New adaptive procedures that control the FDR

Speaker : **Etienne Roquain**

(MIG - INRA, Jouy-en-Josas France)

Joint work with **Gilles Blanchard**

(Fraunhofer FIRST.IDA, Berlin Germany)

Setting

- $(\mathcal{X}, \mathfrak{X}, \mathbf{P})$ probability space.
- $\mathcal{H}$ a finite set of hypotheses for $\mathbf{P}$, $m := |\mathcal{H}|$ (known)

Setting

- $(\mathcal{X}, \mathfrak{X}, \mathbf{P})$ probability space.
- $\mathcal{H}$ a finite set of hypotheses for $\mathbf{P}$, $m := |\mathcal{H}|$ (known)
  - $\mathcal{H}_0$ set of true hypotheses
  - $m_0 := |\mathcal{H}_0|$, $\pi_0 := m_0/m$ (unknown)

Setting

- $(\mathcal{X}, \mathfrak{X}, \mathbf{P})$ probability space.
- $\mathcal{H}$ a finite set of hypotheses for $\mathbf{P}$, $m := |\mathcal{H}|$ (known)
  - $\mathcal{H}_0$ set of true hypotheses
  - $m_0 := |\mathcal{H}_0|$, $\pi_0 := m_0/m$ (unknown)
- For each $h \in \mathcal{H}$, p-value: $p_h : \mathcal{X} \to [0,1]$ (measurable) such that

$$\text{If } h \in \mathcal{H}_0,\ \ \forall t \in [0,1], \mathbf{P}(p_h \le t) \le t$$

$$\text{If } h \notin \mathcal{H}_0, \mathcal{D}(p_h) \text{ unspecified}$$

Setting

- $(\mathcal{X}, \mathfrak{X}, \mathbf{P})$ probability space.
- $\mathcal{H}$ a finite set of hypotheses for $\mathbf{P}$, $m := |\mathcal{H}|$ (known)
  - $\mathcal{H}_0$ set of true hypotheses
  - $m_0 := |\mathcal{H}_0|$, $\pi_0 := m_0/m$ (unknown)
- For each $h \in \mathcal{H}$, p-value: $p_h : \mathcal{X} \to [0,1]$ (measurable) such that

$$\text{If } h \in \mathcal{H}_0, \ \ \forall t \in [0,1], \mathbf{P}(p_h \leq t) \leq t$$

$$\text{If } h \notin \mathcal{H}_0, \mathcal{D}(p_h) \text{ unspecified}$$

- A multiple testing procedure: a (measurable) function

$$R : \mathbf{p} = (p_h)_{h \in \mathcal{H}} \in [0,1]^{\mathcal{H}} \mapsto R(\mathbf{p}) \subset \mathcal{H}$$

(returns the rejected hypotheses)

Type I error

R makes a type I error for h:

$$h \in \mathcal{H}_0 \cap R$$

Type I error

R makes a type I error for h:

$$h \in \mathcal{H}_0 \cap R$$

False Discovery Rate of R [Benjamini and Hochberg (1995)]:

$$\mathrm{FDR}(R) := \mathbf{E}\left[\frac{|\mathcal{H}_0 \cap R|}{|R|}\mathbb{1}\{|R| > 0\}\right]$$

Type I error

R makes a type I error for h:

$$h \in \mathcal{H}_0 \cap R$$

False Discovery Rate of R [Benjamini and Hochberg (1995)]:

$$\mathrm{FDR}(R) := \mathbf{E}\left[\frac{|\mathcal{H}_0 \cap R|}{|R|}\mathbb{1}\{|R| > 0\}\right]$$

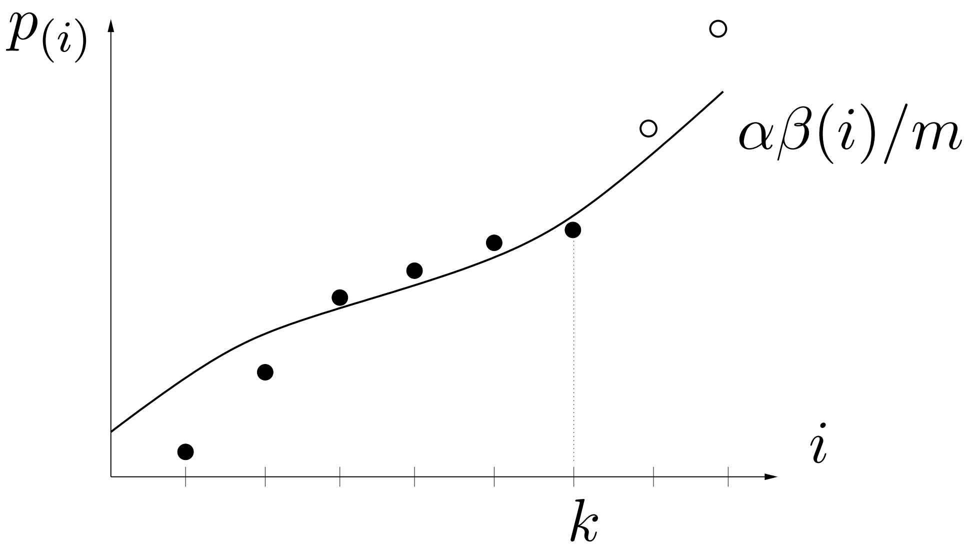
How to build R with $\mathrm{FDR}(R) \leq \alpha$?

Step-up procedure

If $p_{(1)} \leq \cdots \leq p_{(m)}$ are the ordered p-values
and $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ non-decreasing: threshold function

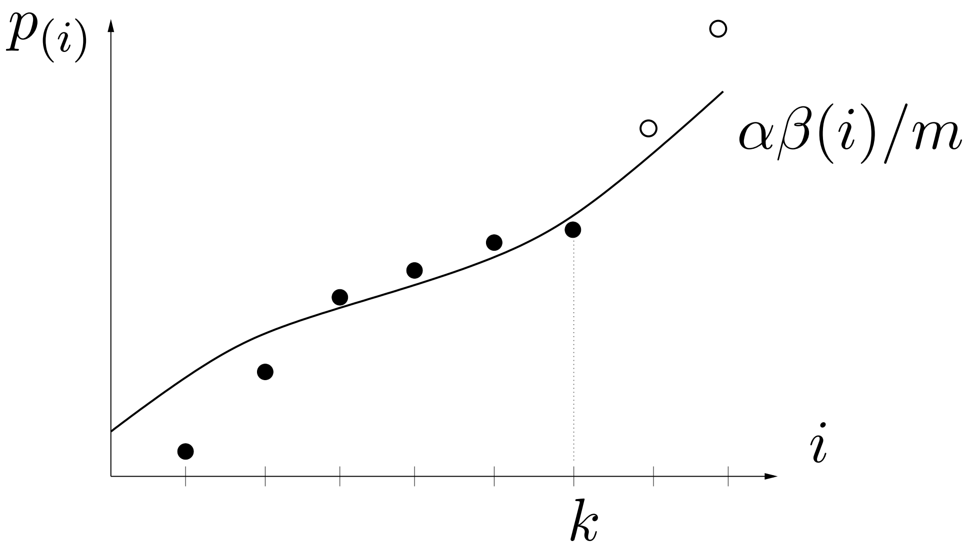
Step-up procedure

If $p_{(1)} \leq \cdots \leq p_{(m)}$ are the ordered p-values
and $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ non-decreasing: threshold function



Step-up procedure

If $p_{(1)} \leq \cdots \leq p_{(m)}$ are the ordered p-values
and $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ non-decreasing: threshold function

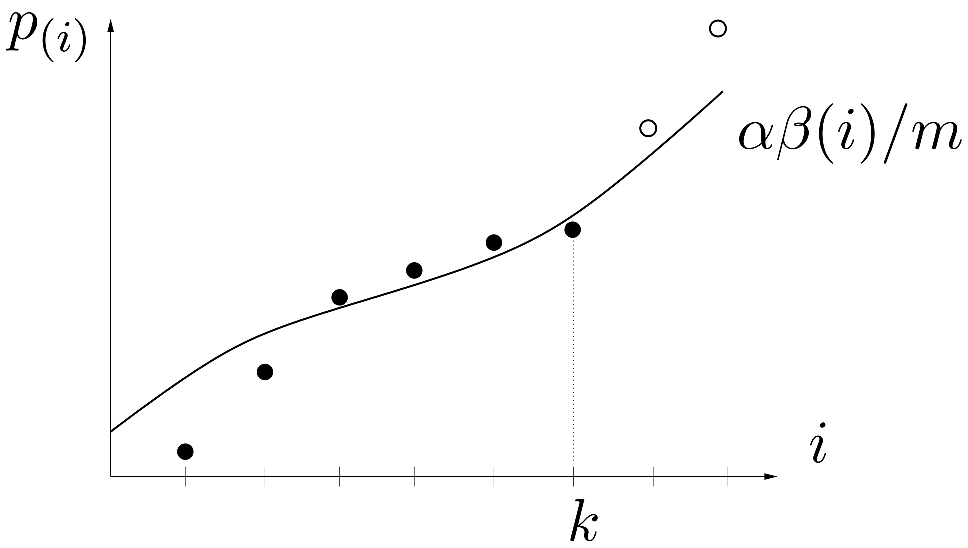


Definition (step-up procedure with threshold function β)

$$R_\beta := \{h \in \mathcal{H} |\, p_h \leq p_{(k)}\} \quad \text{where} \quad k := \max\{i |\, p_{(i)} \leq \alpha\beta(i)/m\}$$

Step-up procedure

If $p_{(1)} \leq \cdots \leq p_{(m)}$ are the ordered p-values
and $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ non-decreasing: threshold function



Definition (step-up procedure with threshold function β)

$$R_\beta := \{h \in \mathcal{H} | \, p_h \leq p_{(k)}\} \quad \text{where} \quad k := \max\{i | \, p_{(i)} \leq \alpha\beta(i)/m\}$$

Linear step-up procedure: R_β where $\beta(i) = i$.

Goal in step-up FDR control

Find a threshold function β such that

$$\mathrm{FDR}(R_\beta) \le \alpha$$

Goal in step-up FDR control

Find a threshold function β such that

$$\mathrm{FDR}(R_\beta) \leq \alpha$$

With β as "large" as possible

Known FDR control results

Recall that $\pi_0 = m_0/m$.

Known FDR control results

Recall that $\pi_0 = m_0/m$.

Benjamini and Hochberg (1995), Benjamini and Yekutieli (2001):

Theorem 1 p-values independent or positively dependent (PRDS):
The linear step-up procedure has a FDR smaller than $\pi_0\alpha$, i.e. for $\beta(i) = i$,

$$\text{FDR}(R_\beta) \leq \pi_0\alpha.$$

Known FDR control results

Recall that $\pi_0 = m_0/m$.

Benjamini and Hochberg (1995), Benjamini and Yekutieli (2001):

Theorem 1 p-values independent or positively dependent (PRDS):
The linear step-up procedure has a FDR smaller than $\pi_0\alpha$, i.e. for $\beta(i) = i$,

$$\mathrm{FDR}(R_\beta) \leq \pi_0\alpha.$$

Benjamini and Yekutieli (2001):

Theorem 2 p-values with general dependencies:
The step-up procedure with threshold function $\beta(i) = i/(1 + 1/2 + \cdots + 1/m)$ satisfies

$$\mathrm{FDR}(R_\beta) \leq \pi_0\alpha.$$

Known FDR control results (2)

Generalization of Blanchard and Fleuret (2007):

Known FDR control results (2)

Generalization of Blanchard and Fleuret (2007):

Theorem 3 p-values with general dependencies:
If β of the form:

$$\beta(i) = \int_0^i u d\nu(u),$$

and ν is some distribution on $(0, \infty)$,

$$\mathrm{FDR}(R_\beta) \leq \pi_0 \alpha.$$

Known FDR control results (2)

Generalization of Blanchard and Fleuret (2007):

Theorem 3 p-values with general dependencies:
If β of the form:

$$\beta(i) = \int_0^i u d\nu(u),$$

and ν is some distribution on $(0, \infty)$,

$$\mathrm{FDR}(R_\beta) \leq \pi_0 \alpha.$$

Remarks : ν prior idea on the final number of rejections

Known FDR control results (2)

Generalization of Blanchard and Fleuret (2007):

Theorem 3 p-values with general dependencies:
If β of the form:

$$\beta(i) = \int_0^i u d\nu(u),$$

and ν is some distribution on $(0, \infty)$,

$$\mathrm{FDR}(R_\beta) \leq \pi_0 \alpha.$$

Remarks : ν prior idea on the final number of rejections

- ν uniform on $\{1, \dots, m\}$ gives $\beta(i) = i(i+1)/(2m)$

Known FDR control results (2)

Generalization of Blanchard and Fleuret (2007):

Theorem 3 *p-values with general dependencies*:
If β of the form:

$$\beta(i) = \int_0^i u d\nu(u),$$

and ν is some distribution on $(0, \infty)$,

$$\mathrm{FDR}(R_\beta) \leq \pi_0 \alpha.$$

Remarks : ν prior idea on the final number of rejections

- ν uniform on $\{1, \ldots, m\}$ gives $\beta(i) = i(i+1)/(2m)$
- $\nu(\{k\}) \propto 1/k$ gives $\beta(i) = i/(1 + 1/2 + \cdots + 1/m)$

Known FDR control results (2)

Generalization of Blanchard and Fleuret (2007):

Theorem 3 p-values with general dependencies:
If β of the form:

$$\beta(i) = \int_0^i u d\nu(u),$$

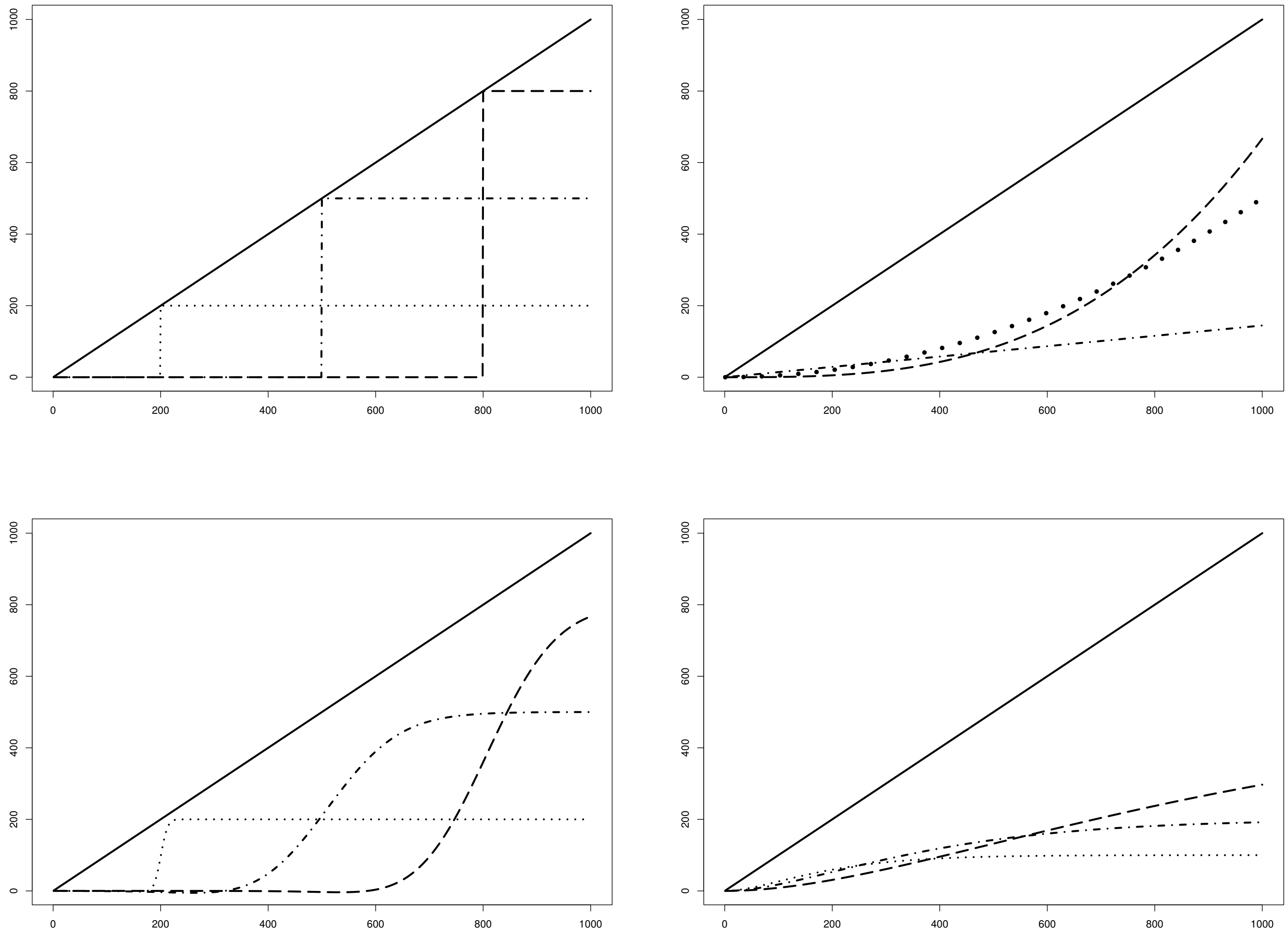
and ν is some distribution on $(0, \infty)$,

$$\mathrm{FDR}(R_\beta) \le \pi_0 \alpha.$$

Remarks : ν prior idea on the final number of rejections

- ν uniform on $\{1, \ldots, m\}$ gives $\beta(i) = i(i+1)/(2m)$
- $\nu(\{k\}) \propto 1/k$ gives $\beta(i) = i/(1 + 1/2 + \cdots + 1/m)$
- $\nu(\{k\}) \propto k$ gives $\beta(i) = \frac{i(i+1)(2i+1)}{3m(m+1)}$

Some choices for β (general dependencies)



$\Rightarrow$ no uniformly-optimal choice (it depends on the data)

To control FDR under gen. dep.: not only BY's procedure !

Introduction to adaptive procedures

Summary: we have $\text{FDR}(R_\beta) \leq \pi_0 \alpha$ when

1. The p-values are independent or positively dependent and $\beta(i) = i$
2. The p-values with general dependencies and $\beta(i) = \int_0^i u d\nu(u) \leq i$, and ν is some prior on $(0, \infty)$

Introduction to adaptive procedures

Summary: we have $\mathrm{FDR}(R_\beta) \leq \pi_0 \alpha$ when

1. The p-values are independent or positively dependent and $\beta(i) = i$
2. The p-values with general dependencies and $\beta(i) = \int_0^i u d\nu(u) \leq i$, and ν is some prior on $(0, \infty)$

Adaptivity problem: [Benjamini & Hochberg (00)], [Black (04)], [Genovese & Wasserman (04)],[Benjamini, Krieger & Yekutieli (06)],...

Introduction to adaptive procedures

Summary: we have $\mathrm{FDR}(R_\beta) \leq \pi_0 \alpha$ when

1. The p-values are independent or positively dependent and $\beta(i) = i$
2. The p-values with general dependencies and $\beta(i) = \int_0^i u d\nu(u) \leq i$, and ν is some prior on $(0, \infty)$

Adaptivity problem: [Benjamini & Hochberg (00)], [Black (04)], [Genovese & Wasserman (04)],[Benjamini, Krieger & Yekutieli (06)],...

Level $\pi_0 \alpha$ smaller than level α
$\Rightarrow$ conservative results (especially when π_0 small)

Introduction to adaptive procedures

Summary: we have $\text{FDR}(R_\beta) \le \pi_0 \alpha$ when

1. The p-values are independent or positively dependent and $\beta(i) = i$
2. The p-values with general dependencies and $\beta(i) = \int_0^i u d\nu(u) \le i$, and ν is some prior on $(0, \infty)$

Adaptivity problem: [Benjamini & Hochberg (00)], [Black (04)], [Genovese & Wasserman (04)],[Benjamini, Krieger & Yekutieli (06)],...

Level $\pi_0 \alpha$ smaller than level α
$\Rightarrow$ conservative results (especially when π_0 small)

Idea: put $\beta^\star = \beta \pi_0^{-1}$, so that $\text{FDR}(R_{\beta^\star}) \le \alpha$

Introduction to adaptive procedures

Summary: we have $\mathrm{FDR}(R_\beta) \leq \pi_0\alpha$ when

1. The p-values are independent or positively dependent and $\beta(i) = i$
2. The p-values with general dependencies and $\beta(i) = \int_0^i u d\nu(u) \leq i$, and ν is some prior on $(0, \infty)$

Adaptivity problem: [Benjamini & Hochberg (00)], [Black (04)], [Genovese & Wasserman (04)],[Benjamini, Krieger & Yekutieli (06)],...

Level $\pi_0\alpha$ smaller than level α

$\Rightarrow$ conservative results (especially when π_0 small)

Idea: put $\beta^\star = \beta\pi_0^{-1}$, so that $\mathrm{FDR}(R_{\beta^\star}) \leq \alpha$

⚠ π_0 unknown $\Rightarrow$ $\beta^\star$ unknown ! $\Rightarrow$ $R_{\beta^\star}$ oracle procedure

Introduction to adaptive procedures

Summary: we have $\mathrm{FDR}(R_\beta) \leq \pi_0 \alpha$ when

1. The p-values are independent or positively dependent and $\beta(i) = i$
2. The p-values with general dependencies and $\beta(i) = \int_0^i u d\nu(u) \leq i$, and ν is some prior on $(0, \infty)$

Adaptivity problem: [Benjamini & Hochberg (00)], [Black (04)], [Genovese & Wasserman (04)],[Benjamini, Krieger & Yekutieli (06)],...

Level $\pi_0 \alpha$ smaller than level α
$\Rightarrow$ conservative results (especially when π_0 small)

Idea: put $\beta^\star = \beta \pi_0^{-1}$, so that $\mathrm{FDR}(R_{\beta^\star}) \leq \alpha$

⚠ π_0 unknown $\Rightarrow \beta^\star$ unknown ! $\Rightarrow R_{\beta^\star}$ oracle procedure

(π_0)-Adaptive step-up procedures: $R_{\hat{\beta}}$ with $\hat{\beta} \simeq \beta^\star$

Our Goal

Find $\hat{\beta} \geq \beta$ such that $\mathrm{FDR}(R_{\hat{\beta}}) \leq \alpha$ and $\hat{\beta}$ "close to" $\beta\pi_0^{-1}$

Our Goal

Find $\hat{\beta} \geq \beta$ such that $\mathrm{FDR}(R_{\hat{\beta}}) \leq \alpha$ and $\hat{\beta}$ "close to" $\beta\pi_0^{-1}$

Under several dependence cases :
- independence ($\beta(i) = i$)
- positive dependencies ($\beta(i) = i$)
- general dependencies ($\beta(i) = \int_0^i u d\nu(u)$ given a prior ν)

Our Goal

Find $\hat{\beta} \geq \beta$ such that $\text{FDR}(R_{\hat{\beta}}) \leq \alpha$ and $\hat{\beta}$ "close to" $\beta\pi_0^{-1}$

Under several dependence cases :
- independence ($\beta(i) = i$)
- positive dependencies ($\beta(i) = i$)
- general dependencies ($\beta(i) = \int_0^i u d\nu(u)$ given a prior ν)

For this:

Two-stage procedure:

1. $\hat{F}$ estimates π_0^{-1}
2. Take $\hat{\beta} = \beta.\hat{F}$

Our Goal

Find $\hat{\beta} \geq \beta$ such that $\text{FDR}(R_{\hat{\beta}}) \leq \alpha$ and $\hat{\beta}$ "close to" $\beta\pi_0^{-1}$

Under several dependence cases :
- independence ($\beta(i) = i$)
- positive dependencies ($\beta(i) = i$)
- general dependencies ($\beta(i) = \int_0^i u d\nu(u)$ given a prior ν)

For this:

Two-stage procedure:

1. $\hat{F}$ estimates π_0^{-1}
2. Take $\hat{\beta} = \beta.\hat{F}$

One-stage procedure: $\hat{\beta}$ is deterministic

Outline

I. When the p-values are independent
- Some existing adaptive procedures
- New adaptive procedures

II. When the p-values are dependent
- New (and first ?) adaptive procedures

I. Existing adaptive procedures with FDR control

[Benjamini, Krieger and Yekutieli (2006)] BKY06:

1. Apply the standard step-up linear procedure R_0 at level $\alpha/(1+\alpha)$ and put $\hat{F} = \frac{m}{m-|R_0|}$
2. Take the step-up procedure R with threshold $\frac{\alpha}{1+\alpha}\frac{i}{m}\hat{F}$

I. Existing adaptive procedures with FDR control

[Benjamini, Krieger and Yekutieli (2006)] BKY06:

1. Apply the standard step-up linear procedure R_0 at level $\alpha/(1+\alpha)$ and put $\hat{F} = \frac{m}{m-|R_0|}$
2. Take the step-up procedure R with threshold $\frac{\alpha}{1+\alpha}\frac{i}{m}\hat{F}$

Using [Storey (2001)] Storey-λ :

1. $\hat{F} = \frac{(1-\lambda)m}{|\{h\in\mathcal{H}|p_h>\lambda\}|+1}$ (slightly modified)
2. Take the step-up procedure R with threshold $\frac{\alpha i}{m}\hat{F}$

I. Existing adaptive procedures with FDR control

[Benjamini, Krieger and Yekutieli (2006)] BKY06:

1. Apply the standard step-up linear procedure R_0 at level $\alpha/(1+\alpha)$ and put $\hat{F} = \frac{m}{m-|R_0|}$
2. Take the step-up procedure R with threshold $\frac{\alpha}{1+\alpha}\frac{i}{m}\hat{F}$

Using [Storey (2001)] Storey-λ :

1. $\hat{F} = \frac{(1-\lambda)m}{|\{h\in\mathcal{H}|p_h>\lambda\}|+1}$ (slightly modified)
2. Take the step-up procedure R with threshold $\frac{\alpha i}{m}\hat{F}$

Classical choice : $\lambda = 1/2$.

I. Existing adaptive procedures with FDR control

[Benjamini, Krieger and Yekutieli (2006)] BKY06:

1. Apply the standard step-up linear procedure R_0 at level $\alpha/(1+\alpha)$ and put $\hat{F} = \frac{m}{m-|R_0|}$
2. Take the step-up procedure R with threshold $\frac{\alpha}{1+\alpha}\frac{i}{m}\hat{F}$

Using [Storey (2001)] Storey-λ :

1. $\hat{F} = \frac{(1-\lambda)m}{|\{h\in\mathcal{H}|p_h>\lambda\}|+1}$ (slightly modified)
2. Take the step-up procedure R with threshold $\frac{\alpha i}{m}\hat{F}$

Classical choice : $\lambda = 1/2$.

Benjamini, Krieger and Yekutieli (2006):

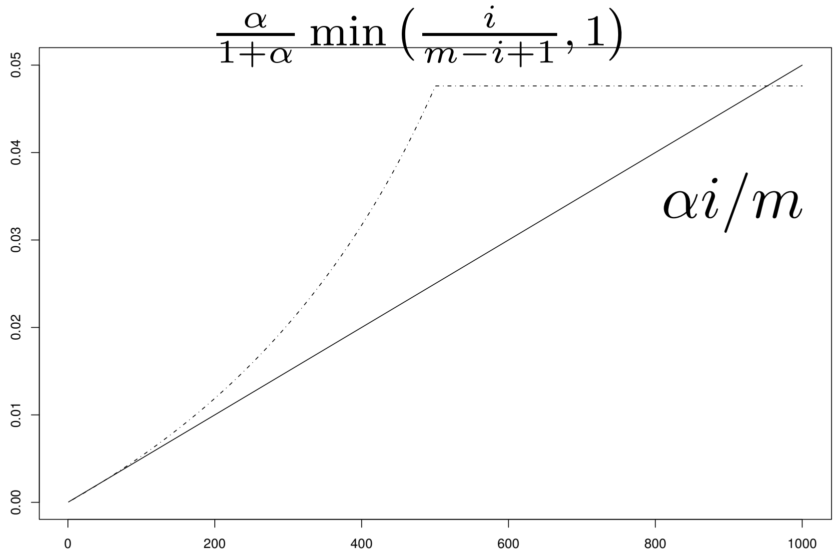
Theorem 4 p-values independent
These two procedures satisfy $\mathrm{FDR}(R) \le \alpha$

I. New one-stage adaptive procedure

Theorem 5 p-values independent:
The step-up procedure with global threshold $\frac{\alpha}{1+\alpha}\min\left(\frac{i}{m-i+1}, 1\right)$ has a FDR smaller than α.

I. New one-stage adaptive procedure

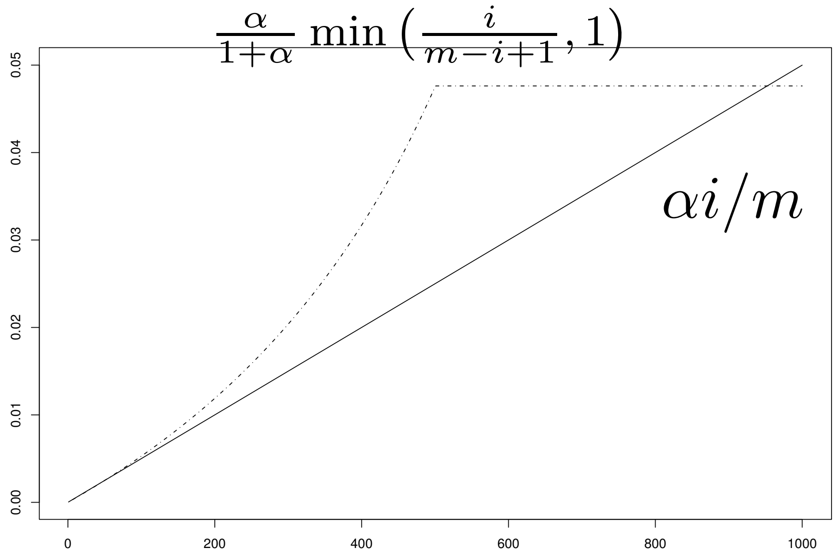
Theorem 5 p-values independent:
The step-up procedure with global threshold $\frac{\alpha}{1+\alpha}\min\left(\frac{i}{m-i+1},1\right)$ has a FDR smaller than α.



$\alpha = 0.05$

I. New one-stage adaptive procedure

Theorem 5 p-values independent:
The step-up procedure with global threshold $\frac{\alpha}{1+\alpha}\min\left(\frac{i}{m-i+1},1\right)$ has a FDR smaller than α.

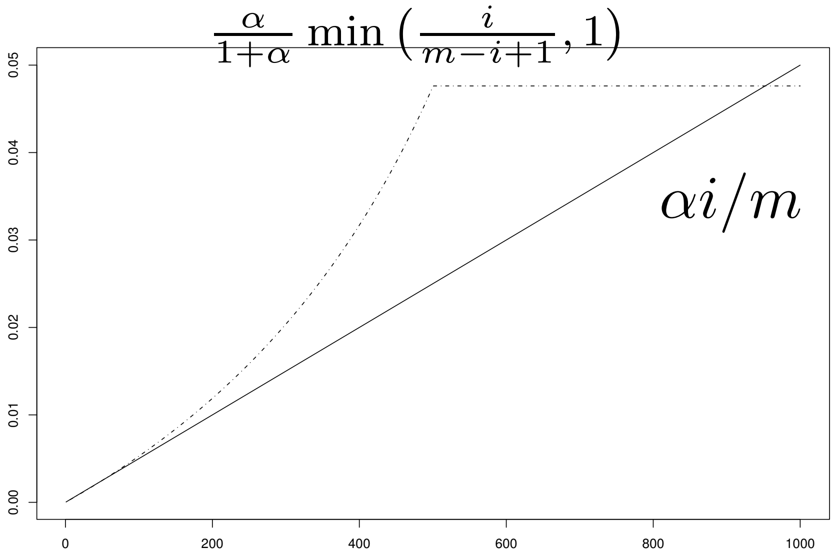


$\alpha = 0.05$

$\Rightarrow$ Explicit threshold!
$\Rightarrow$ better than the linear step-up procedure
(except when rejects $\geq \frac{1}{1+\alpha}$ or rejects $\leq \frac{1}{m}+\frac{\alpha}{1+\alpha}$ % of hypotheses)

I. New one-stage adaptive procedure

Theorem 5 p-values independent:
The step-up procedure with global threshold $\frac{\alpha}{1+\alpha}\min\left(\frac{i}{m-i+1},1\right)$ has a FDR smaller than α.



$\alpha = 0.05$

$\Rightarrow$ Explicit threshold!
$\Rightarrow$ better than the linear step-up procedure
(except when rejects $\geq \frac{1}{1+\alpha}$ or rejects $\leq \frac{1}{m} + \frac{\alpha}{1+\alpha}$ % of hypotheses)

Idea: use this procedure in the first step !

I. New two-stage adaptive procedure

Theorem 6 p-values independent: consider the two-stage procedure :

1. Apply the new one-stage adaptive procedure R_0' at level α and put $\hat{F} = \frac{m}{m-|R_0'|+1}$
2. Take the step-up procedure R with global threshold $\frac{\alpha}{1+\alpha}\frac{i}{m}\hat{F}$

Then $\mathrm{FDR}(R) \leq \alpha$.

I. New two-stage adaptive procedure

Theorem 6 p-values independent: consider the two-stage procedure :

1. Apply the new one-stage adaptive procedure R'_0 at level α and put $\hat{F} = \frac{m}{m-|R'_0|+1}$
2. Take the step-up procedure R with global threshold $\frac{\alpha}{1+\alpha}\frac{i}{m}\hat{F}$

Then $\mathrm{FDR}(R) \leq \alpha$.

$\Rightarrow$ BR07-2S always better than BKY06
(up to the extreme cases in the first step and "$+1$")

I. New two-stage adaptive procedure

Theorem 6 p-values independent: consider the two-stage procedure :

1. Apply the new one-stage adaptive procedure R'_0 at level α and put $\hat{F} = \frac{m}{m-|R'_0|+1}$
2. Take the step-up procedure R with global threshold $\frac{\alpha}{1+\alpha}\frac{i}{m}\hat{F}$

Then $\mathrm{FDR}(R) \leq \alpha$.

$\Rightarrow$ BR07-2S always better than BKY06
(up to the extreme cases in the first step and "$+1$")

On simulations:

I. New two-stage adaptive procedure

Theorem 6 p-values independent: consider the two-stage procedure :

1. Apply the new one-stage adaptive procedure R_0' at level α and put $\hat{F} = \frac{m}{m-|R_0'|+1}$
2. Take the step-up procedure R with global threshold $\frac{\alpha}{1+\alpha}\frac{i}{m}\hat{F}$

Then $\mathrm{FDR}(R) \leq \alpha$.

$\Rightarrow$ BR07-2S always better than BKY06
(up to the extreme cases in the first step and "$+1$")

On simulations:
- independent case: Power?
Storey-$1/2 \gg$ BR07-2S $\gg$ BKY06

I. New two-stage adaptive procedure

Theorem 6 p-values independent: consider the two-stage procedure :

1. Apply the new one-stage adaptive procedure R'_0 at level α and put $\hat{F} = \frac{m}{m-|R'_0|+1}$
2. Take the step-up procedure R with global threshold $\frac{\alpha}{1+\alpha}\frac{i}{m}\hat{F}$

Then $\mathrm{FDR}(R) \leq \alpha$.

$\Rightarrow$ BR07-2S always better than BKY06
(up to the extreme cases in the first step and "$+1$")

On simulations:
- independent case: Power?
Storey-$1/2 \gg$ BR07-2S $\gg$ BKY06
- positively dependent case: FDR control?
Storey-$1/2$ is not robust!
New two-stage procedure seems robust to positive correlations

I I. Under dependence

Recall: FDR control for step-up procedures with $\alpha\beta(i)/m$ in the cases:
- positive dependencies with $\beta(i) = i$
- general dependencies with $\beta(i) = \int_0^i u d\nu(u)$ given a prior ν

I I. Under dependence

Recall: FDR control for step-up procedures with $\alpha\beta(i)/m$ in the cases:
- positive dependencies with $\beta(i) = i$
- general dependencies with $\beta(i) = \int_0^i u d\nu(u)$ given a prior ν

Theorem 7 two-stage adaptive procedure:

1. Non-adaptive step-up procedure R_0 with threshold $\frac{\alpha}{4}\frac{\beta(i)}{m}$ and put $\hat{F} = \frac{1}{1-\sqrt{(2|R_0|/m-1)_+}}$
2. Step-up procedure R with threshold $\frac{\alpha}{2}\frac{\beta(i)}{m}\hat{F}$

I I. Under dependence

Recall: FDR control for step-up procedures with $\alpha\beta(i)/m$ in the cases:
- positive dependencies with $\beta(i) = i$
- general dependencies with $\beta(i) = \int_0^i u d\nu(u)$ given a prior ν

Theorem 7 two-stage adaptive procedure:

1. Non-adaptive step-up procedure R_0 with threshold $\frac{\alpha}{4}\frac{\beta(i)}{m}$ and put $\hat{F} = \frac{1}{1-\sqrt{(2|R_0|/m-1)_+}}$
2. Step-up procedure R with threshold $\frac{\alpha}{2}\frac{\beta(i)}{m}\hat{F}$

Then $\mathrm{FDR}(R) \leq \alpha$ in the two cases:
- positive dependencies (PRDS) with $\beta(i) = i$
- general dependencies with $\beta(i) = \int_0^i u d\nu(u)$ given a prior ν

I I. Under dependence

Recall: FDR control for step-up procedures with $\alpha\beta(i)/m$ in the cases:
- positive dependencies with $\beta(i) = i$
- general dependencies with $\beta(i) = \int_0^i u d\nu(u)$ given a prior ν

Theorem 7 two-stage adaptive procedure:

1. Non-adaptive step-up procedure R_0 with threshold $\frac{\alpha}{4}\frac{\beta(i)}{m}$ and put $\hat{F} = \frac{1}{1-\sqrt{(2|R_0|/m-1)_+}}$
2. Step-up procedure R with threshold $\frac{\alpha}{2}\frac{\beta(i)}{m}\hat{F}$

Then $\mathrm{FDR}(R) \leq \alpha$ in the two cases:
- positive dependencies (PRDS) with $\beta(i) = i$
- general dependencies with $\beta(i) = \int_0^i u d\nu(u)$ given a prior ν

Loss/ indep case: $\frac{\alpha}{2}$, $\frac{\alpha}{4}$ and $\frac{1}{1-\sqrt{(2x-1)_+}} \leq \frac{1}{1-x}$

I I. Under dependence (2)

Remarks :

- Estimation based on Markov’s inequality (conservative device)

I I. Under dependence (2)

Remarks :

- Estimation based on Markov's inequality (conservative device)

- New procedure better than non-adaptive: if $|R_0|/m \geq 62.5\%$.

I I. Under dependence (2)

Remarks :

- Estimation based on Markov's inequality (conservative device)

- New procedure better than non-adaptive: if $|R_0|/m \geq 62.5\%$.

$\Rightarrow$ Useful only if large number of rejections!
(m_0 small and p_h, $h \notin \mathcal{H}_0$ small)

I I. Under dependence (2)

Remarks :

- Estimation based on Markov's inequality (conservative device)

- New procedure better than non-adaptive: if $|R_0|/m \geq 62.5\%$.

$\Rightarrow$ Useful only if large number of rejections!
(m_0 small and p_h, $h \notin \mathcal{H}_0$ small)

$\Rightarrow$ interest more theoretical than practical.

Conclusion and future works

New adaptive procedures that control the FDR:

- when the p-values are independent:
  one-stage explicit and better than LSU
  two-stage $\longrightarrow$ better than BKY06
  $\longrightarrow$ seems robust to positive correlations.
- when the p-values have positive or general dependencies:
  new (and first ?) two-stage procedures
  $\longrightarrow$ only relevant when large number of rejections.

Conclusion and future works

New adaptive procedures that control the FDR:

* when the p-values are independent:
  one-stage explicit and better than LSU
  two-stage $\longrightarrow$ better than BKY06
  $\longrightarrow$ seems robust to positive correlations.
* when the p-values have positive or general dependencies:
  new (and first ?) two-stage procedures
  $\longrightarrow$ only relevant when large number of rejections.

Future works:

- Under independence:

Integrate a more efficient estimator in procedures?

Conclusion and future works

New adaptive procedures that control the FDR:

- ∗ when the p-values are independent:
  one-stage explicit and better than LSU
  two-stage $\longrightarrow$ better than BKY06
  $\longrightarrow$ seems robust to positive correlations.
- ∗ when the p-values have positive or general dependencies:
  new (and first ?) two-stage procedures
  $\longrightarrow$ only relevant when large number of rejections.

Future works:

- Under independence:
Integrate a more efficient estimator in procedures?
- Under dependence:
Find another device than Markov's inequality?

Conclusion and future works

New adaptive procedures that control the FDR:

- $*$ when the p-values are independent:
  one-stage explicit and better than LSU
  two-stage $\longrightarrow$ better than BKY06
  $\longrightarrow$ seems robust to positive correlations.
- $*$ when the p-values have positive or general dependencies:
  new (and first ?) two-stage procedures
  $\longrightarrow$ only relevant when large number of rejections.

Future works:

- Under independence:

Integrate a more efficient estimator in procedures?

- Under dependence:

Find another device than Markov's inequality?

Choose the prior ν from an estimator of π_0?

Thank you for your attention!

A preprint is available on: http://genome.jouy.inra.fr/$\sim$eroquain

Appendix

I. Simulations

For $k = 1, \ldots, m$,

$$Y_k \sim \mathcal{N}(\mu_k, 1)$$

I. Simulations

For $k = 1, \ldots, m$,

$$Y_k \sim \mathcal{N}(\mu_k, 1)$$

For $k \neq k'$,

$$\mathrm{Cov}(Y_k, Y_{k'}) = \rho \quad \text{with} \quad \rho \in [0, 1].$$

$\rho = 0 \Rightarrow$ independent case
$\rho \geq 0 \Rightarrow$ positive dependent case

I. Simulations

For $k = 1, \ldots, m$,

$$Y_k \sim \mathcal{N}(\mu_k, 1)$$

For $k \neq k'$,

$$\mathrm{Cov}(Y_k, Y_{k'}) = \rho \quad \text{with} \quad \rho \in [0, 1].$$

$\rho = 0 \Rightarrow$ independent case
$\rho \geq 0 \Rightarrow$ positive dependent case

One-sided tests: $H_{0,k}$: "$\mu_k \leq 0$", $k = 1, \ldots, m$

I. Simulations

For $k = 1, \ldots, m$,

$$Y_k \sim \mathcal{N}(\mu_k, 1)$$

For $k \neq k'$,

$$\mathrm{Cov}(Y_k, Y_{k'}) = \rho \quad \text{with} \quad \rho \in [0, 1].$$

$\rho = 0 \Rightarrow$ independent case
$\rho \geq 0 \Rightarrow$ positive dependent case

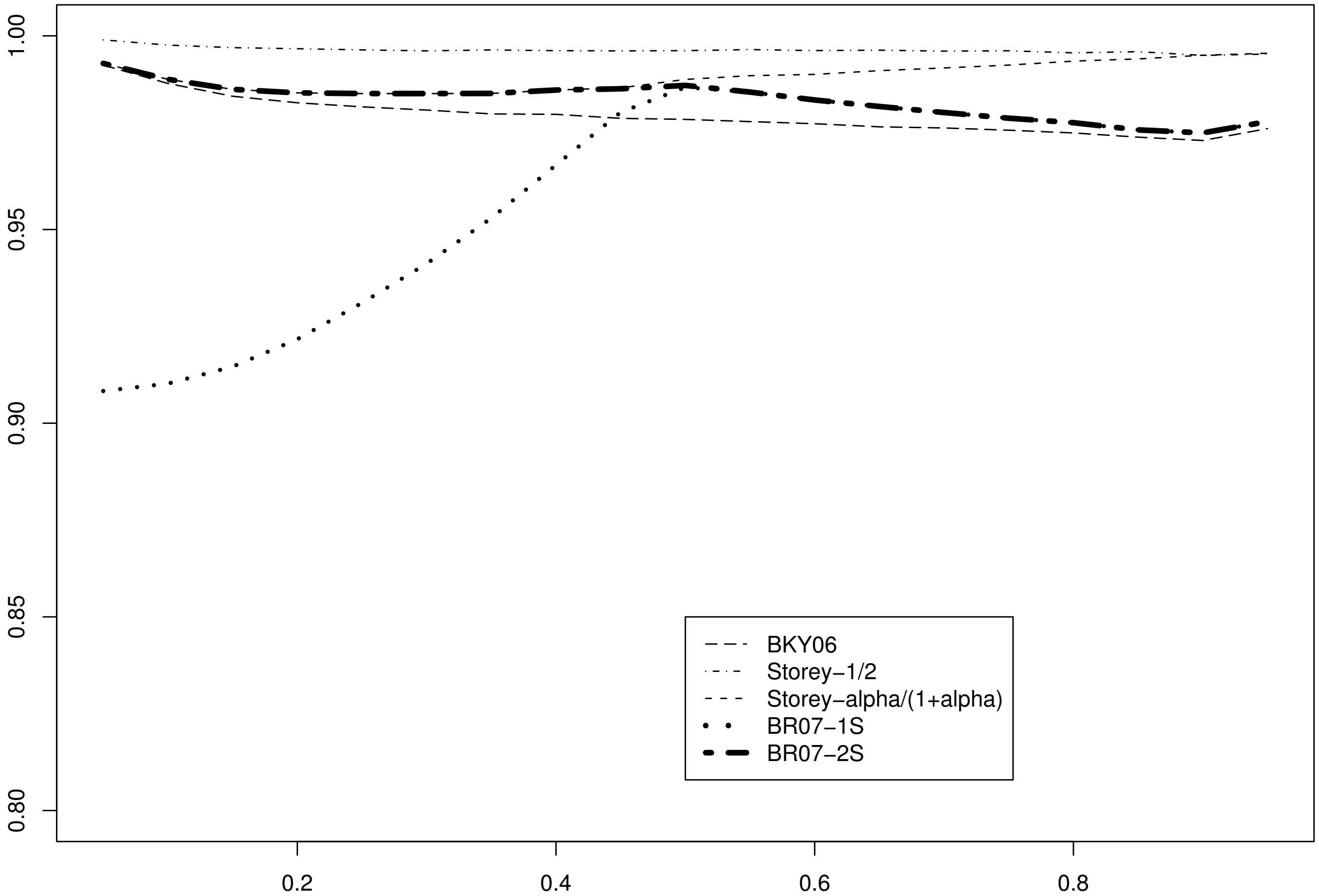
One-sided tests: $H_{0,k}$: "$\mu_k \leq 0$", $k = 1, \ldots, m$

With 10000 simulations, $m = 100$, mean $= 3$:

- Power (in independent case) :
nb of correct rejections / nb of correct rejections of the oracle procedure
- FDR estimation (in the positive dependent case)

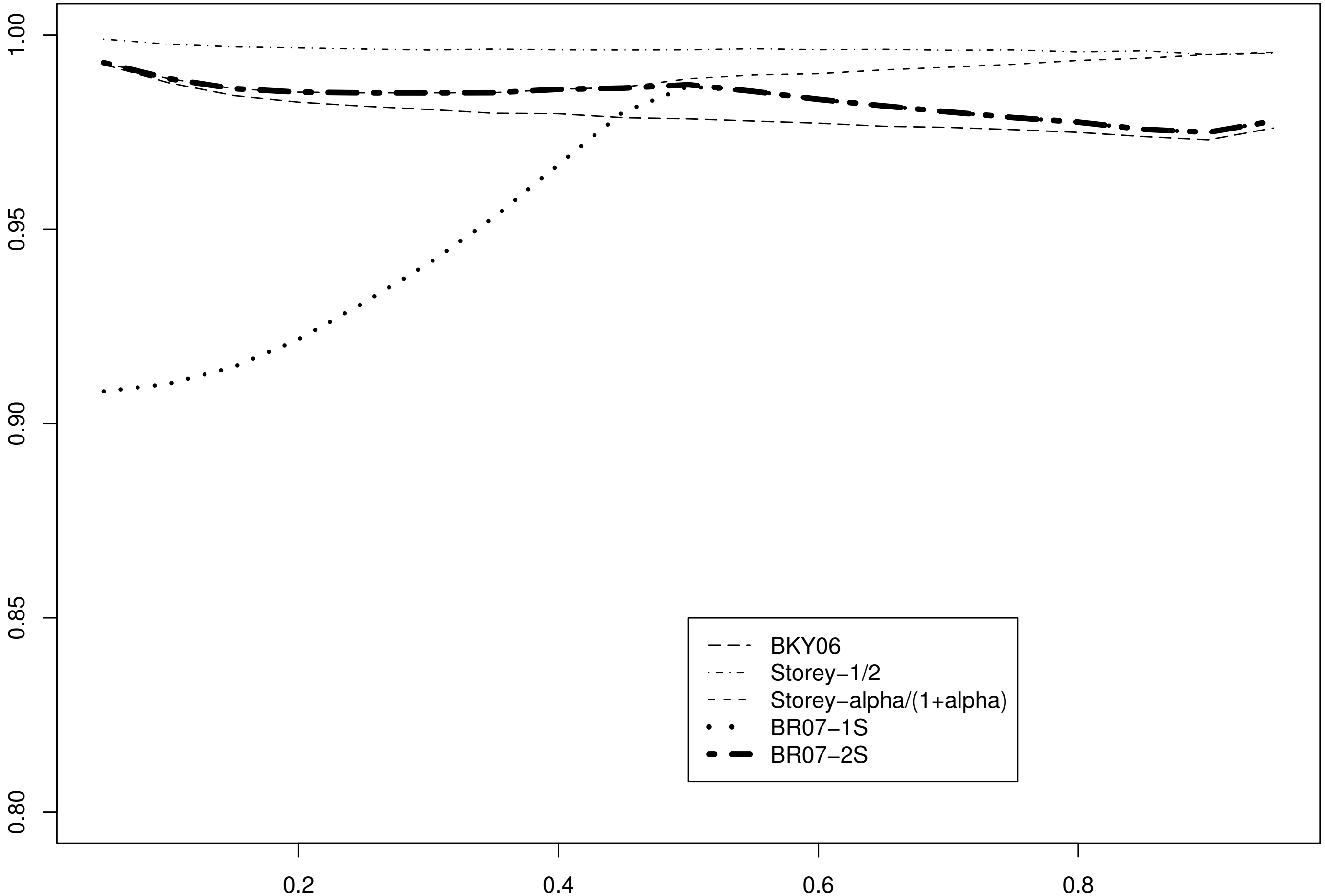
I. Simulations, Power, indep. case

$\rho = 0$ (independent case) :



I. Simulations, Power, indep. case

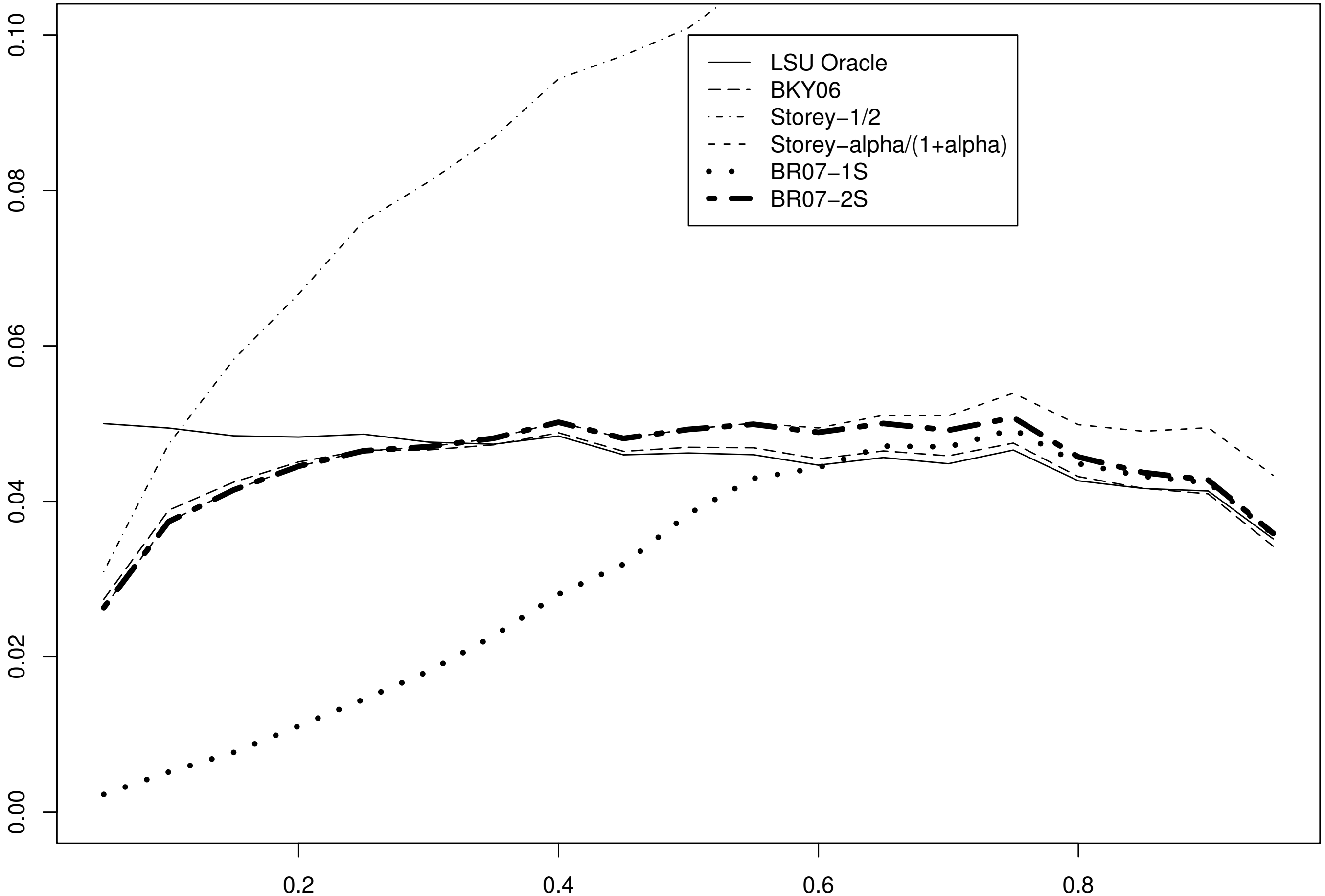
$\rho = 0$ (independent case) :



$$\text{Storey-}1/2 \gg \text{Storey-}\alpha/(1+\alpha) \gg \text{BR07-2S} \gg \text{BKY06}$$

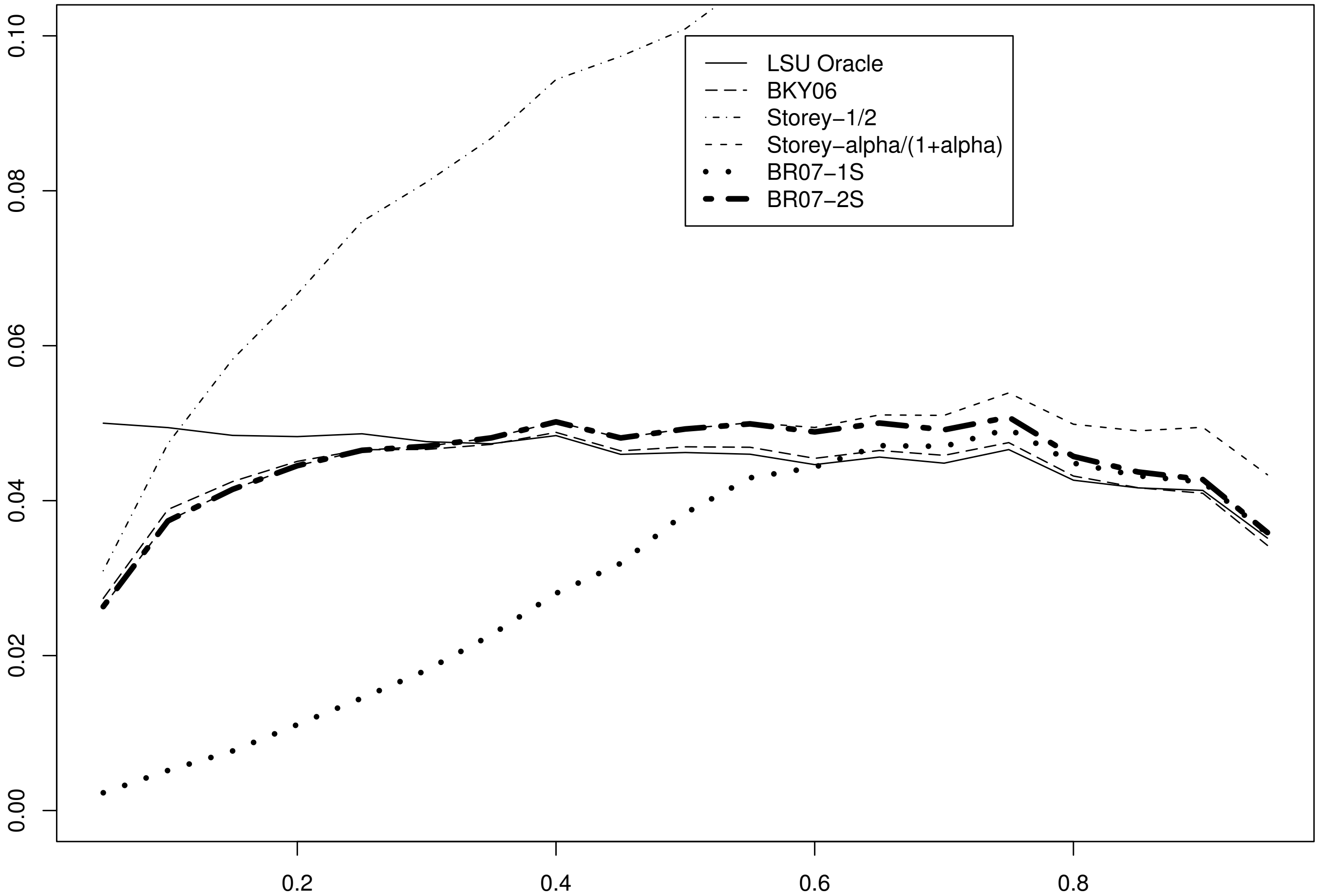
I. Simulations, FDR, with corr

$\rho = 0.5$ (positive correlate case) :



I. Simulations, FDR, with corr

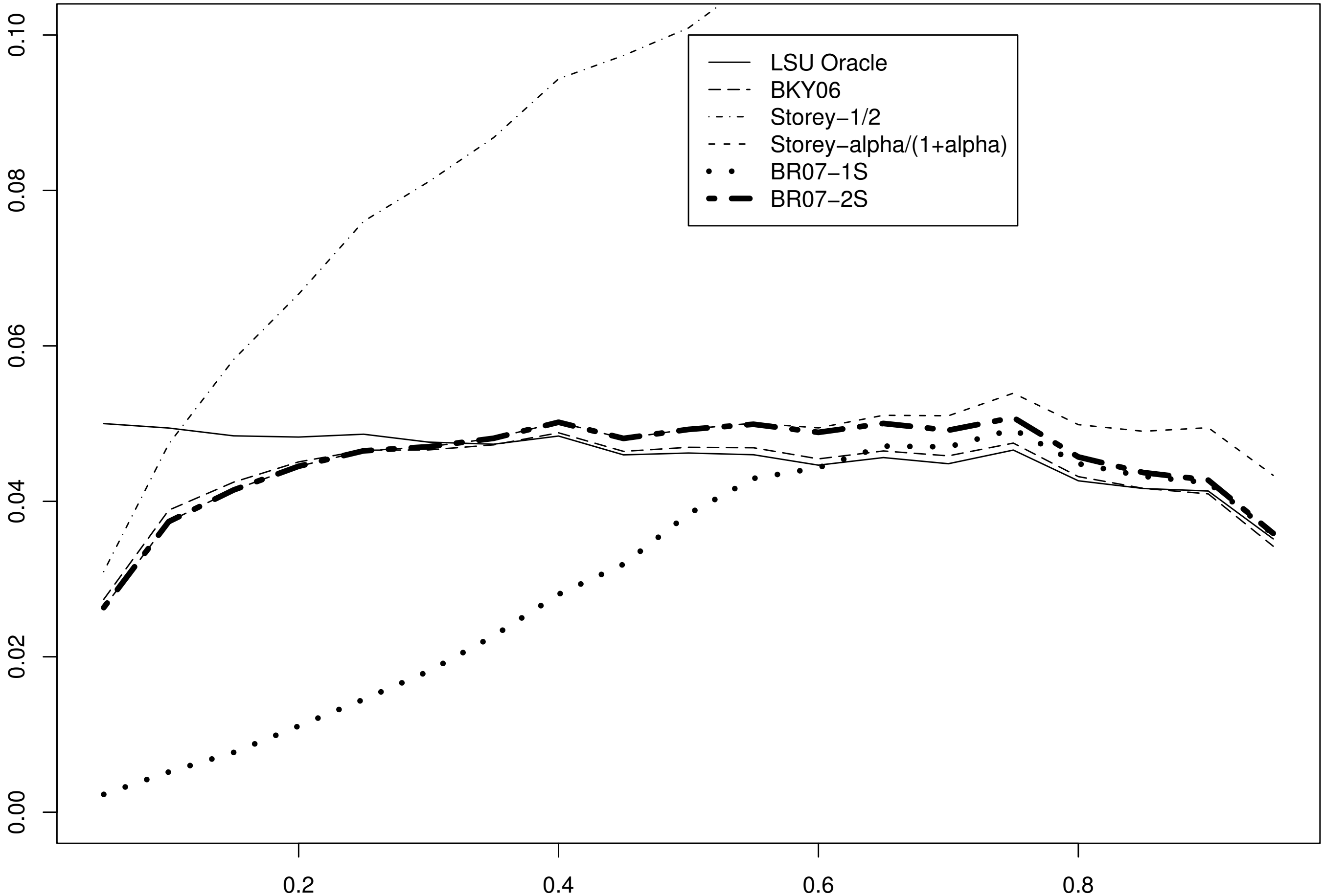
$\rho = 0.5$ (positive correlate case) :



$\Rightarrow$ Storey-$1/2$ is not robust! Storey-$\alpha/(1+\alpha)$ not robust? (max $\simeq 0.054$)

I. Simulations, FDR, with corr

$\rho = 0.5$ (positive correlate case) :



$\Rightarrow$ Storey-$1/2$ is not robust! Storey-$\alpha/(1+\alpha)$ not robust? (max $\simeq 0.054$)

$\Rightarrow$ New procedures seem robust to positive correlations

I I. Simulations

$m = 100$, $m_0 = 5$, $\rho = 0.1$, $2 \leq \text{mean} \leq 5$