Minimally Adaptive BH

a tiny but uniform improvement of the procedure of Benjamini and Hochberg

Aldo Solari

Joint work with Jelle Goeman

Adaptive Designs and Multiple Testing Procedures

Padua, April 29, 2016

Outline

Simes test

- $m \geq 2$ hypotheses, $m_0 \leq m$ true hypotheses
- $p_1 \leq \ldots \leq p_m$ ordered p-values
- Assumption: Simes inequality holds for p-values corresponding to true hypotheses

- Global hypothesis $\mathrm{G}_0 : \pi_0 = \dfrac{m_0}{m} = 1$
- $S(\alpha) = \bigcap_{i=1}^{m} \left\{ p_i > \dfrac{i\alpha}{m} \right\}$: "Simes test doesn't reject G_0 at α"
- $\bar{S}(\alpha) = \bigcup_{i=1}^{m} \left\{ p_i \leq \dfrac{i\alpha}{m} \right\}$: "Simes test rejects G_0 at α"

Benjamini and Hochberg procedure

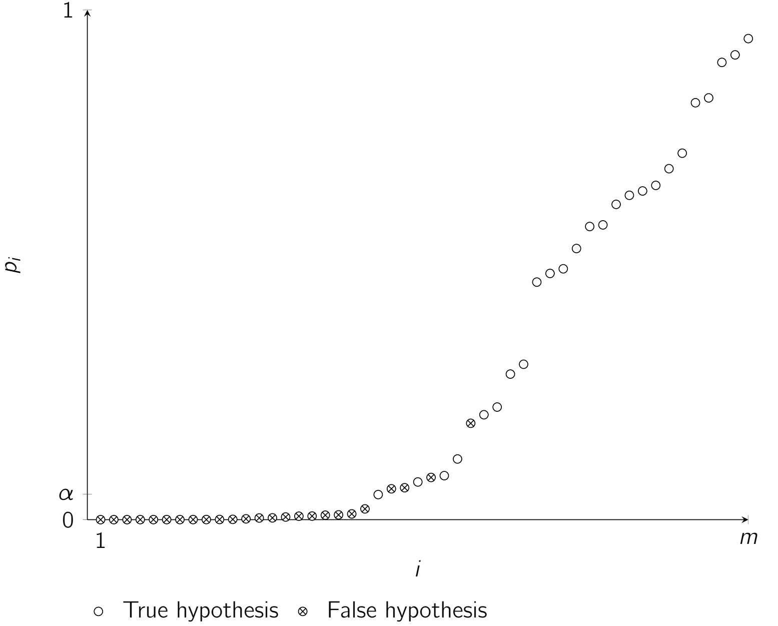
BH(α) rejects the $R(\alpha)$ hypotheses with smallest p-values:

$$R(\alpha) = \begin{cases} 0 & \text{if } S(\alpha) \\ \max\left\{i : p_i \leq \frac{i\alpha}{m}\right\} & \text{if } \bar{S}(\alpha) \end{cases}$$

Very popular: BH95 has $> 32K$ citations

?p.adjust example

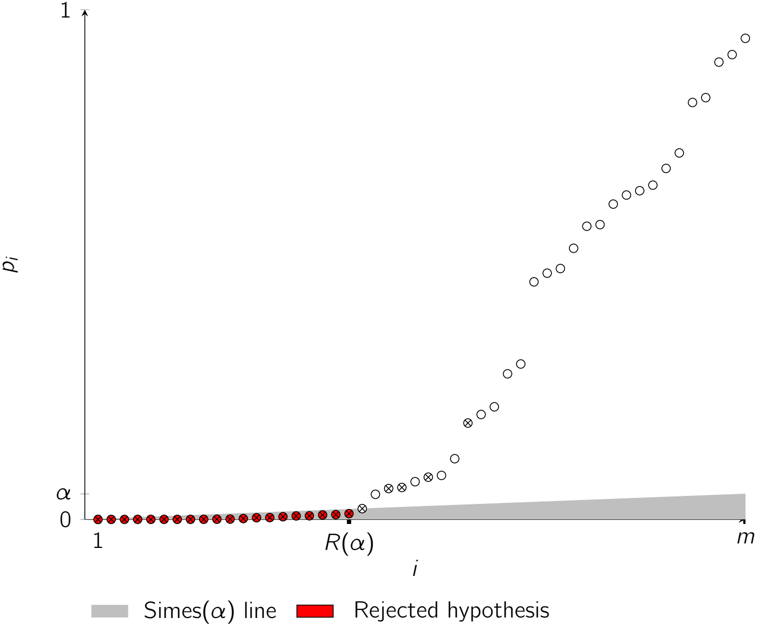
```
set.seed(123)
x <- rnorm(50, mean = c(rep(0, 25), rep(3, 25)))
p <- 2*pnorm(sort(-abs(x)))
```


1
p_i
α
0
1
m
i
True hypothesis
False hypothesis

BH(5%)

```
p.adjust(p, "BH")

 [1] 0.0000 0.0004 0.0004 0.0009 0.0009
 [6] 0.0009 0.0009 0.0010 0.0013 0.0019
[11] 0.0028 0.0067 0.0126 0.0126 0.0175
[16] 0.0208 0.0208 0.0245 0.0247 0.0282
[21] 0.0504 oh,no!
```



1
α
0
p_i
1
R(α)
m
i
Simes(α) line
Rejected hypothesis

FDR control for BH

Under the assumption of independence or PRDS,
BH(α) controls the FDR at level $\pi_0\alpha \leq \alpha$, i.e.

$$\mathrm{E}[Q(\alpha)] \leq \pi_0\alpha \leq \alpha$$

- $Q(\alpha)$ is the FPD of BH(α), i.e.

$$Q(\alpha) = \begin{cases} 0 & \text{if } S(\alpha) \\ \frac{T(\alpha)}{R(\alpha)} & \text{if } \bar{S}(\alpha) \end{cases}$$

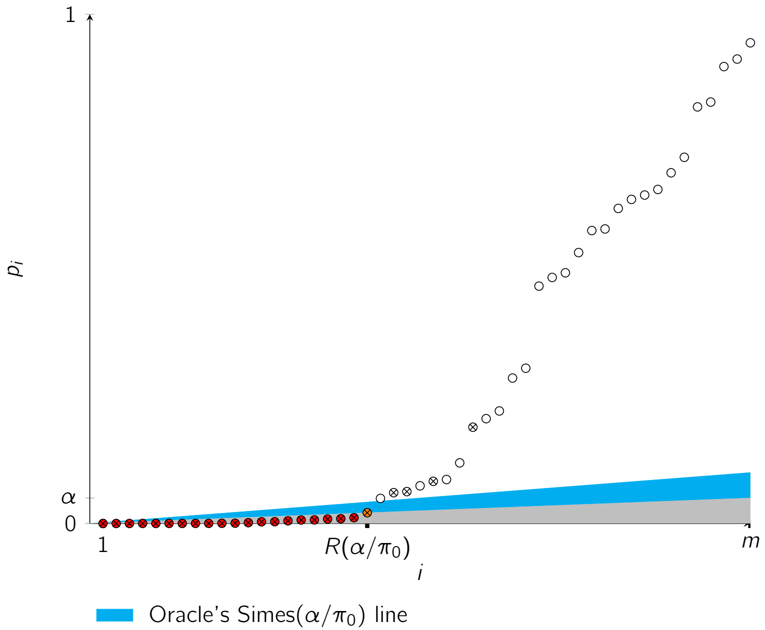
- $T(\alpha)$ is number of true hypotheses rejected by BH(α)

Oracle BH procedure

- The Oracle knows the true value of π_0
- OBH(α) = BH(α/π_0):

$$R(\alpha/\pi_0) = \begin{cases} 0 & \text{if } S(\alpha/\pi_0) \\ \max\left\{i : p_i \leq \frac{i\alpha}{m_0}\right\} & \text{if } \bar{S}(\alpha/\pi_0) \end{cases}$$

- If BH(α) controls FDR at $\pi_0\alpha$ for all α, then OBH(α) controls FDR at α for all α
- OBH(α) is uniformly more powerful than BH(α)



1
α
0
p_i
1
$R(\alpha/\pi_0)$
m
i
Oracle's Simes(α/π_0) line

Two-stages Adaptive BH procedures

1. Estimate π_0 by $\hat{\pi}_0$
2. $\mathrm{ABH}(\alpha) = \mathrm{BH}(\alpha^*/\hat{\pi}_0)$

Existing procedures

- Existing $\mathrm{ABH}(\alpha)$ are not uniformly more powerful than $\mathrm{BH}(\alpha)$ because $\alpha^*/\hat{\pi}_0$ can be strictly smaller than α
- Usually proven FDR control under independence only
- Storey, Taylor and Siegmund (2004)

  $\alpha^* = \alpha$ but $\hat{\pi}_0(\lambda) = \dfrac{\#\{p_i > \lambda\} + 1}{m(1-\lambda)}$ can be larger than 1
- Benjamini, Krieger and Yekutieli (2006)

  $\hat{\pi}_0 = \dfrac{m - R(\alpha^*)}{m}$ with $\alpha^* = \dfrac{\alpha}{1+\alpha} < \alpha$

Outline

1. Introduction
2. **The MABH procedure**
3. The case of two hypotheses

Minimally Adaptive BH procedure

1. $\hat{\pi}_0 = \begin{cases} 1 & \text{if } S(\alpha) \\ (m-1)/m & \text{if } \bar{S}(\alpha) \end{cases}$
2. $\mathsf{MABH}(\alpha) = \mathsf{BH}(\alpha/\hat{\pi}_0) = \begin{cases} \mathrm{BH}(\alpha) & \text{if } S(\alpha) \\ \mathrm{BH}\left(\frac{m\alpha}{m-1}\right) & \text{if } \bar{S}(\alpha) \end{cases}$

MABH(α) rejects the $R'(\alpha)$ hypotheses with smallest p-values:

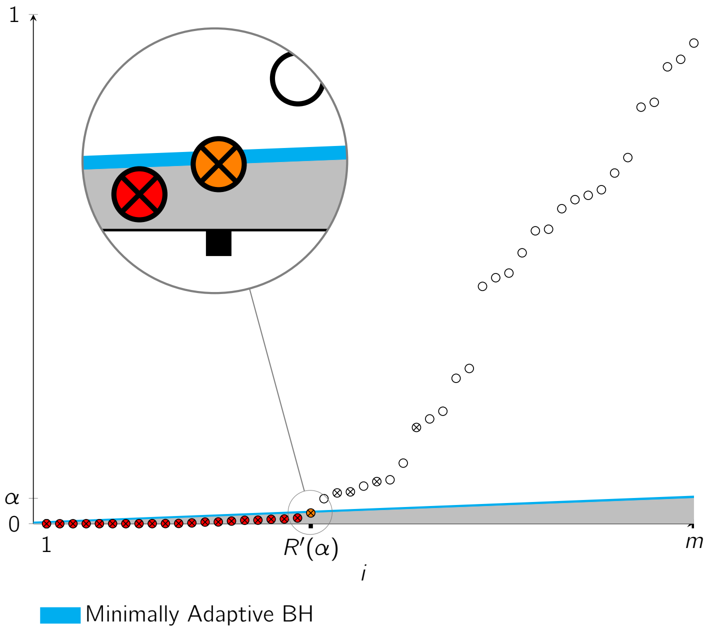
$$R'(\alpha) = \begin{cases} 0 & \text{if } S(\alpha) \\ \max\left\{i : p_i \leq \frac{i\alpha}{m-1}\right\} & \text{if } \bar{S}(\alpha) \end{cases}$$

MABH(5%)

```
p.adjust(p, "MABH")

 [1] 0.0000 0.0004 0.0004 0.0009 0.0009
 [6] 0.0009 0.0009 0.0010 0.0012 0.0019
[11] 0.0027 0.0066 0.0124 0.0124 0.0171
[16] 0.0203 0.0203 0.0240 0.0242 0.0277
[21] 0.0494
```

$$\tilde{p}_i = \max\left(\min_j \frac{mp_j}{j},\ \min_{k \geq i} \frac{(m-1)p_k}{k}\right)$$



1
α
0
1
R′(α)
m
i
Minimally Adaptive BH

FDR control for MABH

If BH(α) controls FDR at $\pi_0\alpha$ for all α, then MABH(α) controls FDR at α for all α, i.e.

$$\mathrm{E}[Q(\alpha)] \leq \pi_0\alpha \Rightarrow \mathrm{E}[Q'(\alpha)] \leq \alpha \quad \forall\alpha$$

- $\pi_0 = 1$

$$\mathrm{E}[Q'(\alpha)] = \mathrm{P}[R'(\alpha) > 0] = \mathrm{P}[\bar{S}(\alpha)] = \mathrm{P}[R(\alpha) > 0] = \mathrm{E}[Q(\alpha)] \leq \alpha$$

- $\pi_0 \leq (m-1)/m$

$$Q'(\alpha) = \begin{cases} 0 \leq Q\left(\frac{m\alpha}{m-1}\right) & \text{if } S(\alpha) \\ Q\left(\frac{m\alpha}{m-1}\right) & \text{if } \bar{S}(\alpha) \end{cases}$$

$$\mathrm{E}[Q'(\alpha)] \leq \mathrm{E}\left[Q\left(\frac{m\alpha}{m-1}\right)\right] \leq \pi_0 \frac{m\alpha}{m-1} \leq \alpha$$

Outline

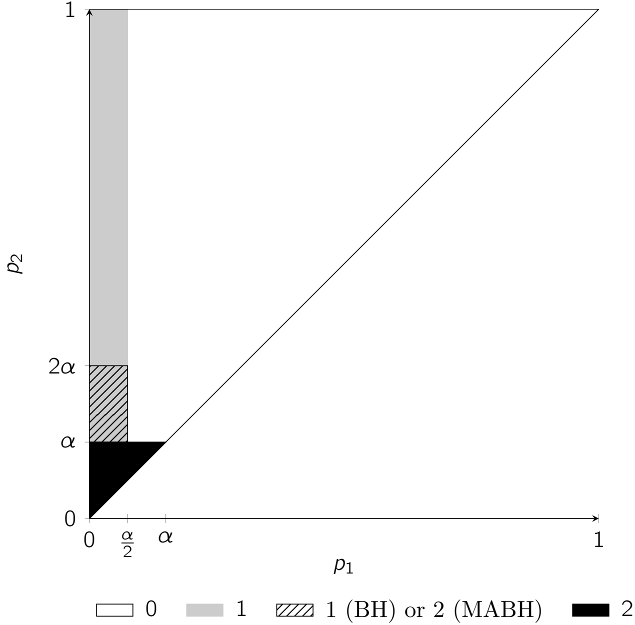
Case $m = 2$

- BH(α)

$$R(\alpha) = \begin{cases} 0 & \text{if } \{p_1 > \alpha/2\} \cap \{p_2 > \alpha\} \\ 1 & \text{if } \{p_1 \leq \alpha/2\} \cap \{p_2 > \alpha\} \\ 2 & \text{if } p_2 \leq \alpha \end{cases}$$

- MABH(α)

$$R'(\alpha) = \begin{cases} 0 & \text{if } \{p_1 > \alpha/2\} \cap \{p_2 > \alpha\} \\ 1 & \text{if } \{p_1 \leq \alpha/2\} \cap \{p_2 > 2\alpha\} \\ 2 & \text{if } \{\{p_1 \leq \alpha/2\} \cap \{p_2 \leq 2\alpha\}\} \cup \{p_2 \leq \alpha\} \end{cases}$$


1
2α
α
0
0
α/2
α
1
p2
p1
0
1
1 (BH) or 2 (MABH)
2

Case $m = 2$: FWER and FDR control

- Hochberg(α) = Hommel(α) controls FWER at α under Simes inequality
- BH(α) = Hochberg(α) = Hommel(α), thus BH(α) controls both FWER and FDR at α under PRDS
- MABH(α) controls FDR at α but not FWER under PRDS
- Counter-example (Dirac-Uniform configuration):
  H_1 false with p-value 0, H_2 true with p-value Uniform
  Then FWER for MABH(α) is $2\alpha > \alpha$

Conclusions

- $\oplus$ MABH is uniformly more powerful than BH
- $\oplus$ When $m = 2$, MABH is the first procedure that controls FDR but not FWER under PRDS
- $\ominus$ Power gain of MABH over BH is negligible when m is large
- $\ominus$ MABH is less elegant than BH

To appear on Biometrical Journal

Solari A and Goeman JJ (2016)
Minimally Adaptive BH: a tiny but uniform improvement of the procedure of Benjamini and Hochberg
`Biometrical Journal, special issue MCP2015`