vartheta(R_n \ge j | p_i \le \alpha_{j:n})] \\
&= \frac{n_0}{n}\alpha. \qquad (3.10)
\end{aligned}$$

Equation (3.7) holds under (T2), and "=" holds in (3.8) if (D3) holds. Inequality (3.9) holds under the assumption (D1) with "=" if $\varphi_{(n)}$ is a step-up test and (D3), (I1) and (I2) hold. Finally, (3.10) is a consequence of (T1). ■

Remark 3.10

The key step in the proof is (3.9), where $\mathbb{P}_\vartheta(R_n \ge j | p_i \le \alpha_{j-1:n})$ is replaced by $\mathbb{P}_\vartheta(R_n \ge j | p_i \le \alpha_{j:n})$ for $j = 2, \ldots, n$ according to assumption (D1). In case of dependency or in case of a non-step-up test the difference between these quantities may sum up to a considerable amount, that is, the FDR may be much smaller than the upper bound $n_0\alpha/n$ in such cases. For a detailed investigation of the latter phenomenon, cf. [86].

One of the main problems in order to ensure FDR-control of a multiple test procedure is to find least favorable parameter configurations (LFCs). Obviously, LFCs are no issue for the LSU procedure if (D3), (I1) and (I2) hold true. To date it looks like that step-up procedures are easier to cope with than step-down or step-up-down procedures. One reason for this is that Dirac-uniform configurations can often be viewed as least favorable for certain step-up procedures. This fact is based on the following important result.

Theorem 3.11 (Benjamini and Yekutieli (2001), cf. [17])
If (D3), (I1) and (I2) are fulfilled, a step-up procedure with critical values $\alpha_{1:n} \le \cdots \le \alpha_{n:n}$ has the following properties:
(1) If the ratio $\alpha_{i:n}/i$ is increasing in i, as $(p_i : i \in I_{n,1})$ increases stochastically, the FDR decreases.

(2) If the ratio $\alpha_{i:n}/i$ is decreasing in i, as $(p_i : i \in I_{n,1})$ increases stochastically, the FDR increases.

Hence, under the assumptions of Theorem 3.11, the Dirac-uniform configurations, where all p-values under alternatives are almost surely 0, can be viewed as LFCs if the ratio $\alpha_{i:n}/i$ is increasing in i. More precisely, on the parameter subspace, where exactly n_0 (n_1) hypotheses are true (false), the FDR becomes largest if the p-values under alternatives are almost surely 0. Therefore, it suffices to consider all Dirac-uniform configurations in order to check whether the FDR is controlled at level α. Notice that the critical values induced by f_α as given in (3.3) fulfill the important ratio condition.

Unfortunately, the method of proof given in [17] does not seem to work for SD and SUD- procedures. However, we show below that Dirac-uniform configurations often provide upper bounds. To this end, we define $q(x) = \rho(x)/x$ for all $x \in (0, 1]$ and assume that

$$q(0) = \limsup_{x \downarrow 0} q(x) < \infty. \tag{3.11}$$

Moreover, we define the (continuous) function $\overline{q}$ by $\overline{q}(x) = \max_{0 \le t \le x} q(t)$, $x \in [0, 1]$. Hence, $\overline{q}$ is the upper isotonic envelope or, in other words, the least isotonic majorant of q. For the derivation of upper FDR bounds, we now introduce the following additional conditions.

(A1) If $(p_1, \ldots, p_n)$ is stochastically not greater under $\vartheta_1 \in \Theta$ than under $\vartheta_2 \in \Theta$, then $\varphi_{(n)}$ is stochastically not greater under $\vartheta_2 \in \Theta$ than under $\vartheta_1 \in \Theta$.

(A2) $q = \overline{q}$, that is, $\rho(x)/x$ is non-decreasing for $x \in (0, 1]$.

Note that $\alpha_{i:n}/i$ is non-decreasing in i if (A2) holds. In case that ρ is differentiable on $(0, 1)$, (A2) is equivalent to $\rho'(x) \ge q(x)$ for $x \in (0, 1)$. Clearly, under (A2), $q(0)$ can and will be defined as $\lim_{x \downarrow 0} q(x)$. In what follows, $\overline{q}$ is essential in deriving upper bounds for the FDR. If (A2) is violated, the bounds for the FDR based on $\overline{q}$ may be not that sharp. For example, this is the case for the truncated step-up procedure introduced in Example 3.8.

The following theorem is the main result of this section and a valuable tool for proving FDR control of SUD-tests. It provides upper bounds for the FDR of stepwise test procedures under independence of the p-values. For SU-tests, these bounds are sharp if Dirac-uniform configurations belong to the set of possible data models.

Theorem 3.12

Let $\vartheta \in \Theta$ such that $n_0 \in \mathbb{N}$ hypotheses are true and the remaining ones are false. Let $i_0 = \min I_0$ (and $I_0' = I_0 \setminus \{i_0\}$ as defined before). If (T1)-(I2) are satisfied, then

$$\begin{aligned} \mathit{FDR}_\vartheta(\varphi_{(n)}) &\le \frac{n_0}{n} \sum_{j=1}^{n} \overline{q}(j/n) \mathbb{P}_{\vartheta^{i_0}}(R_n/n = j/n) && (3.12) \\ &= \frac{n_0}{n} \mathbb{E}_{\vartheta^{i_0}} \overline{q}(R_n/n), && (3.13) \end{aligned}$$

with equality in (3.12) if $\varphi_{(n)}$ is a step-up test and (A2) holds in addition. If (T1)-(I2) and (A1) are fulfilled, then

$$FDR_\vartheta(\varphi_{(n)}) \leq \frac{n_0}{n}\mathbb{E}_{I_0'}\overline{q}(R_n/n). \tag{3.14}$$

Proof: Let $b_j = \mathbb{P}_\vartheta(R_n \geq j | p_{i_0} \leq \alpha_{j:n})$ and $\Delta\overline{q}(j/n) = \overline{q}(j/n) - \overline{q}((j-1)/n)$ for $j = 1, \ldots, n$. Then, proceeding as in the proof of Theorem 3.9 we get for fixed $\vartheta \in \Theta$ under (D1)-(D3), (I1) and (I2)

$$\begin{aligned}
\mathrm{FDR}_\vartheta(\varphi_{(n)}) &= \frac{n_0}{n}\sum_{j=1}^{n} q(j/n)\mathbb{P}_\vartheta(R_n = j | p_{i_0} \leq \alpha_{j:n}) \\
&\leq \frac{n_0}{n}\sum_{j=1}^{n} \overline{q}(j/n)\mathbb{P}_\vartheta(R_n = j | p_{i_0} \leq \alpha_{j:n}) \qquad (3.15) \\
&\leq \frac{n_0}{n}\left[\overline{q}(1/n)b_1 + \sum_{j=2}^{n}\Delta\overline{q}(j/n)b_j\right] \qquad (3.16) \\
&\leq \frac{n_0}{n}\left[\overline{q}(1/n)\mathbb{P}_{\vartheta^{i_0}}(R_n \geq 1) + \sum_{j=2}^{n}\Delta\overline{q}(j/n)\mathbb{P}_{\vartheta^{i_0}}(R_n \geq j)\right] \qquad (3.17) \\
&= \frac{n_0}{n}\sum_{j=1}^{n}\overline{q}(j/n)\mathbb{P}_{\vartheta^{i_0}}(R_n/n = j/n),
\end{aligned}$$

which proves (3.12). In view of $\mathbb{P}_{\vartheta^{i_0}}(R_n > 0) = 1$ according to (T1), (3.13) follows immediately. If $\varphi_{(n)}$ is a step-up test, which implies (3.5) for $\lambda_n = n$, we have equality in (3.16) and (3.17), hence $q = \overline{q}$ yields equality in (3.15). Finally, in order to prove (3.14), we use the same argumentation as in the proof of Theorem 3.11 given in [17], i.e., that stochastic increase in the distribution of the random vector $(p_1, \ldots, p_n)$ can be characterized by the increase of the expectation of all non-decreasing functions (in case the expectation exists). To this end, we note that obviously $R_n = |\{i \in \mathbb{N}_n : \varphi_i = 1\}|$ is a non-decreasing function of $\varphi_{(n)}$ and hence, due to (A1), is stochastically non-increasing in $(p_1, \ldots, p_n)$. The isotonicity of $\overline{q}$ completes the proof. ■

Inequality (3.14) will be a helpful tool in order to calculate upper FDR bounds and to prove FDR control, because it only makes use of the distribution of R_n under Dirac-uniform configurations. Especially for SUD-tests, this distribution can be handled analytically.

3.5 Asymptotic FDR control for procedures based on the AORC

This section deals with conditions for asymptotic FDR control for procedures based on the new rejection curve. A major result will be that the example procedures presented in Section 3.3 control the FDR asymptotically. Theorems 3.13 and 3.16 provide sufficient conditions for asymptotic

FDR control. If the underlying procedure leads to a determinable limiting proportion of rejected hypotheses, Theorems 3.15 and 3.17 even give explicit values for the resulting FDR.

Theorem 3.13

Suppose $\varphi_{(n)}$ is based on $\rho \leq f_\alpha^{-1}$ and that (T1)-(I2) and (A1) are fulfilled. If for all non-empty sets $I_0 \subseteq \mathbb{N}$ and all subsequences $(n_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$ with $\lim_{k\to\infty} \zeta_{n_k} = \zeta$ for some $\zeta \in [0,1]$ it holds

$$\limsup_{k\to\infty} \frac{R_{n_k}}{n_k} \leq f_\alpha(t_\zeta) \;\; [\mathbb{P}_{I_0'}], \tag{3.18}$$

then

$$\limsup_{n\to\infty} \sup_{\vartheta\in\Theta} \mathit{FDR}_\vartheta(\varphi_{(n)}) \leq \alpha. \tag{3.19}$$

Proof: Let, for notational convenience, $\mathbb{P}_{m,n}$ refer to a Dirac-uniform configuration such that the first m p-values are iid uniformly distributed and the remaining ones follow a Dirac distribution with point mass in 0, $0 \leq m \leq n$, $n \in \mathbb{N}$. Then we have from inequality (3.14)

$$\forall n \in \mathbb{N} : \sup_{\vartheta\in\Theta} \mathrm{FDR}_\vartheta(\varphi_{(n)}) \leq \max_{1\leq n_0\leq n} \frac{n_0}{n} \mathbb{E}_{n_0-1,n} \overline{q}(R_n/n).$$

Since for each $n \in \mathbb{N}$ the maximum in this inequality is attained at some value $n_0(n)$ (say), we get

$$\limsup_{n\to\infty} \sup_{\vartheta\in\Theta} \mathrm{FDR}_\vartheta(\varphi_{(n)}) \leq \limsup_{n\to\infty} \zeta_n \mathbb{E}_{n_0(n)-1,n} \overline{q}(R_n/n),$$

where $\zeta_n = n_0(n)/n$, $n \in \mathbb{N}$. We now may extract a subsequence $(n_k)_{k\in\mathbb{N}}$ of $\mathbb{N}$ with $\lim_{k\to\infty} \zeta_{n_k} = \zeta$ for some $\zeta \in [0,1]$ such that

$$\begin{aligned} \limsup_{n\to\infty} \zeta_n \mathbb{E}_{n_0(n)-1,n} \overline{q}(R_n/n) &= \lim_{k\to\infty} \zeta_{n_k} \mathbb{E}_{n_0(n_k)-1,n_k} \overline{q}(R_{n_k}/n_k) \\ &\leq \zeta \limsup_{k\to\infty} \mathbb{E}_{n_0(n_k)-1,n_k} \overline{q}^*(R_{n_k}/n_k), \end{aligned}$$

where $\overline{q}^*$ denotes the $\overline{q}$-function corresponding to the critical value function f_α^{-1}. Similarly as in [91], pp. 1003-1004, we are able to select from $(n_k)_{k\in\mathbb{N}}$ a further subsequence (without loss of generality with the same name) and construct a global set $I \subseteq \mathbb{N}$ with the property $|I \cap I_{n_k}| = n_0(n_k)$ for all $k \in \mathbb{N}$. (At this point it should be noted that the definition of the sets M_k at the bottom of p. 1003 in [91] has a typo at its right end in that the term $k(n_k)$ has to be replaced by n_k.) Now we obtain from (3.18)

$$\begin{aligned} \zeta \limsup_{k\to\infty} \mathbb{E}_{n_0(n_k)-1,n_k} \overline{q}^*(R_{n_k}/n_k) &= \zeta \limsup_{k\to\infty} \mathbb{E}_{I'} \overline{q}^*(R_{n_k}/n_k) \\ &\leq \zeta \mathbb{E}_{I'} \overline{q}^*(\limsup_{k\to\infty} R_{n_k}/n_k) \\ &= \zeta \overline{q}^*(f_\alpha(t_\zeta)) \\ &= \min\{\alpha, \zeta\} \leq \alpha, \end{aligned}$$

hence the assertion of the theorem, i.e., (3.19) follows. ■

Remark 3.14

Asymptotic FDR control for procedures based on f_α in the latter theorem is established under the strong assumption (I1), i.e., independence of the p-values under the null hypotheses. We will present one counterexample which suggests that step-up-down procedures based on f_α fail to control the FDR under PRDS. To this end, we return to the D-EX-EXP model investigated in Section 2.2 and choose (without loss of generality) the scale parameter of the underlying exponential distribution equal to 1. Recall that the representation of the limiting ecdf. $F_\infty(\cdot|z,\zeta)$ of the p-values $p_i(z)$ in the D-EX-EXP(ζ_n) model with $\zeta_n \to \zeta \in (0,1)$ is given by

$$F_\infty(t|z,\zeta) = \begin{cases} (1-\zeta) + 2\zeta\exp(-z)t & \text{for } 0 \le t \le 1/2, \\ (1-\zeta) + \zeta\exp(-z)/(2-2t) & \text{for } 1/2 < t \le u(z), \\ 1 & \text{for } u(z) < t \le 1, \end{cases}$$

where $u(z) = 1 - 1/2\ \exp(-z)$. For the step-up-down procedure, we choose $\lambda = 1/2$. Noting that $f_\alpha(1/2) = 1/(1+\alpha)$ and $F_\infty(1/2|z,\zeta) = (1-\zeta) + \zeta\exp(-z)$, it is immediately clear that $F_\infty(1/2|z,\zeta) > f_\alpha(1/2)$ if and only if $z < z^* = -\ln\left(\left((1+\alpha)^{-1} - (1-\zeta)\right)/\zeta\right)$ and therefore $R_n(z)/n$ tends to 1 in this situation. In the other case, i.e., for $z \ge z^*$, the LCP between f_α and $F_\infty(\cdot|z,\zeta)$ lies in $(0,1/2]$. Equating $f_\alpha(t) = F_\infty(t|z,\zeta)$ in the latter case, we obtain the LCP $t(z|\zeta)$ as

$$t(z|\zeta) = \frac{1}{4}\frac{\zeta(\alpha-1) - \alpha + 2\,\zeta\exp(-z)\alpha + \sqrt{D(\alpha,\zeta,z)}}{\zeta\exp(-z)\,(\alpha-1)},$$

where

$$\begin{aligned} D(\alpha,\zeta,z) &= \left[(1 + 4\exp(-2z) - 4\exp(-z))\,\zeta^2 + (4\exp(-z) - 2)\,\zeta + 1\right]\alpha^2 \\ &\quad + \left[(4\exp(-z) - 2)\,\zeta^2 + (2 - 8\exp(-z))\,\zeta\right]\alpha + \zeta^2 \end{aligned}$$

by straightforward calculation. Now, we can immediately calculate the limiting FDR in case of $\zeta \in (0,1)$ by

$$\text{FDR}_\infty(\zeta) = \zeta\mathbb{P}(Z < z^*) + \zeta\int_{z^*}^{\infty} \frac{t(z|\zeta)}{f_\alpha(t(z|\zeta))}\exp(-z)dz.$$

Setting $\alpha = 0.05$, for example, we obtain by numerical integration

$$\begin{aligned} \text{FDR}_\infty(0.1) &\approx 0.069, & \text{FDR}_\infty(0.2) &\approx 0.085, \\ \text{FDR}_\infty(0.3) &\approx 0.089, & \text{FDR}_\infty(0.5) &\approx 0.093, \\ \text{FDR}_\infty(0.7) &\approx 0.094, & \text{FDR}_\infty(0.9) &\approx 0.095. \end{aligned}$$

These values could be reproduced by corresponding computer simulations.

Simulations also indicate that a step-up-down procedure of order $\lambda = \alpha$ based on the AORC fails to control the FDR at level α in Dirac-exchangeable normal models as well. In a simulation setup analogous to the setup for φ^{LSU} presented in Appendix A.1 below for a D-EX-N(0.5) model, we chose $\alpha = 5.0\%$ and obtained simulated FDR values larger than 5.5% for all $\rho \in [0.4, 0.8]$ and

even FDR values larger than 6.0% for $\rho \in [0.6, 0.7]$. With the crossing point techniques derived in Chapter 2, numerical calculation of the theoretical FDR values is possible in this case as well.

If we sharpen assumption (3.18), we can even give explicit values for the FDR.

Theorem 3.15
Let $\vartheta \in \Theta$, $\varphi_{(n)}$ be based on $\rho \leq f_\alpha^{-1}$ and assume (T2), (D3), (I1) and

$$\lim_{n\to\infty} \zeta_n = \zeta \in [0,1]. \tag{3.20}$$

If $\lim_{n\to\infty} R_n/n = r^* \ [\mathbb{P}_\vartheta]$ for some $r^* \in (0, f_\alpha(t_\zeta)]$, then it holds

$$\lim_{n\to\infty} \mathit{FDR}_\vartheta(\varphi_{(n)}) \ = \ \zeta\rho(r^*)/r^* = \zeta q(r^*) \leq \min\{\alpha, \zeta\}. \tag{3.21}$$

Proof: From (T2) and for $n_0, n \in \mathbb{N}$ we get the representation

$$V_n = n_0 F_{n,0}(\rho(R_n/n))\mathbf{1}_{\{R_n>0\}}.$$

From this we obtain the inequality chain

$$|V_n/n - \zeta_n\rho(R_n/n)| \leq \zeta_n |F_{n,0}(\rho(R_n/n)) - \rho(R_n/n)| \leq \sup_{t\in[0,1]} |F_{n,0}(t) - t|.$$

Hence, using the Glivenko-Cantelli property (3.2) together with the remaining assumptions of the theorem and the continuity of ρ, we finally see that V_n/n converges $\mathbb{P}_\vartheta$-almost surely to $\zeta\rho(r^*)$. Thus, due to $r^* > 0$, we have $\lim_{n\to\infty} \mathbb{E}_\vartheta[V_n/(R_n \vee 1)] = \zeta\rho(r^*)/r^*$. The right-hand side inequality in (3.21) is obtained by noting that $\zeta f_\alpha^{-1}(t)/t$ is increasing in $t \in (0, f_\alpha(t_\zeta)]$ to $\zeta t_\zeta / f_\alpha(t_\zeta) = \min\{\alpha, \zeta\}$ at $t = f_\alpha(t_\zeta)$. ■

The remaining case $r^* = 0$ is treated in the following two theorems.

Theorem 3.16
Let $\vartheta \in \Theta$, $\varphi_{(n)}$ be based on $\rho \leq f_\alpha^{-1}$ and assume (T1)-(I2), (A1), (3.20) and

$$\forall \varepsilon > 0 : \liminf_{n\to\infty} \inf_{\varepsilon \leq t \leq 1} (t - F_n(\rho(t))) > 0 \ \ [\mathbb{P}_\vartheta]. \tag{3.22}$$

Then it holds

$$\limsup_{n\to\infty} \mathit{FDR}_\vartheta(\varphi_{(n)}) \leq \zeta \limsup_{x\downarrow 0} q(x) = \zeta q(0) = \zeta \overline{q}(0) \leq \zeta\alpha. \tag{3.23}$$

Proof: To avoid triviality, we assume $I_0(\vartheta)$ to be non-empty. Then, from (3.12) and (3.13) we have that

$$\limsup_{n\to\infty} \mathrm{FDR}_\vartheta(\varphi_{(n)}) \leq \zeta \limsup_{n\to\infty} \mathbb{E}_{\vartheta^{i_0}} \overline{q}(R_n/n). \tag{3.24}$$

Since due to (T1) and (T2) we have for all $n \in \mathbb{N}$

$$F_n(\rho(R_n/n)) = R_n/n,$$

(3.22) implies that for every fixed $\varepsilon > 0$ we obtain $\limsup_{n\to\infty} R_n/n \leq \varepsilon$ $\mathbb{P}_\vartheta$-almost surely, i.e., $\lim_{n\to\infty} R_n/n = 0$ $\mathbb{P}_\vartheta$-almost surely. Now, since for all $n \in \mathbb{N}$ the maximum absolute difference on the unit interval of the ecdf F_n (corresponding to the sequence of p-values $(p_n)_{n\in\mathbb{N}}$) and the ecdf $F_n^{i_0}$ (say) (corresponding to the sequence of p-values $(p_n^{i_0})_{n\in\mathbb{N}}$ defined in Section 3.1) is at most $1/n$, condition (3.22) also holds $\mathbb{P}_{\vartheta^{i_0}}$-almost surely, which entails that $\lim_{n\to\infty} R_n/n = 0$ $\mathbb{P}_{\vartheta^{i_0}}$-almost surely. Hence, due to the continuity of $\overline{q}$ we have $\lim_{n\to\infty} \mathbb{E}_{\vartheta^{i_0}} \overline{q}(R_n/n) = \overline{q}(0) = q(0) \leq \lim_{t\downarrow 0} f_\alpha^{-1}(t)/t = \alpha$. In view of inequality (3.24), this completes the proof. ■

Theorem 3.17

Under the assumptions of Theorem 3.16 let $\varphi_{(n)}$ be a SUD-test of order λ_n with $\liminf_{n\to\infty} \lambda_n/n > 0$ and the condition (3.22) be replaced by

$$\forall \varepsilon > 0 : \liminf_{n\to\infty} \inf_{\varepsilon \leq t \leq K} (t - F_n(\rho(t))) > 0 \ \ [\mathbb{P}_\vartheta] \tag{3.25}$$

for some $K \in [0,1]$ fulfilling $K > L = \limsup_{n\to\infty} \lambda_n/n$ in case of $L < 1$ and $K = 1$ otherwise. Supposing that $\lim_{x\downarrow 0} q(x)$ exists, we have

$$\lim_{n\to\infty} \mathit{FDR}_\vartheta(\varphi_{(n)}) = \zeta \lim_{x\downarrow 0} q(x) = \zeta q(0) = \zeta \overline{q}(0) \leq \zeta\alpha. \tag{3.26}$$

Proof: Again, to avoid triviality, we assume a non-empty index set $I_0(\vartheta)$. Equation (3.26) can be shown by utilizing the notation introduced in the proof of Theorem 3.12 and the decomposition

$$\begin{aligned} \mathrm{FDR}_\vartheta(\varphi_{(n)}) &= \zeta_n \sum_{j=1}^{\lambda_n} q(j/n) \mathbb{P}_\vartheta(R_n = j | p_{i_0} \leq \alpha_{j:n}) \\ &\quad + \zeta_n \sum_{j=\lambda_n+1}^{n} q(j/n) \mathbb{P}_\vartheta(R_n = j | p_{i_0} \leq \alpha_{j:n}) \\ &= M_n + m_n \ \text{(say)}. \end{aligned}$$

In view of Theorem 3.12 and the structure of a SUD-test, we obtain by applying (3.6) that

$$\begin{aligned} M_n &= \zeta_n \mathbb{E}_{\vartheta^{i_0}} \left[q(R_n/n) \mathbf{1}_{\{R_n/n \leq \lambda_n/n\}} \right], \\ m_n &\leq \zeta_n \mathbb{E}_{\vartheta^{i_0}} \left[\overline{q}(R_n/n) \mathbf{1}_{\{R_n/n > \lambda_n/n\}} \right]. \end{aligned}$$

From (3.25) it follows that $\mathbb{P}_\vartheta$-almost surely $F_n(\rho(\lambda_n/n)) < \lambda_n/n \leq K$ and consequently $R_n/n < \lambda_n/n \leq K$ holds true for eventually all $n \in \mathbb{N}$. Therefore, again due to (3.25), in analogy to the proof of Theorem 3.16 we conclude that $\lim_{n\to\infty} R_n/n = 0$ $\mathbb{P}_{\vartheta^{i_0}}$-almost surely, which finally entails $\lim_{n\to\infty} \mathbf{1}_{\{R_n/n > \lambda_n/n\}} = 0$ $\mathbb{P}_{\vartheta^{i_0}}$-almost surely. Together with the boundedness of $\overline{q}$ this entails that $\lim_{n\to\infty} m_n = 0$. Moreover, exploiting the continuity of q at $x = 0$ we see that $\lim_{n\to\infty} M_n = \zeta q(0) = \zeta \overline{q}(0)$, which altogether yields the desired result. ■

Remark 3.18

One cannot expect to obtain exact values for the limiting FDR under the quite general assumptions of Theorem 3.15 if $r^* = 0$. To see this, consider the case $\zeta_n \equiv 1$ in which the FDR is equal to the family-wise error rate (FWER). For $\zeta_n \equiv 1$ it was shown in [91] that the FWER is equal to α for any $n \in \mathbb{N}$ in case of a linear step-up procedure, while it tends to $1 - \exp(-\alpha) < \alpha$ for a linear step-down procedure. We therefore have to know more about the structure of the underlying procedure in order to compute the limiting FDR in case of $r^* = 0$. The limiting behavior for procedures based on f_α (or its modifications) satisfying the assumptions of Theorem 3.17 is in accordance with the linear SU-procedure and should be expected, since the difference of the critical values $\alpha_{i:n} - i\alpha/n$ tends to zero for $i \in o(n)$. Therefore, the local behavior around zero should not differ much for large n.

Returning to our proposed example procedures, we finally obtain the following properties of these tests.

Corollary 3.19 (Examples 3.5, 3.7, 3.8 continued)
Assume the distributional assumptions (D3), (I1) and (I2) hold. Then the SUD-procedure based on f_α with parameter $\lambda \in [0, 1)$ and the SU-procedures based on $f^{(i)}_{\alpha,\kappa}$, $i = 1, 2$, as well as the truncated SU-procedure asymptotically control the FDR at level α. More precisely, if condition (3.20) is fulfilled, i.e., $\lim_{n\to\infty} \zeta_n = \zeta \in [0, 1]$, then

(i) for the SUD-procedure the upper bound α for the limiting FDR is sharp for $\zeta \in [\alpha, 1]$.

(ii) for the SU-procedures based on $f^{(i)}_{\alpha,\kappa}$, $i = 1, 2$, the upper bound α for the limiting FDR is sharp for $\zeta \geq \zeta^*(\kappa) = \alpha/(\kappa(1-\alpha)+\alpha)$. In case of $\zeta < \zeta^*(\kappa)$, an upper bound for the asymptotic FDR is given by $\zeta\tilde{t}_\zeta/(1-\zeta+\zeta\tilde{t}_\zeta)$, where $\tilde{t}_\zeta$ denotes the unique solution of the equation $F_\infty(t|\zeta) = h_i(t),\ i = 1, 2$, on $(0, t_\zeta)$. For finite n, the upper bound given in (3.14) is sharp.

(iii) for the truncated SU-procedure the upper bound α for the limiting FDR is sharp for $\zeta \geq \zeta^*(\kappa)$. In case of $\zeta < \zeta^*(\kappa)$, an upper bound for the asymptotic FDR is given by $\zeta\kappa/(1-\zeta+\zeta\kappa)$.

Proof: First of all, as mentioned before, a step-up-down test has the structural properties (T1), (T2) and (A1). Moreover, assumptions (D3), (I1) and (I2) imply the crucial monotonicity properties (D1) and (D2) for a step-up-down test. Hence, in order to apply Theorem 3.13, it remains to check the validity of condition (3.18). To this end, for notational convenience and without loss of generality, we work under condition (3.20). We make use of (3.2), i.e., that the ecdf F_n converges $\mathbb{P}_{I_0'}$-almost surely to its limit $F_\infty(\cdot|\zeta)$ uniformly in $t \in [0, 1]$. Since under (T1) and (T2) we have the identity $F_n(\rho(R_n/n)) = R_n/n$ for all $n \in \mathbb{N}$, (3.2) leads to

$\lim_{n\to\infty}(F_\infty(\rho(R_n/n)|\zeta) - R_n/n) = 0$ $\mathbb{P}_{I_0'}$-almost surely. From this we conclude that ($\mathbb{P}_{I_0'}$-almost surely) the only possible accumulation points of the sequence $(R_n/n)_{n\in\mathbb{N}}$ consist of the solutions of the equation $F_\infty(\rho(t)|\zeta) = t$ in $t \in [0,1]$. If, as in Examples 3.7 and 3.8, this solution is unique, then the sequence $(R_n/n)_{n\in\mathbb{N}}$ necessarily converges to this solution $r^* = r^*(\zeta) \in [0,1]$ (say) $\mathbb{P}_{I_0'}$-almost surely.

If, however, as in Example 3.5, the equation $F_\infty(\rho(t)|\zeta) = t$ has the solution $t = 1$ and exactly one further (smaller) solution in $[0,1)$, we have to exclude $t = 1$ as a possible accumulation point of $(R_n/n)_{n\in\mathbb{N}}$ in the latter case in order to prove the $\mathbb{P}_{I_0'}$-almost sure convergence of $(R_n/n)_{n\in\mathbb{N}}$ to the smallest solution $r^* = r^*(\zeta)$ (say) of the aforementioned equation. To this end, we only consider values of ζ leading to the two distinct solutions $r^*(\zeta)$ and 1. For critical value functions ρ with $\rho(t) \le f_\alpha^{-1}(t)$ for all $t \in [0,1]$ it is then evident that $F_\infty(\rho(t)|\zeta) < t$ for all $t \in (r^*, 1)$. Moreover, notice that, by definition of $\lambda_n(\lambda)$, we have the inequalities $\lambda_n(\lambda) - 1 \le nr(\lambda) \le \lambda_n(\lambda)$ for all $n \in \mathbb{N}$. Now, if $\lambda > \rho(r^*)$, this, together with condition (3.2), yields that $\mathbb{P}_{I_0'}$-almost surely $F_n(\rho(\lambda_n(\lambda)/n)) < \lambda_n(\lambda)/n$ and consequently $R_n < \lambda_n(\lambda)$ holds true for eventually all $n \in \mathbb{N}$. This entails $\limsup_{n\to\infty} R_n/n \le \limsup_{n\to\infty} \lambda_n(\lambda)/n = r(\lambda) < 1$ $\mathbb{P}_{I_0'}$-almost surely, which is just what we wanted to show. Finally, if $\lambda \le \rho(r^*)$, we may choose a $\lambda' > \rho(r^*)$ and compare the number of rejections of the corresponding SUD-procedures. Since this number is non-decreasing with increasing parameter we eventually arrive at $\limsup_{n\to\infty} R_n/n \le r(\lambda') < 1$ $\mathbb{P}_{I_0'}$-almost surely. Since for all procedures under investigation it holds $\rho(t) \le f_\alpha^{-1}(t)$ for all $t \in [0,1]$ we conclude that $r^* = r^*(\zeta) \le f_\alpha(t_\zeta)$. Hence, Theorem 3.13 applies. As a consequence, the example procedures asymptotically control the FDR.

In case of the SUD-procedure in (i), we use $\rho = f_\alpha^{-1}$ and obtain $r^* = f_\alpha(t_\zeta)$. Hence, the upper bound α for the asymptotic FDR is sharp in (i) under Dirac-uniform configurations. The sharpness of the upper bound α for the asymptotic FDR in (ii) and (iii) is due to the fact that under Dirac-uniform configurations with $\zeta \ge \zeta^*(\kappa)$ we obtain $r^* = f_\alpha(t_\zeta)$.

Finally, the sharpness of the upper bounds for the finite n FDR in (ii) is a consequence of (A2), which is fulfilled for $f_{\alpha,\kappa}^{(i)}$, $i = 1, 2$. Sharpness here means that the upper bound given in (3.14) is exactly attained for finite n under Dirac-uniform configurations. ■

3.6 Asymptotic optimality of the AORC

The latter Corollary 3.19 means reworded that procedures based on f_α fulfilling the assumptions of Theorem 3.15 asymptotically exhaust the whole FDR level α under Dirac-uniform configurations. Moreover, the rejection curve f_α cannot be improved in the sense of the following theorem, which is another consequence of Theorem 3.13. In order to formalize "optimality", let $\alpha \in (0,1)$, $\lambda \in [0,1]$ and $\mathcal{M}_\lambda$ denote the set of rejection curves r with the property that for all $I_0 \subseteq \mathbb{N}$ with

$\lim_{n\to\infty}\zeta_n = \zeta$ for some $\zeta \in [0,1]$ it holds

$$\limsup_{n\to\infty} \mathrm{FDR}_{I_0}\left(\varphi_{n,\lambda_n}^{\mathrm{SUD}(r)}\right) \le \limsup_{n\to\infty}\sup_{\vartheta\in\Theta} \mathrm{FDR}_{\vartheta}\left(\varphi_{n,\lambda_n}^{\mathrm{SUD}(r)}\right) \le \alpha, \tag{3.27}$$

where $\varphi_{n,\lambda_n}^{\mathrm{SUD}(r)}$ is the step-up-down procedure of order $\lambda_n = \lambda_n(\lambda)$ based on r. It should be noted that the first inequality in (3.27) is not very restrictive since many statistical models satisfy the "model continuity assumptions" (SA) formulated in [91], due to which, at least for SUD-procedures such as $\varphi_{n,\lambda_n}^{\mathrm{SUD}(r)}$, the corresponding FDR values $\mathrm{FDR}_{I_0}\left(\varphi_{n,\lambda_n}^{\mathrm{SUD}(r)}\right)$ can be approximated arbitrarily closely by the values $\mathrm{FDR}_{\vartheta}\left(\varphi_{n,\lambda_n}^{\mathrm{SUD}(r)}\right)$ for some suitably chosen $\vartheta \in \Theta$, $n \in \mathbb{N}$.
In terms of power it is immediately clear that, whenever $r_1, r_2 \in \mathcal{M}_\lambda$ with $r_1 \le r_2$, then $\varphi_{n,\lambda_n}^{\mathrm{SUD}(r_1)} \ge \varphi_{n,\lambda_n}^{\mathrm{SUD}(r_2)}$. Therefore, a smaller rejection curve typically leads to a more powerful test procedure in the sense that more (never less) false hypotheses can be rejected.

Theorem 3.20 (Asymptotic optimality of f_α)

(i) Let $\lambda \in [0,1]$ and $r \in \mathcal{M}_\lambda$. Then

$$\forall t \in [0,\lambda]:\ r(t) \ge f_\alpha(t). \tag{3.28}$$

If $\lambda < 1$, then it holds for any $\tau \in (\lambda, 1]$ that

$$\forall t \in (\lambda,\tau]:\ r(t) \le f_\alpha(t) \Rightarrow \forall t \in (\lambda,\tau]: r(t) = f_\alpha(t). \tag{3.29}$$

(ii) If $\lambda < 1$ and $r \in \mathcal{M}_\lambda$ is such that, for every $\zeta \in (\alpha,1)$, the equation $F_\infty(\rho(t)|\zeta) = 1-\zeta+\zeta\rho(t) = t$ has at most one solution in (0,1), it even holds $r(t) \ge f_\alpha(t)$ for all $t \in [0,1]$.

(iii) If $\lambda = 1$ and assuming (D3), (I1) and (I2), it holds

$$\inf_{r\in\mathcal{M}_1} r = f_\alpha.$$

Moreover, for any $\vartheta \in \Theta_\kappa = \{\vartheta\in\Theta : \liminf_{n\to\infty}\zeta_n(\vartheta) > \alpha/(\kappa(1-\alpha)+\alpha)\}$, $\kappa \in (0,1)$, the power of any $\tilde{f}_\alpha \in \mathcal{M}_1$ with $\tilde{f}_\alpha(t) = f_\alpha(t)$ for all $t \in [0,\kappa]$ is asymptotically not smaller than the power of any other $r \in \mathcal{M}_1$, that is,

$$\liminf_{n\to\infty}\left[\bar{\beta}_\vartheta(\varphi_{n,n}^{SUD(\tilde{f}_\alpha)}) - \bar{\beta}_\vartheta(\varphi_{n,n}^{SUD(r)})\right] \ge 0 \ \textit{for all}\ \vartheta \in \Theta_\kappa. \tag{3.30}$$

Proof: In order to prove part (i), assume that for an arbitrary chosen rejection curve $r \in \mathcal{M}_\lambda$ it holds $r(t^*) < f_\alpha(t^*)$ for some $t^* \in (0,\lambda)$. Consider now a Dirac-unifom configuration $\mathbb{P}_{I_0}$ with $\lim_{n\to\infty}\zeta_n = \zeta$ and $\zeta \in (\alpha,1)$ chosen such that $r(t^*) < F_\infty(t^*|\zeta) < f_\alpha(t^*)$. Then it is obvious that property (3.27) is violated, because (with self-explaining notation) it follows $\mathbb{P}_{I_0}$-almost surely

$$\liminf_{n\to\infty} R_n^{(r)}/n \ge F_\infty(t^*|\zeta) > F_\infty(t_\zeta|\zeta) = f_\alpha(t_\zeta)$$

and consequently

$$\liminf_{n\to\infty} \mathrm{FDR}_{I_0}\left(\varphi_{n,\lambda_n}^{\mathrm{SUD}(r)}\right) \geq \zeta t^*/(1-\zeta+\zeta t^*) > \zeta t_\zeta/(1-\zeta+\zeta t_\zeta) = \alpha,$$

due to the fact that the function $x \to \zeta x/(1-\zeta+\zeta x)$ is strictly increasing in $x \in (0,1)$ and $t^* > t_\zeta$. Hence, for all $t \in (0,\lambda)$ we have $r(t) \geq f_\alpha(t)$, from which the assertion follows.
Now assume that we have $r(t) \leq f_\alpha(t)$ for all $t \in (\lambda,\tau]$ and $r(t^*) < f_\alpha(t^*)$ for some $t^* \in (\lambda,\tau)$. Consider now the Dirac-unifom asymptotic model $\mathrm{DU}_\infty(\zeta^*)$ with $\zeta^* \in (\alpha,1)$ chosen such that $f_\alpha(\lambda) < F_\infty(\lambda|\zeta^*)$, $F_\infty(t^*|\zeta^*) < f_\alpha(t^*)$ and $\inf_{\lambda\leq t\leq t^*}(F_\infty(t|\zeta^*) - r(t)) > 0$, which is possible due to the left-continuity of the rejection curve r. Then the argumentation is the same as before. Part (ii) and the first assertion of part (iii) can be proven similarly.
For the proof of (3.30), we assume (in order to avoid triviality) $n_1(n) > 0$ for all $n \in \mathbb{N}$, define $S_n = R_n - V_n$ and denote the set of all $\tilde{f}_\alpha \in \mathcal{M}_1$ with $\tilde{f}_\alpha(t) = f_\alpha(t)$ for all $t \in [0,\kappa]$ by $\mathcal{S}_\kappa$. Then we have (with self-explaining notation as before) the inequality

$$\forall n \in \mathbb{N}:\ \forall \tilde{f}_\alpha \in \mathcal{S}_\kappa:\ \forall r \in \mathcal{M}_1 : \left(\frac{S_n(\tilde{f}_\alpha)}{n_1} - \frac{S_n(r)}{n_1}\right)\mathbf{1}_{\{t_n^*(r)\leq\kappa\}} \geq 0,$$

which holds true due to (3.28) and the fact that S_n is non-decreasing in t_n^*. Now, for fixed $\vartheta \in \Theta_\kappa$, we utilize the chain of inequalities

$$t_n^*(r|\mathbb{P}_\vartheta) \leq t_n^*(r|\mathrm{DU}_n(\zeta_n(\vartheta))) \leq t_n^*(\tilde{f}_\alpha|\mathrm{DU}_n(\zeta_n(\vartheta))) < \kappa$$

which holds $\mathbb{P}_\vartheta$-almost surely for eventually all $n \in \mathbb{N}$, leading to $\limsup_{n\to\infty} t_n^*(r|\mathbb{P}_\vartheta) < \kappa$ and consequently to $\mathbf{1}_{\{t_n^*(r)\leq\kappa\}} \to 1\ [\mathbb{P}_\vartheta]$ for all $\vartheta \in \Theta_\kappa$. Therefore, we obtain $\mathbb{P}_\vartheta$-almost surely

$$\liminf_{n\to\infty}\left(\frac{S_n(\tilde{f}_\alpha)}{n_1} - \frac{S_n(r)}{n_1}\right) \geq 0 \ \text{ for all } \ \vartheta \in \Theta_\kappa, \tilde{f}_\alpha \in \mathcal{S}_\kappa, r \in \mathcal{M}_1. \tag{3.31}$$

Taking expectation in (3.31) and utilizing Fatou's lemma, we finally arrive at assertion (4.1). ■

Theorem 3.20 shows that in the class of SU-procedures with rejection curve $r \in \mathcal{M}_1$ we always have $r \geq f_\alpha$. In the class of truncated SU-procedures with parameter $\kappa \in (0,1)$, the truncated procedure based on f_α is the best choice. More generally, if we restrict attention to the subspace $\Theta_\kappa \subset \Theta$ described in (iii) of Theorem 3.20, f_α is the asymptotically uniformly best choice on $[0,\kappa]$ for a step-up procedure. For SUD-procedures with parameter $\lambda < 1$, f_α leads to the asymptotically uniformly best choice of critical values on the step-up part, see (3.28). On the step-down part of a SUD-procedure, f_α cannot be uniformly improved by some $r \in \mathcal{M}_\lambda$ whatever r does on the step-up part, see (3.29) with $\tau = 1$. For arbitrary $\tau \in (\lambda,1]$, assertion (3.29) states that a rejection curve $r \in \mathcal{M}_\lambda$ cannot be first smaller and then larger than f_α on the interval $(\lambda,1]$. It seems possible that $\mathcal{M}_\lambda$ contains an r which is first larger and then smaller on the step-down part. But this would imply that the SUD-procedure based on r is asymptotically less powerful than the

SUD-procedure based on f_α on some Θ_κ. If we restrict attention to rejection curves $r \in \mathcal{M}_\lambda$ described in (ii) of Theorem 3.20, then f_α is the best choice. These considerations may justify to call f_α the asymptotically optimal rejection curve (AORC).

In view of these asymptotic results, it is natural to ask how large n has to become in order to achieve a reasonable behavior of the FDR of the proposed procedures. As already mentioned in Example 3.5, the asymptotic exhaustion of the whole FDR level has to be traded off with a slightly liberal behavior of the procedures based on f_α in the finite case. In order to illustrate this effect, we consider the SU-procedures based on $f^{(i)}_{\alpha,\kappa}$, $i = 1, 2$, where the upper bound given in (3.14) is sharp in the $\mathrm{DU}_n(\zeta_n)$-model. Due to the pointwise order of these two rejection curves (cf. Figure 3.4) it is clear that a SU-procedure based on $f^{(2)}_{\alpha,\kappa}$ is more liberal in the $\mathrm{DU}_n(\zeta_n)$-model. We therefore present results for this procedure. Figure 3.5 depicts the behavior of this procedure under DU configurations with varying number of true hypotheses for $n = 100, 500$ and 1000. For $n = 100$, there is a notable violation of the FDR level $\alpha = 5\%$ for $12 \leq n_0 \leq 35$. The largest FDR under Dirac-uniform is attained for $n_0 = 16$ with numerical value 0.05801. For the two larger values of n, the actual level does not exceed α by much.

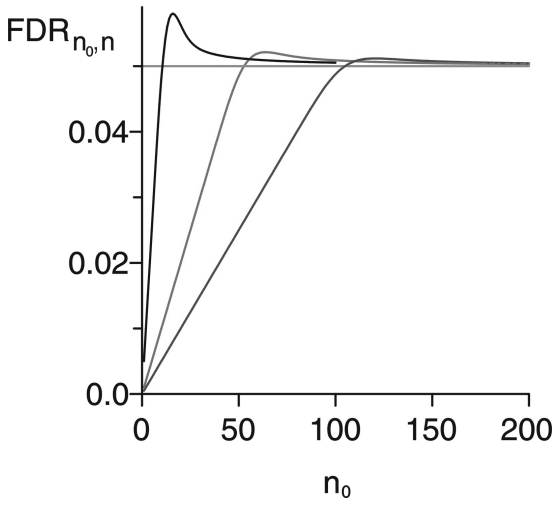

Figure 3.5: Actual DU-FDR of the SU-procedure based on $f^{(2)}_{0.05,\kappa_2}$ depending on n_0

The method of computing the FDR for a SU(D)-procedure in case of an underlying Dirac-uniform configuration will be described in the next section. Moreover, we will give some brief suggestions for modifications of f_α in the finite case. However, this will not be emphasized to much, because on the one hand, the AORC is designed for the asymptotic case and on the other hand, we have to keep in mind that the FDR values under Dirac-uniform reflect an unrealistic worst case scenario. For realistic alternatives, we get much smaller realized FDRs so that the original AORC may safely be used in the finite case for e.g. $n \geq 500$.

A detailed numerical study of the FDR behavior of the example procedures presented in Examples

3.5, 3.7 and 3.8 in case of a finite number of hypotheses at hand is given in Appendix A.2.

3.7 FDR control for a fixed number of hypotheses

In this section we briefly discuss some possibilities to achieve strict FDR control for a finite number of hypotheses for procedures related to the AORC. It would be attractive to find critical values close to (3.3) for step-up-down procedures as described in the previous sections such that the FDR is strictly controlled. As shown before, an upper bound for the FDR of a step-up procedure with critical values satisfying that $\alpha_{i:n}/i$ is non-decreasing in i is obtained in one of the Dirac-uniform configurations. This bound is sharp if the corresponding Dirac-uniform configuration belongs to the model. For step-up-down procedures with parameter $\lambda_n \in \{1, \ldots, n-1\}$ it is not known whether Dirac-uniform configurations are least favorable. However, Dirac-uniform configurations also yield an upper bound for the FDR in this case.
More specifically, under the assumptions of Theorem 3.12, the aforementioned upper bound for a fixed n_0 is given by (see (3.14))

$$b(n_0, n) = \frac{n_0}{n} \mathbb{E}_{I_0'} \left[\overline{q}(R_n/n) \right], \; n_0 = 1, \ldots, n,$$

with I_0' defined as in Theorem 3.9. Hence, the upper bound for the FDR is given by $b_n^* = \max_{1 \le n_0 \le n} b(n_0, n)$.

Lemma 3.21
For a SUD-procedure with critical values $\alpha_{i:n}$ of order λ_n satisfying (T1)-(A2), $b(n_0, n)$ is given by

$$b(n_0, n) = n_0 \sum_{j=1}^{n_0} \frac{\alpha_{n_1+j:n}}{n_1 + j} \mathbb{P}_{n_0-1,n}(V_n = j-1), \tag{3.32}$$

where $\mathbb{P}_{m,n}$ refers to a Dirac-uniform configuration such that m p-values are iid uniformly distributed and the remaining ones follow a Dirac distribution. If $\lambda_n = n$, which corresponds to a SU-procedure, $b(n_0, n)$ can alternatively be calculated by

$$b(n_0, n) = \sum_{j=1}^{n_0} \frac{j}{n_1 + j} \mathbb{P}_{n_0,n}(V_n = j) = \mathit{FDR}_{I_0}\left(\varphi_{n,n}^{\mathit{SUD}(\vec{\alpha})} \right) \tag{3.33}$$

and it even holds equality in every summand, i.e.,

$$\mathbb{P}_{n_0,n}(V_n = j) = \frac{n_0}{j} \alpha_{n_1+j:n} \mathbb{P}_{n_0-1,n}(V_n = j-1) \text{ for } j = 1, \ldots, n_0. \tag{3.34}$$

Proof: In order to prove (3.32), keep in mind that the expectation in (3.14) refers to a Dirac-uniform configuration with $(n_0 - 1)$ true hypotheses and $(n_1 + 1)$ false hypotheses and since $p_j \sim \varepsilon_0$ for all $j \in I_{n,1}$, we have $R_n = V_n + (n+1)$ $\mathbb{P}_{I_0'}$-almost surely. Straightforward

calculation now yields

$$\begin{aligned}\frac{n_0}{n}\mathbb{E}_{I_0'}\left[\overline{q}\left(\frac{R_n}{n}\right)\right] &= \frac{n_0}{n}\mathbb{E}_{I_0'}\left[\frac{\rho(R_n/n)}{R_n/n}\right] = n_0\mathbb{E}_{I_0'}\left[\frac{\alpha_{R_n:n}}{R_n}\right] \\ &= n_0\mathbb{E}_{I_0'}\left[\frac{\alpha_{V_n+n_1+1:n}}{V_n+n_1+1}\right] \\ &= n_0\sum_{k=0}^{n_0-1}\frac{\alpha_{k+n_1+1:n}}{k+n_1+1}\mathbb{P}_{n_0-1,n}(V_n=k) \\ &= n_0\sum_{j=1}^{n_0}\frac{\alpha_{n_1+j:n}}{n_1+j}\mathbb{P}_{n_0-1,n}(V_n=j-1),\end{aligned}$$

which is (3.32). Equality (3.34) and consequently the left-hand side equality of (3.33) are immediate consequences of the representation of the pmf. of V_n for a step-up test given in Corollary 3.23. The right-hand side equality in (3.33) is obtained by noticing that in a Dirac-uniform model with n_0 true hypotheses and n_1 false hypotheses, we have $R_n = V_n + n_1$ $\mathbb{P}_{I_0'}$-almost surely and it therefore holds

$$\begin{aligned}\mathrm{FDR}_{I_0}\left(\varphi_{n,n}^{\mathrm{SUD}(\vec{\alpha})}\right) &= \mathbb{E}_{I_0}\left[\frac{V_n}{R_n\vee 1}\right] = \mathbb{E}_{I_0}\left[\frac{V_n}{(V_n+n_1)\vee 1}\right] \\ &= \sum_{j=1}^{n_0}\frac{j}{j+n_1}\mathbb{P}_{n_0,n}(V_n=j),\end{aligned}$$

according to the discrete expectation formula. ■

Formulas for the pmf. of V_n under Dirac-uniform configurations can be obtained in terms of the joint cdf. of order statistics. For SUD-procedures the computation of the pmf. of V_n becomes numerically difficult for larger values of n. A way out is to simulate the upper bound.
For the derivation of the pmf. of V_n, we use the following considerations. Under the assumption that $0 \le c_{1:n} \le \cdots \le c_{n:n} \le 1,\ n \in \mathbb{N}$, a general recursive formula for the joint cdf. F_n^k of the order statistics $U_{1:n}, \ldots, U_{n-k:n}$, $0 \le k \le n$, of n i.i.d. UNI$[0,1]$-distributed random variables U_i is given by

$$F_n^k(c_{1:n},\ldots,c_{k:n}) = 1 - \sum_{j=0}^{n-k-1}\binom{n}{j}F_j(c_{1:n},\ldots,c_{j:n})(1-c_{j+1:n})^{n-j}, \tag{3.35}$$

with $F_n^0 = F_n$ and $F_0^0 \equiv F_n^n \equiv 1$. This is essentially Bolshev's recursion, which is proved in different ways in [262], pp. 366-367, and in [89].
Formula (3.35) (for $k = 0$) may be used to calculate the pmf. of V_n for a SUD-procedure of order r under Dirac-uniform configurations and yields the following result.

Lemma 3.22

For the pmf. of V_n of a step-up-down procedure of order r based on critical values $0 \leq \alpha_{1:n} \leq \ldots \leq \alpha_{n:n} \leq 1$ under a Dirac-uniform configuration with n_0 true hypotheses and $n_1 = n - n_0$ false hypotheses, we obtain that $\mathbb{P}_{n_0,n}(V_n = j)/\binom{n_0}{j}$ is equal to

$$\begin{cases} F_j(\alpha_{n_1+1:n}, \ldots, \alpha_{n_1+j:n})\bar{\alpha}^{n_0-j}_{n_1+j+1:n}, & \text{if } r \leq n_1, \\ F_{n_0-j}(\underbrace{\bar{\alpha}_{r:n}, \ldots, \bar{\alpha}_{r:n}}_{n-r+1}, \bar{\alpha}_{r-1:n}, \ldots, \bar{\alpha}_{n_1+j+1:n})\alpha^j_{n_1+j:n}, & \text{if } r > n_1 \wedge j \leq r - n_1, \\ F_j(\underbrace{\alpha_{r:n}, \ldots, \alpha_{r:n}}_{r-n_1}, \alpha_{r+1:n}, \ldots, \alpha_{n_1+j:n})\bar{\alpha}^{n_0-j}_{n_1+j+1:n}, & \text{if } r > n_1 \wedge j > r - n_1, \end{cases}$$

where $\bar{\alpha}_j = 1 - \alpha_j$, $j = 1, \ldots, n$.

Proof: For notational convenience, we denote the p-values corresponding to true hypotheses by $p^{(0)}$-values. It is remarkable that the vector of ordered p-values $(p_{1:n}, \ldots, p_{n:n})$ is almost surely of the form $(\underbrace{p_{1:n} = 0 = \ldots = 0 = p_{n_1:n}}_{n_1}, p_{n_1+1:n} = p^{(0)}_{1:n}, \ldots, p^{(0)}_{n_0:n} = p_{n:n})$.

Case 1, $r \leq n_1$: In this case, we necessarily fall into the step-down branch of the test procedure, because at least the first r components of the vector of ordered p-values are 0 such that $p_{r:n} \leq \alpha_{r:n}$ is true with probability 1. Consequently, the event $\{V_n = j\}$ can be expressed as

$$\begin{aligned} \{V_n = j\} &= \{p_{n_1+1:n} \leq \alpha_{n_1+1:n}, \ldots, p_{n_1+j:n} \leq \alpha_{n_1+j:n}\} \\ &\quad \cap \{p_{n_1+j+1:n} > \alpha_{n_1+j+1:n}\}. \end{aligned}$$

Since the second event implies that all ordered p-values with ordered indices $n_1 + j + 1$ or greater are larger than $\alpha_{n_1+j+1:n}$, the event means that $(n_0 - j)$ $p^{(0)}$-values are greater than $\alpha_{n_1+j+1:n}$. Since we have $\binom{n_0}{j}$ possibilities to choose these $p^{(0)}$-values and all p-values are assumed to be independent, we immediately obtain the result.

Case 2, $r > n_1$ and $j \leq r - n_1$: In order to get into this case, we must have $p_{r:n} > \alpha_{r:n}$ and fall into the step-up branch of the procedure. Consequently, we can write

$$\begin{aligned} \{V_n = j\} &= \{p_{r:n} > \alpha_{r:n}, p_{r-1:n} > \alpha_{r-1:n}, \ldots, p_{n_1+j+1:n} > \alpha_{n_1+j+1:n}\} \\ &\quad \cap \{p_{n_1+j:n} \leq \alpha_{n_1+j:n}\}. \end{aligned}$$

Since $p_{n_1+j:n} \leq \alpha_{n_1+j:n}$ automatically implies that $p^{(0)}_{k:n} \leq \alpha_{n_1+j:n}$ for all $k = 1, \ldots, j$, we can again choose j out of the n_0 $p^{(0)}$-values to fulfill this relationship.

Case 3, $r > n_1$ and $j > r - n_1$: In this third case, we fall into the step-down branch of the procedure, resulting in

$$\begin{aligned} \{V_n = j\} &= \{p_{r:n} \leq \alpha_{r:n}, p_{r+1:n} \leq \alpha_{r+1:n}, \ldots, p_{n_1+j:n} \leq \alpha_{n_1+j:n}\} \\ &\quad \cap \{p_{n_1+j+1:n} > \alpha_{n_1+j+1:n}\}. \end{aligned}$$

The assertion then follows in analogy to the step-down considerations under case 1. ■

Corollary 3.23

For an SU-procedure with critical values $\alpha_{1:n} \leq \cdots \leq \alpha_{n:n}$ under a Dirac-uniform configuration with n_0 true hypotheses, we get (see also [92])

$$\mathbb{P}_{n_0,n}(V_n = j) = \binom{n_0}{j} F_{n_0-j}(1-\alpha_{n:n}, \ldots, 1-\alpha_{n-n_0+j+1:n})\alpha^j_{n-n_0+j:n}. \tag{3.36}$$

This result is immediate if we consider the case $r = n$ in Lemma 3.22. Alternatively, the pmf. of V_n in this case can be calculated by iteratively applying (3.34).

3.7.1 Simultaneous β-adjustment

Here, we mention one ad-hoc possibility to obtain a valid set of critical values for a SU- or SUD-procedure guaranteeing strict FDR control, that is, we adjust the critical values given in (3.3) in a suitable way. For example, we can try to find a suitable $\beta_n > 0$ such that the choice

$$\alpha_{i:n} = \frac{\frac{i}{n+\beta_n}\alpha}{1-\frac{i}{n+\beta_n}(1-\alpha)} = \frac{i\alpha}{n+\beta_n-i(1-\alpha)},\ i = 1, \ldots, n, \tag{3.37}$$

yields a SU-procedure (or SUD-procedure) controlling the FDR at level α. The critical values (3.37) correspond to the rejection curve

$$\tilde{f}_\alpha(t) = (1+\frac{\beta_n}{n})f_\alpha(t),\ t \in [0, \alpha/(\alpha+\beta/n)].$$

Technically, the determination of the minimal β_n can be done by a grid search. Starting with $\beta_n = \varepsilon$ for some $\varepsilon > 0$, we evaluate (3.33) for all possible values of n_0 and check if the condition $\max_{1\leq n_0\leq n} b(n_0, n) \leq \alpha$ is fulfilled. If not, we update β_n by iteratively adding ε until no violation of the FDR level α occurs any more. For example, for $\alpha = 0.05$, an SU-procedure with $n = 100$ and the choice $\beta_{100} = 1.76$ leads to strict FDR control (we chose $\varepsilon = 0.01$).

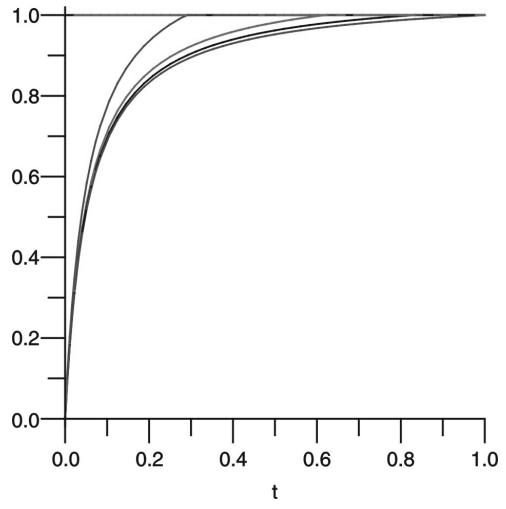

Figure 3.6: $\tilde{f}_\alpha$ for $n = 10, 50, 200$ together with f_α

Figure 3.6 depicts the modified curves $\tilde{f}_\alpha$ for $n = 10, 50$ and 200 together with f_α. The left picture in Figure 3.7 shows the minimum values for β_n that have to be used to ensure strict FDR control for SU-procedures based on f_α and $f^{(i)}_{\alpha,\kappa},\ i = 1, 2$, respectively, for varying n. In the right picture of Figure 3.7, the corresponding factors $1 + \beta_n/n$ are displayed. It is easy to prove that $\lim_{n\to\infty} \beta_n/n = 0$ in all three cases.

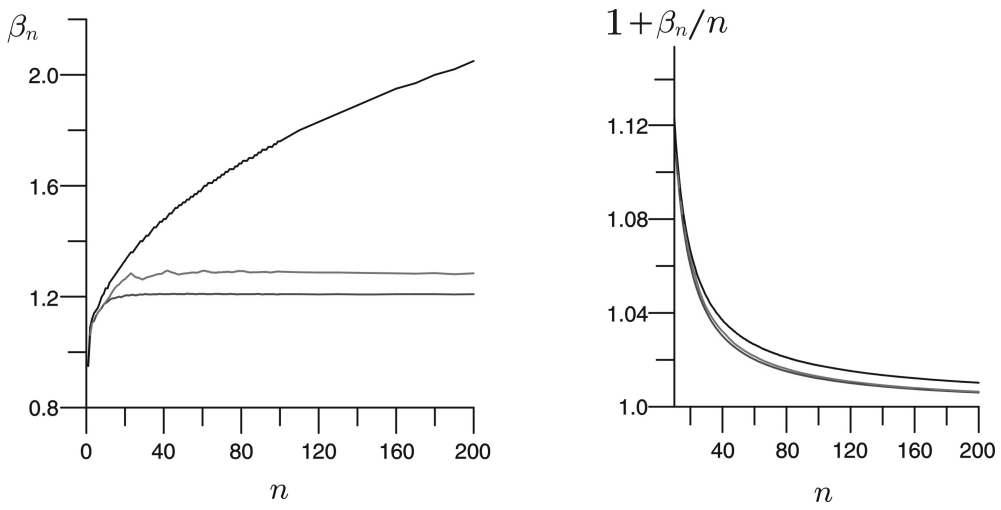

Figure 3.7: β_n and $1 + \beta_n/n$ for SU-procedures based on f_α, $f^{(1)}_{\alpha,\kappa_1}$ and $f^{(2)}_{\alpha,\kappa_2}$

Remark 3.24
In [15] (Remark to Definition 7), an SD-procedure with the universal adjustment constant $\beta_n \equiv 1.0$ was proposed.

3.7.2 Multivariate optimization problem

A more advanced approach towards finding a valid set of critical values $(\gamma_{i:n})_{i=1,\dots,n}$ (say) related to the AORC for a finite number of p-values leading to strict FDR control may consist in comprehending this as a multivariate optimization problem under constraints.
Formally, this problem can be expressed by the task to find the minimum of the target function

$$d((\gamma_{1:n}, \dots, \gamma_{n:n})) = \left(\sum_{i=1}^{n} |f_\alpha^{-1}(i/n) - \gamma_{i:n}|^p \right)^{1/p} \tag{3.38}$$

for some $p \in (0, \infty]$, i.e., minimize the L_p-distance of the set of critical values $(\gamma_{1:n}, \dots, \gamma_{n:n}) \in (0, 1)^n$ and the critical values originating from the AORC under the constraints

$$\mathrm{FDR}_{n_0,n}(\varphi^{\mathrm{SU}(\vec{\gamma})}) \le \alpha \ \ \forall n_0 = 1, \dots, n, \ \text{ and} \tag{3.39}$$

$$\frac{\gamma_{i+1:n}}{i+1} \ge \frac{\gamma_{i:n}}{i} \ \ \forall i = 1, \dots, n-1. \tag{3.40}$$

Of course, this problem can only be solved numerically by employing iteration techniques. In order to find the global minimum of (3.38) under the constraints (3.39) and (3.40), it is necessary to utilize a simplex-typed algorithm which is rather complex and goes beyond the scope of this work. Anyhow, we will at least present one simple algorithm to obtain a valid set of critical values which are close to the $\left(f_\alpha^{-1}(i/n)\right)_{i=1,\ldots,n}$ and works as follows.

Algorithm 3.25

1. We start with $(\gamma_{i:n})_{i=1,\ldots,n} = \left(f_\alpha^{-1}(i/n)\right)_{i=1,\ldots,n}$. These critical values obviously minimize the function $d((\gamma_{1:n},\ldots,\gamma_{n:n}))$ unconditionally, but they violate constraint (3.39).

2. Now we search for the smallest possible positive constants $(\varepsilon_i)_{i=1,\ldots,n}$, such that the critical values

$$(f_\alpha^{-1}(i/n) - \varepsilon_i)_{i=1,\ldots,n} \text{ or} \tag{3.41}$$

$$((1-\varepsilon_i) f_\alpha^{-1}(i/n))_{i=1,\ldots,n}, \tag{3.42}$$

   respectively, fulfill the constraints (3.39) and (3.40). The ε_i's can be found via a grid search.

Under the multiplicative ansatz (3.42) with the special choice $\varepsilon_i = i\varepsilon$ for some $\varepsilon > 0$, it can be shown that constraint (3.40) is fulfilled if

$$\varepsilon \leq \min_i (\Delta q_i / \Delta \alpha_i), \tag{3.43}$$

where

$$\begin{aligned} \Delta q_i &= \frac{f_\alpha^{-1}((i+1)/n)}{i+1} - \frac{f_\alpha^{-1}(i/n)}{i} \text{ and} \\ \Delta \alpha_i &= f_\alpha^{-1}((i+1)/n) - f_\alpha^{-1}(i/n). \end{aligned}$$

Of course, the latter policy is closely linked to the β-adjustment method described before and does not lead to the global solution of the minimization problem. Anyhow, it is simple and easy to implement and the obtained critical values remain pretty close to their initial values $\left(f_\alpha^{-1}(i/n)\right)_{i=1,\ldots,n}$. The following figure depicts the solutions of Algorithm 3.25 with ansatz (3.42) for a step-up test with $n = 25$ (left picture) and $n = 50$ (right picture) together with f_α.

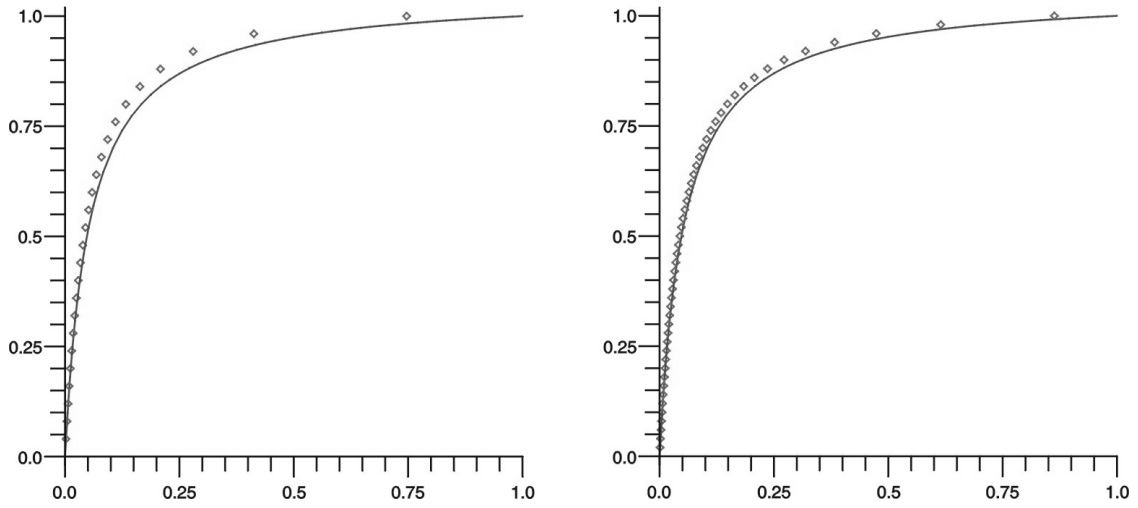

Figure 3.8: Critical values obtained by Algorithm 3.25 for $n = 25$ and $n = 50$.

Remark 3.26

Kwong et al. (cf. [161]) have also developed an algorithm for finding a valid set of critical values for step-up tests under Dirac-uniform configurations based on the distribution of V_n given in (3.36). For small n, the results are very similar to the results obtained by Algorithm 3.25. For larger n, however, their solutions do not converge to $\left(f_\alpha^{-1}(i/n)\right)_{i=1,\ldots,n}$, but the sets of critical values often show some peculiar behavior, for example that they lie on several distinct lines (cf. Figure 3.9).

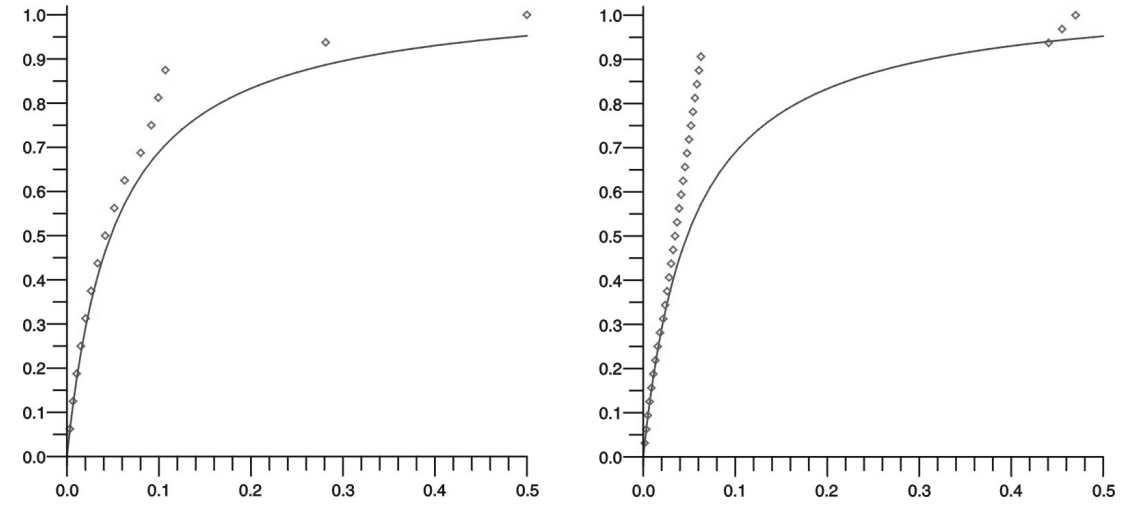

Figure 3.9: Critical values obtained by Kwong et al. (2002) for $n = 16$ and $n = 32$.

This is due to the fact that the underlying target equations of the algorithm are given by

$$\mathrm{FDR}_{n_0,n}(\varphi^{\mathrm{SU}(\vec{\gamma})}) = \alpha \ \ \forall n_0 = 1, \ldots, n, \tag{3.44}$$

which cannot be fulfilled under the constraints (3.40) for larger n. The iterative structure of the algorithm in [161] implies that the critical values in case of nonexistence of a solution of (3.44) under (3.40) are simply linearly increasing in order to fulfill at least (3.40).

3.8 Connection to Storey's approach

The AORC and methods based on it are by far not the only discussed improvements of the original linear step-up procedure. In 2004, John D. Storey et al. (cf. [275]) proposed a method in order to gain power by presenting a data-adaptive testing algorithm relying on an estimator for the proportion of true hypotheses π_0 (in Storey's nomenclature) which works as follows.

Algorithm 3.27 (Storey's method)

1. Choose an FDR level α and a tuning parameter $\lambda \in [0, 1)$.
2. Compute the p-values $p_1, \ldots, p_n$ for each individual test and denote their ecdf. by F_n.
3. Compute the estimator $\hat{\pi}_0(\lambda) = (1 - F_n(\lambda))/(1 - \lambda)$ for the proportion π_0 of true hypotheses.
4. Denote by $\widehat{\mathrm{FDR}}_\lambda(t) = \hat{\pi}_0(\lambda)t/(F_n(t) \vee n^{-1})$ an estimator function for the FDR of a single-step procedure with threshold $t \in [0, 1]$.
5. Choose the value $t_\alpha(\widehat{\mathrm{FDR}}_\lambda) = \sup\{0 \le t \le 1 : \widehat{\mathrm{FDR}}_\lambda(t) \le \alpha\}$ as threshold for the p-values and reject all hypotheses with p-values lower than or equal to $t_\alpha(\widehat{\mathrm{FDR}}_\lambda)$.

This algorithm was quickly implemented into software and is nowadays widely spread. If we only consider such t-values with $F_n(t) \ge 1/n$ for the moment and notice that

$$\begin{aligned} t_\alpha(\widehat{\mathrm{FDR}}_\lambda) &= \sup\{0 \le t \le 1 : \frac{\hat{\pi}_0(\lambda)t}{F_n(t)} \le \alpha\} \\ &= \sup\{0 \le t \le 1 : F_n(t) \ge \frac{\hat{\pi}_0(\lambda)}{\alpha}t\}, \end{aligned}$$

it becomes obvious that Algorithm 3.27 corresponds to the Benjamini-Hochberg method where α is replaced by $\alpha/\hat{\pi}_0(\lambda)$ leading to larger critical values if $\hat{\pi}_0(\lambda) < 1$. For the special choice $\lambda = 0$, we obtain $\hat{\pi}_0(0) \equiv 1$, hence, the original linear step-up procedure. For the remaining t-values with $F_n(t) < 1/n$, Algorithm 3.27 leads to determining

$$\sup\{t \in [0, p_{1:n}) : n\hat{\pi}_0(\lambda)t \le \alpha\} < p_{1:n}.$$

Hence, we reject nothing in this case and this completes the proof that Algorithm 3.27 is equivalent to an adjusted linear step-up procedure.

Now, the question is near at hand whether Storey's approach and the methods and critical values based on the asymptotically optimal rejection curve have something in common. Especially, it is of interest whether the rejection regions induced by the AORC can be derived by Storey's method for a special choice of λ. To see that this is not the case, we may define

$$\hat{\pi}_0^{\text{AORC}}(t) = 1 - F_n(t)(1-\alpha),$$

and can (in analogy to Storey et al. (2004)) construct a point estimate for the FDR of a single-step procedure with given $t \in (0,1)$ by

$$\widehat{\text{FDR}}_{\text{AORC}}(t) = \frac{t}{F_n(t)} \hat{\pi}_0^{\text{AORC}}(t) = t \left[\frac{1}{F_n(t)} - (1-\alpha) \right].$$

Finding a crossing point of F_n and f_α on $(0,1)$ is then (similarly to Storey's approach) equivalent to equating $\widehat{\text{FDR}}_{\text{AORC}}(t) = \alpha$, because

$$\begin{array}{rccc} & \widehat{\text{FDR}}_{\text{AORC}}(t) & = & \alpha \\ \iff & \alpha/t & = & 1/F_n(t) - (1-\alpha) \\ \iff & F_n(t) & = & \frac{1}{\alpha/t + (1-\alpha)} \\ \iff & F_n(t) & = & f_\alpha(t). \end{array}$$

This calculation shows that the rejection regions coming from the AORC cannot be embedded in Storey's framework, because we do not employ a scalar estimator $\hat{\pi}_0(\lambda)$, but an "estimator function" $\hat{\pi}_0(t)$ if we use his notation, i.e., the estimation of π_0 has to be done for every threshold t separately. This characterizes the difference between a fixed rejection curve approach and Storey's variable rejection threshold depending on the data and the resulting estimator $\hat{\pi}_0(\lambda)$. Both approaches are not transferable into another.

The latter observation shows that both test methods are in competition and makes it interesting to discuss assets and drawbacks of both procedures and to compare their performance in various distributional settings. A systematic comparison of both methods and some others under several setups is the topic of the following Chapter 4.

In general, one can summarize the pro's and con's of the two methods described before, especially in comparison with the original linear step-up procedure which both methods claim to improve, as follows.

Pro linear step-up:

- Actual FDR of φ^{LSU} depends only on the proportion of true nulls, not on the values of the alternative parameters
- Theoretically valid under positive dependency
- Easy and intuitive

Contra linear step-up:

- Not α-exhausting for $n_0 < n$
- Often low power for small proportions of true null hypotheses

Pro Storey method:

- Flexible with respect to concrete underlying p-value distribution
- If $\hat{\pi}_0(\lambda)$ is a good estimate for π_0, the resulting procedure is close to the optimal "oracle procedure" (see Chapter 4 below)

Contra Storey method:

- The realized estimate $\hat{\pi}_0(\lambda)$ can be greater than 1 (this happens especially if discrete p-values are involved); then the procedure is less powerful than the linear step-up procedure
- The estimator $\hat{\pi}_0(\lambda)$ introduces a new variance component which makes the FDP of such a procedure even more volatile, especially under dependency
- The choice of a good tuning parameter λ is a sensible issue

Pro AORC-based procedures:

- Works for any ζ and is even theoretically optimal under Dirac-uniform configurations
- No tuning parameter necessary
- No estimation of ζ necessary (implicitly covered)
- Even the adjusted critical values for a step-up test based on f_α are an almost uniform improvement of Simes' critical values. For example, for $n = 100$ and $\alpha = 0.05$ we obtain with the $\beta_{100} = 1.76$ adjustment described in Subsection 3.7.1 that only the smallest critical value $\alpha_{1:n}$ is smaller than its Simes' counterpart α/n, while all other 99 $\alpha_{j:n}$'s are larger than $j\alpha/n$.
- Theoretically valid and often superior to Storey's approach if the underlying p-values are stochastically larger than a UNI$[0, 1]$-distributed random variable (cf. Chapter 4).

Contra AORC-based procedures:

- Fail to control the FDR under positive dependency
- Designed to control the FDR even under the worst case scenario of infinitely large alternative parameters and therefore often not powerful for realistic alternatives / small effect sizes

- Not flexible with regard to the actual data, i.e., no data-dependent critical values (fixed rejection curve drawback)

On the whole, one should expect that Storey's method works better in case of strictly continuous distributions and smaller effect sizes while the AORC-based procedures will behave better if we have extreme parameter configurations. A comparison is not really fair because both methods are designed to achieve different goals or, in other words, have different underlying optimality criterions. However, on a descriptive basis it may be interesting to study their behaviors under various circumstances in more detail. This is the topic of the next chapter.

Chapter 4

Power study for some FDR-controlling test procedures

In Chapter 3, we have seen that there are improvements of the Benjamini-Hochberg method with which we *can* gain power while (at least asymptotically) keeping the FDR level α. Here, we are now concerned with *how much* gain of power is possible with these methods.

Using the nomenclature introduced in Table 1.1, we recall our formal definition of the power of a multiple test procedure $\varphi = (\varphi_1, \ldots, \varphi_n)$, which was given in Definition 1.3 as

$$\text{power}_n(\varphi) = \mathbb{E}_\vartheta \left(\frac{S_n}{n_1 \vee 1} \right). \tag{4.1}$$

Supposing that the proportion $\zeta_n = n_0/n$ of true hypotheses converges with $\lim_{n\to\infty} \zeta_n = \zeta$, we can herewith calculate the asymptotic power of a multiple test procedure φ based on a rejection curve r leading to the rejection region $[0, t^*]$ for $0 < \zeta < 1$ by

$$\begin{aligned} \text{power}_\infty(\varphi) &= \lim_{n\to\infty} \frac{S_n}{n_1 \vee 1} = \lim_{n\to\infty} \frac{R_n/n - V_n/n}{(n_1 \vee 1)/n} \\ &= \frac{r(t^*) - \zeta t^*}{1 - \zeta}. \end{aligned} \tag{4.2}$$

Obviously, under the Dirac-uniform configurations DU(ζ), we have (finite and asymptotic) power equal to 1.

In order to make our comparison as fair as possible, we only consider step-up procedures based on rejection curves, namely

$$\begin{aligned} r^{(1)}(t) &= t/\alpha && \text{Benjamini-Hochberg procedure,} \\ r^{(2)}(t) &= \hat{\pi}_0(\lambda)t/\alpha && \text{Storey procedure,} \\ r^{(3)}(t) &= f^{(2)}_{\alpha,\kappa_2}(t) && \text{adjusted AORC-based procedure,} \\ r^{(4)}(t) &= n_0 t/(n\alpha) && \text{"Oracle" procedure,} \end{aligned}$$

where the tuning parameter λ was chosen equal to 0.5 and $\kappa_2 = (1/2 - \alpha)/(1 - \alpha)$. Both choices imply that the maximum possible rejection threshold for procedures (2) and (3) is equal to $1/2$.

As stated before (cf. Section 3.8), the crucial difference between these two procedures is the fact that the AORC-based procedure utilizes a fixed, deterministic rejection curve while in Storey's method the rejection curve is a random object estimated from the observed data. Of course, the "oracle" procedure is impracticable and only serves as a benchmark, because it has maximum power. Loosely formulated, we want to find out three things:

(i) How much power gain is possible with Storey's method and with the AORC-based method with respect to φ^{LSU} under which parameter configurations ?

(ii) Which of the two improvements is more powerful in which cases ?

(iii) How far are the two improvements away from optimal power under which parameter configurations ?

Our simulation setup is as follows. For three different choices of n, i.e., $n = 40$ (moderate problem size), $n = 400$ (large problem size) and $n = 4000$ (asymptotic case, justification see below), we consider the proportions $\zeta_n = 90\%, 75\%, 50\%, 25\%$ and 10% and investigate the power and the realized FDR of the step-up procedures based on the rejection curves $r^{(i)}, i = 1, \ldots, 4$, under varying parameter constellations. The FDR level is chosen as $\alpha = 5\%$ and the quantities FDR_n and power_n are estimated by their average values in $B = 10,000$ Monte Carlo replications, i.e.,

$$\widehat{\text{FDR}}_n = \frac{1}{B}\sum_{b=1}^{B} q_b \text{ with } q_b = \frac{v_{n,b}}{r_{n,b} \vee 1},$$

$$\widehat{\text{power}}_n = \frac{1}{B}\sum_{b=1}^{B} \frac{s_{n,b}}{n_1},$$

where the entities written in lowercase letters denote the realizations of the corresponding random variables defined in Table 1.1. We subdivide our power study into two parts, namely the simple hypotheses case and the composite hypotheses case.

4.1 Simple hypotheses case

For illustrative purpose, we again study the one-sided normal means problem

$$H_i : \{\mu = 0\} \text{ vs. } K_i : \{\mu > 0\}, i = 1, \ldots, n,$$

with test statistics $T_i \sim \mathcal{N}(0, \sigma^2)$ i.i.d., $i = 1, \ldots, n_0$ for the true hypotheses and $T_j \sim \mathcal{N}(\mu, \sigma^2)$ i.i.d., $j = n_0 + 1, \ldots, n$ for the false hypotheses with $0 < n_0 < n$. For the sake of simplicity and without loss of generality, we assume unit variance, i.e., $\sigma^2 = 1$. Moreover, we add $\mu = \infty$ to the model such that Dirac-uniform configurations can be covered in this framework as well. Suitable p-values for testing H_i versus K_i are then given by $p_i(t) = \mathbb{P}_{H_i}(T_i > t) = 1 - \Phi(t)$ and for

the distribution functions of the $p_i(T_i)$ under H_i and K_i, respectively, we get for $t \in [0,1]$ the representations

$$G_0(t) = t, \tag{4.3}$$

$$G_1(t) = \mathbb{P}_0(T_{n_0+1} > \Phi^{-1}(1-t)) = 1 - \Phi(\Phi^{-1}(1-t) - \mu), \tag{4.4}$$

$$F_\infty(t|\zeta,\mu) = \zeta t + (1-\zeta)(1 - \Phi(\Phi^{-1}(1-t) - \mu)). \tag{4.5}$$

We now investigate the power and the realized FDR of the step-up procedures based on the rejection curves $r^{(i)}, i = 1,\ldots,4$, in case of

$$\begin{aligned} \mu_s &= 0.5 \quad \text{small / minor effect,} \\ \mu_r &= 2.0 \quad \text{realistic / relevant effect, and} \\ \mu_a &= 5.0 \quad p - \text{value approximately} \sim \varepsilon_0. \end{aligned}$$

The following tables list our simulation results for each considered value of ζ_n separately.

$\zeta_n = 0.9$	$\mu_s = 0.5$			$\mu_r = 2.0$			$\mu_a = 5.0$		
n	40	400	4000	40	400	4000	40	400	4000
$\widehat{\text{FDR}}_n^{(1)}$	4.40%	4.42%	4.29%	4.04%	4.49%	4.52%	4.62%	4.51%	4.50%
$\widehat{\text{power}}_n^{(1)}$	0.66%	0.090%	0.012%	19.58%	12.08%	10.36%	99.27%	99.29%	99.27%
$\widehat{\text{FDR}}_n^{(2)}$	4.85%	4.65%	4.57%	4.80%	5.06%	4.98%	5.21%	5.01%	5.00%
$\widehat{\text{power}}_n^{(2)}$	0.71%	0.094%	0.012%	20.80%	13.11%	11.40%	99.34%	99.36%	99.35%
$\widehat{\text{FDR}}_n^{(3)}$	4.53%	4.45%	4.29%	4.37%	4.59%	4.56%	5.25%	5.01%	5.00%
$\widehat{\text{power}}_n^{(3)}$	0.68%	0.090%	0.012%	20.27%	12.28%	10.47%	99.39%	99.36%	99.35%
$\widehat{\text{FDR}}_n^{(4)}$	4.87%	4.96%	4.92%	4.71%	5.09%	5.01%	5.10%	5.00%	5.00%
$\widehat{\text{power}}_n^{(4)}$	0.73%	0.10%	0.013%	20.82%	13.23%	11.46%	99.38%	99.36%	99.35%

Table 4.1: Power study in the simple hypotheses case - Results for $\zeta_n = 0.9$

Utilizing (4.2) and the explicit formula for $F_\infty(t|\zeta,\mu)$, we can additionally calculate the asymptotic power of the four procedures for the interesting case $\mu_r = 2.0$ numerically and get (for $\zeta_n \to \zeta = 0.9$) the values

$$\begin{aligned} \text{power}_\infty^{(1)} &\approx 10.14\%, \quad \text{power}_\infty^{(2)} \approx 11.22\%, \\ \text{power}_\infty^{(3)} &\approx 10.25\%, \quad \text{power}_\infty^{(4)} \approx 11.28\%. \end{aligned}$$

In this sense, it seems justified to call the setup $n = 4000$ already the "asymptotic case". As expected, the linear step-up procedure performs nearly equally well than its competitors here. Storey's method is close to the oracle procedure, because the assumption of a UNI$[0,1]$ distribution for the p-values under the null hypotheses holds true and the choice of λ is appropriate (in the linear part of the limiting ecdf. of the p-values). For small family sizes and large parameter values, Storey's method and the AORC-based step-up test both tend to anti-conservativity.

$\zeta_n = 0.75$	$\mu_s = 0.5$			$\mu_r = 2.0$			$\mu_a = 5.0$		
n	40	400	4000	40	400	4000	40	400	4000
$\widehat{\text{FDR}}_n^{(1)}$	3.65%	3.96%	3.87%	3.80%	3.76%	3.74%	3.70%	3.72%	3.75%
$\widehat{\text{power}}_n^{(1)}$	0.62%	0.098%	0.013%	26.80%	22.93%	22.51%	99.74%	99.73%	99.72%
$\widehat{\text{FDR}}_n^{(2)}$	4.09%	4.28%	4.27%	4.98%	4.95%	4.90%	4.94%	4.98%	5.00%
$\widehat{\text{power}}_n^{(2)}$	0.71%	0.11%	0.015%	31.11%	28.09%	27.79%	99.80%	99.81%	99.81%
$\widehat{\text{FDR}}_n^{(3)}$	3.76%	3.96%	3.88%	4.33%	4.01%	3.97%	5.13%	5.00%	5.00%
$\widehat{\text{power}}_n^{(3)}$	0.65%	0.098%	0.013%	28.68%	24.18%	23.62%	99.81%	99.81%	99.81%
$\widehat{\text{FDR}}_n^{(4)}$	4.92%	5.04%	5.18%	5.02%	5.03%	4.98%	4.94%	4.98%	5.00%
$\widehat{\text{power}}_n^{(4)}$	0.85%	0.13%	0.019%	31.90%	28.48%	28.10%	99.81%	99.81%	99.81%

Table 4.2: Power study in the simple hypotheses case - Results for $\zeta_n = 0.75$

Again, we additionally report the asymptotic power of the four procedures in case of $\mu_r = 2.0$ under the parameter configuration $\zeta_n \to \zeta = 0.75$. We obtain

$$\begin{aligned} \text{power}_\infty^{(1)} &\approx 22.45\%, \quad \text{power}_\infty^{(2)} \approx 27.74\%, \\ \text{power}_\infty^{(3)} &\approx 23.55\%, \quad \text{power}_\infty^{(4)} \approx 28.05\%. \end{aligned}$$

Here, with shrunken ζ_n, the improvements of φ^{LSU} begin to considerably outperform the linear step-up procedure. Again, the Storey procedure shows a superior behavior. The AORC-based step-up method is not much more powerful than φ^{LSU}, because ζ_n is too large to lead to substantial power gain by the procedure based on f_α. It is remarkable that even the fully α-exhausting oracle procedure can hardly detect small effects like $\mu_s = 0.5$ in such a normal means-comparisons problem.

$\zeta_n = 0.5$	$\mu_s = 0.5$			$\mu_r = 2.0$			$\mu_a = 5.0$		
n	40	400	4000	40	400	4000	40	400	4000
$\widehat{\text{FDR}}_n^{(1)}$	2.34%	2.56%	2.46%	2.51%	2.50%	2.50%	2.49%	2.49%	2.50%
$\widehat{\text{power}}_n^{(1)}$	0.68%	0.11%	0.015%	37.55%	36.29%	36.14%	99.89%	99.88%	99.89%
$\widehat{\text{FDR}}_n^{(2)}$	2.86%	3.15%	3.12%	4.79%	4.81%	4.80%	4.98%	4.98%	5.00%
$\widehat{\text{power}}_n^{(2)}$	0.90%	0.14%	0.020%	51.72%	52.46%	52.49%	99.95%	99.96%	99.96%
$\widehat{\text{FDR}}_n^{(3)}$	2.38%	2.57%	2.46%	3.46%	3.19%	3.15%	5.33%	5.01%	5.00%
$\widehat{\text{power}}_n^{(3)}$	0.73%	0.11%	0.015%	44.50%	42.01%	41.62%	99.96%	99.97%	99.96%
$\widehat{\text{FDR}}_n^{(4)}$	4.71%	4.97%	5.13%	5.01%	5.03%	5.02%	5.02%	4.98%	5.00%
$\widehat{\text{power}}_n^{(4)}$	1.5%	0.26%	0.040%	54.35%	53.82%	53.70%	99.96%	99.97%	99.96%

Table 4.3: Power study in the simple hypotheses case - Results for $\zeta_n = 0.5$

Considering only $\mu_r = 2.0$ and letting $\zeta_n \to \zeta = 0.5$, we now get the asymptotic power values

$$\begin{aligned} \text{power}_\infty^{(1)} &\approx 36.14\%, \quad \text{power}_\infty^{(2)} \approx 52.51\%, \\ \text{power}_\infty^{(3)} &\approx 41.62\%, \quad \text{power}_\infty^{(4)} \approx 53.71\%. \end{aligned}$$

It becomes apparent that the improvements with the step-up test based on the AORC grow with decreasing ζ. Moreover, the general detectability of any of the considered procedures increases with increasing proportion of false hypotheses.

$\zeta_n = 0.25$	$\mu_s = 0.5$			$\mu_r = 2.0$			$\mu_a = 5.0$		
n	40	400	4000	40	400	4000	40	400	4000
$\widehat{\text{FDR}}_n^{(1)}$	1.17%	1.43%	1.15%	1.24%	1.25%	1.25%	1.27%	1.25%	1.25%
$\widehat{\text{power}}_n^{(1)}$	0.75%	0.11%	0.017%	46.17%	45.53%	45.48%	99.94%	99.94%	99.94%
$\widehat{\text{FDR}}_n^{(2)}$	1.69%	1.87%	1.62%	4.37%	4.42%	4.39%	4.99%	4.99%	5.00%
$\widehat{\text{power}}_n^{(2)}$	1.15%	0.18%	0.029%	74.65%	77.28%	77.53%	99.99%	100%	100%
$\widehat{\text{FDR}}_n^{(3)}$	1.20%	1.43%	1.14%	3.35%	2.30%	2.24%	6.48%	5.06%	5.00%
$\widehat{\text{power}}_n^{(3)}$	0.80%	0.11%	0.017%	65.17%	60.87%	60.56%	100%	100%	100%
$\widehat{\text{FDR}}_n^{(4)}$	4.89%	5.05%	5.04%	4.94%	5.01%	5.00%	5.00%	4.99%	5.00%
$\widehat{\text{power}}_n^{(4)}$	4.75%	1.20%	0.034%	80.54%	80.50%	80.51%	100%	100%	100%

Table 4.4: Power study in the simple hypotheses case - Results for $\zeta_n = 0.25$

Under $\mu_r = 2.0$ and for $\zeta_n \to \zeta = 0.25$, we obtain

$$\begin{aligned} \text{power}_\infty^{(1)} &\approx 45.47\%, \quad \text{power}_\infty^{(2)} \approx 77.57\%, \\ \text{power}_\infty^{(3)} &\approx 60.52\%, \quad \text{power}_\infty^{(4)} \approx 80.51\% \end{aligned}$$

as asymptotic power values. For such low proportions of ζ_n, the weakness of the linear step-up procedure (does not exhaust α resulting in poor power) can be seen clearly. Both the AORC-based procedure and the Storey procedure show a much more performant behavior, whereby the latter procedure still remains close to the oracle procedure and is the undisputed best choice of the three procedures under investigation that can be carried out in practice.

The last setting $\zeta_n = 0.1$ seems irrelevant for practical considerations. It is mainly included to show that the adjustment of f_α works as expected, because here we obtain crossing points of the limiting ecdf. of the p-values and the resulting rejection curve in the linearly continued part of $\tilde{f}_\alpha$. Indeed, this "protection" works and even under the extreme value $\mu_a = 5.0$ no violation of the FDR-level occurs with the AORC-based test.

$\zeta_n = 0.1$	$\mu_s = 0.5$			$\mu_r = 2.0$			$\mu_a = 5.0$		
n	40	400	4000	40	400	4000	40	400	4000
$\widehat{\text{FDR}}_n^{(1)}$	0.62%	0.58%	0.42%	0.49%	0.50%	0.50%	0.51%	0.50%	0.50%
$\widehat{\text{power}}_n^{(1)}$	0.82%	0.12%	0.019%	50.36%	49.94%	49.84%	99.95%	99.95%	99.95%
$\widehat{\text{FDR}}_n^{(2)}$	0.86%	0.80%	0.67%	3.36%	3.55%	3.55%	4.03%	4.92%	5.00%
$\widehat{\text{power}}_n^{(2)}$	1.44%	0.23%	0.040%	88.78%	92.86%	93.26%	100%	100%	100%
$\widehat{\text{FDR}}_n^{(3)}$	0.65%	0.59%	0.42%	3.87%	1.96%	1.59%	4.83%	4.86%	4.93%
$\widehat{\text{power}}_n^{(3)}$	0.88%	0.12%	0.019%	88.16%	80.58%	78.56%	100%	100%	100%
$\widehat{\text{FDR}}_n^{(4)}$	4.88%	4.98%	5.00%	4.98%	4.99%	5.00%	4.96%	5.01%	5.00%
$\widehat{\text{power}}_n^{(4)}$	32.18%	25.08%	24.02%	97.15%	97.13%	97.12%	100%	100%	100%

Table 4.5: Power study in the simple hypotheses case - Results for $\zeta_n = 0.1$

Finally, the asymptotic power values in case of $\mu_r = 2.0$ and $\zeta_n \to \zeta = 0.1$ are (numerically) given by

$$\begin{aligned} \text{power}_\infty^{(1)} &\approx 49.85\%, \quad \text{power}_\infty^{(2)} \approx 93.33\%, \\ \text{power}_\infty^{(3)} &\approx 78.41\%, \quad \text{power}_\infty^{(4)} \approx 97.13\%. \end{aligned}$$

On the whole, we have shown that in this setting with high practical relevance, the proposed improvements of φ^{LSU} considerably outperform the original Benjamini-Hochberg method if the proportion of true nulls is bounded away from 1. It has to be conceded that Storey's method under these regularity circumstances (strictly continuous distributions, simple null hypotheses) performs very well and is preferable over the AORC-based method considered here.

4.2 Composite hypotheses case

Here, we assume now that the null hypotheses are composite in the one-sided normal means problem. More precisely, we consider the multiple test problem

$$H_i : \{\mu \leq 0\} \text{ vs. } K_i : \{\mu > 0\}, i = 1, \ldots, n,$$

where μ is the expectation of a normal distribution and the test statistics are of the form

$$\begin{aligned} T_i &\sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d.}, \quad i = 1, \ldots, n_0/2, \\ T_j &\sim \mathcal{N}(-\mu, \sigma^2) \quad \text{i.i.d.}, \quad j = n_0/2 + 1, \ldots, n_0, \text{ and} \\ T_k &\sim \mathcal{N}(\mu, \sigma^2) \quad \text{i.i.d.}, \quad k = n_0 + 1, \ldots, n, \end{aligned}$$

with $0 < n_0 < n$. This means that for the first half of the true hypotheses, we are at the boundary of the null hypotheses while for the second half of the true hypotheses we are far in the inside of the corresponding hypothesis H_j. Again, we set $\sigma^2 = 1$. Typically, as described in Section 1.2, the probabilities $p_i(t) = \mathbb{P}_0(T_i > t) = 1 - \Phi(t)$ at the boundary of the H_i's are used as p-values

for testing H_i versus K_i. The distribution functions of the $p_i(T_i)$ under H_i and K_i, respectively, now have for $t \in [0,1]$ the representations

$$G_0(t) = \frac{1}{2}t + \frac{1}{2}(1 - \Phi(\Phi^{-1}(1-t) + \mu)), \tag{4.6}$$

$$G_1(t) = 1 - \Phi(\Phi^{-1}(1-t) - \mu), \tag{4.7}$$

$$F_\infty(t|\zeta,\mu) = \frac{\zeta}{2}\left(t + (1 - \Phi(\Phi^{-1}(1-t) + \mu))\right) + (1-\zeta)(1 - \Phi(\Phi^{-1}(1-t) - \mu)). \tag{4.8}$$

Again, we investigate the power and the realized FDR of the step-up procedures based on the rejection curves $r^{(i)}, i = 1, \ldots, 4$, in case of $\mu_r = 2.0$ (the realistic / relevant effect already used in the previous section). The following table lists our simulation results for the considered values of ζ_n.

$\mu_r = 2.0$	$\widehat{\text{FDR}}_n^{(1)}$ and $\widehat{\text{power}}_n^{(1)}$			$\widehat{\text{FDR}}_n^{(2)}$ and $\widehat{\text{power}}_n^{(2)}$		
n	40	400	4000	40	400	4000
$\zeta_n = 0.1$	0.24%	0.26%	0.25%	1.24%	1.34%	1.34%
	50.17%	49.87%	49.77%	84.57%	88.27%	88.61%
$\zeta_n = 0.25$	0.63%	0.63%	0.62%	1.51%	1.56%	1.55%
	46.02%	45.42%	45.30%	66.33%	68.38%	68.54%
$\zeta_n = 0.5$	1.25%	1.28%	1.25%	1.60%	1.66%	1.64%
	37.57%	35.95%	35.86%	42.45%	42.28%	42.32%
$\zeta_n = 0.75$	1.89%	1.86%	1.88%	1.89%	1.86%	1.88%
	26.62%	22.46%	22.13%	26.64%	22.46%	22.13%
$\zeta_n = 0.9$	2.22%	2.17%	2.29%	2.22%	2.17%	2.29%
	19.13%	11.72%	10.07%	19.13%	11.72%	10.07%

$\mu_r = 2.0$	$\widehat{\text{FDR}}_n^{(3)}$ and $\widehat{\text{power}}_n^{(3)}$			$\widehat{\text{FDR}}_n^{(4)}$ and $\widehat{\text{power}}_n^{(4)}$		
n	40	400	4000	40	400	4000
$\zeta_n = 0.1$	1.56%	0.82%	0.75%	2.60%	2.60%	2.59%
	83.73%	78.00%	77.10%	96.88%	96.93%	96.92%
$\zeta_n = 0.25$	1.35%	1.13%	1.10%	2.49%	2.53%	2.51%
	62.92%	60.12%	59.75%	80.21%	80.01%	79.95%
$\zeta_n = 0.5$	1.71%	1.59%	1.55%	2.46%	2.52%	2.51%
	43.93%	41.31%	41.09%	53.79%	53.11%	53.05%
$\zeta_n = 0.75$	2.08%	2.01%	2.00%	2.48%	2.50%	2.50%
	28.55%	23.59%	23.16%	31.56%	27.80%	27.54%
$\zeta_n = 0.9$	2.34%	2.21%	2.31%	2.42%	2.41%	2.54%
	19.67%	11.89%	10.18%	20.20%	12.74%	11.15%

Table 4.6: Power study in the composite hypotheses case - Results for $\mu_r = 2.0$

Apparently, for $\zeta \geq 0.75$, the Storey method is almost always equivalent to the Benjamini-Hochberg step-up procedure and therefore gets outperformed by the step-up procedure based on $r^{(3)}$. This is due to the fact that the estimator $\hat{\pi}_0$ used in Storey's approach is almost always equal to 1 in such cases.
It may be argued that the numerical power gain by the AORC-based procedure is not very large, but obviously we might have chosen a different parameter configuration such that the results would have been more impressing. However, focus should have been laid on the fact that the data-adaptive method developed by Storey is no improvement of the linear step-up procedure at all in case of composite null hypotheses and large proportion of true nulls.

4.3 Summary

Returning to our initial questions, we can now give the following answers.

(i) In test problems with continuous test statistics, simple null hypotheses and moderate alternative parameter values, Storey's approach seems to be the method of choice from today's perspective.

(ii) If, under the assumptions listed under (i), the alternative parameters grow larger, the AORC-based methods perform equally well with respect to power and FDR control. Moreover, they have the advantage that no tuning parameter is necessary.

(iii) For composite hypotheses, the AORC-based methods are preferable, especially if the proportion of true null hypotheses is not too large and some parameter values belonging to true null hypotheses are assumed to lie inside of the null hypotheses' parameter space.

(iv) If the multiple test problem employs discrete test statistics and the corresponding p-values are not uniformly distributed on the unit interval but stochastically larger, both methods are most likely to have low power. This is due to the fact that crossing points of the ecdf. of the p-values and the underlying rejection curves are decisive objects determining the decision rule. One way to face this problem is to use randomized p-values in such a situation again leading to UNI$[0,1]$-distributed p-values under null hypotheses. This approach has been worked out for the example application of testing for Hardy-Weinberg equilibrium in genetics studies in [95].

Chapter 5

Concluding remarks and outlook

With this work, two main aspects should have been worked out.

First of all, the calculations in Chapter 2 illustrate that the false discovery proportion FDP is a very volatile quantity (a random variable with possibly large variance) under positive dependency. In the i.i.d. case, R_n/n converges almost surely (cf. Lemma A.2 in [91] or Theorem 5 in [275]), while under dependency it typically has positive variance. We have seen this clearly in our D-EX-models, where (depending on $Z = z$) in some cases $R_n(z)/n$ (determined by the ordinate of the largest crossing point) was close to zero while in other cases it was close to 1. We are not the only ones who have discovered this effect. Recently, under researchers working on this field even some doubt has arisen if control of the FDR is a reasonable error handling criterion under dependency at all. Indeed, controlling only the expectation of the FDP when it is highly variable is quite unsatisfactory, because we can only state that the *average* proportion of false significances is bounded by α, but can make no statement about the individual experiment that we evaluate.

Consequently, it was suggested not only to look at the first moment of the FDP, but try to construct confidence intervals taking the second moment into consideration as well or even look at the whole distribution of the FDP. Two remunerative references for this topic are [104] and [170]. In the latter article, alternatively, usage of the so-called k-FWER was proposed. Here, the probability $\mathbb{P}(V_n > k)$ is controlled at level α for some integer $k \geq 0$. The authors argue that this is a good tradeoff between relaxation of the error rate and keeping courtesy to the experimenter in a way that we are able to guarantee only k false significances with statistical certainty $(1 - \alpha)$. Stepwise test procedures for k-FWER control are developed in [170] as well.

In Chapters 3 and 4, we have pointed out that the classical Benjamini-Hochberg procedure φ^{LSU} from 1995 (in some sources even called *the FDR procedure* or something similar) gives room for improvements, but this room is limited. Especially if the proportion ζ_n of true hypotheses is close to 1, it is hard to outperform φ^{LSU} to a considerable amount. This is also reflected in the shape of our asymptotically optimal rejection curve f_α which starts with the same derivative as Simes'

line in 0 and deviates from it only moderately with growing argument. However, with procedures based on f_α it is possible to reject hypotheses with p-values larger than α while keeping the FDR level and this leads to a substantial gain in power for smaller values of ζ_n where the ecdf. of the p-values typically has a concave shape and often intersects f_α at a larger abscissa than α. One may argue that small values of ζ_n (many false hypotheses at hand) seem quite artificial and are not encountered in the relevant application fields where we e.g. scan maybe $100,000$s of SNPs to find a dozen candidates, but even in the microarray analyses framework, situations with $\zeta_n \approx 0.6$ are not unusual if we test for Hardy-Weinberg disequilibrium in strata, for example.

In future research, it may be interesting to study some more complicated dependency concepts than exchangeability in more detail, e.g. investigate two-sided t-type test statistics of the form $T_i = |X_i - X_0|/S$. Also, it may be of interest how the FDR behaves under more realistic configurations of the alternative parameters than the Dirac-models. Determination of least favorable parameter configurations for step-up-down test procedures remains a very challenging issue.

As far as the AORC is concerned, it may be worth to study the behavior of procedures based on f_α in PRDS situations. As shown in Remark 3.14, further adjustments of f_α will be necessary to implement FDR-controlling procedures in such a setup. Exact proofs for FDR control under dependency for any stepwise test procedure based on critical values different from Simes' critical values α_k are very complicated, because the key expression $\mathbb{P}_\vartheta(R_n = j | p_i \leq \alpha_{j:n})$ is hard to handle and the difference $[\alpha_{j:n}/j - \alpha_{j-1:n}/(j-1)]$ appearing in all the proofs in Chapter 3 only vanishes for critical values with the structure of Simes' α_k's. Anyhow, it seems possible that FDR control for AORC-based procedures can be proved under other well-defined dependency structures such as, e. g., block dependence as relevant in GWA analyses. This may be investigated in future research.

Finally, the exact solution of the minimization problem given in Subsection 3.7.2 employing some simplex-type algorithm can lead to an optimal set of critical values even for any finite n.

Appendix A

Numerical simulations and calculations

A.1 Simulations for FDR under dependency

In this section, we present some simulations for the FDR behavior of φ^{LSU} under the D-EX-N models introduced in Section 2.3. The setup for these simulations has been chosen as follows. For given ζ and for varying ρ, we generated $n\zeta = 1000\zeta$ test statistics T_i, $i = 1, \ldots, n\zeta$ as given in the introductory part of Section 2.3 with $\vartheta_i = 0$ for all $i = 1, \ldots, n\zeta$ and computed the corresponding p-values. The $n(1-\zeta)$ remaining p-values were set equal to zero. This was done independently in $m = 10,000$ simulation runs. Denoting the FDP for a linear step-up procedure with such p-values for a particular run by $Q_n = V_n/(R_n \vee 1)$, we estimated $\text{FDR}_\infty(\zeta)$ (for one ρ) by

$$\widehat{\text{FDR}_\infty(\zeta)} = \frac{1}{m}\sum_{j=1}^{m} q_n^{(j)} = \bar{Q}_{n,m},$$

where the index j indicates the j-th of the $m = 10,000$ simulation runs and $q_n^{(j)}$ denotes the j-th realization of Q_n. The following figures show the results for $\zeta = 0.2, 0.4, 0.6, 0.8, 0.9$ and 1.0 depending on ρ.

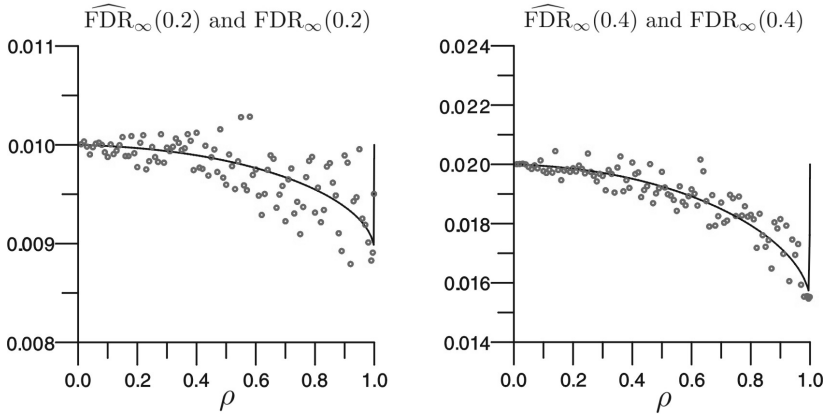

Figure A.1: Simulation for the FDR of φ^{LSU} in the D-EX-N model with $\zeta = 0.2$ and $\zeta = 0.4$

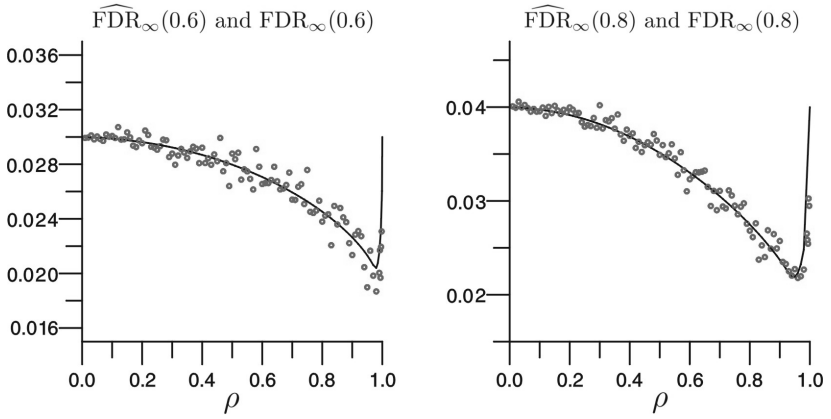

Figure A.2: Simulation for the FDR of φ^{LSU} in the D-EX-N model with $\zeta = 0.6$ and $\zeta = 0.8$

The simulated data points are represented by the small dots while the solid curves reflect the theoretical values. It can be seen that the values obtained by simulation reproduce the shape of the theoretical FDR_∞-curves for $\zeta < 1.0$ remarkably well, especially for larger ζ's where the bulged shape of the FDR_∞-curve becomes more and more distinct.

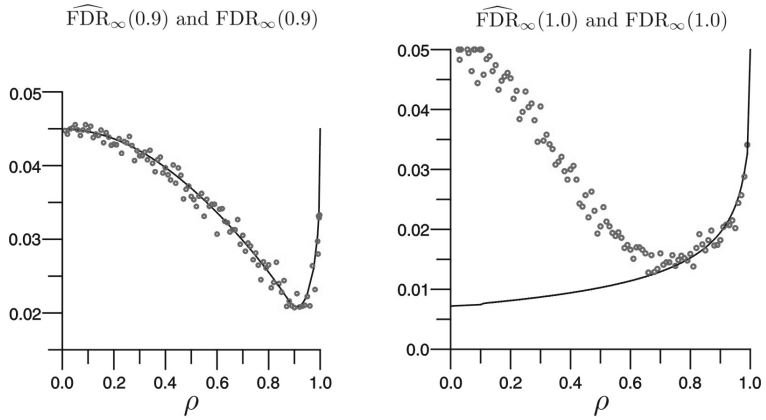

Figure A.3: Simulation for the FDR of φ^{LSU} in the D-EX-N model with $\zeta = 0.9$ and $\zeta = 1.0$

For $\zeta = 1.0$ and larger ρ's, this is still the case while for smaller ρ's the simulated data points have much larger ordinates than the y-coordinates of the $\mathrm{FDR}_\infty(1)$-curve at the corresponding abscissas. This phenomenon is due to the fact that here the FDR behavior in the limiting case differs substantially from the finite case. This becomes clear if we consider the case $\rho \to 0^+$, where the order of limits plays a severe role, because it holds

$$\lim_{n\to\infty}\left(\lim_{\rho\to 0^+}\text{FDR}_n(1)\right) = \alpha, \text{ but} \tag{A.1}$$

$$\lim_{\rho\to 0^+}\left(\lim_{n\to\infty}\text{FDR}_n(1)\right) = \Phi(-\sqrt{-2\ln(\alpha)}) << \alpha. \tag{A.2}$$

Obviously, the limit given in (A.2) cannot be reproduced by any simulation. Consequently, it turns out that n has to be unrealistically large in order to reflect the limiting behavior for smaller ρ's adequately if $\zeta = 1.0$. For $\zeta < 1$, such artefacts do not occur and the simulation gives valid results even for moderate $n \sim 1000$ as relevant in practice.

A.2 Adjusted procedures based on the AORC

In this section, we refer again to the example procedures based on the AORC presented in Examples 3.5, 3.7 and 3.8. In Corollary 3.19, it has been shown that the proposed procedures asymptotically control the FDR under any parameter configuration if the distributional assumptions (D3)-(I2) for the p-values hold and the proportion ζ_n of true hypotheses converges to some $\zeta \in [0, 1]$ for $n \to \infty$. However, in the finite case the procedures do not strictly control the FDR (cf. the discussion around Figure 3.5). Here, we now numerically investigate how liberal the procedures really are for n ranging from 5 to 1000.

A.2.1 SUD-procedure, Example 3.5

We investigate the SUD-procedure based on the original AORC with parameter $\lambda_n = \left\lceil \frac{n}{2-\alpha} \right\rceil$. The FDR level was set to $\alpha = 5\%$. As already mentioned in Section 3.7, the evaluation of the pmf. of V_n for an SUD-procedure using the formula derived in Lemma 3.22 becomes numerically difficult for larger values of n. Therefore, we employed computer simulations for the results for $n \geq 100$ with $M = 10000$ Monte Carlo replications.

In case of an SUD-procedure, it is not clear if Dirac-uniform configurations are least favorable. Therefore, the following table lists the upper bound $b_n^* = b(n_0^*, n) = \max_{n_0\in\mathbb{N}_n} b(n_0, n)$ as well as the maximal FDR under Dirac-uniform configurations, namely $\text{FDR}_{n_0^{**},n}(\varphi_{(n)}) = \max_{n_0\in\mathbb{N}_n} \text{FDR}_{n_0,n}(\varphi_{(n)})$, together with the corresponding values for n_0^* and n_0^{**}.

n	n_0^*	$b(n_0^*, n)$	n_0^{**}	$\mathrm{FDR}_{n_0^{**},n}(\varphi_{(n)})$	n	n_0^*	$b(n_0^*, n)$	n_0^{**}	$\mathrm{FDR}_{n_0^{**},n}(\varphi_{(n)})$
5	2	40.0%	1	20.0%	70	8	7.31%	8	5.92%
10	2	20.0%	1	10.49%	100	10	6.70%	12	5.72%
15	3	15.04%	2	8.72%	150	16	6.18%	18	5.50%
20	3	12.42%	3	7.77%	200	19	5.97%	23	5.41%
25	4	10.97%	3	7.20%	250	23	5.81%	23	5.36%
30	4	10.02%	4	6.90%	300	28	5.70%	31	5.33%
35	5	9.31%	5	6.62%	400	34	5.58%	39	5.26%
40	5	8.84%	5	6.47%	500	39	5.47%	48	5.25%
50	6	8.13%	6	6.21%	750	61	5.35%	63	5.16%
60	7	7.65%	7	6.04%	1000	77	5.28%	81	5.13%

Table A.1: Results for the SUD-procedure, cf. Example 3.5

It turns out that for $n \geq 500$, the FDR level violation is lower than 0.5% for this procedure.

A.2.2 SU-procedure based on $f^{(1)}_{0.05,\kappa_1}$, Example 3.7

In this subsection, we investigate the adjusted AORC $f^{(1)}_{0.05,\kappa_1}$ with κ_1 chosen such that $h_1(1/2) = 1$ (no hypothesis with p-value larger than 0.5 is rejected) and implement this curve into an SU-test. The usage of an SU-test has the advantage that Dirac-uniform configurations are known to be least favorable. Therefore, it suffices to calculate $\mathrm{FDR}_{n_0^*,n}(\varphi_{(n)}) = \max_{n_0 \in \mathbb{N}_n} \mathrm{FDR}_{n_0,n}(\varphi_{(n)})$ in order to achieve an upper bound for the FDR of such a procedure. All results were obtained by numerical evaluation of the pmf. formula given in Corollary 3.23.

n	n_0^*	$\mathrm{FDR}_{n_0^*,n}(\varphi_{(n)})$	n	n_0^*	$\mathrm{FDR}_{n_0^*,n}(\varphi_{(n)})$
10	3	9.66%	150	27	5.34%
20	5	7.38%	200	36	5.26%
30	7	6.60%	250	44	5.21%
40	9	6.20%	300	53	5.18%
50	10	5.98%	400	70	5.13%
60	12	5.83%	500	86	5.11%
70	14	5.71%	750	128	5.07%
100	19	5.51%	1000	170	5.06%

Table A.2: Results for the SU-procedure based on $f^{(1)}_{0.05,\kappa_1}$, cf. Example 3.7

If we accept a FDR level violation of 0.5% under the quite unrealistic worst case scenario of a

Dirac-uniform configuration, we are already "on the safe side" for $n > 100$.

A.2.3 SU-procedure based on $f^{(2)}_{0.05,\kappa_2}$, Example 3.7

In analogy to Subsection A.2.2, we now focus on the second possible AORC adjustment discussed in Example 3.7, namely $f^{(2)}_{0.05,\kappa_2}$. Again, κ_2 is chosen such that $h_2(1/2) = 1$ and we use the resulting critical values in an SU-test.
Table A.3 again displays the maximum FDR values, i.e., $\text{FDR}_{n_0^*,n}(\varphi_{(n)}) = \max\limits_{n_0 \in \mathbb{N}_n} \text{FDR}_{n_0,n}(\varphi_{(n)})$ for different n making use of the result in Corollary 3.23.

n	n_0^*	$\text{FDR}_{n_0^*,n}(\varphi_{(n)})$	n	n_0^*	$\text{FDR}_{n_0^*,n}(\varphi_{(n)})$
10	3	9.71%	150	22	5.58%
20	5	8.09%	200	28	5.45%
30	6	7.02%	250	34	5.38%
40	8	6.71%	300	40	5.33%
50	9	6.38%	400	52	5.26%
60	11	6.21%	500	64	5.21%
70	12	6.06%	750	92	5.15%
100	16	5.80%	1000	120	5.12%

Table A.3: Results for the SU-procedure based on $f^{(2)}_{0.05,\kappa_2}$, cf. Example 3.7

As expected ($\kappa_2 > \kappa_1$ so that more critical values originate from the original AORC), this procedure shows a more liberal behavior with respect to FDR control and the number n of tests has to be approximately 200 in order to guarantee an FDR less than 0.055 under Dirac-uniform configurations.

A.2.4 SU-procedure with truncated curve, Example 3.8

Finally, we use the truncated version of f_α, discussed in Example 3.8, with $\kappa = 1/2$ in an SU test and choose the FDR level as $\alpha = 5\%$.
As before, the following table displays $\text{FDR}_{n_0^*,n}(\varphi_{(n)}) = \max\limits_{n_0 \in \mathbb{N}_n} \text{FDR}_{n_0,n}(\varphi_{(n)})$ together with the value n_0^* leading to this maximum.

n	n_0^*	$\text{FDR}_{n_0^*,n}(\varphi_{(n)})$	n	n_0^*	$\text{FDR}_{n_0^*,n}(\varphi_{(n)})$
10	3	9.71%	150	22	5.63%
20	5	8.25%	200	28	5.49%
30	6	7.12%	250	33	5.42%
40	8	6.81%	300	39	5.36%
50	9	6.46%	400	50	5.28%
60	10	6.30%	500	62	5.23%
70	12	6.13%	750	89	5.17%
100	16	5.86%	1000	116	5.13%

Table A.4: Results for the truncated SU-procedure with $\kappa = 1/2$, cf. Example 3.8

All values are very close to the corresponding values in Table A.3. This is due to the fact that κ_2 is very close to $1/2$ such that the truncated curve and $f^{(2)}_{0.05,\kappa_2}$ do only differ on the small interval $(\kappa_2, 1/2)$ as depicted in Figure 3.4.

Appendix B

Concepts of positive dependency

Proofs for FDR control often rely on assumptions about the correlation structure of the test statistics. Some special structures have been systematically covered in the past and especially concepts of *positive dependency* of some kind have useful properties. This section presents some of them. The following definition has preparing character.

Definition B.1 (Increasing / decreasing set)
Let $(\Omega, \leq)$ *be an ordered set and assume that* D, C *are subsets of* Ω.
The set D *is called* ***increasing*** $:\Leftrightarrow \forall y \in \Omega : [x \in D \wedge y \geq x \Rightarrow y \in D]$.
Analogously, the set C *is called* ***decreasing*** $:\Leftrightarrow \forall y \in \Omega : [x \in C \wedge y \leq x \Rightarrow y \in C]$.

Examples for increasing sets according to this definition as subsets of the set of real numbers $\mathbb{R}$ are rays of the form $D = [u, \infty[$ or $D =]u, \infty[$, respectively or correspondent constructs in higher dimensions as subsets of an $\mathbb{R}^n$ for $n > 1$ if we define the underlying order relations componentwise. Decreasing sets are therefore $C =]-\infty, o]$ or $C =]-\infty, o[$ in $\mathbb{R}$ and their higher-dimensional analogues, for example.

Remark B.2
If D is a closed, increasing subset of the target space of a real valued random variable $X : \Omega \to \mathbb{R}$, so it obviously exists a minimal subset $b(D)$ of D (the bound of D), so that for all $\omega \in \Omega$ with $X(\omega) \in D$ it exists an $x \in b(D)$ with $X(\omega) \geq x$. In the same way we can express the event that X maps to a decreasing subset C of its target space: $\{\omega : X(\omega) \in C\} = \{\omega : \exists\, x \in b(C) : X(\omega) \leq x\}$. In the special case of ray- or rectangle-shaped sets mentioned above, $b(D)$ or $b(C)$, respectively, only consist of one single point (a vertex). Inversely formulated, this consideration yields that every event of the form $\{X \geq u\}$ can be interpreted as an entry of X into a ray-shaped, increasing set bounded by u and an analogue event $\{X \leq o\}$ gets the meaning that X maps into a ray-shaped decreasing set bounded by o.

After these preparing considerations, it is possible to put the concepts of positive dependency

developed by Lehmann in 1966 (see [171]), T. K. Sarkar in 1969 ([244]) and Benjamini and Yekutieli in 2001 ([17]) into a common context. First, we recall Lehmann's definition of "positive regression dependency" in the bivariate case:

Definition B.3 (Positive regression dependency in the bivariate case, Lehmann (1966))
Let X and Y be real-valued random variables.
*Y is called **positively regression dependent on** X, if*

$$\mathbb{P}(Y \leq y \mid X = x) \text{ is non-increasing in } x.$$

This concept can be generalized to the multivariate case $n \geq 2$ in the following way:

Definition B.4 (Positive regression dependency in the multivariate case (PRD))
Let X be a random vector with $n \geq 2$ components $X_i : \Omega_i \to \mathbb{R}$.
*The multivariate distribution of X is called **positive regression dependent (PRD)**, if for all indices $i = 1, \ldots, n$ and for every increasing set $D \subset im(X)$*

$$\mathbb{P}(X \in D \mid X_i = x) \text{ is non-decreasing in } x. \tag{B.1}$$

Remark B.5
Since the complement $\complement D$ of an increasing set D is a decreasing set, (B.1) is equivalent to the postulation that the conditional probability $\mathbb{P}(X \in \complement D \mid X_i = x)$ has to be non-increasing in x. Thereby, Lehmann's Definition B.3 (together with the preparing remarks dealing with decreasing sets) is contained in Definition B.4 by setting $n = 2$. However, Definition B.4 is a little more general, because (in contrast to Lehmann) it makes no further assumptions about the explicit shape of the increasing (or decreasing, respectively) sets into which X maps.

Remark B.6
In [244], Theorem 2.3, a similar generalization of Lehmann's definition is given in the context of reliability analysis. It is shown that the property

$$\mathbb{P}(X_j > x_j, j = 1, \ldots, i-1 \mid X_i = u_i) \text{ is non-decreasing in } u_i, i = 2, \ldots, n \tag{B.2}$$

implies

$$\mathbb{P}\left(\bigcap_{i=1}^{n}\{X_i > x_i\}\right) \geq \prod_{i=1}^{n}\mathbb{P}(X_i > x_i),$$

which is a useful relation in order to compute lower bounds for the reliability of e. g. technical systems. It is clear that (B.2) is somewhat weaker than (B.1).

A further attenuation of the PRD property of a multivariate distribution is presented in the article [17] by Benjamini and Yekutieli from 2001. The authors introduce the concept of "positive regression dependency on subsets" (PRDS) as follows.

Definition B.7 (Positive regression dependency on subsets (PRDS), see [17])
Let X be a random vector with $n \geq 2$ components $X_i : \Omega_i \to \mathbb{R}$.
*The multivariate distribution of X is called **positive regression dependent on a subset** I_0 of the set of indices $\mathbb{N}_n = \{1, \ldots, n\}$, if for every increasing set $D \subset im(X)$ and for every index $i \in I_0$*

$$\mathbb{P}(X \in D | X_i = x) \text{ is non-decreasing in } x. \tag{B.3}$$

Therefore, condition (B.1) does not need to hold for all indices $i \in \mathbb{N}_n$, but only for those in the subset I_0 of indices. The authors point out that the PRDS condition is an appropriate technical tool to prove FDR-control of various stepwise test procedures if the underlying test statistics are dependent. For this purpose, the subset I_0 typically consists of the indices of test statistics corresponding to the true null hypotheses.
However, the verification of the PRDS property for a special distribution occurring in practice may lead to technical difficulties due to the structure of condition (B.3) which is defined via conditional probabilities. The following condition of "multivariate total positivity of order 2" (MTP_2) is easier to manage.

Definition B.8 (Multivariate total positivity of order 2 (MTP_2))
For $n \geq 2$, let X be a random vector with components $X_i : \Omega_i \to \mathcal{X}_i \subseteq \mathbb{R}$, $i = 1, \ldots, n$, and let f denote the joint n-dimensional probability density function of the variables $X_1, \ldots, X_n$ with respect to some product measure $\bigotimes_{i=1}^n \mu_i$.
*The multivariate distribution of X is called **multivariate totally positive of order 2 (MTP_2)**, if for all $x, y \in im(X)$*

$$f(x) \cdot f(y) \quad \leq \quad f(\min(x, y)) \cdot f(\max(x, y)), \tag{B.4}$$

where the minimum or maximum, respectively, is being taken component-wise.

This condition of multivariate total positivity of order 2 is the strictest of all concepts of dependency introduced here, because it holds:

Theorem B.9 (Implications of concepts of dependency)
For the properties of multivariate distributions introduced in Definitions B.4, B.7 and B.8, it holds:

(i) Multivariate total positivity of order 2 implies positive regression dependency.

(ii) Positive regression dependency implies positive regression dependency on subsets.

Proof: Part (ii) is obvious.
In order to prove part (i), let the increasing set $D \subset \mathcal{X} = \text{im}(X)$ be arbitrary chosen but fixed. Without loss of generality, we set $i = 1$ in (B.1) and write $X = (X_1, X^{(1)})$, i.e., we set $X^{(1)} = (X_2, \ldots, X_n)$ and, analogously, $\mathcal{X} = \mathcal{X}_1 \otimes \mathcal{X}^{(1)}$. The joint pdf. $f(X_1, X^{(1)})$ will be regarded as

a $\mu_1 \otimes \mu^{(1)}$-density of $(X_1, X^{(1)})$. Finally, let $u_1, u_2 \in \mathcal{X}_1$ with $u_1 < u_2$ be arbitrary chosen but fixed and describe the u_1- and u_2-cuts of D and $\mathcal{X}$, respectively, by

$$\begin{aligned} D_{u_j} &= \{x^{(1)} \in \mathcal{X}^{(1)} : (u_j, x^{(1)}) \in D\},\ j = 1, 2, \\ \mathcal{X}_{u_j} &= \{x^{(1)} \in \mathcal{X}^{(1)} : (u_j, x^{(1)}) \in \mathcal{X}\},\ j = 1, 2. \end{aligned}$$

With these definitions, we now have to show that

$$\mathbb{P}(X \in D \mid X_1 = u_1) \leq \mathbb{P}(X \in D \mid X_1 = u_2). \tag{B.5}$$

To this end, we make use of the fact that for all $x \in D_{u_1}$ and for all $y \in \mathcal{X}_{u_2} \setminus D_{u_2}$ we have the relationship $(u_2, x)^T \geq (u_1, y)^T$, which holds true due to the property that D is increasing. Therefore, the MTP$_2$ condition for the joint pdf. f yields $f(u_1, x)f(u_2, y) \leq f(u_2, x)f(u_1, y)$. Integrating with respect to $\mu^{(1)}$, we obtain

$$\begin{aligned} &\int_{D_{u_1}} f(u_1, x)d\mu^{(1)}(x) \cdot \int_{\mathcal{X}_{u_2} \setminus D_{u_2}} f(u_2, y)d\mu^{(1)}(y) \\ \leq &\int_{D_{u_1}} f(u_2, x)d\mu^{(1)}(x) \cdot \int_{\mathcal{X}_{u_2} \setminus D_{u_2}} f(u_1, y)d\mu^{(1)}(y). \end{aligned} \tag{B.6}$$

Due to the subset relation $D_{u_1} \subseteq D_{u_2}$, we have that the right-hand side of (B.6) is lower than or equal to

$$\int_{D_{u_2}} f(u_2, x)d\mu^{(1)}(x) \cdot \int_{\mathcal{X}_{u_1} \setminus D_{u_1}} f(u_1, y)d\mu^{(1)}(y).$$

If we now divide the resulting inequality chain by

$$\int_{\mathcal{X}^{(1)}} f(u_1, \xi)d\mu^{(1)}(\xi) \cdot \int_{\mathcal{X}^{(1)}} f(u_2, \eta)d\mu^{(1)}(\eta),$$

we arrive at

$$\begin{aligned} &\int_{D_{u_1}} \frac{f(u_1, x)}{\int_{\mathcal{X}^{(1)}} f(u_1, \xi)d\mu^{(1)}(\xi)}d\mu^{(1)}(x) \cdot \int_{\mathcal{X}_{u_2} \setminus D_{u_2}} \frac{f(u_2, y)}{\int_{\mathcal{X}^{(1)}} f(u_2, \eta)d\mu^{(1)}(\eta)}d\mu^{(1)}(y) \\ \leq &\int_{D_{u_2}} \frac{f(u_2, x)}{\int_{\mathcal{X}^{(1)}} f(u_2, \eta)d\mu^{(1)}(\eta)}d\mu^{(1)}(x) \cdot \int_{\mathcal{X}_{u_1} \setminus D_{u_1}} \frac{f(u_1, y)}{\int_{\mathcal{X}^{(1)}} f(u_1, \xi)d\mu^{(1)}(\xi)}d\mu^{(1)}(y). \end{aligned} \tag{B.7}$$

Inequality (B.7) can be re-written in terms of conditional expectations yielding

$$\begin{aligned} &\mathbb{E}\left(\mathbf{1}_{D_{u_1}} \circ X^{(1)} | X_1 = u_1\right) \mathbb{E}\left(\mathbf{1}_{\mathcal{X}_{u_2} \setminus D_{u_2}} \circ X^{(1)} | X_1 = u_2\right) \\ \leq &\mathbb{E}\left(\mathbf{1}_{D_{u_2}} \circ X^{(1)} | X_1 = u_2\right) \mathbb{E}\left(\mathbf{1}_{\mathcal{X}_{u_1} \setminus D_{u_1}} \circ X^{(1)} | X_1 = u_1\right). \end{aligned} \tag{B.8}$$

Because of

$$\begin{aligned} \mathbb{E}\left(\mathbf{1}_{\mathcal{X}_{u_1} \setminus D_{u_1}} \circ X^{(1)} | X_1 = u_1\right) &= \mathbb{E}\left(\mathbf{1}_{\mathcal{X}_{u_1}} \circ X^{(1)} | X_1 = u_1\right) - \mathbb{E}\left(\mathbf{1}_{D_{u_1}} \circ X^{(1)} | X_1 = u_1\right), \\ \mathbb{E}\left(\mathbf{1}_{\mathcal{X}_{u_2} \setminus D_{u_2}} \circ X^{(1)} | X_1 = u_2\right) &= \mathbb{E}\left(\mathbf{1}_{\mathcal{X}_{u_2}} \circ X^{(1)} | X_1 = u_2\right) - \mathbb{E}\left(\mathbf{1}_{D_{u_2}} \circ X^{(1)} | X_1 = u_2\right), \end{aligned}$$

relation (B.8) is equivalent to

$$\mathbb{E}\left(\mathbf{1}_{D_{u_1}} \circ X^{(1)} | X_1 = u_1\right) \leq \mathbb{E}\left(\mathbf{1}_{D_{u_2}} \circ X^{(1)} | X_1 = u_2\right). \tag{B.9}$$

Rewriting (B.9) in terms of conditional probabilities yields

$$\mathbb{P}^{X^{(1)}|X_1=u_1}(D_{u_1}) \leq \mathbb{P}^{X^{(1)}|X_1=u_2}(D_{u_2}),$$

which is equivalent to (B.5), hence, the PRD property of the distribution of X. ■

Remark B.10

(i) If in Definition B.8 especially $n = 2$, i.e., X underlies a bivariate totally positive distribution of order 2, this property is denoted by TP_2. In this special case, the defining condition for the bivariate probability density function of X simplifies to

$$f(x_1, y_2)f(x_2, y_1) \leq f(x_1, y_1)f(x_2, y_2) \tag{B.10}$$

for all $(x_1, y_1)^T \leq (x_2, y_2)^T \in \text{im}(X) \subseteq \mathbb{R}^2$.

(ii) The TP_2 property may also be employed for the characterization of one-parametric families of probability density functions $\{f_\vartheta, \vartheta \in \Theta\}$, interpreting the real-valued parameter ϑ as second argument and re-writing the pdf.'s as $f_\vartheta(x) = f(x, \vartheta)$. If then the TP_2 property (B.10) holds for $f(x, \vartheta)$, this is equivalent to the fact that $f_\vartheta(x)$ has an isotone likelihood ratio with respect to $\vartheta \in \Theta$ in the identity. Furthermore, the corresponding family $\{\mathbb{P}_\vartheta, \vartheta \in \Theta\}$ of probability measures underlies a stochastic ordering, because for the cdf.'s it holds

$$F_{\vartheta_1}(x) \geq F_{\vartheta_2}(x) \ \ \forall x \in \mathbb{R}, \ \text{ if } \vartheta_1 < \vartheta_2.$$

Example B.11 (TP_2 property on subintervals)
Let $\mathcal{T} = \{\mathbb{P}_\nu \mid \nu \in (0, \infty]\}$ be the family of Student's t-distributions with ν degrees of freedom for $0 < \nu < \infty$ (for short: t_ν-distributions) together with its limiting distribution $\mathbb{P}_\infty$, namely the standard normal distribution. Then the corresponding pdf.'s have the form

$$f_\nu(x) = \begin{cases} \frac{\Gamma(\nu/2+1/2)}{\Gamma(\nu/2)} \frac{1}{\sqrt{\nu\,\pi}} \left(1 + \frac{x^2}{\nu}\right)^{-\nu/2-1/2} & \text{for } 0 < \nu < \infty,\ x \in \mathbb{R} \\ \frac{1}{\sqrt{2\pi}} \exp\left(-x^2/2\right) & \text{for } \nu = \infty,\ x \in \mathbb{R}. \end{cases}$$

$\mathcal{T}$ is no TP_2-family, but it has a monotone likelihood ratio for $x \in I_1 = (-\infty, -1)$ as well as for $x \in I_2 = (0, 1)$. The case $0 < \nu_1 < \nu_2 < \infty$ is treated and proven in [270]. In the case $\nu_2 = \infty$, i.e., the combination of the standard normal distribution with a t_ν-distribution with $0 < \nu < \infty$, we obtain the corresponding likelihood ratio function

$$q(x) := \frac{f_\infty(x)}{f_\nu(x)} = \sqrt{\frac{\nu}{2}} \frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2} + \frac{1}{2}\right)} \exp\left(-\frac{x^2}{2}\right) \left(\frac{\nu + x^2}{\nu}\right)^{\nu/2+1/2}$$

and because of

$$\frac{d}{dx}\, q(x) = \sqrt{\frac{\nu}{2}} \frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}+\frac{1}{2}\right)}\ \exp\left(-\frac{x^2}{2}\right) \frac{(\nu+x^2)^{\nu/2-1/2}}{\nu^{(\nu/2+1/2)}}\ x\,(1-x^2),$$

$q(x)$ is isotone in $x \in I_1$ and $x \in I_2$. The TP_2 property is therefore valid for $\mathcal{T}$ on the subintervals I_1 and I_2.

Some deeper investigations concerning structural properties of the t_ν-distributions can be found in [93]. They are useful for some of the derivations in Section 2.4.

Example B.12 (MTP_2 for binary variables)

In [9], the MTP_2 property is investigated for binary variables. The authors establish a connection between MTP_2 and the odds ratio as follows.

Let $X = (X_1, \ldots, X_n)$ be a random vector with $n > 2$ components taking only binary values, i.e., $X : \Omega \to \{0,1\}^n$. Then X is MTP_2 if and only if for all $\{j_1, j_2\} \subset \{1, \ldots, n\}$, it holds

$$\ln\left(\frac{\mathbb{P}(X_{j_1}=0, X_{j_2}=0|\{X_k\}_{k\in K}) \cdot \mathbb{P}(X_{j_1}=1, X_{j_2}=1|\{X_k\}_{k\in K})}{\mathbb{P}(X_{j_1}=0, X_{j_2}=1|\{X_k\}_{k\in K}) \cdot \mathbb{P}(X_{j_1}=1, X_{j_2}=0|\{X_k\}_{k\in K})}\right) \geq 0, \qquad \text{(B.11)}$$

where $K \equiv K(j_1, j_2) = \{1, \ldots, n\} \setminus \{j_1, j_2\}$. It should be mentioned that this is a special application of a more general result given in [181]. The meaning of equation (B.11) can be interpreted in an epidemiological setup. If for example X_{j_1} denotes the status of a certain disease, X_{j_2} an exposition status and $\{X_k\}_{k\in K}$ the states of a set of binary covariates, then (B.11) means that the logarithmic odds ratio for the exposure-disease relationship is always non-negative, i.e., we have no protective factors or all covariates are coded in that way that their effect on the disease is harmful, respectively.

Remark B.13

The concept of total positivity is a widely studied, important issue with various statistical applications. For a deeper study, the works of S. Karlin ([147]), Karlin and Rinott ([148], [149]), and Cohen and Sackrowitz ([45]) are recommendable. One special application field consists in reliability and life testing (cf. [8] and [244]).

The dissertation of Astrid Heinicke ([121]) gives a good overview of various concepts of dependency.

List of Tables

List of Figures

Bibliography

(The following bibliographic list attempts to provide a comprehensive overview of mainly theoretical articles on the FDR. Not all references given here are cited in the main text. Cited references are marked with an asterix.)

[1] ABRAMOVICH, F. AND ANGELINI, C. (2006). Bayesian Maximum a posteriori Multiple Testing Procedure. *Sankhyā* **68**, Part 3, 436–460.

[2] ABRAMOVICH, F. AND BENJAMINI, Y. (1996). Adaptive thresholding of wavelet coefficients. *Comput. Stat. Data Anal.* **22**, 4, 351–361.

[3] ABRAMOVICH, F., BENJAMINI, Y., DONOHO, D. L., AND JOHNSTONE, I. M. (2006). Adapting to Unknown Sparsity by controlling the False Discovery Rate. *Ann. Stat.* **34**, 2, 584–653.

[4] ABRAMOVICH, F., GRINSHTEIN, V., AND PENSKY, M. (2007). On optimality of Bayesian testimation in the normal means problem. *Ann. Stat., to appear*.

[5] ABRAMOVICH, F., SAPATINAS, T., AND SILVERMAN, B. W. (1998). Wavelet thresholding via a Bayesian appoach. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **60**, 4, 725–749.

[6] AUBERT, J., BAR-HEN, A., DAUDIN, J. J., AND ROBIN, S. (2004). Determination of the differentially expressed genes in microarray experiments using local FDR. *BMC Bioinformatics 5*, 125.

[7] BAR-HEN, A. AND ROBIN, S. (2003). An iterative procedure for differential analysis of gene expression. *C. R., Math., Acad. Sci. Paris* **337**, 5, 343–346.

[8] *BARLOW, R. E. AND PROSCHAN, F. (1975). *Statistical theory of reliability and life testing.* International Series in Decision Processes. Holt, Rinehart and Winston, Inc., New York.

[9] *BARTOLUCCI, F. AND FORCINA, A. (2000). A likelihood ratio test for MTP_2 within binary variables. *Ann. Stat.* **28**, 4, 1206–1218.

[10] BASFORD, K. E. AND TUKEY, J. W. (1997). Graphical profiles as an aid to understanding plant breeding experiments. *J. Stat. Plann. Inference* **57**, 1, 93–107.

[11] BAY, S. D. AND PAZZANI, M. J. (2001). Detecting group differences: Mining contrast sets. *Data Mining and Knowledge Discovery* **5**, 3, 213–246.

[12] BENJAMINI, Y., DRAI, D., ELMER, G., KAFKAFI, N., AND GOLANI, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* **125**, 1-2, 279–284.

[13] *BENJAMINI, Y. AND HOCHBERG, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **57**, 1, 289–300.

[14] BENJAMINI, Y. AND HOCHBERG, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J. Edu. and Behav. Stat. 25*, 60–83.

[15] *BENJAMINI, Y., KRIEGER, A. M., AND YEKUTIELI, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* **93**, 3, 491–507.

[16] BENJAMINI, Y. AND LIU, W. (1999). A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence. *J. Stat. Plann. Inference* **82**, 1-2, 163–170.

[17] *BENJAMINI, Y. AND YEKUTIELI, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**, 4, 1165–1188.

[18] BENJAMINI, Y. AND YEKUTIELI, D. (2005a). False Discovery Rate-Adjusted Multiple Confidence Intervals for Selected Parameters. *J. Am. Stat. Assoc.* **100**, 469, 71–93.

[19] BENJAMINI, Y. AND YEKUTIELI, D. (2005b). Quantitative Trait Loci Analysis Using the False Discovery Rate. *Genetics 171*, 783–790.

[20] BENNEWITZ, J., REINSCH, N., GUIARD, V., FRITZ, S., THOMSEN, H., LOOFT, C., KUHN, C., SCHWERIN, M., WEIMANN, C., ERHARDT, G., REINHARDT, F., REENTS, R., BOICHARD, D., AND KALM, E. (2004). Multiple Quantitative Trait Loci Mapping With Cofactors and Application of Alternative Variants of the False Discovery Rate in an Enlarged Granddaughter Design. *Genetics* **168**, 2, 1019–1027.

[21] BERNARDO, R. (2004). What proportion of declared QTL in plants are false? *Theor. Appl. Genet.* **109**, 2, 419–424.

[22] BICKEL, D. R. (2004). Degrees of differential gene expression: detecting biologically significant expression differences and estimating their magnitudes. *Bioinformatics* **20**, 5, 682–688.

[23] *BICKEL, D. R. (2004). Error-rate and decision-theoretic methods of multiple testing: which genes have high objective probabilities of differential expression? *Stat. Appl. Genet. Mol. Biol.* **3**, 1.

[24] BICKEL, D. R. (2005). Probabilities of spurious connections in gene networks: application to expression time series. *Bioinformatics* **21**, 7, 1121–1128.

[25] BLACK, M. (2004). A note on the adaptive control of false discovery rates. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **66**, 2, 297–304.

[26] BLEAKLEY, K., BIAU, G., AND VERT, J.-P. (2007). Supervised reconstruction of biological networks with local models. *Bioinformatics* **23**, 13, i57–65.

[27] BLÜTHGEN, N., BRAND, K., CAJAVEC, B., SWAT, M., HERZEL, H., AND BEULE, D. (2005). Biological profiling of gene groups utilizing Gene Ontology. *Genome Inform.* **16**, 1, 106–115.

[28] BOCHKINA, N. AND RICHARDSON, S. (2007). Tail posterior probability for inference in pairwise and multiclass gene expression data. *Biometrics, to appear*.

[29] BOLTON, R. J., HAND, D. J., AND ADAMS, N. M. (2002). Determining hit rate in pattern search. In *Hand, David J. (ed.) et al., Pattern detection and discovery. ESF exploratory workshop, London, UK, September 16-19, 2002. Proceedings.* Berlin: Springer. Lect. Notes Comput. Sci. 2447, 36-48 .

[30] BRATCHER, T. AND HAMILTON, C. (2005). A Bayesian multiple comparison procedure for ranking the means of normally distributed data. *J. Stat. Plann. Inference* **133**, 1, 23–32.

[31] BRETZ, F., LANDGREBE, J., AND BRUNNER, E. (2005). Multiplicity issues in microarray experiments. *Methods Inf. Med.* **44**, 3, 431–437.

[32] BROBERG, P. (2005). A comparative review of estimates of the proportion unchanged genes and the false discovery rate. *BMC Bioinformatics 6*, 199.

[33] BROET, P., LEWIN, A., RICHARDSON, S., DALMASSO, C., AND MAGDELENAT, H. (2004). A mixture model-based strategy for selecting sets of genes in multiclass response microarray experiments. *Bioinformatics* **20**, 16, 2562–2571.

[34] BROET, P. AND RICHARDSON, S. (2006). Detection of gene copy number changes in CGH microarrays using a spatially correlated mixture model. *Bioinformatics* **22**, 8, 911–918.

[35] BUNEA, F., WEGKAMP, M. H., AND AUGUSTE, A. (2006). Consistent variable selection in high dimensional regression via multiple testing. *J. Stat. Plann. Inference* **136**, 12, 4349–4364.

[36] CAI, G. AND SARKAR, S. K. (2006). Modified Simes' critical values under positive dependence. *J. Stat. Plann. Inference* **136**, 12, 4129–4146.

[37] CARLBORG, O., DE KONING, D. J., MANLY, K. F., CHESLER, E., WILLIAMS, R. W., AND HALEY, C. S. (2005). Methodological aspects of the genetic dissection of gene expression. *Bioinformatics* **21**, 10, 2383–2393.

[38] CHAN, I., HILLMAN, D., AND LOUIS, T. (1998). Treatment comparisons with screenable endpoints. *Comput. Stat. Data Anal.* **27**, 4, 401–419.

[39] CHEN, J. AND SARKAR, S. K. (2006). A Bayesian determination of threshold for identifying differentially expressed genes in microarray experiments. *Stat. Med.* **25**, 18, 3174–3189.

[40] CHEN, L., LIU, N., WANG, S., OH, C., CARRIERO, N. J., AND ZHAO, H. (2005). Whole-genome association studies on alcoholism comparing different phenotypes using single-nucleotide polymorphisms and microsatellites. *BMC Genet. 6 Suppl. 1*, S130.

[41] CHENG, C. (2006). An adaptive significance threshold criterion for massive multiple hypotheses testing. *IMS Lecture Notes–Monograph Series 49*, 51–76.

[42] CHENG, C., POUNDS, S. B., BOYETT, J. M., PEI, D., KUO, M.-L., AND ROUSSEL, M. F. (2004). Statistical significance threshold criteria for analysis of microarray gene expression data. *Stat. Appl. Genet. Mol. Biol.* **3**, 36.

[43] *CHI, Z. (2007). On the performance of fdr control: constraints and a partial solution. *Ann. Stat.,* **35**, 4, 1409–1431.

[44] CHI, Z. (2007). Sample size and positive false discovery rate control for multiple testing. *Electronic Journal of Statistics 1*, 77–118.

[45] *COHEN, A. AND SACKROWITZ, H. B. (1993). Some remarks on a notion of positive dependence, association, and unbiased testing. *In: Stochastic Inequalities, IMS Lecture Notes - Monograph Series 22*, 33–37.

[46] COHEN, A. AND SACKROWITZ, H. B. (2005). Decision theory results for one-sided multiple comparison procedures. *Ann. Stat.* **33**, 1, 126–144.

[47] COHEN, A. AND SACKROWITZ, H. B. (2007). More on the inadmissibility of step-up. *J. Multivariate Anal.* **98**, 3, 481–492.

[48] COX, D. AND WONG, M. Y. (2004). A simple procedure for the selection of significant effects. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **66**, 2, 395–400.

[49] CUPPLES, L. A., BAILEY, J., CARTIER, K. C., FALK, C. T., LIU, K.-Y., YE, Y., YU, R., ZHANG, H., AND ZHAO, H. (2005). Data mining. *Genet. Epidemiol. 29 Suppl. 1*, 103–109.

[50] CURRAN-EVERETT, D. (2000). Multiple comparisons: philosophies and illustrations. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **279**, 1, 1–8.

[51] DALMASSO, C., BROET, P., AND MOREAU, T. (2004). Multiple comparison procedures: principles, limits. Applications to microarray phenotype-genotype analysis. *Rev. Epidemiol. Sante Publique* **52**, 6, 523–537.

[52] DALMASSO, C., BROET, P., AND MOREAU, T. (2005). A simple procedure for estimating the false discovery rate. *Bioinformatics* **21**, 5, 660–668.

[53] DASGUPTA, A. AND ZHANG, T. (2006). On the false discovery rates of a frequentist: Asymptotic expansions. *IMS Lecture Notes–Monograph Series 50*, 190–212.

[54] DATSON, N. A., MEIJER, L., STEENBERGEN, P. J., MORSINK, M. C., VAN DER LAAN, S., MEIJER, O. C., AND DE KLOET, E. R. (2004). Expression profiling in laser-microdissected hippocampal subregions in rat brain reveals large subregion-specific differences in expression. *Eur. J. Neurosci.* **20**, 10, 2541–2554.

[55] DATTA, S. AND DATTA, S. (2004). An Empirical Bayes Adjustment to Multiple p-values For the Detection of Differentially Expressed Genes in Microarray Experiments. *2nd Asia-Pacific Bioinformatics Conference (APBC2004), Dunedin, New Zealand. Conferences in Research and Practice in Information Technology 29*.

[56] DATTA, S. AND DATTA, S. (2005). Empirical Bayes screening of many p-values with applications to microarray studies. *Bioinformatics* **21**, 9, 1987–1994.

[57] DATTA, S., SATTEN, G. A., BENOS, D. J., XIA, J., HESLIN, M. J., AND DATTA, S. (2004). An empirical bayes adjustment to increase the sensitivity of detecting differentially expressed genes in microarray experiments. *Bioinformatics* **20**, 2, 235–242.

[58] DE LA FUENTE, A., BING, N., HOESCHELE, I., AND MENDES, P. (2004). Discovery of meaningful associations in genomic data using partial correlation coefficients. *Bioinformatics* **20**, 18, 3565–3574.

[59] DELONGCHAMP, R. R., BOWYER, J. F., CHEN, J. J., AND KODELL, R. L. (2004). Multiple-testing strategy for analyzing cDNA array data on gene expression. *Biometrics* **60**, 3, 774–782.

[60] DIETERICH, C., RAHMANN, S., AND VINGRON, M. (2004). Functional inference from non-random distributions of conserved predicted transcription factor binding sites. *Bioinformatics* **20**, Suppl.1, i109–115.

[61] DOBBIN, K. AND SIMON, R. (2005). Sample size determination in microarray experiments for class comparison and prognostic classification. *Biostatistics* **6**, 1, 27–38.

[62] DONOHO, D. AND JIN, J. (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Stat.* **32**, 3, 962–994.

[63] DONOHO, D. AND JIN, J. (2006). Asymptotic Minimaxity of False Discovery Rate Thresholding for Sparse Exponential Data. *Ann. Stat.,* **34**, 6, 2980–3018.

[64] DRIGALENKO, E. I. AND ELSTON, R. C. (1997). False discoveries in genome scanning. *Genet. Epidemiol.* **14**, 6, 779–784.

[65] DUDBRIDGE, F., GUSNANTO, A., AND KOELEMAN, B. P. C. (2006). Detecting multiple associations in genome-wide studies. *Hum. Genomics* **2**, 5, 310–317.

[66] DUDBRIDGE, F. AND KOELEMAN, B. P. (2003). Rank truncated product of P-values, with application to genomewide association scans. *Genet. Epidemiol.* **25**, 4, 360–366.

[67] *DUDOIT, S., SHAFFER, J. P., AND BOLDRICK, J. C. (2003). Multiple hypothesis testing in microarray experiments. *Stat. Sci.* **18**, 1, 71–103.

[68] *DUDOIT, S., YANG, Y. H., CALLOW, M. J., AND SPEED, T. P. (2002). Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Stat. Sin.* **12**, 1, 111–139.

[69] *EFRON, B. (2003). Robbins, empirical Bayes and microarrays. *Ann. Stat.* **31**, 2, 366–378.

[70] EFRON, B. (2004). Large-Scale Simultaneous Hypothesis Testing: The Choice of a Null Hypothesis. *J. Am. Stat. Assoc.* **99**, 465, 96–104.

[71] EFRON, B. (2006). Correlation and Large-Scale Simultaneous Significance Testing. *J. Am. Stat. Assoc.* **102**, 477, 93–103.

[72] *EFRON, B. (2007). Size, Power, and False Discovery Rates. *Ann. Stat.,* **35**, 4, 1351–1377.

[73] *EFRON, B. AND TIBSHIRANI, R. (2002). Empirical Bayes methods and false discovery rates for microarrays. *Genet. Epidemiol.* **23**, 1, 70–86.

[74] EFRON, B. AND TIBSHIRANI, R. (2007). On testing the significance of sets of genes. *Annals of Applied Statistics* **1**, 1.

[75] *EFRON, B., TIBSHIRANI, R., STOREY, J. D., AND TUSHER, V. (2001). Empirical Bayes analysis of a microarray experiment. *J. Am. Stat. Assoc.* **96**, 456, 1151–1160.

[76] EKLUND, G. (1961-1963). *Massignifikansproblemet.* Unpublished seminar papers, Uppsala University Institute of Statistics.

[77] EKLUND, G. AND SEEGER, P. (1965). Massignifikansanalys. *Statistisk Tidskrift Stockholm, third series 4*, 355–365.

[78] ERICKSON, S. AND SABATTI, C. (2005). Empirical Bayes estimation of a sparse vector of gene expression changes. *Stat. Appl. Genet. Mol. Biol.* **4**, 1.

[79] FADILI, M. J. AND BULLMORE, E. T. (2004). A comparative evaluation of wavelet-based methods for hypothesis testing of brain activation maps. *Neuroimage* **23**, 3, 1112–1128.

[80] FAN, J. AND LI, R. (2006). Statistical Challenges with High Dimensionality: Feature Selection in Knowledge Discovery. In *Sanz-Solé, Marta (ed.) et al., Proceedings of the international congress of mathematicians (ICM), Madrid, Spain, August 22–30, 2006. Volume III: Invited lectures. Zürich: European Mathematical Society (EMS). 595-622.*

[81] FAN, J., MUELLER, K.-R., AND SPOKOINY, V. (2005). New inference concepts for analysing complex data. *Mathematisches Forschungsinstitut Oberwolfach 52*.

[82] FARCOMENI, A. (2006). More powerful control of the false discovery rate under dependence. *Stat. Methods Applications 15*, 43–73.

[83] FERNANDO, R. L., NETTLETON, D., SOUTHEY, B. R., DEKKERS, J. C. M., ROTHSCHILD, M. F., AND SOLLER, M. (2004). Controlling the proportion of false positives in multiple dependent tests. *Genetics* **166**, 1, 611–619.

[84] FERREIRA, J. A. AND ZWINDERMAN, A. H. (2006). On the Benjamini-Hochberg method. *Ann. Stat.* **34**, 4, 1827–1849.

[85] FINNER, H. (2004). S35.2: Pre-ordered hypotheses and false discovery rate. *Biom. J.* **46**, S1, 74.

[86] *FINNER, H., DICKHAUS, T., AND ROTERS, M. (2007a). Dependency and false discovery rate: Asymptotics. *Ann. Stat.,* **35**, 4, 1432–1455.

[87] *FINNER, H., DICKHAUS, T., AND ROTERS, M. (2007b). On the false discovery rate and an asymptotically optimal rejection curve. *Ann. Stat., to appear*.

[88] *FINNER, H., DICKHAUS, T., AND ROTERS, M. (2008). Asymptotic Tail Properties of Student's t-Distribution. *Commun. Stat., Theory Methods* **37**, 2, to appear.

[89] *FINNER, H., HAYTER, A. J., AND ROTERS, M. (1993). *On the joint distribution function of order statistics with reference to step-up multiple test procedures.* Technical Report 93-19, 1-11, Universität Trier.

[90] FINNER, H. AND ROTERS, M. (1998). Asymptotic comparison of step-down and step-up multiple test procedures based on exchangeable test statistics. *Ann. Stat.* **26**, 2, 505–524.

[91] *FINNER, H. AND ROTERS, M. (2001). On the false discovery rate and expected type I errors. *Biom. J.* **43**, 8, 985–1005.

[92] *FINNER, H. AND ROTERS, M. (2002). Multiple hypotheses testing and expected number of type I errors. *Ann. Stat.* **30**, 1, 220–238.

[93] *FINNER, H., ROTERS, M., AND DICKHAUS, T. (2007). Characterizing Density Crossing Points. *Amer. Statist.* **61**, 1, 28–33.

[94] *FINNER, H. AND STRASSBURGER, K. (2002). The partitioning principle: a powerful tool in multiple decision theory. *Ann. Stat.* **30**, 4, 1194–1213.

[95] *FINNER, H., STRASSBURGER, K., HEID, I. M., HERDER, C., RATHMANN, W., GIANI, G., DICKHAUS, T., LICHTNER, P., MEITINGER, T., WICHMANN, H.-E., ILLIG, T., AND GIEGER, C. (2007). How to link call rate and Hardy-Weinberg equilibrium as measures of genome-wide SNP data quality - results from the KORA 500K project. *Submitted for publication*.

[96] FINOS, L. AND SALMASO, L. (2007). FDR- and FWE-controlling methods using data-driven weights. *J. Stat. Plann. Inference* **137**, 12, 3859–3870.

[97] FUTSCHIK, A. AND POSCH, M. (2005). On the optimum number of hypotheses to test when the number of observations is limited. *Stat. Sin.* **15**, 3, 841–855.

[98] GAO, X. (2006). Construction of null statistics in permutation based multiple testing for multi-factorial microarray experiments. *Bioinformatics* **22**, 12, 1486 – 1494.

[99] GARCIA, L. (2003). Controlling the false discovery rate in ecological research. *Trends Ecol. Evol.* **18**, 11, 553–554.

[100] *GE, Y., DUDOIT, S., AND SPEED, T. P. (2003). Resampling-based multiple testing for microarray data analysis (With comments). *Test* **12**, 1, 1–77.

[101] GENOVESE, C., LAZAR, N., AND NICHOLS, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* **15**, 4, 870–878.

[102] *GENOVESE, C. AND WASSERMAN, L. (2002). Operating characteristics and extensions of the false discovery rate procedure. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **64**, 3, 499–517.

[103] *GENOVESE, C. AND WASSERMAN, L. (2004). A stochastic process approach to false discovery control. *Ann. Stat.* **32**, 3, 1035–1061.

[104] *GENOVESE, C. AND WASSERMAN, L. (2006). Exceedance Control of the False Discovery Proportion. *J. Am. Stat. Assoc.* **101**, 476, 1408–1417.

[105] GENOVESE, C. R., ROEDER, K., AND WASSERMAN, L. (2006). False discovery control with p-value weighting. *Biometrika* **93**, 3, 509–524.

[106] GHOSH, D. (2006). Shrunken p-values for assessing differential expression with applications to genomic data analysis. *Biometrics* **62**, 4, 1099–1106.

[107] GHOSH, D., CHEN, W., AND RAGHUNATHAN, T. (2006). The false discovery rate: a variable selection perspective. *J. Stat. Plann. Inference* **136**, 8, 2668–2684.

[108] GHOSH, D. AND CHINNAIYAN, A. M. (2007). Empirical Bayes Identication of Tumor Progression Genes from Microarray Data. *Biom. J.* **49**, 1, 68–77.

[109] GIBBONS, R. D., BHAUMIK, D. K., COX, D. R., GRAYSON, D. R., DAVIS, J. M., AND SHARMA, R. P. (2005). Sequential prediction bounds for identifying differentially expressed genes in replicated microarray experiments. *J. Stat. Plann. Inference* **129**, 1-2, 19–37.

[110] GILBERT, P. B. (2005). A modified false discovery rate multiple-comparisons procedure for discrete data, applied to human immunodeficiency virus genetics. *Appl. Statist.* **54**, 1, 143–158.

[111] GLYNN, E. F., CHEN, J., AND MUSHEGIAN, A. R. (2006). Detecting periodic patterns in unevenly spaced gene expression time series using Lomb-Scargle periodograms. *Bioinformatics* **22**, 3, 310–316.

[112] GOLL, A. AND BAUER, P. (2007). Two-stage designs applying methods differing in costs. *Bioinformatics* **23**, 12, 1519–1526.

[113] GORDON, A., GLAZKO, G., QIU, X., AND YAKOVLEV, A. (2007). Control of the Mean Number of False Discoveries, Bonferroni, and Stability of Multiple Testing. *Annals of Applied Statistics* **1**, 1.

[114] GORDON, A. Y. (2007). Explicit formulas for generalized family-wise error rates and unimprovable step-down multiple testing procedures. *J. Stat. Plann. Inference* **137**, 11, 3497–3512.

[115] GOULD, A. L. (2006). Accounting for Multiplicity in the Evaluation of "Signals" Obtained by Data Mining from Spontaneous Report Adverse Event Databases. *Biom. J.* **48**, 6, 1–15.

[116] GRANT, G. R., LIU, J., AND STOECKERT, C. J. J. (2005). A practical false discovery rate approach to identifying patterns of differential expression in microarray data. *Bioinformatics* **21**, 11, 2684–2690.

[117] GUINDON, S., BLACK, M., AND RODRIGO, A. (2006). Control of the False Discovery Rate Applied to the Detection of Positively Selected Amino Acid Sites. *Mol. Biol. Evol.* **23**, 5, 919–926.

[118] GUO, X. AND PAN, W. (2005). Using weighted permutation scores to detect differential gene expression with microarray data. *J. Bioinform. Comput. Biol.* **3**, 4, 989–1006.

[119] GUO, X., QI, H., VERFAILLIE, C. M., AND PAN, W. (2003). Statistical significance analysis of longitudinal gene expression data. *Bioinformatics* **19**, 13, 1628–1635.

[120] GYORFFY, B., GYORFFY, A., AND TULASSAY, Z. (2005). The problem of multiple testing and solutions for genome-wide studies. *Orv. Hetil.* **146**, 12, 559–563.

[121] *HEINICKE, A. (1999). *Statistische Modellierung und Erfassung multivariater Assoziationsstrukturen anhand von KS-Verteilungen.* Logos-Verlag, Berlin.

[122] HERO, A. O., FLEURY, G., MEARS, A. J., AND SWAROOP, A. (2004). Multicriteria gene screening for analysis of differential expression with DNA microarrays. *EURASIP J. Appl. Signal Process.* **2004**, 1, 43–52.

[123] HOCHBERG, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* **75**, 4, 800–802.

[124] HOCHBERG, Y. AND HOMMEL, G. (1998). Step-up multiple testing procedures. *Encyc. Statist. Sc. Supplementary 2*.

[125] HOCHBERG, Y. AND ROM, D. (1995). Extensions of multiple testing procedures based on Simes' test. *J. Stat. Plann. Inference* **48**, 2, 141–152.

[126] HOLLAND, B. AND CHEUNG, S. H. (2002). Familywise robustness criteria for multiple-comparison procedures. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **64**, 1, 63–77.

[127] HOMMEL, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika* **75**, 2, 383–386.

[128] HOMMEL, G. AND KROPF, S. (2004). S36.1: Tests for differentiation in gene expression using a data-driven order or weights for hypotheses. *Biom. J.* **46**, S1, 76.

[129] HOMMEL, G. AND KROPF, S. (2005). Tests for Differentiation in Gene Expression Using a Data-Driven Order or Weights for Hypotheses. *Biom. J.* **47**, 4, 554–562.

[130] HONG, F. AND LI, H. (2006). Functional hierarchical models for identifying genes with different time-course expression profiles. *Biometrics* **62**, 2, 534–544.

[131] HOPKINS, A. M., MILLER, C. J., CONNOLLY, A. J., GENOVESE, C., NICHOL, R. C., AND WASSERMAN, L. (2002). A New Source Detection Algorithm Using the False-Discovery Rate. *A.J. 123*, 1086–1094.

[132] HSUEH, H.-M., CHEN, J. J., AND KODELL, R. L. (2003). Comparison of Methods for Estimating the Number of True Null Hypotheses in Multiplicity Testing. *J. Biopharm. Stat.* **13**, 4, 675–689.

[133] HU, J. AND WRIGHT, F. A. (2007). Assessing differential gene expression with small sample sizes in oligonucleotide arrays using a mean-variance model. *Biometrics* **63**, 1, 41–49.

[134] HU, J., ZOU, F., AND WRIGHT, F. A. (2005). Practical FDR-based sample size calculations in microarray experiments. *Bioinformatics* **21**, 15, 3264–3272.

[135] HUANG, J., GUSNANTO, A., O'SULLIVAN, K., STAAF, J., BORG, A., AND PAWITAN, Y. (2007). Robust smooth segmentation approach for array CGH data analysis. *Bioinformatics*, btm359.

[136] HUANG, T., WU, B., LIZARDI, P., AND ZHAO, H. (2005). Detection of DNA copy number alterations using penalized least squares regression. *Bioinformatics* **21**, 20, 3811–3817.

[137] HUANG, X. AND ZHU, Q. (2002). A pseudo-nearest-neighbor approach for missing data recovery on Gaussian random data sets. *Pattern Recognit. Lett.* **23**, 13, 1613–1622.

[138] ISHWARAN, H. AND RAO, J. (2003). Detecting differentially expressed genes in microarrays using Bayesian model selection. *J. Am. Stat. Assoc.* **98**, 462, 438–455.

[139] IYER, V. AND SARKAR, S. (2007). An Adaptive Single-step FDR Procedure with Applications to DNA Microarray Analysis. *Biom. J.* **49**, 1, 127–135.

[140] JAIN, N., CHO, H., O'CONNELL, M., AND LEE, J. (2005). Rank-invariant resampling based estimation of false discovery rate for analysis of small sample microarray data. *BMC Bioinformatics* **6**, 187.

[141] JAIN, N., THATTE, J., BRACIALE, T., LEY, K., O'CONNELL, M., AND LEE, J. K. (2003). Local-pooled-error test for identifying differentially expressed genes with a small number of replicated microarrays. *Bioinformatics* **19**, 15, 1945–1951.

[142] JI, H. AND WONG, W. H. (2005). TileMap: create chromosomal map of tiling array hybridizations. *Bioinformatics* **21**, 18, 3629–3636.

[143] JUNG, K., QUAST, K., GANNOUN, A., AND URFER, W. (2006). A Renewed Approach to the Nonparametric Analysis of Replicated Microarray Experiments. *Biom. J.* **48**, 2, 245–254.

[144] JUNG, S.-H. (2005). Sample size for FDR-control in microarray data analysis. *Bioinformatics* **21**, 14, 3097–3104.

[145] JUNG, Y.-Y., OH, M.-S., SHIN, D. W., KANG, S.-H., AND OH, H. S. (2006). Identifying Differentially Expressed Genes in Meta-Analysis via Bayesian Model-Based Clustering. *Biom. J.* **48**, 3, 435–450.

[146] KAFKAFI, N., BENJAMINI, Y., SAKOV, A., AND ET AL. (2005). Genotype-environment interactions in mouse behavior: A way out of the problem. *Proc. Natl. Acad. Sci. USA* **102**, 12, 4619–4624.

[147] *KARLIN, S. (1968). *Total Positivity.* Stanford University Press, Stanford, California.

[148] *KARLIN, S. AND RINOTT, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Anal. 10*, 467–498.

[149] *KARLIN, S. AND RINOTT, Y. (1981). Total positivity properties of absolute value multinormal variables with applications to confidence interval estimates and related probabilistic inequalities. *Ann. Stat.* **9**, 5, 1035–1049.

[150] KATHLEEN KERR, M. (2003). Design considerations for efficient and effective microarray studies. *Biometrics* **59**, 4, 822–828.

[151] KELES, S., VAN DER LAAN, M. J., DUDOIT, S., AND CAWLEY, S. E. (2006). Multiple Testing Methods For ChIP-Chip High Density Oligonucleotide Array Data. *J. Comput. Biol.* **13**, 3, 579–613.

[152] *KELKER, D. (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhyā Ser. A 32*, 419–430.

[153] KENDZIORSKI, C. M., NEWTON, M. A., LAN, H., AND GOULD, M. N. (2003). On parametric empirical Bayes methods for comparing multiple groups using replicated gene expression profiles. *Stat. Med.* **22**, 24, 3899–3914.

[154] KESELMAN, H., CRIBBIE, R., AND HOLLAND, B. (2002). Controlling the rate of Type I error over a large set of statistical tests. *Br. J. Math. Stat. Psychol.* **55**, 1, 27–39.

[155] KOOPERBERG, C., SIPIONE, S., LEBLANC, M., STRAND, A. D., CATTANEO, E., AND OLSON, J. M. (2002). Evaluating test statistics to select interesting genes in microarray experiments. *Hum. Mol. Genet.* **11**, 19, 2223–2232.

[156] KORN, E. L. AND FREIDLIN, B. (2007). A Note on Controlling the Number of False Positives. *Biometrics, to appear*.

[157] KORN, E. L., TROENDLE, J. F., MCSHANE, L. M., AND SIMON, R. (2004). Controlling the number of false discoveries: application to high-dimensional genomic data. *J. Stat. Plann. Inference* **124**, 2, 379–398.

[158] KOSOROK, M. R. AND MA, S. (2007). Marginal asymptotics for the 'Large P, Small N' Paradigm: with applications to microarray data. *Ann. Stat.,* **35**, 4, 1456–1486.

[159] KROHN, K., ESZLINGER, M., PASCHKE, R., ROEDER, I., AND SCHUSTER, E. (2005). Increased power of microarray analysis by use of an algorithm based on a multivariate procedure. *Bioinformatics* **21**, 17, 3530–3534.

[160] KWONG, K. S., HOLLAND, B., AND CHEUNG, S. H. (2002). A modified Benjamini-Hochberg multiple comparisons procedure for controlling the false discovery rate. *J. Stat. Plann. Inference* **104**, 2, 351–362.

[161] *KWONG, K.-S. AND WONG, E.-H. (2002). A more powerful step-up procedure for controlling the false discovery rate under independence. *Stat. Probab. Lett.* **56**, 2, 217–225.

[162] LAI, Y. AND ZHAO, H. (2005). A statistical method to detect chromosomal regions with DNA copy number alterations using SNP-array-based CGH data. *Comput. Biol. Chem.* **29**, 1, 47–54.

[163] LAIRD, A. R., FOX, P. M., PRICE, C. J., GLAHN, D. C., UECKER, A. M., LANCASTER, J. L., TURKELTAUB, P. E., KOCHUNOV, P., AND FOX, P. T. (2005). ALE meta-analysis: controlling the false discovery rate and performing statistical contrasts. *Hum. Brain Mapp.* **25**, 1, 155–164.

[164] *LANDSMAN, Z. M. AND VALDEZ, E. A. (2003). Tail conditional expectations for elliptical distributions. *N. Am. Actuar. J.* **7**, 4, 55–71.

[165] LANGAAS, M., LINDQVIST, B., AND FERKINGSTAD, E. (2005). Estimating the Proportion of True Null Hypotheses, with Application to DNA Microarray Data. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **67**, 4, 555–572.

[166] *LEE, E. T. (1992). *Statistical Methods for Survival Data Analysis. Second Edition.* John Wiley & Sons, Inc., New York.

[167] LEE, H., DEKKERS, J. C. M., SOLLER, M., MALEK, M., FERNANDO, R. L., AND ROTHSCHILD, M. F. (2002). Application of the false discovery rate to quantitative trait loci interval mapping with multiple traits. *Genetics* **161**, 2, 905–914.

[168] LEE, M.-L. T. AND WHITMORE, G. A. (2002). Power and sample size for DNA microarray studies. *Stat. Med.* **21**, 23, 3543–3570.

[169] LEHMANN, E., ROMANO, J., AND SHAFFER, J. (2005). On optimality of stepdown and stepup multiple test procedures. *Ann. Stat.* **33**, 3, 1084–1108.

[170] *LEHMANN, E. AND ROMANO, J. P. (2005). Generalizations of the familywise error rate. *Ann. Stat.* **33**, 3, 1138–1154.

[171] *LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Stat. 37*, 1137–1153.

[172] LEWIS, C. AND THAYER, D. T. (2004). A loss function related to the FDR for random effects multiple comparisons. *J. Stat. Plann. Inference* **125**, 1-2, 49–58.

[173] LI, H., WOOD, C. L., LIU, Y., GETCHELL, T. V., GETCHELL, M. L., AND STROMBERG, A. J. (2006). Identification of gene expression patterns using planned linear contrasts. *BMC Bioinformatics 7*, 245.

[174] LI, J. AND JI, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* **95**, 3, 221–227.

[175] LI, S. S., BIGLER, J., LAMPE, J. W., POTTER, J. D., AND FENG, Z. (2005). FDR-controlling testing procedures and sample size determination for microarrays. *Stat. Med.* **24**, 15, 2267–2280.

[176] *LIANG, C.-L., RICE, J., DE PATER, I., ALCOCK, C., AXELROD, T., WANG, A., AND MARSHALL, S. (2004). Statistical methods for detecting stellar occultations by Kuiper Belt objects: The Taiwanese-American Occultation Survey. *Statist. Sci. 19*, 265–274.

[177] LIAO, J. AND CHIN, K.-V. (2007). Logistic regression for disease classification using microarray data: model selection in a large p and small n case. *Bioinformatics* **23**, 15, 1945–1951.

[178] LIAO, J. G., LIN, Y., SELVANAYAGAM, Z. E., AND SHIH, W. J. (2004). A mixture model for estimating the local false discovery rate in DNA microarray analysis. *Bioinformatics* **20**, 16, 2694–2701.

[179] LIN, D. Y. (2005). An efficient Monte Carlo approach to assessing statistical significance in genomic studies. *Bioinformatics* **21**, 6, 781–787.

[180] LIU, P. AND HWANG, J. T. G. (2007). Quick calculation for sample size while controlling false discovery rate with application to microarray analysis. *Bioinformatics* **23**, 6, 739–746.

[181] *LORENTZ, G. (1953). An inequality for rearrangements. *Am. Math. Mon. 60*, 176–179.

[182] LU, Y., ZHU, J., AND LIU, P. (2005). A two-step strategy for detecting differential gene expression in cDNA microarray data. *Curr. Genet.* **47**, 2, 121–131.

[183] LUAN, Y. AND LI, H. (2004). Model-based methods for identifying periodically expressed genes based on time course microarray gene expression data. *Bioinformatics* **20**, 3, 332–339.

[184] LÄUTER, J. (2005). *Hochdimensionale Statistik - Anwendung in der Genexpressionsanalyse. Vorlesung am Interdisziplinären Zentrum für Bioinformatik 2003/ 2005.* Leipzig Bioinformatics Working Papers No. 7.

[185] MA, J., LI, F., AND LIU, J. (2005). Non-parametric statistical tests for informative gene selection. In *Wang, Jun (ed.) et al., Advances in neural networks - ISNN 2005. Second international symposium on neural networks, Chongqing, China, May 30 - June 1, 2005. Proceedings, Part III. Berlin: Springer. Lecture Notes in Computer Science 3498, 697-702* .

[186] MARCHINI, J. AND PRESANIS, A. (2004). Comparing methods of analyzing fMRI statistical parametric maps. *Neuroimage* **22**, 3, 1203–1213.

[187] MARTIN, D. E., DEMOUGIN, P., HALL, M. N., AND BELLIS, M. (2004). Rank Difference Analysis of Microarrays (RDAM), a novel approach to statistical analysis of microarray expression profiling data. *BMC Bioinformatics 5*, 148.

[188] MEHROTRA, D. V. AND HEYSE, J. F. (2004). Use of the false discovery rate for evaluating clinical safety data. *Stat. Methods Med. Res.* **13**, 3, 227–238.

[189] MEINSHAUSEN, N. (2006). False discovery control for multiple tests of association under general dependence. *Scand. J. Statist.* **33**, 5, 227–237.

[190] MEINSHAUSEN, N. AND BÜHLMANN, P. (2005). Lower bounds for the number of false null hypotheses for multiple testing of associations under general dependence structures. *Biometrika* **92**, 4, 893–907.

[191] MEINSHAUSEN, N. AND RICE, J. (2006). Estimating the proportion of false null hypotheses among a large number of independently tested hypotheses. *Ann. Stat.* **34**, 1, 373–393.

[192] MEUWISSEN, T. H. E. AND GODDARD, M. E. (2004). Bootstrapping of gene-expression data improves and controls the false discovery rate of differentially expressed genes. *Genet. Sel. Evol.* **36**, 2, 191–205.

[193] *MILLS, J. P. (1926). Table of the ratio: Area to bounding ordinate, for any portion of the normal curve. *Biometrika 18*, 395–400.

[194] MOTULSKY, H. J. AND BROWN, R. E. (2006). Detecting outliers when fitting data with nonlinear regression - a new method based on robust nonlinear regression and the false discovery rate. *BMC Bioinformatics 7*, 123.

[195] MUELLER, P., PARMIGIANI, G., ROBERT, C., AND ROUSSEAU, J. (2004). Optimal sample size for multiple testing: the case of gene expression microarrays. *J. Am. Stat. Assoc.* **99**, 468, 990–1001.

[196] MUKHOPADHYAY, N. D. AND CHATTERJEE, S. (2007). Causality and pathway search in microarray time series experiment. *Bioinformatics* **23**, 4, 442–449.

[197] NETTLETON, D., HWANG, J. T. G., CALDO, R. A., AND WISE, R. P. (2006). Estimating the Number of True Null Hypotheses From a Histogram of p Values. *J. Agric. Biol. Environ. Stat.* **11**, 3, 337–356.

[198] NEWTON, M. A., NOUEIRY, A., SARKAR, D., AND AHLQUIST, P. (2004). Detecting differential gene expression with a semiparametric hierarchical mixture method. *Biostatistics* **5**, 2, 155–176.

[199] NGUYEN, D. (2004). On estimating the proportion of true null hypotheses for false discovery rate controlling procedures in exploratory DNA microarray studies. *Comput. Stat. Data Anal.* **47**, 3, 611–637.

[200] NGUYEN, D. V. N. (2005). A Unified Computational Framework to Compare Direct and Sequential False Discovery Rate Algorithms for Exploratory DNA Microarray Studies. *J. Data Sci. 3*, 331–352.

[201] NICHOLS, T. AND HAYASAKA, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat. Methods Med. Res.* **12**, 5, 419–446.

[202] NORRIS, A. W. AND KAHN, C. R. (2006). Analysis of gene expression in pathophysiological states: balancing false discovery and false negative rates. *Proc. Natl. Acad. Sci. USA* **103**, 3, 649–653.

[203] *NOTTERMAN, D. A., ALON, U., AND SIERK, A. J. (2001). Transcriptional Gene Expression Profiles of Colorectal Adenoma, Adenocarcinoma, and Normal Tissue Examined by Oligonucleotide Arrays. *Cancer Research 61*, 3124–3130.

[204] OEHLERT, G. W. (2001a). Correction on "Student-Newman-Kuels controls the false discovery rate". *Stat. Probab. Lett.* **51**, 4, 431.

[205] OEHLERT, G. W. (2001b). Student-Newman-Kuels controls the false discovery rate. *Stat. Probab. Lett.* **46**, 4, 381–383.

[206] OWEN, A. B. (2005). Variance of the number of false discoveries. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **67**, 3, 411–426.

[207] PACIFICO, M., GENOVESE, C., VERDINELLI, I., AND WASSERMAN, L. (2004). False discovery control for random fields. *J. Am. Stat. Assoc.* **99**, 468, 1002–1014.

[208] PAN, W. (2005). Incorporating Biological Information as a Prior in an Empirical Bayes Approach to Analyzing Microarray Data. *Stat. Appl. Genet. Mol. Biol.* **4**, 1.

[209] PAWITAN, Y., CALZA, S., AND PLONER, A. (2006). Estimation of false discovery proportion under general dependence. *Bioinformatics* **22**, 24, 3025–3031.

[210] PAWITAN, Y., MICHIELS, S., KOSCIELNY, S., GUSNANTO, A., AND PLONER, A. (2005). False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics* **21**, 13, 3017–3024.

[211] PAWITAN, Y., MURTHY, K. R. K., MICHIELS, S., AND PLONER, A. (2005). Bias in the estimation of false discovery rate in microarray studies. *Bioinformatics* **21**, 20, 3865–3872.

[212] PAWLUK-KOLC, M., ZIEBA-PALUS, J., AND PARCZEWSKI, A. (2006). Application of false discovery rate procedure to pairwise comparisons of refractive index of glass fragments. *Forensic Sci. Int.* **160**, 1, 53–58.

[213] PENNELLO, G. A. (2006). Duncan's k-Ratio Bayes Rule Approach to Multiple Comparisons: An Overview. *Biom. J.* **48**, 6, 1–16.

[214] PLONER, A., CALZA, S., GUSNANTO, A., AND PAWITAN, Y. (2006). Multidimensional local false discovery rate for microarray studies. *Bioinformatics* **22**, 5, 556–565.

[215] POUNDS, S. AND CHENG, C. (2004). Improving false discovery rate estimation. *Bioinformatics* **20**, 11, 1737–1745.

[216] POUNDS, S. AND CHENG, C. (2005). Sample size determination for the false discovery rate. *Bioinformatics* **21**, 23, 4263–4271.

[217] POUNDS, S. AND CHENG, C. (2006). Robust estimation of the false discovery rate. *Bioinformatics* **22**, 16, 1979–1987.

[218] POUNDS, S. B. (2006). Estimation and control of multiple testing error rates for microarray studies. *Brief. Bioinform.* **7**, 1, 25–36.

[219] PYNE, S., FUTCHER, B., AND SKIENA, S. (2003). Meta-analysis based on control of false discovery rate: combining yeast ChIP-chip datasets. *Bioinformatics* **22**, 20, 2516–2522.

[220] QIAN, H.-R. AND HUANG, S. (2005). Comparison of false discovery rate methods in identifying genes with differential expression. *Genomics* **86**, 4, 495–503.

[221] RAJICIC, N., FINKELSTEIN, D. M., SCHOENFELD, D. A., AND THE INFLAMMATION HOST RESPONSE TO INJURY RESEARCH PROGRAM INVESTIGATORS. (2006). Survival analysis of longitudinal microarrays. *Bioinformatics* **22**, 21, 2643–2649.

[222] REINER, A., YEKUTIELI, D., AND BENJAMINI, Y. (2003). Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* **19**, 3, 368–375.

[223] REINER-BENAIM, A. (2007). FDR Control by the BH Procedure for Two-Sided Correlated Tests with Implications to Gene Expression Data Analysis. *Biom. J.* **49**, 1, 107–126.

[224] REPSILBER, D., MIRA, A., LINDROOS, H., ANDERSSON, S., AND ZIEGLER, A. (2005). Data Rotation Improves Genomotyping Efficiency. *Biom. J.* **47**, 4, 585–598.

[225] ROMANO, J. AND SHAIKH, A. (2006a). On stepdown control of the false discovery proportion. *IMS Lecture Notes–Monograph Series 49*, 33–50.

[226] ROMANO, J. AND SHAIKH, A. (2006b). Stepup procedures for control of generalizations of the familywise error rate. *Ann. Stat.* **34**, 4, 1850–1873.

[227] ROMANO, J. P., SHAIKH, A. M., AND WOLF, M. (2005). Formalized Data Snooping Based on Generalized Error Rates. *Institute for Empirical Research in Economics, University of Zurich, Working Paper Series, Working Paper No. 259*.

[228] ROMANO, J. P. AND WOLF, M. (2005). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica* **73**, 4, 1237–1282.

[229] ROSENBERG, P. S., CHE, A., AND CHEN, B. E. (2005). Multiple hypothesis testing strategies for genetic case-control association studies. *Stat. Med.* **25**, 18, 3134–3149.

[230] *RÜSCHENDORF, L. (1981). Characterization of dependence concepts in normal distributions. *Ann. Inst. Statist. Math.* **33**, Part A, 347–359.

[231] RUPPERT, D., NETTLETON, D., AND HWANG, J. T. G. (2007). Exploring the information in p-values for the analysis and planning of multiple-test experiments. *Biometrics,* **63**, 2, 483–495.

[232] SABATTI, C., KARSTEN, S. L., AND GESCHWIND, D. H. (2002). Thresholding rules for recovering a sparse signal from microarray experiments. *Math. Biosci.* **176**, 1, 17–34.

[233] SABATTI, C., SERVICE, S., AND FREIMER, N. (2003). False discovery rate in linkage and association genome screens for complex disorders. *Genetics* **164**, 2, 829–833.

[234] *SACKROWITZ, H. AND SAMUEL-CAHN, E. (1999). P Values as Random Variables - Expected P Values. *Amer. Statist.* **53**, 4, 326–331.

[235] SAMUEL-CAHN, E. (1999). A note about a curious generalization of Simes' theorem. *J. Stat. Plann. Inference* **82**, 1-2, 147–149.

[236] *SARKAR, S. K. (1998). Some probability inequalities for ordered MTP_2 random variables: A proof of the Simes conjecture. *Ann. Stat.* **26**, 2, 494–504.

[237] *SARKAR, S. K. (2002). Some results on false discovery rate in stepwise multiple testing procedures. *Ann. Stat.* **30**, 1, 239–257.

[238] SARKAR, S. K. (2004). FDR-controlling stepwise procedures and their false negatives rates. *J. Stat. Plann. Inference* **125**, 1-2, 119–137.

[239] SARKAR, S. K. (2006). False discovery and false nondiscovery rates in single-step multiple testing procedures. *Ann. Stat.* **34**, 1, 394–415.

[240] SARKAR, S. K. (2007a). Stepup Procedures Controlling Generalized FWER and Generalized FDR. *Ann. Stat., to appear*.

[241] SARKAR, S. K. (2007b). Two-stage stepup procedures controlling FDR. *J. Stat. Plann. Inference, to appear*.

[242] SARKAR, S. K. AND CHANG, C.-K. (1997). The Simes method for multiple hypothesis testing with positively dependent statistics. *J. Amer. Statist. Assoc. 92*, 1601–1608.

[243] SARKAR, S. K. AND ZHOU, T. (2007). Controlling Bayes directional false discovery rate in random effects model. *J. Stat. Plann. Inference, to appear*.

[244] *SARKAR, T. K. (1969). *Some lower bounds of reliability.* Technical Report No. 124, Department of Statistics, Stanford University, Stanford, California.

[245] SCHAFER, J. AND STRIMMER, K. (2005). An empirical Bayes approach to inferring large-scale gene association networks. *Bioinformatics* **21**, 6, 754–764.

[246] SCHEID, S. AND SPANG, R. (2005). twilight; a Bioconductor package for estimating the local false discovery rate. *Bioinformatics* **21**, 12, 2921–2922.

[247] SCHWENDER, H. (2003). *Assessing the False Discovery Rate in a Statistical Analysis of Gene Expression Data. Diploma thesis.* University of Dortmund, Department of Statistics.

[248] SEEGER, P. (1966). Variance Analysis of Complete Designs. Some Practical Aspects. *Almquist & Wiksell, Stockholm*.

[249] SEEGER, P. (1968). A Note on a Method for the Analysis of Significances en masse. *Technometrics* **10**, 3, 586–593.

[250] *SEN, P. K. (1999). Some remarks on Simes-type multiple test of significance. *J. Statist. Plann. Inference 82*, 139–145.

[251] SEN, P. K. (2007). Burden of bioinformatics in medical research: Statistical perspectives and controversies. *J. Stat. Plann. Inference,* **138**, 2, 450–463.

[252] SHAFFER, J. P. (1995). Multiple hypothesis testing. *Annual Rev. Psych. 46*, 561–584.

[253] SHAFFER, J. P. (1999). A semi-Bayesian study of Duncan's Bayesian multiple comparison procedure. *J. Stat. Plann. Inference* **82**, 1-2, 197–213.

[254] SHAFFER, J. P. (2002). Multiplicity, directional (type III) errors, and the null hypothesis. *Psychol. Methods* **7**, 3, 356–369.

[255] SHAFFER, J. P. (2007). Controlling the False Discovery Rate with Constraints: The Newman-Keuls Test Revisited. *Biom. J.* **49**, 1, 1–8.

[256] SHAROV, A. A., DUDEKULA, D. B., AND KO, M. S. H. (2005). A web-based tool for principal component and significance analysis of microarray data. *Bioinformatics* **21**, 10, 2548–2549.

[257] SHEDDEN, K., CHEN, W., KUICK, R., GHOSH, D., MACDONALD, J., CHO, K. R., GIORDANO, T. J., GRUBER, S. B., FEARON, E. R., TAYLOR, J. M. G., AND HANASH, S. (2005). Comparison of seven methods for producing Affymetrix expression scores based on False Discovery Rates in disease profiling data. *BMC Bioinformatics* **6**, 1, 26.

[258] SHEFFLER, W., UPFAL, E., SEDIVY, J., AND NOBLE, W. S. (2005). A Learned Comparative Expression Measure for Affymetrix GeneChip DNA Microarrays. *Proc. IEEE Comput. Syst. Bioinform. Conf.*, 144–154.

[259] SHEN, C., LI, L., AND CHEN, J. Y. (2005). Discover True Association Rates in Multiprotein Complex Proteomics Data Sets. *Proc. IEEE Comput. Syst. Bioinform. Conf.*, 167–174.

[260] SHEN, C., LI, L., AND CHEN, J. Y. (2006). A statistical framework to discover true associations from multiprotein complex pull-down proteomics data sets. *Proteins* **64**, 2, 436–443.

[261] SHEN, X., HUANG, H.-C., AND CRESSIE, N. (2002). Nonparametric hypothesis testing for a spatial signal. *J. Am. Stat. Assoc.* **97**, 460, 1122–1140.

[262] *SHORACK, G. P. AND WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

[263] SIEGEL, E. R., SPENCER, T., AND PARRISH, R. S. (2003). Controlling False Discovery when planning microarray experiments. *Joint Statistical Meetings - Biometrics Section-to include ENAR & WNAR*, 3897–3899.

[264] *SIMES, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika 73*, 751–754.

[265] SIMONSEN, K. L. AND MCINTYRE, L. M. (2004). Using alpha wisely: Improving power to detect multiple QTL. *Stat. Appl. Genet. Mol. Biol.* **3**, 1.

[266] SINHA, R., SINHA, M., MATHEW, G., ELSTON, R., AND LUO, Y. (2005). Local false discovery rate and minimum total error rate approaches to identifying interesting chromosomal regions. *BMC Genet. 6 Suppl. 1*, S23.

[267] SOMERVILLE, P. N. (2004). FDR step-down and step-up procedures for the correlated case. *Recent developments in multiple comparisons. IMS Lecture Notes–Monograph Series 47*, 100–118.

[268] SONG, J. S., MAGHSOUDI, K., LI, W., FOX, E., QUACKENBUSH, J., AND SHIRLEY LIU, X. (2007). Microarray blob-defect removal improves array analysis. *Bioinformatics* **23**, 8, 966–971.

[269] SORIC, B. (1989). Statistical "Discoveries" and Effect-Size Estimation. *J. Am. Stat. Assoc.* **84**, 406, 608–610.

[270] *STADER, A. (1992). *Analytische Eigenschaften nichtzentraler Verteilungen.* Dissertation, Fachbereich IV, Universität Trier.

[271] *STOREY, J. D. (2002a). A direct approach to false discovery rates. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **64**, 3, 479–498.

[272] *STOREY, J. D. (2002b). *False Discovery Rates. Theory and Applications to DNA microarrays. Ph. D. dissertation.* Stanford University, Department of Statistics.

[273] *STOREY, J. D. (2003). The positive false discovery rate: A Bayesian interpretation and the q-value. *Ann. Stat.* **31**, 6, 2013–2035.

[274] STOREY, J. D. (2007). The optimal discovery procedure: a new approach to simultaneous significance testing. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **69**, 3, 347–368.

[275] *STOREY, J. D., TAYLOR, J. E., AND SIEGMUND, D. (2004). Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **66**, 1, 187–205.

[276] STOREY, J. D., TAYLOR, J. E., AND SIEGMUND, D. (2005). Acknowledgement of priority: Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **67**, 1, 197–197.

[277] STOREY, J. D. AND TIBSHIRANI, R. (2003a). SAM thresholding and false discovery rates for detecting differential gene expression in DNA microarrays. *In: The Analysis of Gene Expression Data: Methods and Software, by G. Parmigiani, E. S. Garrett, R. A. Irizarry and S. L. Zeger (editors). Springer, New York*.

[278] STOREY, J. D. AND TIBSHIRANI, R. (2003b). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA* **100**, 16, 9440–9445.

[279] TADESSE, M. G., IBRAHIM, J. G., VANNUCCI, M., AND GENTLEMAN, R. (2005). Wavelet thresholding with Bayesian false discovery rate control. *Biometrics* **61**, 1, 25–35.

[280] *TAMHANE, A. C., LIU, W., AND DUNNETT, C. W. (1998). A generalized step-up-down multiple test procedure. *Can. J. Stat. 26*, 353–363.

[281] TAN, C. S., PLONER, A., QUANDT, A., LEHTIO, J., AND PAWITAN, Y. (2006). Finding regions of significance in SELDI measurements for identifying protein biomarkers. *Bioinformatics* **22**, 12, 1515–1523.

[282] TANG, Y., GHOSAL, S., AND ROY, A. (2007). Nonparametric bayesian estimation of positive false discovery rates. *Biometrics , to appear*.

[283] TAYLOR, J. AND TIBSHIRANI, R. (2006). A tail strength measure for assessing the overall univariate significance in a dataset. *Biostatistics* **7**, 2, 167–181.

[284] TAYLOR, J., TIBSHIRANI, R., AND EFRON, B. (2005). The 'miss rate' for the analysis of gene expression data. *Biostatistics* **6**, 1, 111–117.

[285] TAYLOR, J. E., WORSLEY, K. J., AND GOSSELIN, F. (2007). Maxima of discretely sampled random fields, with an application to 'bubbles'. *Biometrika* **94**, 1, 1–18.

[286] TESCHENDORFF, A. E., NADERI, A., BARBOSA-MORAIS, N. L., AND CALDAS, C. (2006). PACK: Profile Analysis using Clustering and Kurtosis to find molecular classifiers in cancer. *Bioinformatics* **22**, 18, 2269–2275.

[287] TIBSHIRANI, R. (2006). A simple method for assessing sample sizes in microarray experiments. *BMC Bioinformatics 7*, 106.

[288] TROENDLE, J. F. (2000). Stepwise normal theory multiple test procedures controlling the false discovery rate. *J. Stat. Plann. Inference* **84**, 1-2, 139–158.

[289] TSAI, C.-A. AND CHEN, J. J. (2004). Significance analysis of ROC indices for comparing diagnostic markers: applications to gene microarray data. *J. Biopharm. Stat.* **14**, 4, 985–1003.

[290] TSAI, C.-A., HSUEH, H.-M., AND CHEN, J. J. (2003). Estimation of False Discovery Rates in Multiple Testing: Application to Gene Microarray Data. *Biometrics 59*, 1071–1081.

[291] TSUI, K.-W. AND TANG, S. (2007). Simultaneous testing of multiple hypotheses using generalized p-values. *Stat. Probab. Lett.* **77**, 12, 1362–1370.

[292] TUSHER, V. G., TIBSHIRANI, R., AND CHU, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* **98**, 9, 5116–5121.

[293] TZENG, J.-Y., BYERLEY, W., DEVLIN, B., ROEDER, K., AND WASSERMAN, L. (2003). Outlier detection and false discovery rates for whole-genome DNA matching. *J. Am. Stat. Assoc.* **98**, 461, 236–246.

[294] UNGER, T., KORADE, Z., LAZAROV, O., TERRANO, D., SISODIA, S. S., AND MIRNICS, K. (2005). True and false discovery in DNA microarray experiments: transcriptome changes in the hippocampus of presenilin 1 mutant mice. *Methods* **37**, 3, 261–273.

[295] VAN DE WIEL, M. A. AND IN KIM, K. (2007). Estimating the false discovery rate using nonparametric deconvolution. *Biometrics 63*, 806–815.

[296] VAN DEN OORD, E. J. C. G. AND SULLIVAN, P. F. (2003). False discoveries and models for gene discovery. *Trends Genet.* **19**, 10, 537–542.

[297] *VAN DER LAAN, M. J., DUDOIT, S., AND POLLARD, K. S. (2004). Augmentation procedures for control of the generalized family-wise error rate and tail probabilities for the proportion of false positives. *Stat. Appl. Genet. Mol. Biol.* **3**, 1.

[298] VICTOR, A., FALDUM, A., AND HOMMEL, G. (2004). S36.4: Control of false discovery rate in adaptive designs. *Biom. J.* **46**, S1, 77.

[299] VICTOR, A. AND HOMMEL, G. (2007). Combining Adaptive Designs with Control of the False Discovery Rate - A Generalized Definition for a Global p -Value. *Biom. J.* **49**, 1, 94–106.

[300] WANG, P., KIM, Y., POLLACK, J., NARASIMHAN, B., AND TIBSHIRANI, R. (2005). A method for calling gains and losses in array CGH data. *Biostatistics* **6**, 1, 45–58.

[301] WANG, S. AND ETHIER, S. (2004). A generalized likelihood ratio test to identify differentially expressed genes from microarray data. *Bioinformatics* **20**, 1, 100–104.

[302] WEATHERLY, D. B., ASTWOOD, J. A. R., MINNING, T. A., CAVOLA, C., TARLETON, R. L., AND ORLANDO, R. (2005). A Heuristic method for assigning a false-discovery rate

for protein identifications from Mascot database search results. *Mol. Cell. Proteomics* **4**, 6, 762–772.

[303] WEI, Z. AND LI, H. (2007). A Markov random field model for network-based analysis of genomic data. *Bioinformatics* **23**, 12, 1537–1544.

[304] WELLER, J. I., SONG, J. Z., HEYEN, D. W., LEWIN, H. A., AND RON, M. (1998). A new approach to the problem of multiple comparisons in the genetic dissection of complex traits. *Genetics* **150**, 4, 1699–1706.

[305] WICHERT, S., FOKIANOS, K., AND STRIMMER, K. (2004). Identifying periodically expressed transcripts in microarray time series data. *Bioinformatics* **20**, 1, 5–20.

[306] *WIJSMAN, R. A. (1985). A useful inequality on ratios of integrals, with applications to maximum likelihood estimation. *J. Amer. Statist. Assoc. 80*, 472–475.

[307] WILLE, A., HOH, J., AND OTT, J. (2003). Sum statistics for the joint detection of multiple disease loci in case-control association studies with SNP markers. *Genet. Epidemiol.* **25**, 4, 350–359.

[308] WILLIAMS, V. S. L., JONES, L. V., AND TUKEY, J. W. (1994). *Controlling error in multiple comparisons with special attention to the National Assessment of Educational Progress.* National Institute of Statistical Sciences, P. O. Box 14162, Research Triangle Park, N. C. 27709, Technical Report # 33.

[309] WILLIAMS, V. S. L., JONES, L. V., AND TUKEY, J. W. (1999). Controlling error in multiple comparisons, with examples from state-to-state differences in educational achievement. *J. Edu. Behav. Statist. 24*, 42–69.

[310] WU, B., GUAN, Z., AND ZHAO, H. (2006). Parametric and Nonparametric FDR Estimation Revisited. *Biometrics* **62**, 3, 735–744.

[311] WU, S. S. AND WANG, W. (2007). Step-up simultaneous tests for identifying active effects in orthogonal saturated designs. *Ann. Stat.,* **35**, 1, 449–463.

[312] XIE, Y., PAN, W., AND KHODURSKY, A. B. (2005). A note on using permutation-based false discovery rate estimates to compare different analysis methods for microarray data. *Bioinformatics* **21**, 23, 4280–4288.

[313] YANG, J. J. AND YANG, M. C. K. (2006). An improved procedure for gene selection from microarray experiments using false discovery rate criterion. *BMC Bioinformatics 7*, 15.

[314] YANG, M. C. K., YANG, J. J., MCINDOE, R. A., AND SHE, J. X. (2003). Microarray experimental design: power and sample size considerations. *Physiol. Genomics* **16**, 1, 24–28.

[315] YANG, Q., CUI, J., CHAZARO, I., CUPPLES, L. A., AND DEMISSIE, S. (2005). Power and type I error rate of false discovery rate approaches in genome-wide association studies. *BMC Genet. 6 Suppl. 1*, S134.

[316] YANG, Y., HOH, J., BROGER, C., NEEB, M., EDINGTON, J., LINDPAINTNER, K., AND OTT, J. (2003). Statistical methods for analyzing microarray feature data with replications. *J. Comput. Biol.* **10**, 2, 157–169.

[317] YEKUTIELI, D. (2007). False discovery rate control for non-positively regression dependent test statistics. *J. Stat. Plann. Inference,* **138**, 2, 405–415.

[318] YEKUTIELI, D. AND BENJAMINI, Y. (1999). Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *J. Stat. Plann. Inference* **82**, 1-2, 171–196.

[319] YEKUTIELI, D., REINER-BENAIM, A., BENJAMINI, Y., ELMER, G., N., K., LETWIN, N., AND LEE, N. (2007). Approaches to multiplicity issues in complex research in microarray analysis. *Stat. Neerl.* **60**, 4, 414–437.

[320] YUAN, M. AND KENDZIORSKI, C. (2006). A Unified Approach for Simultaneous Gene Clustering and Differential Expression Identification. *Biometrics* **62**, 4, 1089–1098.

[321] ZAYKIN, D. V., YOUNG, S. S., AND WESTFALL, P. H. (2000). Using the false discovery rate approach in the genetic dissection of complex traits: a response to Weller et al. *Genetics* **154**, 4, 1917–1918.

[322] ZAYKIN, D. V., ZHIVOTOVSKY, L. A., WESTFALL, P. H., AND WEIR, B. S. (2002). Truncated product method for combining P-values. *Genet. Epidemiol.* **22**, 2, 170–185.

[323] *ZEHETMAYER, S., BAUER, P., AND POSCH, M. (2005). Two-stage designs for experiments with a large number of hypotheses. *Bioinformatics* **21**, 19, 3771–3777.

[324] ZHENG, Y., CAI, T., AND FENG, Z. (2006). Application of the time-dependent ROC curves for prognosric accuracy with multiple biomarkers. *Biometrics.* **62**, 1, 279–287.

[325] ZHONG, S., TIAN, L., LI, C., STORCH, K.-F., AND WONG, W. H. (2004). Comparative analysis of gene sets in the Gene Ontology space under the multiple hypothesis testing framework. *Proc. IEEE Comput. Syst. Bioinform. Conf.*, 425–435.

[326] ZHONG, W., ZENG, P., MA, P., LIU, J. S., AND ZHU, Y. (2005). RSIR: regularized sliced inverse regression for motif discovery. *Bioinformatics* **21**, 22, 4169–4175.

[327] ZHOU, Y., CRAS-MENEUR, C., OHSUGI, M., STORMO, G. D., AND PERMUTT, M. (2007). A global approach to identify differentially expressed genes in cDNA (two-color) microarray experiments. *Bioinformatics*, btm292.

[328] ZHU, D., HERO, A. O., QIN, Z. S., AND SWAROOP, A. (2005). High throughput screening of co-expressed gene pairs with controlled false discovery rate (FDR) and minimum acceptable strength (MAS). *J. Comput. Biol.* **12**, 7, 1029–1045.

[329] ZOU, G. AND ZUO, Y. (2006). On the sample size requirement in genetic association tests when the proportion of false positives is controlled. *Genetics* **172**, 1, 687–691.

Abstract

The false discovery rate (FDR) is a rather young error control criterion in multiple testing problems. Initiated by the pioneering paper by Benjamini and Hochberg from 1995, it has become popular in the 1990ies as an alternative to the strong control of the family-wise error rate, especially if a large system of hypotheses is at hand and the analysis has mainly explorative character. Instead of controlling the probability of one or more false rejections, the FDR controls the expected proportion of falsely rejected hypotheses among all rejections. One typical application with strong impact on the development of the FDR is the first step (screening phase) of a microarray experiment where the experimenter aims at detecting a few candidate genes or SNPs potentially associated with a disease, which are than further analyzed using more stringent error handling methods. Especially due to such nowadays' applications with families of ten thousands or even some hundred thousands of hypotheses at hand, asymptotic considerations (with the number of hypotheses to be tested simultaneously tending to infinity) become more and more relevant.

In this work, the behavior of the FDR is mainly studied from a theoretical point of view. After some fundamental issues as a preparation in Chapter 1, focus is laid in Chapter 2 on the asymptotic behaviour of the linear step-up procedure originally introduced by Benjamini and Hochberg. Since it is well known that this procedure strongly controls the FDR under positive dependency, we investigate the asymptotic conservativeness of this procedure under various distributional settings in depth. The results imply that, depending on the strength of positive dependence among the test statistics and the proportion of true nulls, the FDR can be close to the pre-specified error level or can be very small. Typically, the latter case leads to low power of the linear step-up procedure which raises the possibility for improvements of the algorithm.

One improvement of Benjamini and Hochberg's procedure is presented and discussed in Chapter 3. Instead of using critical values increasing linearly (or, in other words, a linear rejection curve), we derive a non-linear and in some sense asymptotically optimal rejection curve leading to the full exhaustion of the FDR level under some extreme parameter configurations. This curve is then implemented into some stepwise multiple test procedures which control the FDR asymptotically or (with slight modifications) for a finite number of hypotheses. For the proof of FDR control for procedures employing non-linear critical values, some new methodology of proof is worked out. Chapter 4 then compares the newly derived methods with the original linear step-up procedure and other improved procedures with respect to multiple power. The results in this comparisons section are based on computer simulations. It turns out that certain procedures perform better in certain distributional setups or in other words that one can choose the appropriate FDR controlling algorithm to serve the purpose of detecting the most relevant alternatives most properly.

Besides all these theoretical and methodological topics, we are also concerned with some practical aspects of FDR. We apply FDR controlling procedures to real life data and illustrate the functionality, assets and drawbacks of the different methods using these data sets.

Kurzfassung

Die "False Discovery Rate" (FDR) ist ein recht junges Fehlerkontrollkriterium in multiplen Testproblemen. Beginnend mit dem Artikel von Benjamini und Hochberg aus dem Jahre 1995 wurde es in den 1990er Jahren als Alternative zur Kontrolle des multiplen Niveaus beliebt, insbesondere bei Vorliegen eines sehr mächtigen Hypothesensystems und vornehmlich explorativem Charakter der Analyse. Die FDR kontrolliert nicht die Wahrscheinlichkeit einer einzigen fälschlichen Verwerfung einer Nullhypothese, sondern den erwarteten Anteil fälscherlicherweise verworfener Hypothesen an allen Verwerfungen. Ein typisches Beispiel mit starkem Einfluss auf die Entwicklung der FDR ist der erste Schritt (die "Screeningphase") eines Microarray-Experimentes, in dem der Experimentator einige potenziell mit einer Erkrankung assoziierten Kandidatengene oder SNPs detektieren möchte, welche dann unter stringenterer statistischer Fehlerkontrolle weiter analysiert werden. Wegen aktueller Anwendungen mit aus mehreren zehntausenden oder gar einigen hunderttausenden simultan zu prüfender Hypothesen bestehender Familien gewinnen asymptotische Überlegungen (gegen unendlich strebende Hypothesenzahl) immer mehr an Relevanz.

In dieser Arbeit wird das Verhalten der FDR vornehmlich vom theoretischen Standpunkt aus untersucht. Nach einigen Vorüberlegungen in Kapitel 1 wird der Fokus in Kapitel 2 auf das asymptotische Verhalten der von Benjamini und Hochberg eingeführten linearen step-up Prozedur gelegt. Da bekannt ist, dass sie die FDR unter positiver Abhängigkeit kontrolliert, untersuchen wir, wie konservativ sich die Prozedur in entsprechenden Verteilungsmodellen asymptotisch verhält. Die Resultate zeigen, dass (je nach Grad der Abhängigkeit und Anteil wahrer Nullhypothesen) die FDR nahe dem vorgegebenen Niveau, aber auch sehr klein sein kann. Letzterer Fall hat eine geringe Güte der Prozedur zur Folge und eröffnet Raum für Verbesserungen des Algorithmus'.

Eine Verbesserung der Benjamini-Hochberg Prozedur wird in Kapitel 3 eingeführt und diskutiert. Anstatt linear wachsende kritische Werte (oder anders ausgedrückt eine Ablehngerade) zu benutzen, entwickeln wir eine nichtlineare und in gewissem Sinne asymptotisch optimale Ablehnkurve, um das FDR-Niveau unter extremen Modellannahmen ganz auszuschöpfen. Die Kurve dient zur Herleitung schrittweiser Tests, die die FDR asymptotisch oder (mit leichten Modifikationen) für eine finite Anzahl an Hypothesen kontrollieren. Zum Beweis der FDR-Kontrolle für Prozeduren, die auf nicht-linearen kritischen Werten basieren, wird eine neue Beweistechnik ausgearbeitet. Kapitel 4 vergleicht die neu entwickelten Methoden mit der ursprünglichen step-up Prozedur und anderen Verbesserungen hinsichtlich eines multiplen Gütemaßes. Die Aussagen dieser Vergleichsstudie basieren auf Computersimulationen. Es zeigt sich, dass bestimmte Tests unter gewissen Verteilungsannahmen Vorteile besitzen bzw. ein geeignetes FDR- kontrollierendes Verfahren ausgewählt werden kann, um gewisse Alternativen bestmöglich zu erkennen.

Neben diesen theoretisch-methodischen Aspekten beschäftigen sich einige Anwendungsbeispiele auch mit der praktischen Seite der FDR. Wir wenden FDR-kontrollierende Prozeduren auf Realdaten an und diskutieren Funktionsweise sowie Vor- und Nachteile der jeweiligen Testprozeduren anhand dieser Datensätze.

Danksagungen

Danken möchte ich zu allererst meinem Betreuer Helmut Finner für sein Vertrauen in mich, für die interessante Themenstellung und für die vielen fruchtbaren (und manchmal auch kontroversen) Diskussionen. Seine Tür stand ohne Einschränkung stets für mich offen und von ihm habe ich viele Dinge über Statistik gelernt, von denen ich vorher noch nicht einmal wusste, dass es sie gibt.

Herrn Markus Roters möchte ich für die Nachhilfestunden in Stochastik, seine vielen handschriftlich angefertigten Fax-Notizen und seine große Geduld und Nachsicht mit meinem manchmal etwas aufbrausenden Gemüt danken. Auch danke ich Herrn Arnold Janssen für die Übernahme des Erstgutachtens und die vielen guten Lehrveranstaltungen, die ich bei ihm miterleben durfte.

Des Weiteren danke ich meinen KollegInnen Silke Bruckhaus, Guido Giani, Veronika Gontscharuk, Burkhard Haastert, Paul Hamann-Hensell, Andrea Icks, Norman Kluthausen, Wolfgang Rathmann, Joachim Rosenbauer, Barbara Schön, Klaus Straßburger, Anni Wachowiak, Bernd Wiemann und Jürgen Willems im Institut für Biometrie und Epidemiologie des Deutschen Diabetes-Zentrums Düsseldorf für die angenehme Arbeitsatmosphäre, den warmherzigen, hilfsbereiten Umgang und viele anregende Gespräche sowie der Deutschen Forschungsgemeinschaft, die weite Teile der in dieser Arbeit dargestellten Forschungsarbeiten gefördert hat.

Mein besonderer Dank gilt meiner Familie, die mich auf meinem bisherigen Lebens- und Berufsweg begleitet und unterstützt hat: Meinem Vater Hans Joachim, dem wir für immer ein liebevoll ehrendes Andenken bewahren werden, meiner Mutter Ursula, die auch in schweren Zeiten immer froh und optimistisch alles Menschenmögliche für das Gelingen unseres Zusammenlebens eingesetzt hat und meinem Bruder Sven, auf den ich genauso stolz bin wie er auf mich.

Schließlich lobpreise ich Gott meinen Schöpfer, unseren Vater im Himmel, und seinen eingeborenen Sohn, unseren Herrn und Erretter Jesus Christus und erweise mich dankbar dafür, dass es mich gibt und dass ich in meinem Leben einen Sinn sehen kann.
"Wenn Gott für uns ist, wer gegen uns?" (Römer 8, 31)

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 16. Januar 2008

Thorsten-Ingo Dickhaus