Lecture 6
Post-selection inference

29 April 2020

Aldo Solari
University of Milano-Bicocca
Statistical Inference II
PhD in Economics and Statistics

fMRI analysis

Significance map

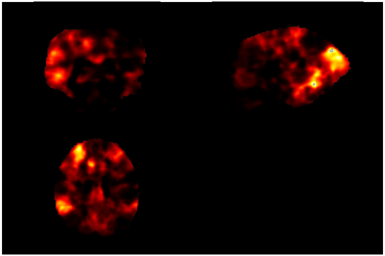
Significance (*p*-value) for brain activity at each location (voxel)

Goal

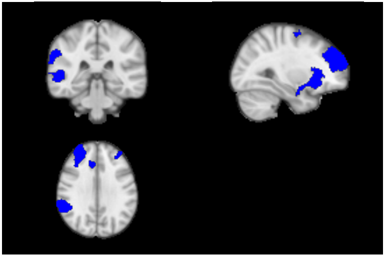
Find brain *regions* of brain activity

Aggregation

Micro-inferences (voxels) $\to$ larger-scale inferences (regions)



Significance map



Selection

Post-selection inference

How to assess the significance of selected regions?
Regions are both selected and tested with the same data

Need to correct for overoptimism in inference due to data-driven selection

Null voxels

$M = \{1, \ldots, m\}$ collection of m voxels
$M_0 \subseteq M$ null voxels with $|M_0| = m_0$
$M_1 = M \setminus M_0$ non-null voxels with $|M_1| = m_1$

Voxel null hypothesis

$H_i : i \in N$ with p-value p_i $i \in M$

Selection

$S \subseteq M$ selected voxels

$m_1(S) = |M_1 \cap S|$ number of true discoveries in the selection
$m_0(S) = |S| - m_1(S), \pi_1(S) = m_1(S)/|S|, \pi_0(S) = 1 - \pi_1(S) = \mathrm{FDP}(S)$

Cluster-size inference

Poline and Mazoyer (1993)

A cluster is significant if its size is larger than chance
It controls the familywise error rate at α

1. Cluster $C \equiv$ contiguous voxels with p-values $< t$

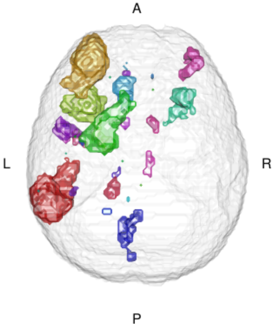
   where t is the cluster-forming threshold
   This gives K clusters $C_1, \ldots, C_K$ with $K \in \mathbb{N}_0$

2. C_k is significant if its size $|C_k| > s_\alpha$

   where s_α is $1 - \alpha$ quantile of the null distribution of $\max\limits_{k=1,\ldots,K} |C_k|$

Clusters

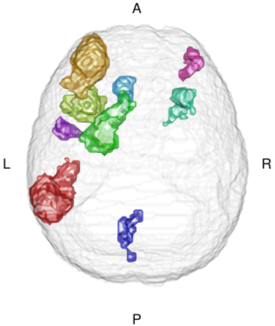
cluster $\equiv$ contiguous voxels with $p < 0.001$



32 clusters of size 2191, 1835, 1400, 698, 421, 304, 245, 232, 187, 82, 69, 43, 43, 28, 14, 12, 12, 10, 10, 6, 5, 4, 3, 2, 1, 1, 1, 1, 1, 1, 1, 1

Significant clusters

size > 161



9 significant clusters of size 2191, 1835, 1400, 698, 421, 304, 245, 232, 187

Cluster null hypothesis

In cluster-size inference we are testing a random number of cluster null hypotheses of the form

$$H_C = \bigcap_{i \in C} H_i$$

Rejection of H_C implies 'at least one voxel is non-null' : $m_1(C) \geq 1$

All-Resolutions Inference

Goeman and Solari (2011), Rosenblatt et al. (2019)

Simultaneous confidence lower bounds for the # of true discoveries

$$\mathrm{P}\Big(\underline{m}_1(S) \leq m_1(S)\ \forall\, S \subseteq M \Big) \geq 1 - \alpha$$

ARI legitimates the practice of exploratory data analysis with strong inferential guarantees (exploratory inference):

- the user looks at the data and selects interesting $S_1, S_2, \ldots$
- ARI informs the user about $\underline{m}_1(S_1), \underline{m}_1(S_2), \ldots$
- based on the results, the user may consider others S_1', S_2'
- ARI informs the user about $\underline{m}_1(S_1'), \underline{m}_1(S_2'), \ldots$
- $\ldots$

Cluster-size inference

S	$\|S\|$	$m_1(S)$
C_1	2191	≥ 1
C_2	1835	≥ 1
C_3	1400	≥ 1
C_4	698	≥ 1
C_5	421	≥ 1
C_6	304	≥ 1
C_7	245	≥ 1
C_8	232	≥ 1
C_9	187	≥ 1

ARI

S	$\lvert S \rvert$	$m_1(S)$	$\pi_1(S)$
C_1	2191	≥ 624	≥ 29 %
C_2	1835	≥ 847	≥ 46 %
C_3	1400	≥ 454	≥ 32 %
C_4	698	≥ 0	≥ 0 %
C_5	421	≥ 25	≥ 6 %
C_6	304	≥ 33	≥ 11 %
C_7	245	≥ 0	≥ 0 %
C_8	232	≥ 0	≥ 0 %
C_9	187	≥ 0	≥ 0 %

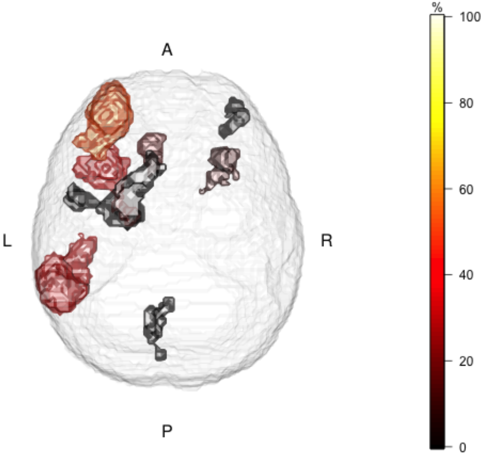
Bonferroni inference

S	$\lvert S \rvert$	$m_1(S)$	$\pi_1(S)$
C_1	2191	≥ 7	≥ 0.3 %
C_2	1835	≥ 86	≥ 4 %
C_3	1400	≥ 82	≥ 6 %
C_4	698	≥ 0	≥ 0 %
C_5	421	≥ 0	≥ 0 %
C_6	304	≥ 0	≥ 0 %
C_7	245	≥ 0	≥ 0 %
C_8	232	≥ 0	≥ 0 %
C_9	187	≥ 0	≥ 0 %

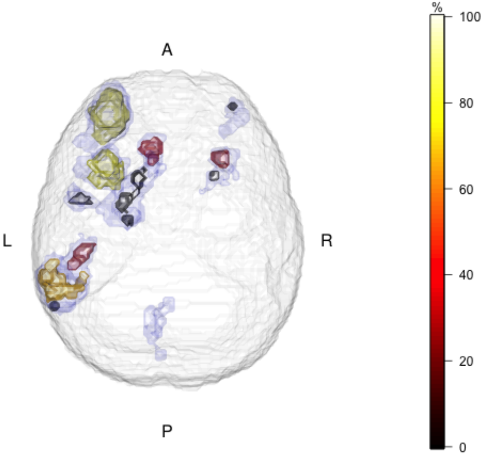
In a selected region S

	Detecting	*Counting*	*Identifying*
	Is there at least one active voxel?	How many voxels are active?	Which voxels are active?
Precision	*	**	***
Power	***	**	*
Method	Cluster-size	ARI	Voxel-wise

$p < t$



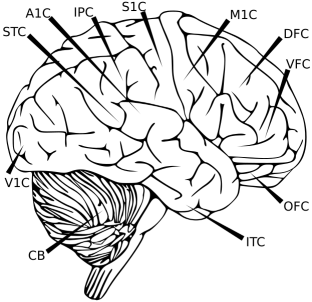
$p < t'$

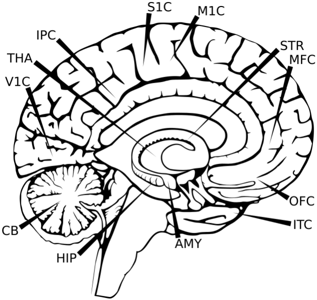


Sub-clusters

cluster	*threshold*	*size*	*# active*	*% active*
C_1	$p < t$	2191	624	29 %
1	$p < t'$	405	267	66 %
2	$p < t'$	133	31	23 %
3	$p < t'$	6	0	0 %
C_2	$p < t$	1835	847	46 %
1	$p < t'$	963	826	86 %
C_3	$p < t$	1400	454	32 %
1	$p < t'$	583	449	77 %
2	$p < t'$	4	0	0 %
3	$p < t'$	1	0	0 %
$\vdots$				

Anatomical regions





Closed testing

Elementary hypotheses

$$H_1, \ldots, H_m$$

Intersection hypotheses

$$H_S = \bigcap_{i \in S} H_i \qquad \forall\ S \subseteq M$$

Local tests

$$\phi_S = \mathbb{1}\{H_S \text{ rejected at level } \alpha\}$$

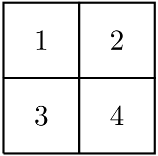
Closed testing adjustment

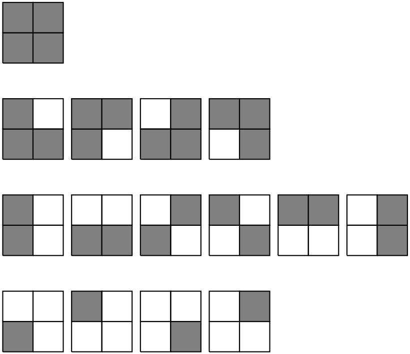
$$\tilde{\phi}_S = \min\left\{\phi_K : S \subseteq K \subseteq M\right\}$$

Lower bound

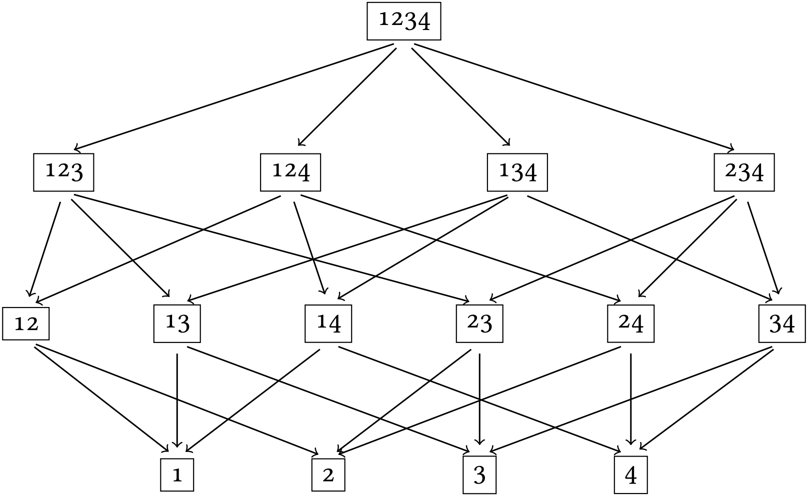
$$\underline{m}_1(S) = \min_{K \in 2^S}\left\{|S \setminus K| : \tilde{\phi}_K = 0\right\}$$

4-pixels brain

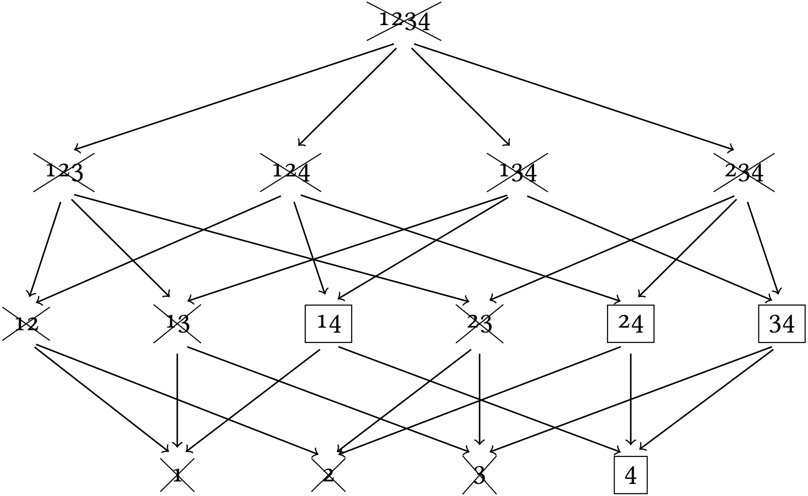




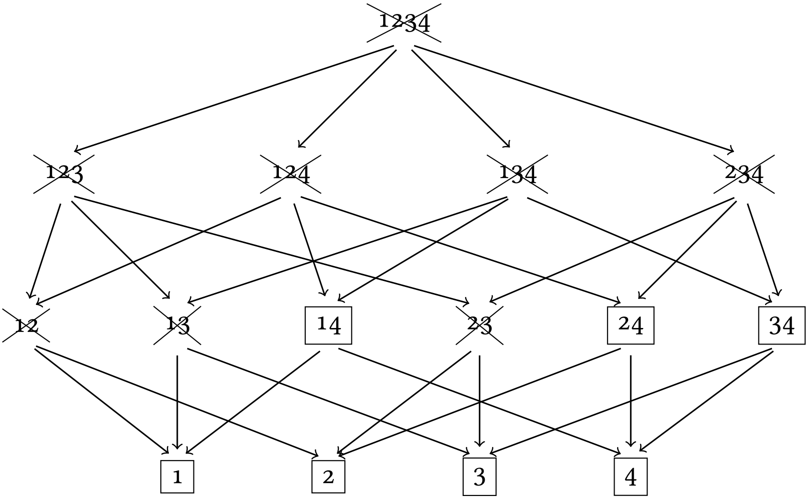
Hypotheses



Rejections



Closed testing rejections



Lower bounds

S	$m_1(S)$	$\pi_1(S)$
$\{1\}$	≥ 0	≥ 0 %
$\{2\}$	≥ 0	≥ 0 %
$\{3\}$	≥ 0	≥ 0 %
$\{4\}$	≥ 0	≥ 0 %
$\{1,2\}$	≥ 1	≥ 50 %
$\{1,3\}$	≥ 1	≥ 50 %
$\{1,4\}$	≥ 0	≥ 0 %
$\{2,3\}$	≥ 1	≥ 50 %
$\{2,4\}$	≥ 0	≥ 0 %
$\{3,4\}$	≥ 0	≥ 0 %
$\{1,2,3\}$	≥ 2	≥ 66.6 %
$\{1,2,4\}$	≥ 1	≥ 33.3 %
$\{1,3,4\}$	≥ 1	≥ 33.3 %
$\{2,3,4\}$	≥ 1	≥ 33.3 %
$\{1,2,3,4\}$	≥ 2	≥ 50 %

Closed testing bottleneck

Computation time is exponential in m

Shortcut

Computation time can be reduced to polynomial time
by specific choice of local tests

Shortcut with Simes local tests

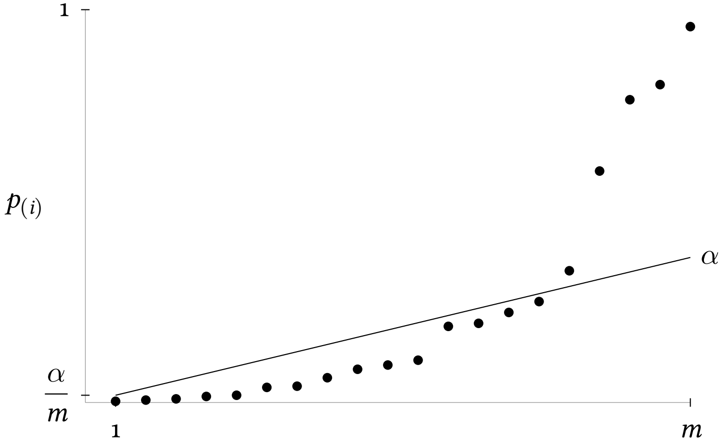
Simes local test for H_S

$$\phi_S = \mathbb{1}\left\{p_{(i\,:\,S)} \leq \frac{i\alpha}{|S|}\right\}$$

where $p_{(i\,:\,S)}$ is the ith smallest p-value in $\{p_i : i \in S\}$

Assumption: Simes inequality (1986) holds for null p-values

$$\mathrm{P}\left(\bigcap_{i=1}^{m_0}\left\{p_{(i\,:\,M_0)} > \frac{i\alpha}{m_0}\right\}\right) \geq 1 - \alpha$$



1
$p_{(i)}$
α
$\frac{\alpha}{m}$
1
m

Hommel (1988)

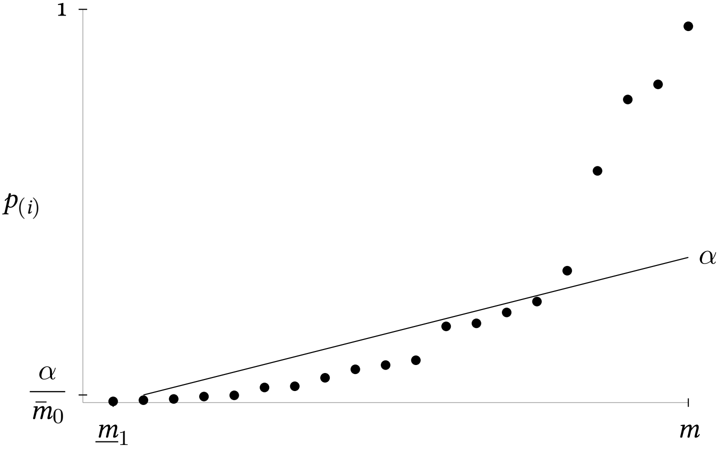
$$\bar{m}_0 = \max\left\{0 \leq k \leq m : \bigcap_{i=1}^{k} \left\{p_{(m-k+i)} > \frac{i\alpha}{k}\right\}\right\}$$

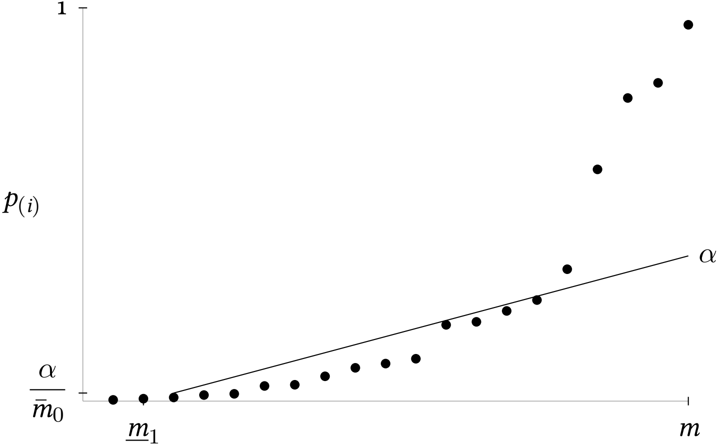
is a confidence upper bound for m_0: $\mathrm{P}(m_0 \leq \bar{m}_0) \geq 1 - \alpha$

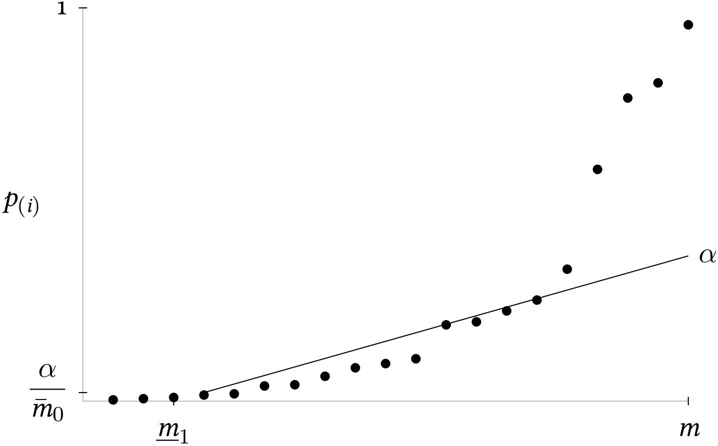
Also, the familywise error rejections are

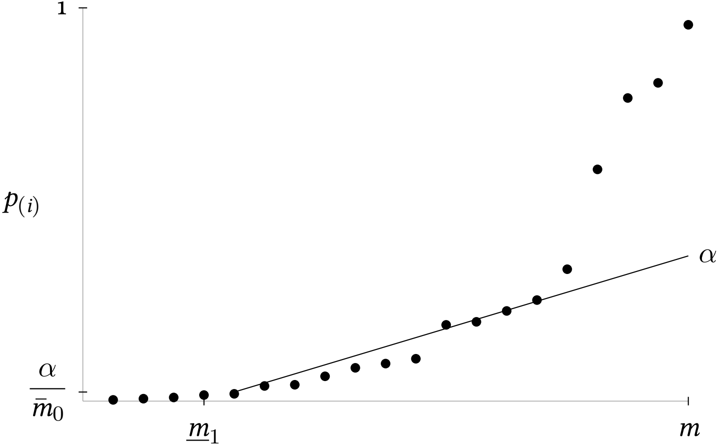
$$R = \left\{i : p_i \leq \frac{\alpha}{\bar{m}_0}\right\}$$

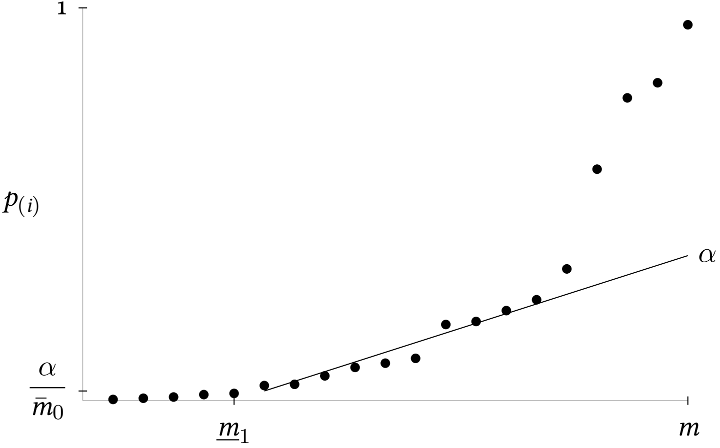
such that $\mathrm{P}(m_1(R) = |R|) \geq 1 - \alpha$


1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
$\underline{m}_1$
m

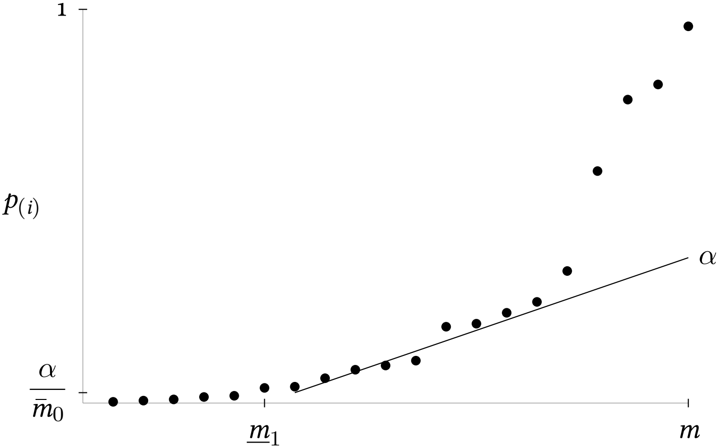

1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
$\underline{m}_1$
m


1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
$\underline{m}_1$
m

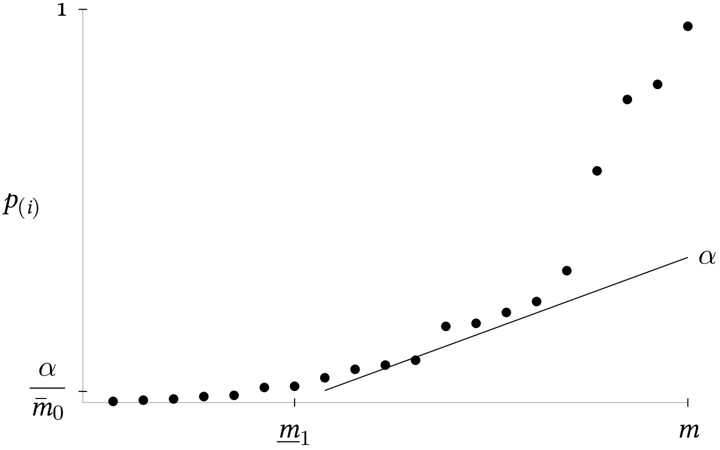

1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
$\underline{m}_1$
m

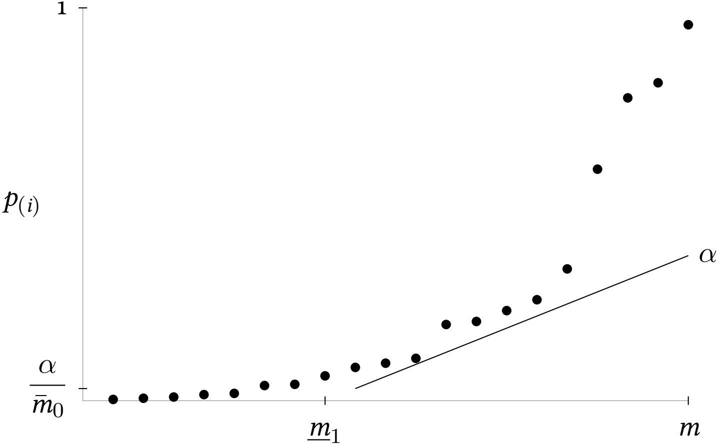


1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
$\underline{m}_1$
m

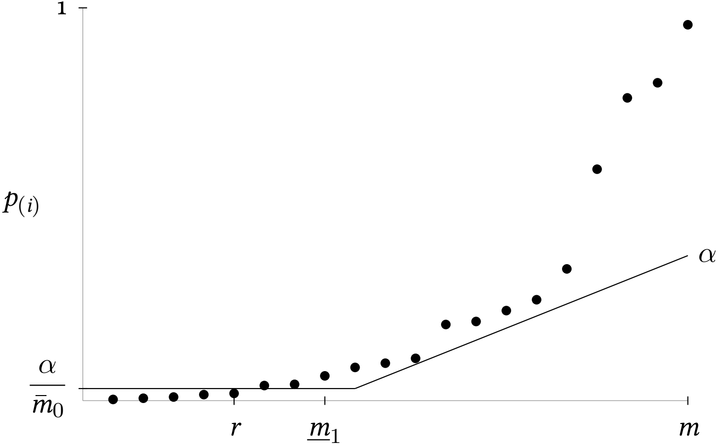


1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
$\underline{m}_1$
m


1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
$\underline{m}_1$
m

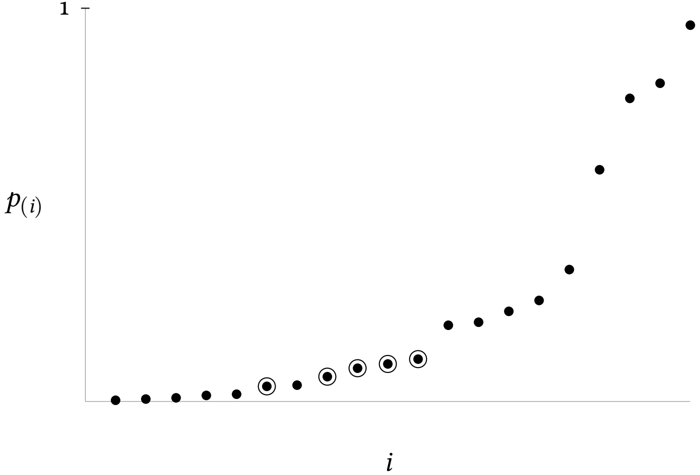


1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
$\underline{m}_1$
m

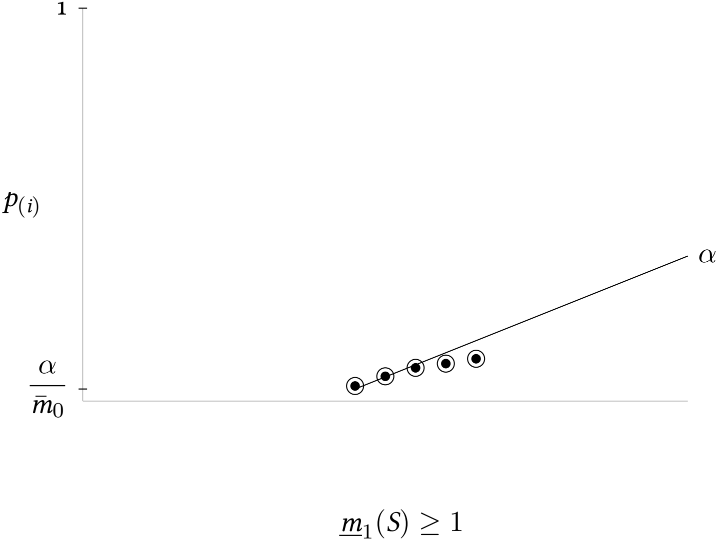

1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
r
$\underline{m}_1$
m

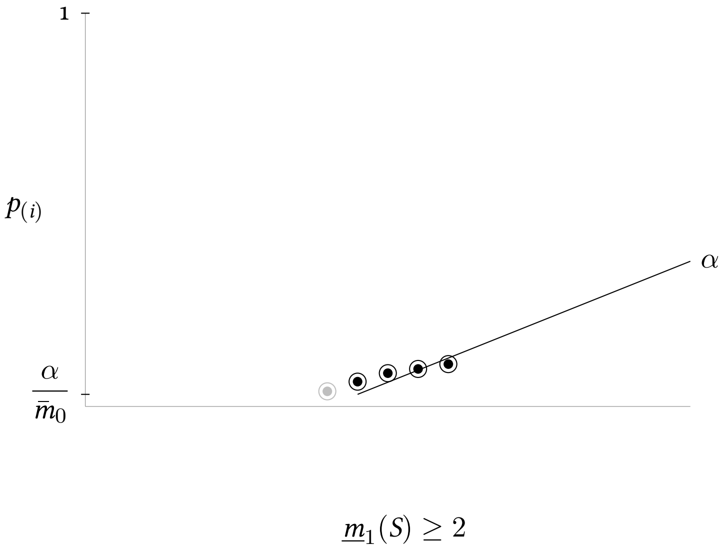
Goeman, Meijer, Krebs, Solari (2019)

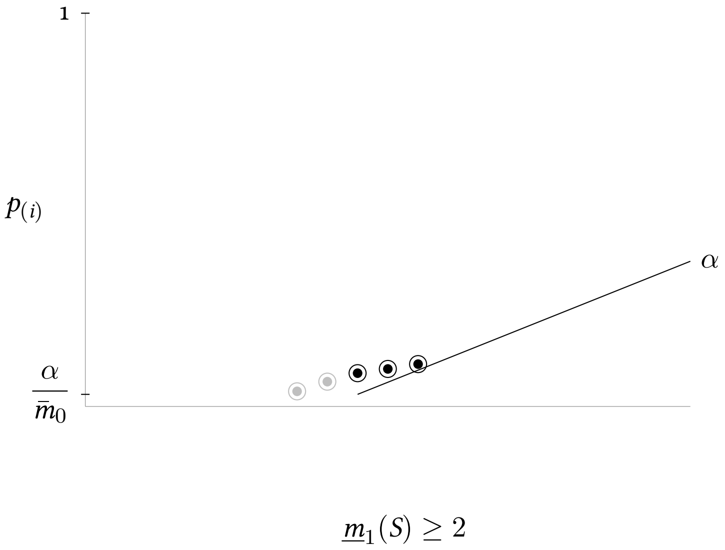
$$\underline{m}_1(S) = \min\left\{ 0 \leq k \leq |S| : \bigcap_{i=1}^{|S|-k} \left\{ p_{(k+i:S)} > \frac{i\alpha}{\bar{m}_0} \right\} \right\}$$



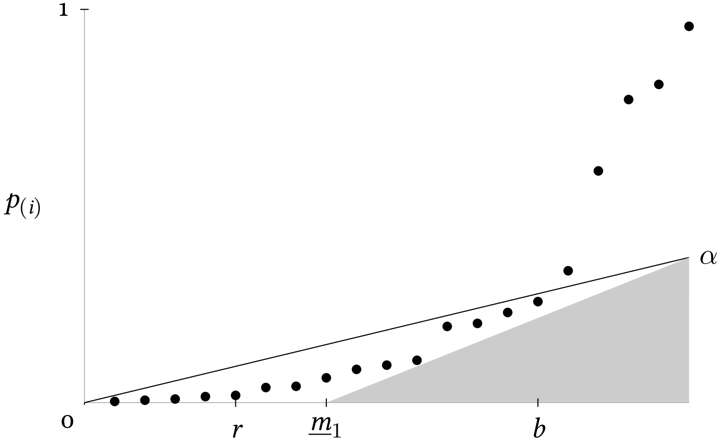
$S \subseteq B$


1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
$\underline{m}_1(S) \geq 1$


1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
$\underline{m}_1(S) \geq 2$


1
$p_{(i)}$
α
$\frac{\alpha}{\bar{m}_0}$
$\underline{m}_1(S) \geq 2$

Relationship to Benjamini-Hochberg



Let $b_q = \max\{1 \le b \le m : p_{(i)} \le \dfrac{iq}{m}\}$ and $B_q = \{i \in M : p_i \le p_{(b_q)}\}$ be the number and the index set of Benjamini-Hochberg rejections at level q

False discovery rate control at level q

$$\mathrm{FDR}(B_q) = \mathrm{E}\Big(\frac{|B_q \cap M_0|}{b_q}\Big) \le q$$

By changing the level from q to $\tilde{q} = q \cdot \frac{\alpha}{\bar{\pi}_0}$ we guarantee that $\mathrm{FDP}(B_{\tilde{q}})$ is bounded by q with $1-\alpha$ confidence, i.e.

$$\mathrm{P}(\pi_0(B_{\tilde{q}}) \le q) \ge 1 - \alpha$$

Scalability of power

ARI is between Bonferroni and Benjamini-Hochberg:

$$r \leq \underline{m}_1 \leq b$$

As the number of hypotheses $m \to \infty$,

$$\operatorname*{plim}_{m\to\infty} \frac{r}{m} = 0 \quad \text{and} \quad \operatorname*{plim}_{m\to\infty} \frac{b}{m} = c > 0$$

What about $\operatorname*{plim}_{m\to\infty} \underline{\pi}_1 = ?$

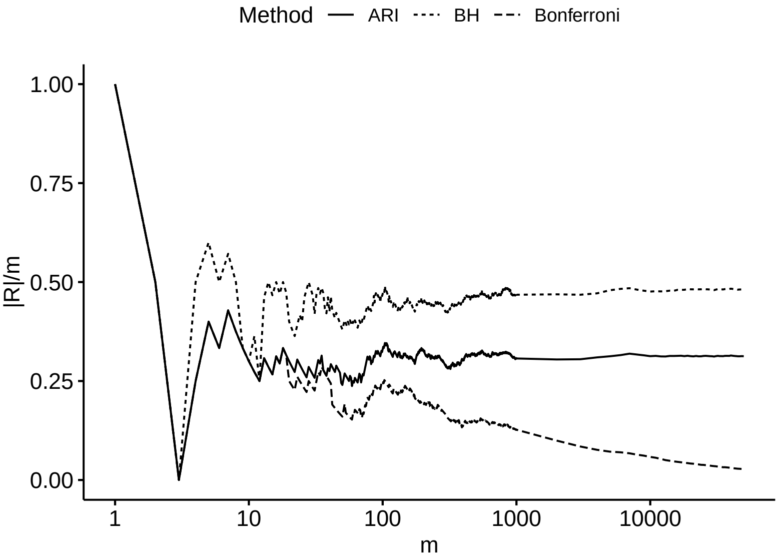
Assume that p-values are drawn from a mixture distribution

$$F(u) = \pi_0 u + \pi_1 F_1(u)$$

Lemma

$$\operatorname*{plim}_{m\to\infty} \underline{\pi}_1 = k > 0$$

if $F(u\alpha) > u$ for at least one $0 \leq u < 1$


Method
ARI
BH
Bonferroni
1.00
0.75
0.50
0.25
0.00
|R|/m
1
10
100
1000
10000
m

Consistency

Is $\underline{\pi}_1$ a consistent estimator of π_1?

Theorem

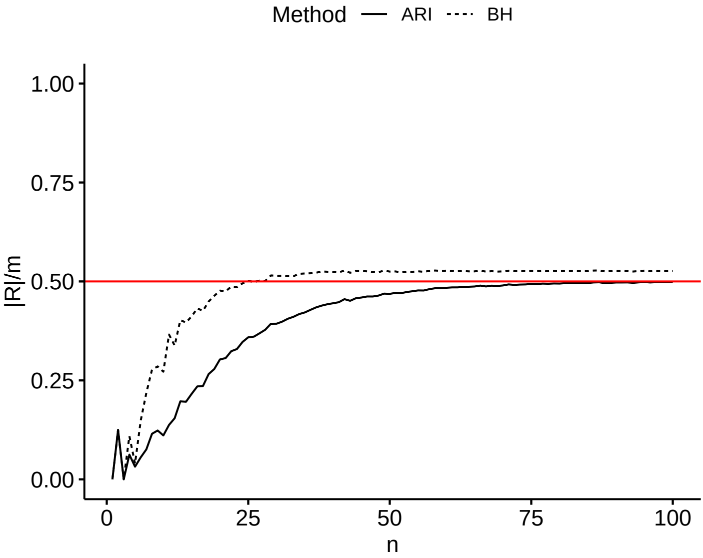
Letting $n \to \infty$ and $m \to \infty$, we have, for fixed $0 < \alpha < 1$:

$$\operatorname*{plim}_{(m,n)\to\infty} \underline{\pi}_1 = \pi_1$$

Compare to Benjamini-Hochberg:

$$\operatorname*{plim}_{(m,n)\to\infty} \frac{b}{m} = \frac{\pi_1}{1 - \alpha(1 - \pi_1)} > \pi_1$$

$m = n^3$



References

Multiple Testing for Exploratory Research
Goeman, Solari
Statistical Science 2011 26:584-597

All-Resolutions Inference for Brain Imaging
Rosenblatt, Finos, Weeda, Solari, Goeman
NeuroImage 2018 181:786-796

Simultaneous Control of All False Discovery Proportions in Large-Scale Multiple Hypothesis Testing
Goeman, Meijer, Krebs, Solari
Biometrika 2019 106:841-856