Implicitly adaptive FDR control based on the asymptotically optimal rejection curve

Thorsten Dickhaus  
(joint work with H. Finner and V. Gontscharuk)

Department of Mathematics  
Humboldt-University Berlin

Seminar on the theory of multiple comparisons  
HU Berlin, 01.11.2010
Outline

Multiple Testing and the False Discovery Rate

Explicitly adaptive FDR control

Asymptotically Optimal Rejection Curve (AORC)
Multiple statistical decision problems

- Multiple comparisons (multiple tests)
- Simultaneous confidence regions
- Multiple power / sample size calculation
- Selection problems
- Ranking problems
- Partitioning problems
Prologue

We assume a statistical model (statistical experiment) 
\( (\Omega, \mathcal{F}, (P_{\vartheta})_{\vartheta \in \Theta}) \)

More concrete scenario (chosen for exemplary purposes):

Balanced ANOVA1 model:

\[
X = (X_{ij})_{i=1,...,k, j=1,...,n}, \quad X_{ij} \sim \mathcal{N}(\mu_i, \sigma^2), \\
\]

\( X_{ij} \) stochastically independent random variables on \( \mathbb{R} \),

\( \mu_i \in \mathbb{R} \ \forall 1 \leq i \leq k, \ \sigma^2 > 0 \) (known or unknown) variance

\( k \geq 3, n \geq 2, \ \nu = k(n - 1) \) (degrees of freedom)

\[
\Omega = \mathbb{R}^{k \cdot n}, \ \mathcal{F} = \mathcal{B}^{k \cdot n} \\
\vartheta = (\mu_1, \ldots, \mu_k, \sigma^2) \in \mathbb{R}^k \times [0, \infty) = \Theta
\]
Multiple comparisons (multiple tests)

\[(\Omega, \mathcal{F}, (\mathbb{P}_\vartheta)_{\vartheta \in \Theta})\]

\[\mathcal{H}_m = (H_i)_{i=1, \ldots, m}\]

Family of null hypotheses with \(\emptyset \neq H_i \subset \Theta\) and alternatives \(K_i = \Theta \setminus H_i\)

\[(\Omega, \mathcal{F}, (\mathbb{P}_\vartheta)_{\vartheta \in \Theta}, \mathcal{H}_m)\]

Multiple testing problem

\[\varphi = (\varphi_i : i = 1, \ldots, m)\]

multiple test for \(\mathcal{H}_m\)

\[
\begin{array}{c|cc}
\text{Hypotheses} & 0 & 1 \\
\hline
\text{true} & U_m & V_m \\
\text{false} & T_m & S_m \\
m - R_m & R_m & m \\
\end{array}
\]
Multiple tests (II)

Wanted features of $\phi$:

No (or only minor) contradictions among the individual test decisions

Control of the probability of erroneous decisions

(Type I) error measures / concepts:

\[
\text{FWER}_m(\phi) = \mathbb{P}_\theta (V_m > 0) \leq \alpha \quad \forall \theta \in \Theta
\]

(Strong) control of the Family-Wise Error Rate (FWER)

\[
\text{FDR}_m(\phi) = \mathbb{E}_\theta \left[ \frac{V_m}{R_m} \right] \leq \alpha \quad \forall \theta \in \Theta
\]

Control of the False Discovery Rate (FDR)
FWER control

FWER-controlling multiple testing principles:

- Historical single-step tests
  (Bonferroni, Sidak, Tukey, Scheffé, Dunnett)
- Holm’s (1979) stepwise rejective procedure
- Closed test principle (Marcus, Peritz, Gabriel (1969))
- Intersection-union principle (generalized closed testing)
- Partitioning principle
  (Finner and Straßburger 2001, Hsu 1996)
- Projection methods under asymptotic normality
  (Bretz, Hothorn and Westfall)
- Resampling-based multiple testing
  (Westfall / Young 1992, Dudoit / van der Laan 2008)
Multiplicity of current applications

Due to rapid technical developments in many scientific fields, the number $m$ of hypotheses to be tested simultaneously can nowadays become almost arbitrary large:

- Genetics, microarrays: $m \sim 30000$ genes / hypotheses
- Genetics, SNPs: $m \sim 500000$ SNPs / hypotheses
- Proteomics: $m \sim 5000$ proteine spots per gele sheet
- Cosmology: Signal detection, $m \sim 10^6$ pixels / hypotheses
- Neurology: Identification of active brain loci, $m \sim 10^3$ voxels
- Biometry: pairwise comparisons of many means, tests for correlations

Analyses have (in a first step) typically explorative character.

\[\longrightarrow\] Control of the FWER much too conservative goal!
Definition of the False Discovery Rate

\( \Theta \)

Null hypotheses

\( H_1, \ldots, H_m \)

Multiple test procedure

\( \varphi = (\varphi_1, \ldots, \varphi_m) \)

Number of falsely rejected, true nulls

\( V_m = |\{ i : \varphi_i = 1 \text{ and } H_i \text{ true } \}| \)

Total number of rejections

\( R_m = |\{ i : \varphi_i = 1 \}| \)

False Discovery Rate (FDR) given \( \vartheta \in \Theta \)

\[
\text{FDR}_\vartheta(\varphi) = \mathbb{E}_\vartheta\left[ \frac{V_m}{R_m \vee 1} \right]
\]

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

The multiple test \( \varphi \) controls the FDR at level \( \alpha \) if

\[
\text{FDR}(\varphi) = \sup_{\vartheta \in \Theta} \text{FDR}_\vartheta(\varphi) \leq \alpha.
\]
The classical FDR theorem
Benjamini and Hochberg (1995)

\( p_1, \ldots, p_m \) (marginal) \( p \)-values for hypotheses \( H_1, \ldots, H_m \)

\( H_i \) true for \( i \in I_{m,0} \), \( H_i \) false for \( i \in I_{m,1} \)

\( I_{m,0} + I_{m,1} = \mathbb{N}_m = \{1, \ldots, m\} \), \( m_0 = |I_{m,0}| \)

\( p_i \sim \text{UNI}[0, 1], \ i \in I_{m,0} \) \quad \text{stochastically independent} \quad (I1)

\( (p_i : i \in I_{m,0}), (p_i : i \in I_{m,1}) \) \quad \text{stoch. independent vectors} \quad (I2)

\( p_{1:m} \leq \cdots \leq p_{m:m} \) \quad \text{ordered } \( p \)-values

Linear step-up procedure \( \varphi^{LSU} \) with Simes’ crit. values \( \alpha_{i:m} = i\alpha/m \):

Reject all \( H_i \) with \( p_i \leq \alpha_{\overline{m}:m} \), where \( \overline{m} = \max\{j : p_{j:m} \leq j\alpha/m\} \).

Then it holds:

\[
\text{FDR}_\theta(\varphi^{LSU}) = \mathbb{E}_\theta \left[ \frac{V_m}{R_m \vee 1} \right] = \frac{m_0}{m} \alpha \quad \forall \theta \in \Theta.
\]
Linear step-up in terms of Simes’ rejection line

\[ R_n/n \]

\[ t \quad t^* \]
FDR control under positive dependency


Proofs for FDR control in presence of special dependency structures:

\[ \text{FDR}_\vartheta(\varphi) \leq \frac{m_0}{m} \alpha \quad \forall \vartheta \in \Theta \]

for stepwise test procedures employing Simes’ critical values.

Model assumptions: \textbf{MTP}_2 \textit{oder} PRDS

Examples:
Multivariate normal distributions with non-negative correlations, multivariate (absolute) \textit{t}-distributions

Explicit adaptation

Since under positive dependency the FDR of the LSU-procedure is bounded by

\[ \frac{m_0}{m} \alpha \]

for any \( m > 1 \) and given \( \alpha \in (0, 1) \), \( \varphi_{\text{LSU}} \) does not exhaust the FDR level \( \alpha \) in case of \( m_0 < m \).

Many modern FDR-controlling methods:

Pre-estimation of \( m_0 \) or \( \pi_0 = m_0/m \) aiming at tighter \( \alpha \)-exhaustion and gain of power

(Explicit adaptation)
Empirical stochastic processes

Interpret the number of rejections of (true / false) null hypotheses and the FDR of a single-step test with threshold $t \in [0, 1]$ for the $p$-values as empirical processes in $t$:

$$V(t) = \sum_{i \in I_{m,0}} 1_{[0,t]}(p_i),$$

$$S(t) = \sum_{i \in I_{m,1}} 1_{[0,t]}(p_i),$$

$$R(t) = V(t) + S(t) = \sum_{i=1}^{m} 1_{[0,t]}(p_i),$$

$$\text{FDR}(t) = \mathbb{E} \left[ \frac{V(t)}{R(t) \lor 1} \right].$$
The adaptive procedure of Storey, Taylor and Siegmund (2004)

For a tuning parameter $\lambda \in (0, 1)$ is the following estimator $\hat{\pi}_0$ of $\pi_0$ reasonable (Schweder and Spjøtvoll (1982)):

$$
\hat{\pi}_0 \equiv \hat{\pi}_0(\lambda) = \frac{m - R(\lambda) + 1}{m(1 - \lambda)} = \frac{1 - \hat{F}_m(\lambda) + 1/m}{1 - \lambda}
$$

Under (I1) and (I2), we additionally obtain an estimator for the FDR of a single-step test with threshold $t \in [0, 1]$:

$$
\widehat{\text{FDR}}_{\lambda}(t) = \frac{\hat{\pi}_0(\lambda)t}{(R(t) \lor 1)/m}
$$

Resulting adaptive thresholding:

$$
t_{\alpha}^{\lambda} = \sup\{0 \leq t \leq \lambda : \widehat{\text{FDR}}_{\lambda}(t) \leq \alpha\}
The estimator $\hat{\pi}_0$

Schweder and Spjøtvoll (1982)

\[
\hat{F}_k(t) = \# \{ p_i \leq t \} / k
\]
The estimator $\hat{\pi}_0$

Schweder and Spjøtvoll (1982)

$$\hat{F}_k(t) = \frac{\# \{ p_i \leq t \}}{k}$$
Lemma: (Storey, Taylor and Siegmund (2004))

Assuming stochastically independent $p$-values under the $m_0$ null hypotheses, $V(t)/t$ is for $0 \leq t < 1$ a reverse martingal with respect to the filtration $\mathcal{F}_t = \sigma(1_{[0,s]}(p_i), t \leq s \leq 1, i = 1, \ldots, m)$, i.e., for $s \leq t$ it holds $\mathbb{E}[V(s)/s | \mathcal{F}_t] = V(t)/t$.

The random threshold $t^\lambda_\alpha$ is a stopping time with respect to $\mathcal{F}_{t^\land \lambda}$.

$\implies$ FDR-proofs utilizing theory of optimal stopping
Proof of FDR-control of the adaptive test procedure by Storey et al.

Assumption: $\lambda$ chosen such that $\hat{\text{FDR}}_\lambda(\lambda) \geq \alpha$.

It follows: $\hat{\text{FDR}}_\lambda(t^\lambda_\alpha) = \alpha \iff R(t^\lambda_\alpha) = m\hat{\pi}_0(\lambda)t^\lambda_\alpha/\alpha$.

Moreover, the process $V(t)/t$ stoppt at $t^\lambda_\alpha$ is bounded and

$$\text{FDR}(t^\lambda_\alpha) = \mathbb{E} \left[ \frac{V(t^\lambda_\alpha)}{R(t^\lambda_\alpha)} \right] = \mathbb{E} \left[ \frac{\alpha}{m\hat{\pi}_0(\lambda)} \frac{V(t^\lambda_\alpha)}{t^\lambda_\alpha} \right]$$

$$= \mathbb{E} \left[ \frac{1 - \lambda}{m - R(\lambda) + 1} \frac{V(t^\lambda_\alpha)}{t^\lambda_\alpha} \right]$$

$$= \mathbb{E} \left[ \frac{1 - \lambda}{m - R(\lambda) + 1} \mathbb{E} \left[ \frac{V(t^\lambda_\alpha)}{t^\lambda_\alpha} | \mathcal{F}_\lambda \right] \right]$$

$$= \mathbb{E} \left[ \frac{1 - \lambda}{m - R(\lambda) + 1} \frac{V(\lambda)}{\lambda} \right].$$
Noticing $V(\lambda) \sim \text{Bin}(n_0, \lambda)$, yields:

\[
\text{FDR}(t_\alpha^\lambda) = \mathbb{E} \left[ \alpha \frac{1 - \lambda}{m - R(\lambda) + 1} \frac{V(\lambda)}{\lambda} \right]
\leq \mathbb{E} \left[ \alpha \frac{1 - \lambda}{m_0 - V(\lambda) + 1} \frac{V(\lambda)}{\lambda} \right]
= \sum_{k=0}^{m_0} \alpha \frac{1 - \lambda}{m_0 - k + 1} \frac{k}{\lambda} \cdot \mathbb{P}(V(\lambda) = k)
= \alpha \frac{1 - \lambda}{\lambda} \sum_{k=0}^{m_0} \frac{k}{m_0 - k + 1} \binom{m_0}{k} \lambda^k (1 - \lambda)^{m_0-k}
= \alpha \frac{1 - \lambda}{\lambda} \cdot \frac{\lambda - \lambda^{m_0+1}}{1 - \lambda} = \alpha (1 - \lambda^{m_0})
\leq \alpha \text{ for all } \lambda \in (0, 1) \text{ and } m_0 \geq 0. \quad \square
Implicit adaptation

Alternatively to explicit adaptation, it may be asked:

Is it possible to derive a better rejection curve circumventing the factor $m_0/m$ appearing in the FDR of $\varphi^{\text{LSU}}$?

First step:
Identification of least favorable parameter configurations (LFCs) for the FDR.
Dirac-uniform models as LFCs

Theorem: (Benjamini & Yekutieli (2001))
If \( p_i \sim U([0, 1]), \ i \in I_{m,0}, \) stochastically independent (I1) and
\((p_i : i \in I_{m,0}), (p_i : i \in I_{m,1})\) stoch. independent vectors (I2),
then a step-up procedure \( \varphi_{(m)}^{SU} \) with critical values \( \alpha_{1:m} \leq \cdots \leq \alpha_{m:m} \)
has the following properties: If

\[
\alpha_{i:m}/i \text{ is increasing (decreasing) in } i
\]

and the distribution of \((p_i : i \in I_{m,1})\) decreases stochastically, then the FDR of \( \varphi_{(m)}^{SU} \) increases (decreases).

If \( \alpha_{i:m}/i \text{ is increasing in } i \), it follows that the FDR becomes largest for \( p_i \sim \delta_0 \ \forall i \in I_{m,1} \) (Dirac-uniform model).

In DU-models, **analytic calculations** are possible!
Asymptotic Dirac-uniform model: $DU(\zeta)$

**Assumptions:**

Independent $p$-values $p_1, \ldots, p_m$;

$m_0 = m_0(m)$ null hypotheses true with

$$\lim_{m \to \infty} \frac{m_0(m)}{m} = \zeta \in (0, 1],$$

$m_0$ $p$-values UNI([0, 1])-distributed (corresp. hypotheses true)

$m_1 = m - m_0$ $p$-values $\delta_0$-distributed (corresp. hypotheses false)

Then the ecdf of the $p$-values $F_m$ (say) converges (Glivenko-Cantelli) for $m \to \infty$ to

$$G_{\zeta}(x) = (1 - \zeta) + \zeta x \text{ for all } x \in [0, 1].$$
Heuristic for an asymptotically optimal rejection curve

Assume we reject all $H_i$ with $p_i \leq x$ for some $x \in (0, 1)$. Then the FDR (depending on $\zeta$ and $x$) under DU($\zeta$) is asymptotically given by

$$FDR_\zeta(x) = \frac{\zeta x}{(1 - \zeta) + \zeta x}.$$ 

**Aim:** Find an optimal threshold $x_\zeta$ (say), such that

$$FDR \equiv \alpha \text{ for all } \zeta \in (\alpha, 1).$$

We obtain:

$$FDR_\zeta(x_\zeta) = \alpha \iff x_\zeta = \frac{\alpha(1 - \zeta)}{\zeta(1 - \alpha)}.$$
Asymptotically optimal rejection curve

**Ansatz:** Rejection curve $f_{\alpha}$ and $G_{\zeta}$ shall cross each other in $x_{\zeta}$, i.e., $f_{\alpha}(x_{\zeta}) = G_{\zeta}(x_{\zeta})$.

Plugging in $x_{\zeta}$ derived above yields

$$f_{\alpha} \left( \frac{\alpha(1 - \zeta)}{\zeta(1 - \alpha)} \right) = \frac{1 - \zeta}{1 - \alpha}.$$

Substituting $t = \frac{\alpha(1 - \zeta)}{\zeta(1 - \alpha)} \iff \zeta = \frac{\alpha}{(1 - \alpha)t + \alpha}$,

we get that

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

is the curve solving the problem!
Asymptotically optimal rejection curve for $\alpha = 0.1$
Critical values, step-up-down procedure

The critical values induced by $f_\alpha$ are given by

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

Due to $\alpha_{m:m} = 1$, a step-up procedure based on $f_\alpha$ cannot work.

One possible solution:

**Step-up-down procedure** with parameter $\lambda \in (0, 1)$:

- $F_m(\lambda) \geq f_\alpha(\lambda) \Rightarrow t^* = \inf\{p_i > \lambda : F_m(p_i) < f_\alpha(p_i)\}$ (SD-branch),

- $F_m(\lambda) < f_\alpha(\lambda) \Rightarrow t^{**} = \sup\{p_i < \lambda : F_m(p_i) \geq f_\alpha(p_i)\}$ (SU-branch).

Reject all $H_i$ with $p_i < t^*$ or $p_i \leq t^{**}$, respectively.
SUD-procedure for $\lambda = 0.3, 0.6$

$(m = 50, \alpha = 0.1)$
Refined results for SUD tests

Finner, Dickhaus, Roters (2009), Annals of Statistics 37, 596-618

- AORC-based stepwise test procedures asymptotically keep the FDR level under (I1) und (I2).
- In the class of stepwise test procedures with fixed rejection curve asymptotically keeping the FDR level, AORC-based tests have largest power for $\zeta \in (\alpha, 1)$.

Technical results:

- New methodology of proof for stepwise test procedures with non-linear rejection curves resp. critical values
- Upper FDR bounds for step-up-down tests (asymptotically and finite)
SU-test with modified version of $f_\alpha$, finite case

$(m = 100, 500, 1000, \alpha = 0.05)$

For $m = 100$, maximum DU FDR is $FDR_{16,100} \approx 0.05801$.

$\Rightarrow$ Adjustment of critical values for finite cases necessary!
FDR control for step-up implies FDR control for step-up-down

Theorem:
Consider a SU test $\varphi^m$ and a SUD($\lambda$) test $\varphi^\lambda$ for $\lambda \in \{1, \ldots, m - 1\}$ with the same set of critical values $(\alpha_{i:m})_{i=1}^m$ belonging to $\mathcal{M}_m = \{(c_{i:m})_{i=1}^m : 0 \leq c_{1:m} \leq \ldots \leq c_{m:m} \leq 1, c_{i:m}/i \text{ increases in } i\}$.
Then, under (I1) and (I2), it holds

$$\text{FDR}_\vartheta(\varphi^\lambda) \leq \text{FDR}_\vartheta(\varphi^m) \text{ for all } \vartheta \in \Theta.$$ 

Hence, if the FDR is controlled by the SU test, then the SUD($\lambda$) test also controls the FDR.

**Sketch of Proof:** $\{R^\lambda_m \geq j, p_{i_0} \leq \alpha_{j:m}\} \subseteq \{R^\lambda_2 \geq j, p_{i_0} \leq \alpha_{j:m}\}$ for any $1 \leq \lambda_1 \leq \lambda_2 \leq m$, which implies that $P_\vartheta(R^\lambda_m \geq j | p_{i_0} \leq \alpha_{j:m})$ is non-decreasing in $\lambda$ for each $j \in \mathbb{N}_m$. 
Exact finite adjustment (for step-up)

(Slight) modification of $f_\alpha$ or its critical values, e.g.

$$\alpha_{i:m} = \frac{i\alpha}{m + \beta_m - i(1 - \alpha)}, \quad i = 1, \ldots, m,$$

for a suitable adjustment constant $\beta_m > 0$.

(Same as: Use $\tilde{f}_\alpha(t) = (1 + \beta_m/m) f_\alpha(t), \quad t \in [0, \alpha/(\alpha + \beta/m)]$.)

$m = 100$ leads to $\beta_{100} \approx 1.76$.

Ray of light:


SD-procedure with $\beta_m \equiv 1.0$ controls FDR for every $m \in \mathbb{N}$.
Adjustment with three parameters

For $i = 1, \ldots, m$, utilize adjusted critical values of the form

$$\alpha_{i:m} = \frac{i\alpha}{m + \beta_{1,m} + \beta_{2,m} \left( \frac{i}{m} \right)^{\beta_{3,m}} - i(1 - \alpha)}$$

Resulting FDRs for adjusted critical values with one (black) and three parameters, respectively, for $m = 100, 200, 400$, depending on $m_0$. 
Iterative method

Let $J \in \mathbb{N}$ be fixed and $\alpha_{1:m}^{(0)}, \ldots, \alpha_{m:m}^{(0)} \in \mathcal{M}_m$ be start critical values, for instance adjusted AORC-based critical values, fulfilling that $FDR_{m,m_0} \approx \alpha$ for all $m_0 \geq k$.

Now, try to iteratively modify crit. values most influencing $FDR_{m,m_0}$ for $k \leq m_0 \leq m$ to reduce $d(m_0, m) = \alpha - FDR_{m,m_0}$.

One possible iteration scheme:

For $j$ from 1 to $J$ do:
  For $i$ from 1 to $i^*(k)$ do:
    1. Determine $\alpha_{i}^{(j-1)}$ from $\alpha_{i:m}^{(j-1)} = i\alpha_{i}^{(j-1)}/(m - i(1 - \alpha_{i}^{(j-1)}))$.
    2. Put $\alpha_{i}^{j} = \alpha\alpha_{i}^{(j-1)}/FDR_{m,m_0}(i)(\alpha_{i}^{(j-1)})$.

Motivation: Fixed-point iteration for

\[ f(\alpha_i) = \alpha\alpha_i/FDR_{m,m_0}(i)(\alpha_i). \]
Stepwise finding

Subsequently solve (for $\alpha_{m_1+1:m}$) the target equations

$$FDR_{m,m_0}(\alpha_{m_1+1:m}, \ldots, \alpha_{m:m}) = \alpha, \quad m_0 \in \mathbb{N}_m. \quad (2)$$

$m_0 = 1 \Rightarrow \alpha_{m:m} = m\alpha$ for a fixed $\alpha$, because $FDR_{m,1} = \alpha_{m:m}/m$.

It follows that $\alpha_{m:m} \geq 1$ for each $m \geq 1/\alpha$, which is unacceptable. We can try to fix that by replacing (2) by

$$FDR_{m,m_0}(\alpha_{m_1+1:m}, \ldots, \alpha_{m:m}) = \min \left( \frac{m_0}{m}, \alpha \right), \quad m_0 \in \mathbb{N}_m. \quad (3)$$

If even (3) cannot be solved in $\mathcal{M}_m$, we further generalize the right-hand-side and solve

$$FDR_{m,m_0}(\alpha_{m_1+1:m}, \ldots, \alpha_{m:m}) = g(\alpha, \zeta_m), \quad m_0 \in \mathbb{N}_m.$$
Flexible FDR-control

Candidates for $g(\alpha, \zeta_m)$:

$$g(\zeta|\gamma, \eta) = \begin{cases} \alpha(1 - (1 - \zeta/\gamma)^\eta), & 0 \leq \zeta < \gamma, \\ \alpha, & \gamma \leq \zeta \leq 1. \end{cases}$$

$g(\zeta|\gamma, \eta)$ for $\alpha = 0.05$ and $\gamma = 1$, $\eta = 20.0, 18.0, 16.0$ together with $\min(\alpha, \zeta)$ for $\zeta$ ranging from 0 to 0.3.
Acknowledgments:

- Helmut Finner
- Veronika Gontscharuk
- Markus Roters

References: