Introduction aux tests multiples

Guillermo Durand

M2 Maths & IA
Notes de cours

This version : January 28, 2025

Table of contents

Simple testing

- Data: $X = (X_1, \dots X_n)$ i.i.d.$\sim \mathcal{N}(\mu, 1)$, $\mu \geq 0$ unknown
- Question: is $\mu = 0$ (no signal) or > 0 (signal) ?
- Null hypothesis H_0: "$\mu = 0$" versus alternative H_1: "$\mu > 0$"
- Test statistic: $T(X) = n^{-1/2} \sum_{i=1}^{n} X_i$, under H_0 $T(X) \sim \mathcal{N}(0, 1)$
- $T(X)$ in the right tail of $\mathcal{N}(0, 1)$ $\Rightarrow$ unrealistic $\Rightarrow$ reject H_0
- So we reject H_0 if $T(X)$ is "large": the rejection region $\mathcal{R}$ is the event $\{T(X) > c\}$ with c to be determined

Test of level α

Choice of the rejecting threshold

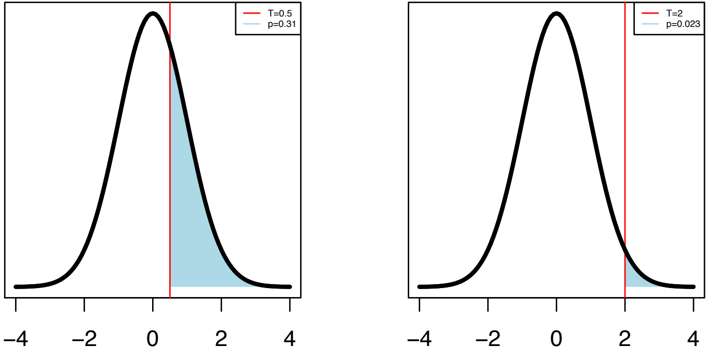
Goal:

Control the type I error = $\mathbb{P}$ of a wrong rejection = $\mathbb{P}$ of a false positive

- "level α" means type I error $= \mathbb{P}_{H_0}(T(X) > c) \leq \alpha$
- $\Rightarrow c \geq q^*_{1-\alpha}$ the $1-\alpha$ quantile of $\mathcal{N}(0,1)$

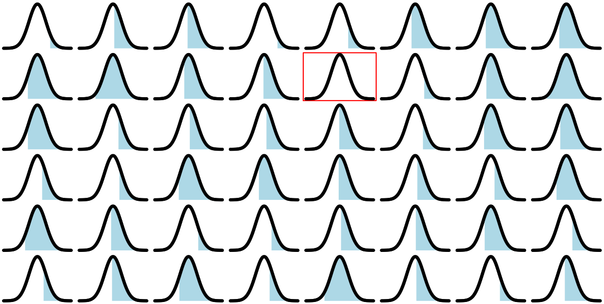
- Given type I control, how to reduce type II error?
  - Take the smallest c
  - $\Rightarrow \mathcal{R} = \{T(X) > q^*_{1-\alpha}\}$
- $\Leftrightarrow$ if the p-value $p(X) = \bar{\Phi}\left(T(X)\right) = 1 - \Phi\left(T(X)\right)$ is $\leq \alpha$
- "Proof": $\mathbb{P}_{H_0}\left(p(X) \leq \alpha\right) = \mathbb{P}_{H_0}\left(T(X) \geq q^*_{1-\alpha}\right) \leq \alpha$
- $p(X)$ is super-uniform under H_0, $p(X) = \mathbb{P}$ of observing an event at least as extreme as the one observed under the null

Test of level α



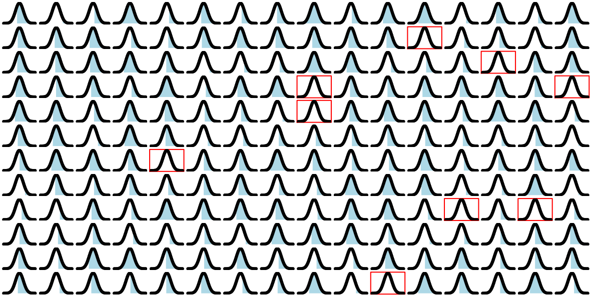
Multiple testing

- Now each X_i is a vector $(X_{i1}, \dots, X_{im}) \sim \mathcal{N}(\boldsymbol{\mu}, \mathrm{Id}_m)$ with $\boldsymbol{\mu} = (\mu_1, \dots, \mu_m) \in \mathbb{R}_+^m$
- m null hypotheses $H_{0,j}$: "$\mu_j = 0$" versus $H_{1,j}$: "$\mu_j > 0$"
- Because of independence, at least one false positive with $\mathbb{P} = 1 - (1-\alpha)^{m_0} \underset{m_0 \to \infty}{\longrightarrow} 1$
- $\mathbb{E}[|\,\mathrm{FP}\,|] = \alpha m_0$, $m_0 = |\{j : H_{0,j} \text{ is true}\}|$
- Example if $m = m_0 = 48$, $\alpha = 0.05$:



Multiple testing

- False positives explosion with m
- $m = m_0 = 192$, $\alpha = 0.05$:



Modern applications

- "Omic data": genomic, proteomic... but also fMRI, exoplanet detection...
- $m = 10^4, 10^5, 10^6$
- Too many false positives without correction

Table of contents

Formal setting

- Let $(\mathbb{X}, \mathcal{X}, \mathfrak{F})$ a statistical model and $(\Omega, \mathcal{A}, \mathbb{P})$ a probability space
- Hence $(\mathbb{X}, \mathcal{X})$ is a measurable space and $\mathfrak{F}$ is a family of probability measures defined on $\mathcal{X}$
- Data is a measurable $X : (\Omega, \mathcal{A}) \to (\mathbb{X}, \mathcal{X})$ with $X \sim P \in \mathfrak{F}$
- Other notation frequently used: $\mathbb{P}_X = \mathcal{L}(X) = X_{\#\mathbb{P}} = P$
- P unknown $\Rightarrow$ everything has to be valid $\forall P \in \mathfrak{F}$

Formal setting

- m null hypotheses $H_{0,i}$ and alternatives $H_{1,i}$ which are subsets of $\mathfrak{F}$
  - $H_{0,i} \cap H_{1,i} = \varnothing$
- $\mathcal{H}_0 = \mathcal{H}_0(P) = \{i : P \in H_{0,i}\}$: $i \in \mathcal{H}_0 \Leftrightarrow H_{0,i}$ is true
  - $\mathcal{H}_1(P) = \mathcal{H}_0(P)^c = \{i : P \in H_{1,i}\}$
- m p-values $p_i = p_i(X)$ such that $\mathcal{L}(p_i) \succeq \mathcal{U}([0,1])$ if $i \in \mathcal{H}_0$
  - Each p_i provides an α level test :

$$\forall \alpha \in [0,1], \forall P \in H_{0,i}, \forall X \sim P, \mathbb{P}(p_i \leq \alpha) \leq \alpha,$$

    or, in short, $\mathbb{P}_{X \sim P \in H_{0,i}}(p_i \leq \alpha) \leq \alpha$
  - 2 points of view: measurable application $p_i(\cdot) : \mathcal{X} \to [0,1]$, then applied to X, or random variable $p_i(X)$
- For every subset of hypotheses S, let $V(S) = |S \cap \mathcal{H}_0|$ the # of false positives (FP) in S

Formal setting

- Formally, a rejection procedure R is a measurable function $(\mathbb{X}, \mathcal{X}) \longrightarrow (\mathcal{P}(\llbracket 1, m \rrbracket), \mathcal{P}(\mathcal{P}(\llbracket 1, m \rrbracket)))$
- For a data point $X \sim P \in \mathfrak{F}$, the associated rejection set is $R(X)$ or R in short ($\Rightarrow$ small ambiguity), the rejected hypotheses are the $H_{0,i}$ such that $i \in R(X)$
- "Classic" MT goal: construct a rejection procedure R with a statistical guarantee on $V(R) \Leftrightarrow$ control of an error rate related to # of FP

A toy example

In this formal setting

- $(\Omega, \mathcal{A}, \mathbb{P})$ a probability space
- $(\mathbb{X}, \mathcal{X}) = (\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$: $X = (X_1, \dots, X_m)$
- $\mathcal{V}(1)$ the set of positive semidefinite matrices with 1's on the diagonal
- $\mathfrak{F} = \{\mathcal{N}(\boldsymbol{\mu}, \Sigma) : \forall j \in [\![1, m]\!], \mu_j \geq 0, \Sigma \in \mathcal{V}(1)\}$
- $H_{0,i} = \{\mathcal{N}(\boldsymbol{\mu}, \Sigma) : \mu_i = 0, \forall j \in [\![1, m]\!] \setminus \{i\}, \mu_j \geq 0, \Sigma \in \mathcal{V}(1)\}$
- $H_{1,i} = \{\mathcal{N}(\boldsymbol{\mu}, \Sigma) : \mu_i > 0, \forall j \in [\![1, m]\!] \setminus \{i\}, \mu_j \geq 0, \Sigma \in \mathcal{V}(1)\}$
- $p_i(X) = \bar{\Phi}(X_i) = 1 - \Phi(X_i)$

Generic construction of p-values

Following the idea of "the probability of an event at least as extreme as"

- Assume we have at hand m test statistics $T_1, \dots, T_m : \mathcal{X} \longrightarrow \mathbb{R}$
- for all $i \in [\![1, m]\!]$, we can let
  - $\hat{p}_i(X) = \sup_{P \in H_{0,i}} \mathbb{P}_{Z \sim P, Z \perp X}(T_i(Z) \geq T_i(X) | X) = \sup_{P \in H_{0,i}} P(T_i^{-1}([T_i(X), \infty[)) = \sup_{P \in H_{0,i}} (T_i)_{\#P}([T_i(X), \infty[)$, or
  - $\bar{p}_i(X) = \sup_{P \in H_{0,i}} \mathbb{P}_{Z \sim P, Z \perp X}(T_i(Z) \leq T_i(X) | X) = \sup_{P \in H_{0,i}} P(T_i^{-1}(]-\infty, T_i(X)])) = \sup_{P \in H_{0,i}} (T_i)_{\#P}(]-\infty, T_i(X)])$, or
  - $\breve{p}_i(X) = 2 \min(\hat{p}_i(X), \bar{p}_i(X))$
- Classical constructions for unilateral and bilateral tests, equivalent to UMP or UMP unbiased tests from Neyman-Pearson and Lehmann's theory for the appropriate choice of test statistics.
- Knowledge of $P \in H_{0,i}$, is required to compute $\hat{p}_i$, $\bar{p}_i$ or $\breve{p}_i$

Generic construction of p-values

Following the idea of "the probability of an event at least as extreme as"

Theorem

$\hat{p}_i$, $\bar{p}_i$, $\breve{p}_i$ all are appropriate p-values, that is, they are super-uniform under the null:
Denote by u the c.d.f. of $\mathcal{U}([0,1])$: $u(x) = 0 \vee (x \wedge 1)$. Let $Q \in H_{0,i}$, $X \sim Q$, then

$$\forall x \in \mathbb{R}, \mathbb{P}\left(\hat{p}_i(X) \leq x\right) \leq u(x), \tag{1}$$

$$\forall x \in \mathbb{R}, \mathbb{P}\left(\bar{p}_i(X) \leq x\right) \leq u(x), \tag{2}$$

$$\forall x \in \mathbb{R}, \mathbb{P}\left(\breve{p}_i(X) \leq x\right) \leq u(x). \tag{3}$$

Generic construction of p-values

Proof

- Only for (1), (2) and (3) left as an exercise
- $\hat{p}_i(X) \in [0,1]$ a.s. so we only need to check (1) for $x \in [0,1[$.
- (1) for $x \in]0,1[$ implies (1) for $x = 0$ by right-continuity of the c.d.f
- Let $x \in]0,1[$

$$\begin{aligned}
\mathbb{P}\left(\hat{p}_i(X) \leq x\right) &= \mathbb{P}\left(\sup_{P \in H_{0,i}} P(T_i^{-1}([T_i(X), \infty[)) \leq x\right) \\
&= \mathbb{P}\left(\bigcap_{P \in H_{0,i}} \left\{P(T_i^{-1}([T_i(X), \infty[)) \leq x\right\}\right) \\
&\leq \mathbb{P}\left(Q(T_i^{-1}([T_i(X), \infty[)) \leq x\right)
\end{aligned}$$

Generic construction of p-values

Proof

- Let F_i the c.d.f. of $T_i(X)$ and F_i^- its left-limit: $F_i^-(x) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} F_i(x - \varepsilon) = \mathbb{P}\,(T_i(X) < x)$
- $$\begin{aligned} \mathbb{P}\left(Q(T_i^{-1}([T_i(X), \infty[)) \leq x\right) &= \mathbb{P}\left(1 - Q(T_i^{-1}(] - \infty, T_i(X)[)) \leq x\right) \\ &= \mathbb{P}\left(1 - x \leq F_i^-(T_i(X))\right) \\ &= \mathbb{P}\left(T_i(X) \in \left(F_i^-\right)^{-1}([1 - x, 1])\right) \end{aligned}$$
- F_i^- is nondecreasing with limits 0 in $-\infty$ and 1 in ∞ so $\left(F_i^-\right)^{-1}([1 - x, 1])$ is an interval: $]a, \infty[$ of $[a, \infty[$ for some a.

Generic construction of p-values

Proof

- Case 1: $a \in \left(F_i^-\right)^{-1}([1-x,1])$ then

$$\begin{aligned}\mathbb{P}\left(Q(T_i^{-1}([T_i(X),\infty[)) \leq x\right) &= \mathbb{P}\left(T_i(X) \geq a\right)\\ &= 1 - F_i^-(a)\\ &\leq 1-(1-x)\\ &= x.\end{aligned}$$

- Case 2:

$$\begin{aligned}\mathbb{P}\left(Q(T_i^{-1}([T_i(X),\infty[)) \leq x\right) &= \mathbb{P}\left(T_i(X) > a\right)\\ &= 1 - F_i(a)\end{aligned}$$

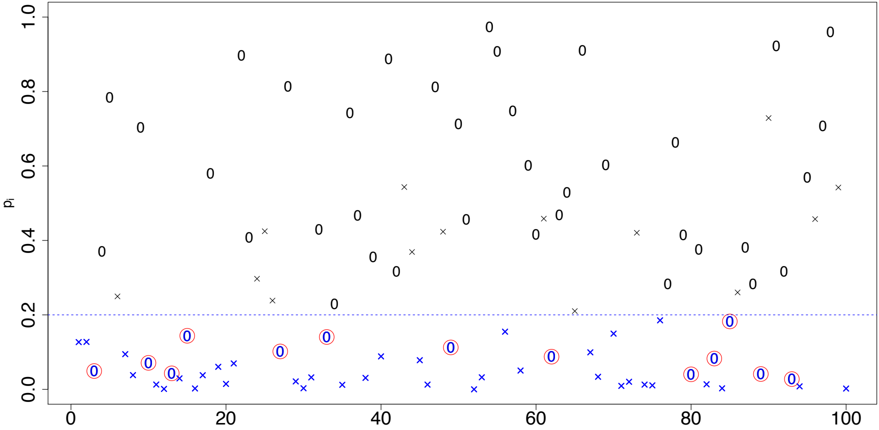
Generic construction of p-values
Proof

- $F_i(a)$ is the right-limit of F_i^- in a: $F_i(a) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} F_i^-(a+\varepsilon)$
- Note that $a + \varepsilon \in (F_i^-)^{-1}([1-x, 1])$ hence $F_i^-(a+\varepsilon) \geq 1 - x$
- $\varepsilon \to 0$: $F_i(a) \geq 1 - x$, which concludes case 2 and the proof for $\hat{p}_i(X)$
- For $\bar{p}_i(X)$, just use $-T_i$ as statistic and go back to previous case
- $\forall x \in \mathbb{R}, \mathbb{P}\left(2\min(\hat{p}_i(X), \bar{p}_i(X)) \leq x\right) \leq \left(2u\left(\frac{x}{2}\right)\right) \wedge 1 = u(x)$ □

Rejection set

Thresholding

- Main idea: small p-values = signal $(\mathcal{H}_1)$
- In the remainder of this course, $R(X) = \{i : p_i(X) \leq \hat{t}(X)\}, = R\left(\hat{t}\right)$ in short, with $\hat{t} = \hat{t}(X) = \hat{t}\left(p_1(X), \ldots, p_m(X)\right)$ a random threshold



Thresholding

Sorting p-values

- Sorted p-values: $p_{(1)} \leq \cdots \leq p_{(m)}$, $p_{(0)} = 0$ by convention
- $R(\hat{t}) = \{i : p_i \leq \hat{t}\} = \left\{i : p_i \leq p_{(\hat{k})}\right\}$, $\hat{k} = \max\{k : p_{(k)} \leq \hat{t}\}$

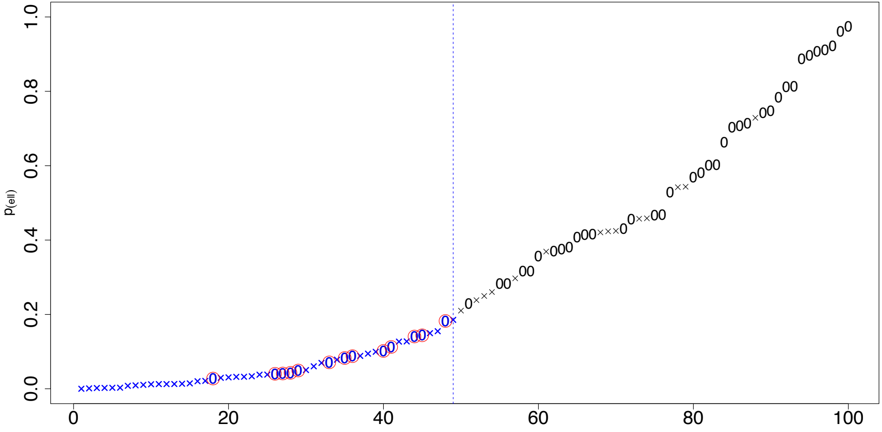


Table of contents

Family-Wise Error Rate (FWER)

- Probability to make at least one false positive
$$\mathrm{FWER}(R) = \mathbb{P}(V(R) > 0)$$
$$\mathrm{FWER}(R(\hat{t})) = \mathbb{P}(\exists i, i \in \mathcal{H}_0 : p_i \leq \hat{t})$$
- The probability is taken with respect to $P \in \mathfrak{F}$

- Philosophy : we don't want any false positive
- Choose $\hat{t}_\alpha$ such that $\mathrm{FWER}(R(\hat{t}_\alpha)) \leq \alpha$? $(\forall P \in \mathfrak{F})$

Bonferroni method

- Bonferroni method: $\hat{t}_\alpha^{\text{Bonf}} = \frac{\alpha}{m}$ and $R^{\text{Bonf}} = R\left(\hat{t}_\alpha^{\text{Bonf}}\right)$ [Bonferroni (1936)]

Theorem

For all $P \in \mathfrak{F}$,

$$\text{FWER}\left(R^{\text{Bonf}}\right) \leq \alpha$$

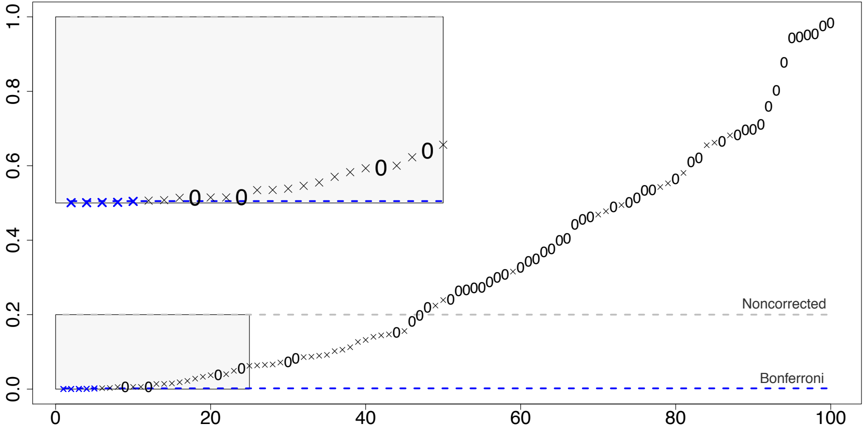
- Proof by union bound:

$$\begin{aligned}
\text{FWER}\left(R^{\text{Bonf}}\right) &= \mathbb{P}\left(\exists i, i \in \mathcal{H}_0 : p_i \leq \hat{t}_\alpha^{\text{Bonf}}\right) \\
&= \mathbb{P}\left(\bigcup_{i \in \mathcal{H}_0} \left\{p_i \leq \frac{\alpha}{m}\right\}\right) \leq \sum_{i \in \mathcal{H}_0} \mathbb{P}\left(p_i \leq \frac{\alpha}{m}\right) \\
&\leq \alpha \frac{m_0}{m} \leq \alpha \quad \square
\end{aligned}$$

- Adjusted p-value p_i^{adj}: smallest level that rejects $H_{0,i}$. For Bonferroni, $p_i^{adj} = 1 \wedge m p_i$

Illustration of Bonferroni method

$\alpha = 0.2$, $m = 100$



k-Family-Wise Error Rate (k-FWER)

A variant

$$k\text{-FWER}(R) = \mathbb{P}(V(R) \geq k)$$

- The probability is taken with respect to $P \in \mathfrak{F}$

- k-Bonferroni method: $\hat{t}_\alpha^{k\text{-Bonf}} = \frac{\alpha k}{m}$, $R^{k\text{-Bonf}} = R\left(\hat{t}_\alpha^{k\text{-Bonf}}\right)$

Theorem [Lehmann and Romano (2005)]

For all $P \in \mathfrak{F}$,

$$\text{FWER}\left(R^{k\text{-Bonf}}\right) \leq \alpha$$

k-Family-Wise Error Rate (k-FWER)

A variant

- Proof by Markov inequality:

$$\begin{aligned}
\mathbb{P}\left(V\left(R^{k\text{-Bonf}}\right) \geq k\right) &\leq \frac{\mathbb{E}\left[V\left(R\left(\frac{\alpha k}{m}\right)\right)\right]}{k} \\
&= \frac{\sum_{i \in \mathcal{H}_0} \mathbb{E}\left[\mathbb{1}_{\left\{p_i \leq \frac{\alpha k}{m}\right\}}\right]}{k} \\
&\leq \frac{1}{k} \sum_{i \in \mathcal{H}_0} \frac{\alpha k}{m} \\
&\leq \alpha \frac{m_0}{m} \leq \alpha \quad \square
\end{aligned}$$

False Discovery Rate (FDR)

FWER too stringent

Especially for some settings where:

- m is large
- we want a lot of detections,
- and we can allow some false positive to do so

False Discovery Proportion (FDP) and FDR

$$\mathrm{FDP}(R) = \frac{V(R)}{|R| \vee 1} \text{ (random variable)}$$

$$\mathrm{FDR}(R) = \mathbb{E}\left[\mathrm{FDP}(R)\right] \ (\in [0,1])$$

- The expectation is taken with respect to $P \in \mathfrak{F}$

- Choose a $\hat{t}_\alpha$ such that $\mathrm{FDR}(R(\hat{t}_\alpha)) \leq \alpha$?

False Discovery Rate (FDR)

Estimating the FDP to derive a procedure

$$
\begin{aligned}
\mathrm{FDP}(R(t)) &= \frac{\sum_{i\in\mathcal{H}_0} \mathbb{1}_{\{p_i\leq t\}}}{|R(t)|\vee 1}\\
&= m\frac{\frac{1}{m}\sum_{i\in\mathcal{H}_0} \mathbb{1}_{\{p_i\leq t\}}}{|R(t)|\vee 1}\\
&\leq m\frac{\frac{1}{m_0}\sum_{i\in\mathcal{H}_0} \mathbb{1}_{\{p_i\leq t\}}}{|R(t)|\vee 1}
\end{aligned}
$$

- Main idea: if m_0 large, $\frac{1}{m_0}\sum_{i\in\mathcal{H}_0} \mathbb{1}_{\{p_i\leq t\}} \lesssim t$ by law of large numbers and super-uniformity
- $\Rightarrow \widehat{\mathrm{FDP}}^{\mathrm{BH}}(t) = \frac{mt}{|R(t)|\vee 1}$
- $\Rightarrow \hat{t}_\alpha^{heur} = \sup\left\{t\in[0,1] : \widehat{\mathrm{FDP}}^{\mathrm{BH}}(t)\leq \alpha\right\} =$ $\sup\left\{t\in[0,1] : \frac{\alpha}{m}\left(|R(t)|\vee 1\right)\geq t\right\}$

Benjamini-Hochberg procedure (BH)

[Benjamini and Hochberg (1995)]

- Sorted p-values: $p_{(1)} \leq \cdots \leq p_{(m)}$, $p_{(0)} = 0$ by convention
- Traditional definition : $\hat{k}^{\text{BH}} = \max\left\{k \in [\![1, m]\!] : p_{(k)} \leq \alpha \frac{k}{m}\right\}$, $\hat{k}^{\text{BH}} = 0$ if set empty, $\hat{t}_\alpha^{\text{BH}} = \alpha \frac{\hat{k}^{\text{BH}}}{m}$
- Slightly equivalent modification: $\hat{k}^{\text{BH}} = \max\left\{k \in [\![0, m]\!] : p_{(k)} \leq \alpha \frac{k \vee 1}{m}\right\}$, $\hat{t}_\alpha^{\text{BH}} = \alpha \frac{\hat{k}^{\text{BH}} \vee 1}{m}$, $R^{\text{BH}} = R\left(\hat{t}_\alpha^{\text{BH}}\right)$
- (really the same except $\hat{t}_\alpha^{\text{BH}} = \frac{\alpha}{m}$ if $\hat{k}^{\text{BH}} = 0$, gives the same R^{BH})
- Adjusted p-values : $p_{(i)}^{adj} = 1 \wedge \min_{j \geq i} \frac{m p_{(j)}}{j}$

Lemma

$\left|R^{\text{BH}}\right| = \hat{k}^{\text{BH}}$ and $\hat{t}_\alpha^{heur} = \hat{t}_\alpha^{\text{BH}}$

This lemma generalizes in more complex settings where it is useful, see [Roquain and Wiel (2009)], [Durand (2019)], and the following

Illustration of BH method

$\alpha = 0.2$, $m = 100$

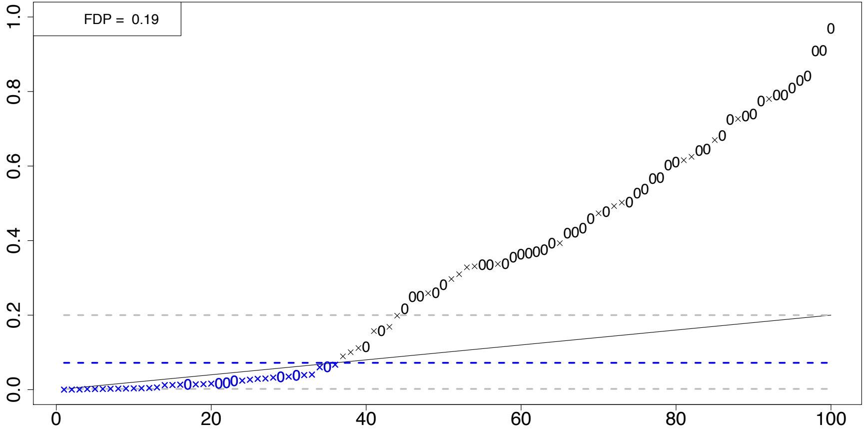
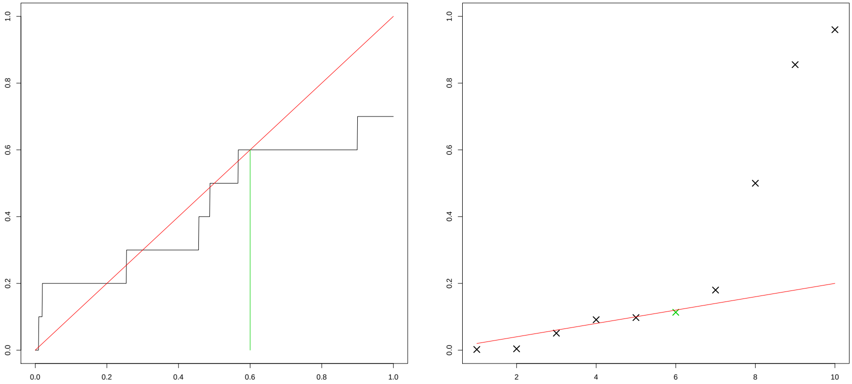


Illustration of BH method

$m = 10$



Benjamini-Hochberg procedure (BH)

Proof of the Lemma

- If $\hat{k}^{\text{BH}} \geq 1$,

$$p_{(\hat{k}^{\text{BH}})} \leq \alpha \frac{\hat{k}^{\text{BH}}}{m} \Rightarrow \forall i \in \left[\!\left[1, \hat{k}^{\text{BH}} \right]\!\right], p_{(i)} \leq \alpha \frac{\hat{k}^{\text{BH}}}{m}$$

$$\Rightarrow \left| R\left(\hat{t}_\alpha^{\text{BH}}\right) \right| = \left| R\left(\alpha \frac{\hat{k}^{\text{BH}} \vee 1}{m}\right) \right| = \left| R\left(\alpha \frac{\hat{k}^{\text{BH}}}{m}\right) \right| \geq \hat{k}^{\text{BH}}$$

- Obvious if $\hat{k}^{\text{BH}} = 0$
- Reductio ad absurdum: if $\left| R\left(\hat{t}_\alpha^{\text{BH}}\right) \right| \geq \hat{k}^{\text{BH}} + 1$ then necessarily $p_{(\hat{k}^{\text{BH}}+1)} \leq \hat{t}_\alpha^{\text{BH}} = \alpha \frac{\hat{k}^{\text{BH}} \vee 1}{m} \leq \alpha \frac{\left(\hat{k}^{\text{BH}}+1\right) \vee 1}{m}$ which contradicts the definition of $\hat{k}^{\text{BH}}$

Benjamini-Hochberg procedure (BH)

Proof of the Lemma

- Supremum well-defined because non-empty set, $0 \in$ it
- Let $\widehat{G}(t) = \frac{\alpha}{m}(|R(t)| \vee 1)$: nondecreasing and $[0,1] \to [0,1]$
- Let $t_n \nearrow \hat{t}_\alpha^{heur}$, such that $\widehat{G}(t_n) \geq t_n$,

$$\widehat{G}\left(\hat{t}_\alpha^{heur}\right) \geq \widehat{G}(t_n) \geq t_n \underset{n\to\infty}{\longrightarrow} t_\alpha^{heur}$$

  so $\hat{t}_\alpha^{heur}$ is a max
- So $\widehat{G}\left(\widehat{G}\left(\hat{t}_\alpha^{heur}\right)\right) \geq \widehat{G}\left(\hat{t}_\alpha^{heur}\right)$ so by def $\widehat{G}\left(\hat{t}_\alpha^{heur}\right) \leq \hat{t}_\alpha^{heur}$
- $\Rightarrow \hat{t}_\alpha^{heur} = \widehat{G}\left(\hat{t}_\alpha^{heur}\right)$

Benjamini-Hochberg procedure (BH)

Proof of the Lemma

- First note that $p_{(|R(t)|)} \leq t$ always
- $p_{\left(\left|R\left(\hat{t}_\alpha^{heur}\right)\right|\right)} \leq \hat{t}_\alpha^{heur} = \widehat{G}\left(\hat{t}_\alpha^{heur}\right) = \frac{\alpha}{m}\left(\left|R\left(\hat{t}_\alpha^{heur}\right)\right| \vee 1\right)$
- $\Rightarrow \left|R\left(\hat{t}_\alpha^{heur}\right)\right| \leq \hat{k}^{\mathrm{BH}} \Rightarrow \hat{t}_\alpha^{heur} \leq \hat{t}_\alpha^{BH}$ by def of $\hat{k}^{\mathrm{BH}}$ and nondecreasing composition
- $\widehat{G}\left(\hat{t}_\alpha^{\mathrm{BH}}\right) = \frac{\alpha}{m}\left(\left|R\left(\hat{t}_\alpha^{\mathrm{BH}}\right)\right| \vee 1\right) = \frac{\alpha}{m}\left(\hat{k}^{\mathrm{BH}} \vee 1\right) = \hat{t}_\alpha^{\mathrm{BH}}$ by previous result
- $\Rightarrow \hat{t}_\alpha^{\mathrm{BH}} \leq \hat{t}_\alpha^{heur}$ by def of $\hat{t}_\alpha^{heur}$ □

Benjamini-Hochberg procedure (BH)

Proof of the adjusted p-value formula

$$
\begin{aligned}
p_{(i)} \leq \alpha \frac{\hat{k}^{\mathrm{BH}}}{m} &\Leftrightarrow \hat{k}^{\mathrm{BH}} \geq i \\
&\Leftrightarrow \exists j \geq i, p_{(j)} \leq \alpha \frac{j}{m} \\
&\Leftrightarrow \exists j \geq i, \frac{m p_{(j)}}{j} \leq \alpha \\
&\Leftrightarrow \min_{j \geq i} \frac{m p_{(j)}}{j} \leq \alpha \quad \square
\end{aligned}
$$

Benjamini-Hochberg procedure (BH)

- What about FDR control?

Theorem [Benjamini and Hochberg (1995)]

Assume that for all $P \in \mathfrak{F}$, the $(p_i)_{i \in \mathcal{H}_0}$ are independent, and they are independent from the $(p_i)_{i \in \mathcal{H}_1}$.
Then for all $P \in \mathfrak{F}$,

$$\mathrm{FDR}\left(R^{\mathrm{BH}}\right) \leq \alpha \frac{m_0}{m} \leq \alpha$$

Interlude

Step-up and step-down procedures

- Given a nondecreasing nonnegative sequence $\tau = (\tau_1, \ldots, \tau_m)$, the respective step-up and step-down procedures associated with τ are:

$$\begin{array}{rcccl} R^{\mathrm{SU}}(\tau) & = & R\left(\tau_{\hat{k}^{\mathrm{SU}}}\right) & = & \left\{i \in [\![1, m]\!] : p_i \leq \tau_{\hat{k}^{\mathrm{SU}}}\right\} \\ R^{\mathrm{SD}}(\tau) & = & R\left(\tau_{\hat{k}^{\mathrm{SD}}}\right) & = & \left\{i \in [\![1, m]\!] : p_i \leq \tau_{\hat{k}^{\mathrm{SD}}}\right\} \end{array}$$

  with

$$\begin{array}{rcl} \hat{k}^{\mathrm{SU}} & = & \max\left\{0 \leq k \leq m : p_{(k)} \leq \tau_k\right\} \\ \hat{k}^{\mathrm{SD}} & = & \max\left\{0 \leq k \leq m : \forall k' \leq k,\ p_{(k')} \leq \tau_{k'}\right\} \end{array}$$

  - Where we let $\tau_0 = \tau_1$ by convention
  - So $\tau_k = \tau_{k \vee 1}$, $\forall 0 \leq k \leq m$
  - Recall that $p_{(0)} = 0$ by convention too
- The τ_k are called the critical values
- The τ_k can be random as long as they stay nondecreasing nonnegative

Interlude

Step-up and step-down procedures

- With same proof as before:
  - $\left|R^{\mathrm{SU}}(\tau)\right| = \hat{k}^{\mathrm{SU}}$
  - $\left|R^{\mathrm{SD}}(\tau)\right| = \hat{k}^{\mathrm{SD}}$
- $R\left(\hat{t}_\alpha^{\mathrm{Bonf}}\right) = R^{\mathrm{SU}}(\tau) = R^{\mathrm{SD}}(\tau)$ with $\tau_i = \frac{\alpha}{m}$
- $R\left(\hat{t}_\alpha^{k\text{-Bonf}}\right) = R^{\mathrm{SU}}(\tau) = R^{\mathrm{SD}}(\tau)$ with $\tau_i = \alpha\frac{k}{m}$
- $R\left(\hat{t}_\alpha^{\mathrm{BH}}\right) = R^{\mathrm{SU}}(\tau)$ with $\tau_i = \alpha\frac{i}{m}$
- Remark: for a fixed τ, $R^{\mathrm{SU}}(\tau)$ is uniformly better than $R^{\mathrm{SD}}(\tau)$, but sometimes SD allows FDR control for some larger τ than SU, see [Döhler, Durand, and Roquain (2018)] and the following

Benjamini-Hochberg procedure (BH)

Proof of the Theorem

- First lemma on SU procedures: let $i \in [\![1, m]\!]$ and the SU procedure applied to all p-values except p_i, with $\tau^{-i} = (\tau_1^{-i}, \ldots, \tau_{m-1}^{-i}) = (\tau_2, \ldots, \tau_m)$
- Let $p_{(1)}^{-i} \leq \ldots \leq p_{(m-1)}^{-i}$ be the ordered p-values of this procedure
- Let $\hat{k}^{-i} = \max \left\{ k : p_{(k)}^{-i} \leq \tau_k^{-i} \right\}$ be the number of rejections of this procedure
- Then $\hat{k}^{-i} \geq \hat{k}^{\mathrm{SU}} - 1$ and the three following assertions are equivalent:

  (i) $p_i \leq \tau_{\hat{k}^{\mathrm{SU}}}$.
  (ii) $p_i \leq \tau_{\hat{k}^{-i}+1}$.
  (iii) $\hat{k}^{-i} = \hat{k}^{\mathrm{SU}} - 1$.

Benjamini-Hochberg procedure (BH)

Proof of the first Lemma

- Assume $\hat{k}^{\mathrm{SU}} \geq 2$, otherwise $\hat{k}^{-i} \geq \hat{k}^{\mathrm{SU}} - 1$ is trivial
- Note that $p^{-i}_{(\hat{k}^{\mathrm{SU}}-1)}$ is always equal to $p_{(\hat{k}^{\mathrm{SU}}-1)}$ or $p_{(\hat{k}^{\mathrm{SU}})}$, so $p^{-i}_{(\hat{k}^{\mathrm{SU}}-1)} \leq \tau_{\hat{k}^{\mathrm{SU}}} = \tau^{-i}_{\hat{k}^{\mathrm{SU}}-1}$ and $\hat{k}^{-i} \geq \hat{k}^{\mathrm{SU}} - 1$, by def of $\hat{k}^{-i}$
- *(i)*$\Rightarrow$*(ii)* τ nondecreasing and $\hat{k}^{\mathrm{SU}} \leq \hat{k}^{-i} + 1$
- *(ii)*$\Rightarrow$*(iii)* By def of $\hat{k}^{-i}$, $p^{-i}_{(1)}, \ldots, p^{-i}_{(\hat{k}^{-i})} \leq \tau^{-i}_{\hat{k}^{-i}} = \tau_{\hat{k}^{-i}+1}$. So if $p_i \leq \tau_{\hat{k}^{-i}+1}$, at least $\hat{k}^{-i} + 1$ p-values are $\leq \tau_{\hat{k}^{-i}+1}$, so $p_{(\hat{k}^{-i}+1)} \leq \tau_{\hat{k}^{-i}+1}$ and $\hat{k}^{\mathrm{SU}} \geq \hat{k}^{-i} + 1$ by def of $\hat{k}^{\mathrm{SU}}$
- *(iii)*$\Rightarrow$*(i)*
  - If $\hat{k}^{-i} = \hat{k}^{\mathrm{SU}} - 1$ then $\tau_{\hat{k}^{\mathrm{SU}}} = \tau^{-i}_{\hat{k}^{-i}}$
  - $p^{-i}_{(\hat{k}^{-i}+1)} > \tau^{-i}_{(\hat{k}^{-i}+1)} \geq \tau^{-i}_{(\hat{k}^{-i})} = \tau_{\hat{k}^{\mathrm{SU}}}, \cdots, p^{-i}_{(m-1)} > \tau_{\hat{k}^{\mathrm{SU}}}$
  - $\Rightarrow m - 1 - \hat{k}^{-i} = m - \hat{k}^{\mathrm{SU}}$ p-values that are not p_i are $> \tau_{\hat{k}^{\mathrm{SU}}}$, there must be $m - \hat{k}^{\mathrm{SU}}$ in total that are $> \tau_{\hat{k}^{\mathrm{SU}}}$, hence $p_i \leq \tau_{\hat{k}^{\mathrm{SU}}}$ $\square$

Benjamini-Hochberg procedure (BH)

Proof of the Theorem

- Second lemma on SU procedures:
  $\{p_i \leq \tau_{\hat{k}^{\text{SU}}},\ \hat{k}^{\text{SU}} = k\} = \{p_i \leq \tau_k,\ \hat{k}^{-i} = k-1\}$
- Decorrelates p_i and the rest of the p-values! Allows to use the independence assumption favorably
- Proof:

$$\begin{aligned}
p_i \leq \tau_{\hat{k}^{\text{SU}}},\ \hat{k}^{\text{SU}} = k &\iff p_i \leq \tau_{\hat{k}^{\text{SU}}},\ \hat{k}^{-i} = \hat{k}^{\text{SU}} - 1,\ \hat{k}^{\text{SU}} = k && \textit{(i)}\Rightarrow\textit{(iii)} \\
&\iff p_i \leq \tau_{\hat{k}^{\text{SU}}},\ \hat{k}^{-i} = k - 1 && \textit{(i)}\Rightarrow\textit{(iii)} \\
&\iff p_i \leq \tau_{\hat{k}^{-i}+1},\ \hat{k}^{-i} = k - 1 && \textit{(i)}\Leftrightarrow\textit{(ii)} \\
&\iff p_i \leq \tau_k,\ \hat{k}^{-i} = k - 1.
\end{aligned}$$

□

Benjamini-Hochberg procedure (BH)

- Let $X \sim P \in \mathfrak{F}$
- For $i \in \mathcal{H}_0$ let $\hat{k}^{-i}$ as in the Lemmas with $\tau = \left(\frac{\alpha k}{m}\right)_{k \in [\![1,m]\!]}$

$$
\begin{aligned}
\mathrm{FDR}\left(R^{\mathrm{BH}}\right) &= \mathbb{E}\left[\frac{\sum_{i \in \mathcal{H}_0} \mathbb{1}_{\left\{p_i \leq \alpha \frac{\hat{k}^{\mathrm{BH}}}{m}\right\}}}{\hat{k}^{\mathrm{BH}} \vee 1}\right] \\
&= \mathbb{E}\left[\sum_{k=1}^{m} \frac{1}{k} \sum_{i \in \mathcal{H}_0} \mathbb{1}_{\left\{p_i \leq \alpha \frac{\hat{k}^{\mathrm{BH}}}{m}\right\}} \mathbb{1}_{\left\{\hat{k}^{\mathrm{BH}} = k\right\}}\right] \\
&= \sum_{i \in \mathcal{H}_0} \sum_{k=1}^{m} \frac{1}{k} \mathbb{P}\left(p_i \leq \alpha \frac{\hat{k}^{\mathrm{BH}}}{m}, \hat{k}^{\mathrm{BH}} = k\right) \\
&= \sum_{i \in \mathcal{H}_0} \sum_{k=1}^{m} \frac{1}{k} \mathbb{P}\left(p_i \leq \alpha \frac{k}{m}, \hat{k}^{-i} = k-1\right) \\
&= \sum_{i \in \mathcal{H}_0} \sum_{k=1}^{m} \frac{1}{k} \mathbb{P}\left(p_i \leq \alpha \frac{k}{m}\right) \mathbb{P}\left(\hat{k}^{-i} = k-1\right)
\end{aligned}
$$

Benjamini-Hochberg procedure (BH)

Proof of the Theorem

$$\begin{aligned}
\mathrm{FDR}\left(R^{\mathrm{BH}}\right) &\leq \sum_{i \in \mathcal{H}_0} \sum_{k=1}^{m} \frac{1}{k} \alpha \frac{k}{m} \mathbb{P}\left(\hat{k}^{-i} = k-1\right) \\
&= \frac{\alpha}{m} \sum_{i \in \mathcal{H}_0} \sum_{k=1}^{m} \mathbb{P}\left(\hat{k}^{-i} = k-1\right) \\
&= \frac{\alpha}{m} \sum_{i \in \mathcal{H}_0} 1 \\
&= \alpha \frac{m_0}{m} \quad \square
\end{aligned}$$

- Note that the only inequality is an equality if $p_i \sim \mathcal{U}([0,1])$ for all $i \in \mathcal{H}_0 \Rightarrow$ a stronger result when uniformity

Benjamini-Hochberg procedure (BH)

Can we do better than the independent case?

Some dependence conditions [Benjamini and Yekutieli (2001)][Blanchard and Roquain (2008)]

- $D \subseteq [0,1]^m$ is nondecreasing if $(x_1, \dots, x_m) \in D$ and $x_i \leq y_i \ \forall i \in [\![1, m]\!]$ imply $(y_1, \dots, y_m) \in D$
- Positive Regression Dependent on each one from a Subset (PRDS) : let $S \subseteq [\![1, m]\!]$ the subset,

$$\forall D \subseteq [0,1]^m \nearrow, \forall i \in S, \exists f_{i,D} \nearrow, \mathbb{P}\left(\boldsymbol{p} \in D | p_i\right) = f_{i,D}(p_i) \text{ a.s.}$$

- weak Positive Regression Dependent on each one from a Subset (wPRDS) : let $S \subseteq [\![1, m]\!]$ the subset,

$$\forall D \subseteq [0,1]^m \nearrow, \forall i \in S, g_{i,D} : u \mapsto \mathbb{P}\left(\boldsymbol{p} \in D | p_i \leq u\right)$$

is nondecreasing on $\{u \in [0,1] : \mathbb{P}\left(p_i \leq u\right) > 0\}$

wPRDS is indeed weaker than PRDS

Proposition [Blanchard and Roquain (2008)]

If the p-values are PRDS on S, they are wPRDS on S.

- Fix D and $i \in S$ once and for all
- Notation: $\forall B \in \mathcal{A}, \mathbb{P}(B) > 0$, $\mathbb{P}_B : A \mapsto \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ and $\mathbb{P}_u = \mathbb{P}_{\{p_i \leq u\}}, \forall u \in [0,1] : \mathbb{P}(p_i \leq u) > 0$
- Likewise, $\mathbb{E}_B$ and $\mathbb{E}_u$
- 2 lemmas:
    - $\mathbb{P}_B \ll \mathbb{P}$ and $\frac{\mathrm{d}\mathbb{P}_B}{\mathrm{d}\mathbb{P}} : \omega \mapsto \frac{\mathbb{1}_B(\omega)}{\mathbb{P}(B)}$
    - $\mathbb{P}_u(\boldsymbol{p} \in D | p_i) = \mathbb{P}(\boldsymbol{p} \in D | p_i) = f_{i,D}(p_i)$ a.s.
    - (The second one is also true if conditioning on $B \in \sigma(p_i)$ instead of $\{p_i \leq u\}$)

wPRDS is indeed weaker than PRDS

Proof of first Lemma

$$\begin{aligned}
\mathbb{P}_B(A) &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \\
&= \int \frac{\mathbb{1}_{A \cap B}(\omega)}{\mathbb{P}(B)} \mathrm{d}\mathbb{P}(\omega) \\
&= \int \mathbb{1}_A(\omega) \frac{\mathbb{1}_B(\omega)}{\mathbb{P}(B)} \mathrm{d}\mathbb{P}(\omega) \quad \square
\end{aligned}$$

wPRDS is indeed weaker than PRDS

Proof of second Lemma

- $\mathbb{P}_u(\boldsymbol{p} \in D | p_i) = \mathbb{E}_u\left[\mathbb{1}_{\{\boldsymbol{p}\in D\}} \middle| p_i\right]$ and $\mathbb{P}(\boldsymbol{p} \in D | p_i) = \mathbb{E}\left[\mathbb{1}_{\{\boldsymbol{p}\in D\}} \middle| p_i\right]$
- Let X $\sigma(p_i)$-measurable

$$
\begin{aligned}
\mathbb{E}_u\left[X\mathbb{1}_{\{\boldsymbol{p}\in D\}}\right] &= \int X(\omega)\mathbb{1}_{\{\boldsymbol{p}\in D\}}(\omega)\mathrm{d}\mathbb{P}_u(\omega) \\
&= \int X(\omega)\mathbb{1}_{\{\boldsymbol{p}\in D\}}(\omega)\frac{\mathbb{1}_{\{p_i\leq u\}}(\omega)}{\mathbb{P}(p_i \leq u)}\mathrm{d}\mathbb{P}(\omega) \\
&= \mathbb{E}\left[X\frac{\mathbb{1}_{\{p_i\leq u\}}}{\mathbb{P}(p_i \leq u)}\mathbb{1}_{\{\boldsymbol{p}\in D\}}\right] \\
&= \mathbb{E}\left[X\frac{\mathbb{1}_{\{p_i\leq u\}}}{\mathbb{P}(p_i \leq u)}f_{i,D}(p_i)\right] \quad \left(X\frac{\mathbb{1}_{\{p_i\leq u\}}}{\mathbb{P}(p_i \leq u)} \ \ \sigma(p_i)\text{-measurable}\right) \\
&= \int X(\omega)f_{i,D}(p_i(\omega))\frac{\mathbb{1}_{\{p_i\leq u\}}(\omega)}{\mathbb{P}(p_i \leq u)}\mathrm{d}\mathbb{P}(\omega) \\
&= \int X(\omega)f_{i,D}(p_i(\omega))\mathrm{d}\mathbb{P}_u(\omega) \\
&= \mathbb{E}_u\left[Xf_{i,D}(p_i)\right] \quad \square
\end{aligned}
$$

wPRDS is indeed weaker than PRDS

Proof of the proposition

- Let $u < u'$ with $\mathbb{P}(p_i \leq u) > 0$

$$
\begin{aligned}
g_{i,D}(u') &= \mathbb{P}_{u'}(\boldsymbol{p} \in D) \\
&= \mathbb{E}_{u'}\left[\mathbb{1}_{\{\boldsymbol{p} \in D\}}\right] \\
&= \mathbb{E}_{u'}\left[\mathbb{E}_{u'}\left[\mathbb{1}_{\{\boldsymbol{p} \in D\}}\middle| p_i\right]\right] \\
&= \mathbb{E}_{u'}\left[\mathbb{P}_{u'}(\boldsymbol{p} \in D | p_i)\right] \\
&= \mathbb{E}_{u'}\left[f_{i,D}(p_i)\right] \\
&= \int f_{i,D}(p_i(\omega)) \frac{\mathbb{1}_{\{p_i \leq u'\}}(\omega)}{\mathbb{P}(p_i \leq u')} \mathrm{d}\mathbb{P}(\omega) \\
&= \int f_{i,D}(p_i(\omega)) \frac{\mathbb{1}_{\{p_i \leq u\}}(\omega)}{\mathbb{P}(p_i \leq u')} \mathrm{d}\mathbb{P}(\omega) + \int f_{i,D}(p_i(\omega)) \frac{\mathbb{1}_{\{u < p_i \leq u'\}}(\omega)}{\mathbb{P}(p_i \leq u')} \mathrm{d}\mathbb{P}(\omega
\end{aligned}
$$

wPRDS is indeed weaker than PRDS

Proof of the proposition

- Let $\gamma = \mathbb{P}_{u'}(p_i \leq u) = \frac{\mathbb{P}(p_i \leq u)}{\mathbb{P}(p_i \leq u')} \in]0,1]$

$$\int f_{i,D}(p_i(\omega)) \frac{\mathbb{1}_{\{p_i \leq u\}}(\omega)}{\mathbb{P}(p_i \leq u')} \mathrm{d}\mathbb{P}(\omega) = \gamma \int f_{i,D}(p_i(\omega)) \frac{\mathbb{1}_{\{p_i \leq u\}}(\omega)}{\mathbb{P}(p_i \leq u)} \mathrm{d}\mathbb{P}(\omega)$$
$$= \gamma g_{i,D}(u)$$

- If $\gamma = 1 \Leftrightarrow \mathbb{P}(u < p_i \leq u') = 0$ then $g_{i,D}(u') = g_{i,D}(u)$
- Else,

$$\int f_{i,D}(p_i(\omega)) \frac{\mathbb{1}_{\{u < p_i \leq u'\}}(\omega)}{\mathbb{P}(p_i \leq u')} \mathrm{d}\mathbb{P}(\omega) = \frac{\mathbb{P}(u < p_i \leq u')}{\mathbb{P}(p_i \leq u')} \int f_{i,D}(p_i(\omega)) \frac{\mathbb{1}_{\{u < p_i \leq u'}}{\mathbb{P}(u < p_i \leq}$$
$$= (1-\gamma) \int f_{i,D}(p_i(\omega)) \frac{\mathbb{1}_{\{u < p_i \leq u'\}}(\omega)}{\mathbb{P}(u < p_i \leq u')} \mathrm{d}\mathbb{P}($$
$$= (1-\gamma) \mathbb{E}_{\{u < p_i \leq u'\}}\left[f_{i,D}(p_i)\right]$$

wPRDS is indeed weaker than PRDS

Proof of the proposition

- $g_{i,D}(u') = \gamma g_{i,D}(u) + (1-\gamma)\mathbb{E}_{\{u<p_i\leq u'\}}\left[f_{i,D}(p_i)\right]$
- $f_{i,D} \nearrow$ so:
  $\mathbb{E}_{\{u<p_i\leq u'\}}\left[f_{i,D}(p_i)\right] \geq f_{i,D}(u) \geq \mathbb{E}_{\{p_i\leq u\}}\left[f_{i,D}(p_i)\right] = g_{i,D}(u) \quad \square$

What is wPRDS?

Proposition [Giraud (2021)]

Assume that for all $P \in \mathfrak{F}$, the $(p_i)_{i \in \mathcal{H}_0}$ are independent, and they are independent from the $(p_i)_{i \in \mathcal{H}_1}$.
Then, for all $P \in \mathfrak{F}$, the (p_i) are wPRDS with $\mathcal{H}_0$ as the subset.

- Other examples in [Benjamini and Yekutieli (2001)], [Roquain (2015)],[Giraud (2021)]
  - Like one-sided Gaussian p-values with $\Sigma_{ij} \geq 0 \; \forall 1 \leq i, j \leq m$
  - Or with $\Sigma = \rho \mathbb{1}_m \mathbb{1}_m^\top + (1-\rho)\mathrm{Id}_m, \rho \in \left[-\frac{1}{m-1}, 1\right]$

What is wPRDS?

Proof of the Proposition

- Fix $P \in \mathfrak{F}$, D nondecreasing and $i \in \mathcal{H}_0$
- Key point: p_i is independent from $(p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_m)$ so for appropriate u:

$$\mathbb{P}\left(\boldsymbol{p} \in D | p_i \leq u\right) = \frac{\mathbb{P}\left(\boldsymbol{p} \in D \text{ and } p_i \leq u\right)}{\mathbb{P}\left(p_i \leq u\right)} = \frac{\mathbb{E}\left[\mathbb{1}_{\{\boldsymbol{p} \in D\}} \mathbb{1}_{\{p_i \leq u\}}\right]}{\mathbb{P}\left(p_i \leq u\right)}$$

$$= \int \mathbb{1}_{(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m) \in D} \frac{\mathbb{1}_{x_i \leq u}}{\mathbb{P}\left(p_i \leq u\right)} \mathrm{d}\mathbb{P}_{\boldsymbol{p}}(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m)$$

(transfer formula)

$$= \int \mathbb{1}_{(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m) \in D} \frac{\mathbb{1}_{x_i \leq u}}{\mathbb{P}\left(p_i \leq u\right)} \mathrm{d}\mathbb{P}_{p_i}(x_i) \mathrm{d}\mathbb{P}_{\boldsymbol{p}_{-i}}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x$$

(key observation)

$$= \int \mathbb{P}\left((x_1, \dots, x_{i-1}, p_i, x_{i+1}, \dots, x_m) \in D | p_i \leq u\right) \mathrm{d}\mathbb{P}_{\boldsymbol{p}_{-i}}(x_1, \dots, x_{i-1}, x_{i+1}, \dots$$

(Fubini) with $\mathbb{P}_{\boldsymbol{p}}$ the law of $\boldsymbol{p}$, $\mathbb{P}_{p_i} = \mathcal{L}(p_i)$ the law of p_i and $\mathbb{P}_{\boldsymbol{p}_{-i}}$ the law of $\boldsymbol{p}_{-i} = (p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_m)$

What is wPRDS?

Proof of the Proposition

- $\Rightarrow$ only need to show that $u \mapsto \mathbb{P}\left((x_1, \dots, x_{i-1}, p_i, x_{i+1}, \dots, x_m) \in D | p_i \leq u\right)$ nondecreasing for any fixed $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_m)$
- $= \mathbb{E}_u\left[g(p_i)\right]$ with $g : x_i \mapsto \mathbb{1}_{\{(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m) \in D\}}$ nondecreasing because D is
- same proof as before for proving that $u \mapsto \mathbb{E}_u\left[f_{i,D}(p_i)\right]$ was nondecreasing □

Benjamini-Hochberg procedure (BH)

Can we do better than the independent case?

Theorem [Benjamini and Yekutieli (2001)]

Assume that for all $P \in \mathfrak{F}$, the (p_i) are wPRDS with $\mathcal{H}_0$ as the subset. Then for all $P \in \mathfrak{F}$,

$$\mathrm{FDR}\left(R^{\mathrm{BH}}\right) \leq \alpha \frac{m_0}{m} \leq \alpha$$

- Previous Theorem is not useless because of:
  - the equality case
  - the proof ideas (and Lemmas) that are reused in more complex procedures [Roquain and Wiel (2009)], [Döhler, Durand, and Roquain (2018)]

Benjamini-Hochberg procedure (BH)

Proof of the Theorem

- As before,

$$\begin{aligned}
\mathrm{FDR}\left(R^{\mathrm{BH}}\right) &= \sum_{i\in\mathcal{H}_0}\sum_{k=1}^{m}\frac{1}{k}\mathbb{P}\left(p_i \le \alpha\frac{\hat{k}^{\mathrm{BH}}}{m}, \hat{k}^{\mathrm{BH}} = k\right)\\
&= \sum_{i\in\mathcal{H}_0}\sum_{k=1}^{m}\frac{1}{k}\mathbb{P}\left(p_i \le \alpha\frac{k}{m}, \hat{k}^{\mathrm{BH}} = k\right)\\
&= \sum_{i\in\mathcal{H}_0}\sum_{k=k_i}^{m}\frac{1}{k}\mathbb{P}_{\alpha\frac{k}{m}}\left(\hat{k}^{\mathrm{BH}} = k\right)\mathbb{P}\left(p_i \le \alpha\frac{k}{m}\right)\\
&\le \frac{\alpha}{m}\sum_{i\in\mathcal{H}_0}\sum_{k=k_i}^{m}\mathbb{P}_{\alpha\frac{k}{m}}\left(\hat{k}^{\mathrm{BH}} = k\right)
\end{aligned}$$

with $k_i = \min\left\{k \in [\![1,m]\!] : \mathbb{P}\left(p_i \le \alpha\frac{k}{m}\right) > 0\right\}$, for all $i \in \mathcal{H}_0$ ($k_i = +\infty$ and empty sum $= 0$ if empty set)

Benjamini-Hochberg procedure (BH)

Proof of the Theorem

$$
\begin{aligned}
\mathrm{FDR}\left(R^{\mathrm{BH}}\right) &\leq \frac{\alpha}{m} \sum_{i \in \mathcal{H}_0} \sum_{k=k_i}^{m} \left(\mathbb{P}_{\alpha \frac{k}{m}}\left(\hat{k}^{\mathrm{BH}} \leq k\right) - \mathbb{P}_{\alpha \frac{k}{m}}\left(\hat{k}^{\mathrm{BH}} \leq k-1\right)\right) \\
&\leq \frac{\alpha}{m} \sum_{i \in \mathcal{H}_0} \sum_{k=k_i}^{m} \left(\mathbb{P}_{\alpha \frac{k+1}{m}}\left(\hat{k}^{\mathrm{BH}} \leq k\right) - \mathbb{P}_{\alpha \frac{k}{m}}\left(\hat{k}^{\mathrm{BH}} \leq k-1\right)\right) \\
&\leq \frac{\alpha}{m} \sum_{i \in \mathcal{H}_0} \mathbb{P}_{\alpha \frac{m+1}{m}}\left(\hat{k}^{\mathrm{BH}} \leq m\right) \\
&\leq \alpha \frac{m_0}{m} \leq \alpha
\end{aligned}
$$

by wPRDS: $\forall k \in \mathbb{N}$, $\{\hat{k}^{\mathrm{BH}} \leq k\} = \{\boldsymbol{p} \in D\}$ with D the preimage of $]-\infty, k]$ under the function that maps $\boldsymbol{p}$ to $\hat{k}^{BH}$ which is coordinate-wise nonincreasing, hence D is nondecreasing □

Step-up procedures

Can we go even beyond, to any dependency?

Theorem [Giraud (2021)]

Let $\tau = (\tau_1, \ldots, \tau_m)$ a nondecreasing nonnegative sequence and consider the step-up procedure associated with τ.
Then for all $P \in \mathfrak{F}$,

$$\mathrm{FDR}\left(R^{\mathrm{SU}}(\tau)\right) \leq m_0 \sum_{j \geq 1} \frac{\tau_{j \wedge m}}{j(j+1)}$$

Step-up procedures

Proof of the Theorem

- As before,

$$\mathrm{FDR}\left(R^{\mathrm{SU}}(\tau)\right) = \mathbb{E}\left[\frac{\sum_{i\in\mathcal{H}_0} \mathbb{1}_{\left\{p_i \le \tau_{\hat{k}^{\mathrm{SU}}}\right\}}}{\hat{k}^{\mathrm{SU}} \vee 1}\right]$$

$$= \sum_{i\in\mathcal{H}_0} \mathbb{E}\left[\mathbb{1}_{\left\{p_i \le \tau_{\hat{k}^{\mathrm{SU}}}\right\}} \frac{1}{\hat{k}^{\mathrm{SU}} \vee 1}\right]$$

- For $k \geq 1$,
$\frac{1}{k} = \frac{1}{k} - \frac{1}{k+1} + \frac{1}{k+1} - \frac{1}{k+2} + \cdots = \sum_{j\geq k} \frac{1}{j(j+1)} = \sum_{j\geq 1} \frac{\mathbb{1}_{j\geq k}}{j(j+1)}$ so

$$\mathrm{FDR}\left(R^{\mathrm{SU}}(\tau)\right) = \sum_{i\in\mathcal{H}_0} \mathbb{E}\left[\mathbb{1}_{\left\{p_i \le \tau_{\hat{k}^{\mathrm{SU}}}\right\}} \sum_{j\geq 1} \frac{\mathbb{1}_{j\geq \hat{k}^{\mathrm{SU}}\geq 1}}{j(j+1)}\right]$$

Step-up procedures

Proof of the Theorem

- By Fubini,

$$\begin{aligned}
\mathrm{FDR}\left(R^{\mathrm{SU}}(\tau)\right) &= \sum_{i\in\mathcal{H}_0}\sum_{j\geq 1}\mathbb{E}\left[\mathbb{1}_{\left\{p_i\leq\tau_{\hat{k}^{\mathrm{SU}}}\right\}}\frac{\mathbb{1}_{j\geq\hat{k}^{\mathrm{SU}}\geq 1}}{j(j+1)}\right] \\
&\leq \sum_{i\in\mathcal{H}_0}\sum_{j\geq 1}\mathbb{E}\left[\mathbb{1}_{\{p_i\leq\tau_{j\wedge m}\}}\frac{\mathbb{1}_{j\geq\hat{k}^{\mathrm{SU}}\geq 1}}{j(j+1)}\right] \\
&\leq \sum_{i\in\mathcal{H}_0}\sum_{j\geq 1}\frac{1}{j(j+1)}\mathbb{E}\left[\mathbb{1}_{\{p_i\leq\tau_{j\wedge m}\}}\right] \\
&\leq \sum_{i\in\mathcal{H}_0}\sum_{j\geq 1}\frac{\tau_{j\wedge m}}{j(j+1)} = m_0\sum_{j\geq 1}\frac{\tau_{j\wedge m}}{j(j+1)} \qquad \square
\end{aligned}$$

Benjamini-Yekutieli procedure (BY)

FDR control under any dependency

- The Benjamini-Yekutieli procedure (BY) is the step-up procedure using $\tau_k = \frac{\alpha k}{m H_m}$, $H_m = \sum_{j=1}^m \frac{1}{j}$: uniformly worst than BH
- $R^{\mathrm{BY}} = R^{SU}\left(\left(\frac{\alpha k}{m H_m}\right)_{k \in [\![1,m]\!]}\right)$
- Adjusted p-values : $p_{(i)}^{adj} = 1 \wedge \min_{j \geq i} \frac{m H_m p_{(j)}}{j}$

Corollary [Benjamini and Yekutieli (2001)]

For all $P \in \mathfrak{F}$,

$$\mathrm{FDR}\left(R^{\mathrm{BY}}\right) \leq \alpha \frac{m_0}{m} \leq \alpha$$

$$m_0 \sum_{j \geq 1} \frac{\tau_{j \wedge m}}{j(j+1)} = \frac{\alpha m_0}{m H_m}\left(\sum_{j=1}^{m-1} \frac{1}{j+1} + m \sum_{j=m}^{\infty} \frac{1}{j(j+1)}\right)$$

$$= \frac{\alpha m_0}{m H_m}\left(\sum_{j=2}^{m} \frac{1}{j} + m \frac{1}{m}\right) = \frac{\alpha m_0}{m H_m} H_m \quad \square$$

Table of contents

Adaptivity to π_0

- $\pi_0 = \frac{m_0}{m}$
- Previous guarantees hold with "oracle" versions of the procedures using $m_0 = |\mathcal{H}_0|$ instead of m ($\Leftrightarrow$ using $\frac{\alpha}{\pi_0}$ instead of α)
- Ex: "oracle" Bonferroni: $\mathbb{P}(\exists i, i \in \mathcal{H}_0 : p_i \leq \frac{\alpha}{m_0}) \leq \alpha$
- Ex: "oracle BH", SU with $\tau_k = \frac{\alpha k}{m_0}$
- $\Rightarrow$ Core idea: estimate m_0 or π_0 and somehow plug $\hat{m}_0$ or $\hat{\pi}_0$ in the procedure

Holm-Bonferroni procedure (HB)

[Holm (1979)]

- Core idea: if $p_{(1)} \leq \frac{\alpha}{m}$, $(1) \in \mathcal{H}_1$, $m_0 \leq m-1$, and we could re-apply Bonferroni but with $m-1$ instead of m
- Repeat this sequentially until stop
- This formalizes as a SD procedure with $\tau_k = \frac{\alpha}{m-k+1}$
- $R^{\text{HB}} = R^{\text{SD}}\left(\left(\frac{\alpha}{m-k+1}\right)_{k\in[\![1,m]\!]}\right) = R\left(\frac{\alpha}{m-\hat{k}^{\text{HB}}+1}\right)$
- $\hat{k}^{\text{HB}} = \max\left\{k \in [\![0,m]\!] : \forall k' \leq k, p_{(k')} \leq \frac{\alpha}{m-k'+1}\right\}$
- Implicit estimation of m_0 by $\hat{m}_0 = m \wedge \left(m - \hat{k}^{\text{HB}} + 1\right)$
- Adjusted p-values : $p_{(i)}^{adj} = 1 \wedge \max_{j\leq i}(m-j+1)p_{(j)}$
- Uniformly rejects more than Bonferroni

Holm-Bonferroni procedure (HB)

FWER control under any dependency

Theorem [Holm (1979)]

For all $P \in \mathfrak{F}$,

$$\text{FWER}\left(R^{\text{HB}}\right) \leq \alpha$$

- $\Rightarrow$ HB has same guarantees than Bonferroni (and is almost as easy, computationally) $\Rightarrow$ Bonferroni should never be used [Aickin and Gensler (1996)]

Holm-Bonferroni procedure (HB)

Proof of the Theorem

- As discussed before, $\mathrm{FWER}\left(R\left(\frac{\alpha}{m_0}\right)\right) \leq \alpha$ so $\mathbb{P}\left(\left|R\left(\frac{\alpha}{m_0}\right) \cap \mathcal{H}_0\right| = 0\right) \geq 1 - \alpha$
- Assume $\left|R\left(\frac{\alpha}{m_0}\right) \cap \mathcal{H}_0\right| = 0 \Leftrightarrow R\left(\frac{\alpha}{m_0}\right) \subseteq \mathcal{H}_1$ $\Leftrightarrow \mathcal{H}_0 \subseteq= \left\{i : p_i > \frac{\alpha}{m_0}\right\}$
- Goal: to show that $m_0 \leq m - \hat{k}^{\mathrm{HB}} + 1$, we'll then have $R^{\mathrm{HB}} \subseteq R\left(\frac{\alpha}{m_0}\right) \subseteq \mathcal{H}_1$ which will conclude

Holm-Bonferroni procedure (HB)

Proof of the Theorem

- 2 proofs, first by recursion:
  - $\forall k \leq \hat{k}^{\text{HB}}$, $m_0 \leq m - k + 1$, $k = 0, 1$ obvious, assume it's true for $k < \hat{k}^{\text{HB}}$
  - For all $k' \leq k$ we have:

$$p_{(k')} \leq p_{(k)} \leq \frac{\alpha}{m - k + 1} \leq \frac{\alpha}{m_0},$$

  - So $\left| R\left(\frac{\alpha}{m_0}\right) \right| \geq k$ so $|\mathcal{H}_1| \geq k$ so $m_0 \leq m - k = m - (k+1) - 1$
- The original proof:
  - There's a true null ($> \frac{\alpha}{m_0}$) in the first $m - m_0 + 1$ hypotheses, so $p_{(m-m_0+1)} > \frac{\alpha}{m_0} = \frac{\alpha}{m-(m-m_0+1)+1}$
  - So by definition $\hat{k}^{\text{HB}} \leq m - m_0$ □

The proof actually gives a better estimator of m_0: just $m - \hat{k}^{\text{HB}}$, but that won't reject more hypotheses (or else contradiction with the def of $\hat{k}^{\text{HB}}$)

Holm-Bonferroni procedure (HB)

Need for step-down

- What about the step-up procedure with same critical values?
- $m = 2$, $\mathcal{H}_0 = [\![1, m]\!]$:
  $\mathrm{FWER}\left(R^{SU}\left(\left(\frac{\alpha}{2}, \alpha\right)\right)\right) = \mathbb{P}\left(p_{(1)} \leq \frac{\alpha}{2} \text{ or } p_{(2)} \leq \alpha\right)$
- $p_1 = p \sim \mathcal{U}([0,1])$ and $p_2 = 1 - p$: extreme negative correlation
- $\mathrm{FWER}\left(R^{SU}\left(\left(\frac{\alpha}{2}, \alpha\right)\right)\right) = \mathbb{P}\left(p_{(1)} \leq \frac{\alpha}{2} \text{ or } 1 - \alpha \leq p_{(1)}\right)$
- $\mathcal{L}\left(p_{(1)}\right) = \mathcal{U}\left(\left[0, \frac{1}{2}\right]\right)$: $\forall x \in \left[0, \frac{1}{2}\right]$,

$$\begin{aligned}
\mathbb{P}\left(p_{(1)} \leq x\right) &= \mathbb{P}\left(\left(p \leq x \text{ and } p \leq \frac{1}{2}\right) \text{ or } \left(1 - p \leq x \text{ and } p \geq \frac{1}{2}\right)\right) \\
&= \mathbb{P}\left(p \leq x \wedge \frac{1}{2}\right) + \mathbb{P}\left(p \geq (1-x) \vee \frac{1}{2}\right) \\
&= \mathbb{P}(p \leq x) + \mathbb{P}(p \geq 1 - x) \\
&= \mathbb{P}(p \leq x) + \mathbb{P}(p \geq 1 - x) = 2x
\end{aligned}$$

Holm-Bonferroni procedure (HB)

Need for step-down

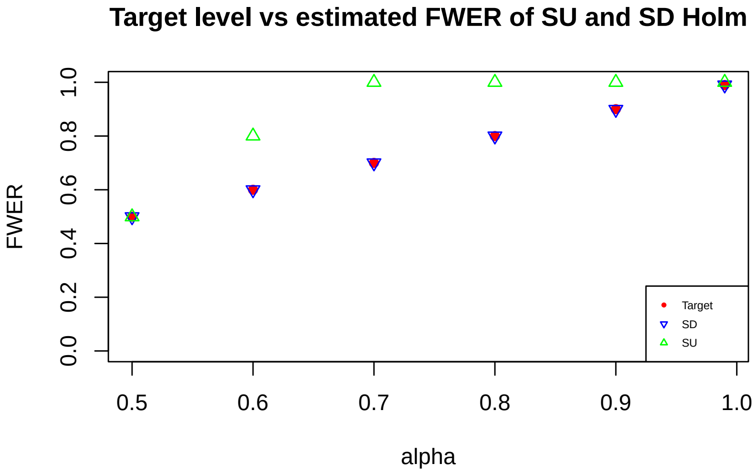
- If $\frac{\alpha}{2} \leq 1-\alpha \Leftrightarrow \alpha \leq \frac{2}{3}$, $\text{FWER}\left(R^{SU}\left(\left(\frac{\alpha}{2},\alpha\right)\right)\right) =$
  $\mathbb{P}\left(p_{(1)} \leq \frac{\alpha}{2}\right) + \mathbb{P}\left(1-\alpha \leq p_{(1)}\right) = \alpha + \mathbb{P}\left(1-\alpha \leq p_{(1)}\right)$
- $\mathbb{P}\left(1-\alpha \leq p_{(1)}\right) = 0$ if $\alpha \leq \frac{1}{2}$,
  $= 1 - \mathbb{P}\left(1-\alpha \geq p_{(1)}\right) = 1 - 2(1-\alpha) = 2\alpha - 1$ if $\alpha \geq \frac{1}{2}$
- If $\alpha \geq \frac{2}{3}$, $\text{FWER}\left(R^{SU}\left(\left(\frac{\alpha}{2},\alpha\right)\right)\right) = 1$

$$\Rightarrow \text{FWER}\left(R^{SU}\left(\left(\frac{\alpha}{2},\alpha\right)\right)\right) = \begin{cases} \alpha & \text{if} \quad \alpha \in \left]0, \frac{1}{2}\right[\\ 3\alpha - 1 & \text{if} \quad \alpha \in \left[\frac{1}{2}, \frac{2}{3}\right] \\ 1 & \text{if} \quad \alpha \in \left[\frac{2}{3}, 1\right[\end{cases}$$

- Remark: for this model, FWER is saturated for Bonf and HB:
  $\text{FWER}\left(R^{\text{Bonf}}\right) = \text{FWER}\left(R^{\text{HB}}\right) = \mathbb{P}\left(p_{(1)} \leq \frac{\alpha}{2}\right) = \alpha$

Holm-Bonferroni procedure (HB)

2.10^5 replications



Storey-BH

Adaptive FDR control

- [Storey, Taylor, and Siegmund (2004)]
- Fix $\lambda \in]0,1[$, $\hat{m}_0 = \frac{\sum_{i=1}^m \mathbb{1}_{\{p_i>\lambda\}}+1}{1-\lambda} = \frac{m-|R(\lambda)|+1}{1-\lambda}$
- Idea : large p-values are mostly null, and nulls are super-uniform, so $\sum_{i=1}^m \mathbb{1}_{\{p_i>\lambda\}} \approx \sum_{i=1}^{m_0} \mathbb{1}_{\{p_i>\lambda\}} \gtrsim (1-\lambda)m_0$
- "+1" for $\hat{m}_0 > 0$ and for technical reasons
- Storey-BH is the SU procedure with $\tau_k = \min\left(\alpha\frac{k}{\hat{m}_0}, \lambda\right)$, $k \geq 1$ (recall $\tau_0 = \tau_1$ so that $\tau_k = \tau_{k\vee 1}$)
- $\hat{k}^{\text{St-BH}} = \max\left\{k \in [\![0,m]\!] : p_{(k)} \leq \min\left(\alpha\frac{k\vee 1}{\hat{m}_0}, \lambda\right)\right\}$
- $\hat{t}_\alpha^{\text{St-BH}} = \tau_{\hat{k}^{\text{St-BH}}} = \min\left(\alpha\frac{\hat{k}^{\text{St-BH}}\vee 1}{\hat{m}_0}, \lambda\right)$, $R^{\text{St-BH}} = R\left(\hat{t}_\alpha^{\text{St-BH}}\right)$
- $\min(\cdot,\lambda)$ above to avoid overfitting: you don't look at the same p-values for estimating m_0 and for rejecting hypotheses
- Up to this, Storey-BH is BH but with $\hat{m}_0$ instead of m

Storey-BH

- As before, a link with FDP estimation:
- $\widehat{\mathrm{FDP}}^{\text{St-BH}}(t) = \frac{\hat{m}_0 t}{|R(t)| \vee 1}$ if $t \leq \lambda$, $= 1$ if $t > \lambda$
- $\hat{t}_\alpha^{St\text{-}heur} = \sup\left\{t \in [0,1] : \widehat{\mathrm{FDP}}^{\text{St-BH}}(t) \leq \alpha\right\} =$
  $\sup\left\{t \in [0,1] : \widehat{G}_\lambda(t) \geq t\right\}$ with $\widehat{G}_\lambda(t) = \frac{\alpha}{\hat{m}_0}(|R(t)| \vee 1) \wedge \lambda$

Lemma

$\hat{t}_\alpha^{St\text{-}heur} = \hat{t}_\alpha^{\text{St-BH}}$

Storey-BH

Proof of the Lemma

- Similar to before, supremum well-defined, and $\widehat{G}_\lambda(t)$ nondecreasing and $[0,1] \to [0,1] \Rightarrow \hat{t}_\alpha^{St\text{-}heur}$ is a max and $\hat{t}_\alpha^{St\text{-}heur} = \widehat{G}_\lambda\left(\hat{t}_\alpha^{St\text{-}heur}\right)$
- Remember that $p_{(|R(t)|)} \leq t$ and note that, here, $\widehat{G}_\lambda(t) = \tau_{|R(t)|}$, combine this:

$$p_{\left(\left|R\left(\hat{t}_\alpha^{St\text{-}heur}\right)\right|\right)} \leq \hat{t}_\alpha^{St\text{-}heur} = \widehat{G}_\lambda\left(\hat{t}_\alpha^{St\text{-}heur}\right) = \tau_{\left|R\left(\hat{t}_\alpha^{St\text{-}heur}\right)\right|}$$

$$\Rightarrow \left|R\left(\hat{t}_\alpha^{St\text{-}heur}\right)\right| \leq \hat{k}^{\text{St-BH}}$$

$$\Rightarrow \hat{t}_\alpha^{St\text{-}heur} = \widehat{G}_\lambda\left(\hat{t}_\alpha^{St\text{-}heur}\right) = \tau_{\left|R\left(\hat{t}_\alpha^{St\text{-}heur}\right)\right|} \leq \tau_{\hat{k}^{\text{St-BH}}}$$

Storey-BH
Proof of the Lemma

- Conversely, using that $\left|R\left(\tau_{\hat{k}^{\text{St-BH}}}\right)\right| = \hat{k}^{\text{St-BH}}$ (property of SU),

$$\widehat{G}_\lambda\left(\tau_{\hat{k}^{\text{St-BH}}}\right) = \frac{\alpha}{\hat{m}_0}\left(\left|R\left(\tau_{\hat{k}^{\text{St-BH}}}\right)\right| \vee 1\right) \wedge \lambda = \alpha\frac{\hat{k}^{\text{St-BH}} \vee 1}{\hat{m}_0} \wedge \lambda = \tau_{\hat{k}^{\text{St-BH}}}$$

- So $\tau_{\hat{k}^{\text{St-BH}}} \leq \hat{t}_\alpha^{St\text{-}heur}$ by definition of $\hat{t}_\alpha^{St\text{-}heur}$ □

Storey-BH

FDR control

Theorem [Storey, Taylor, and Siegmund (2004)]

Assume that for all $P \in \mathfrak{F}$, the $(p_i)_{i \in \mathcal{H}_0}$ are independent and $\sim \mathcal{U}([0,1])$, and they are independent from the $(p_i)_{i \in \mathcal{H}_1}$.
Then for all $P \in \mathfrak{F}$,

$$\mathrm{FDR}\left(R^{\text{St-BH}}\right) \leq \alpha(1 - \lambda^{m_0}) \leq \alpha$$

- Proof by martingale techniques: the stochastic process is important
- Need true uniformity under $\mathcal{H}_0$!

Three Lemmas

- $V(R(t)) \sim \mathcal{B}(m_0, t)$ for all $t \in [0,1]$
- The process $\left(\frac{V(R(t))}{t}\right)_{t \in]0,1]}$ is a reverse-time martingale w.r.t. the filtration $(\mathcal{F}_t)_{t \in]0,1]}$ with $\mathcal{F}_t = \sigma\left(\left(\mathbb{1}_{\{p_i \leq t'\}}\right)_{i \in [\![1,m]\!], t \leq t' \leq 1}\right)$, $t \in [0,1]$
- $\hat{t}_\alpha^{\text{St-BH}}$ is a stopping time (in reverse) w.r.t. $(\mathcal{F}_t)_{t \in [0,1]}$

Storey-BH

Proof of the First Lemma

- $V(R(t)) = \sum_{i \in \mathcal{H}_0} \mathbb{1}_{\{p_i \leq t\}}$
- $\mathbb{1}_{\{p_i \leq t\}}$, $i \in \mathcal{H}_0$, are i.i.d. $\sim \mathcal{B}(t)$ □

Storey-BH

Proof of the Second Lemma

- We want to prove that for $0 < s \leq t$, $\mathbb{E}\left[\frac{V(R(s))}{s}\middle|\mathcal{F}_t\right] = \frac{V(R(t))}{t}$
- $\mathbb{E}\left[V(R(s))|\mathcal{F}_t\right] = \sum_{i\in\mathcal{H}_0} \mathbb{E}\left[\mathbb{1}_{\{p_i\leq s\}}\middle|\mathcal{F}_t\right]$ so proving $\frac{\mathbb{E}\left[\mathbb{1}_{\{p_i\leq s\}}\middle|\mathcal{F}_t\right]}{s} = \frac{\mathbb{1}_{\{p_i\leq t\}}}{t}$ for $i \in \mathcal{H}_0$ is sufficient
- By independence, for $i \in \mathcal{H}_0$,

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}_{\{p_i\leq s\}}\middle|\mathcal{F}_t\right] &= \mathbb{E}\left[\mathbb{1}_{\{p_i\leq s\}}\middle|\sigma\left(\left(\mathbb{1}_{\{p_j\leq t'\}}\right)_{j\in[\![1,m]\!],t\leq t'\leq 1}\right)\right]\\ &= \mathbb{E}\left[\mathbb{1}_{\{p_i\leq s\}}\middle|\sigma\left(\left(\mathbb{1}_{\{p_i\leq t'\}}\right)_{t\leq t'\leq 1}\right)\right]\end{aligned}$$

- To have $\mathbb{E}\left[\frac{\mathbb{1}_{\{p_i\leq s\}}}{s}\middle|\mathcal{F}_t\right] = \frac{\mathbb{1}_{\{p_i\leq t\}}}{t}$ we need $\mathbb{E}\left[\mathbb{1}_A\frac{\mathbb{1}_{\{p_i\leq s\}}}{s}\right] = \mathbb{E}\left[\mathbb{1}_A\frac{\mathbb{1}_{\{p_i\leq t\}}}{t}\right]$ for all $A \in \sigma\left(\left(\mathbb{1}_{\{p_i\leq t'\}}\right)_{t\leq t'\leq 1}\right)$

Storey-BH

Proof of the Second Lemma

- Which is $\mathbb{P}_s(A) = \mathbb{P}_t(A)$ for all $A \in \sigma\left(\left(\mathbb{1}_{\{p_i \leq t'\}}\right)_{t \leq t' \leq 1}\right)$
- By the Sierpiński–Dynkin's π-λ theorem ("lemme des classes monotones"), for all A in a π-system that generates $\sigma\left(\left(\mathbb{1}_{\{p_i \leq t'\}}\right)_{t \leq t' \leq 1}\right)$ is sufficient

$$\sigma\left(\left(\mathbb{1}_{\{p_i \leq t'\}}\right)_{t \leq t' \leq 1}\right) = \sigma\left(\bigcup_{t' \geq t} \sigma\left(\mathbb{1}_{\{p_i \leq t'\}}\right)\right) = \sigma\left(\bigcup_{t' \geq t} \mathbb{1}_{\{p_i \leq t'\}}^{-1}\left(\mathcal{B}(\mathbb{R})\right)\right)$$

- $\mathbb{1}_{\{p_i \leq t'\}}^{-1}\left(\mathcal{B}(\mathbb{R})\right) = \{\varnothing, \{p_i \leq t'\}, \{p_i \leq t'\}^{\mathsf{c}}, \Omega\}$

Storey-BH

Proof of the Second Lemma

- $\sigma\left(\left(\mathbb{1}_{\{p_i \leq t'\}}\right)_{t \leq t' \leq 1}\right) = \sigma\left(\{\{p_i \leq t'\}, t' \geq t\}\right)$
- $\{\{p_i \leq t'\}, t' \geq t\}$ is a π-system:
  $\{p_i \leq t'\} \cap \{p_i \leq t''\} = \{p_i \leq t' \wedge t''\}$
- $\mathbb{P}_s\left(\{p_i \leq t'\}\right) = \mathbb{P}_s\left(\{p_i \leq s\}\right) = 1 = \mathbb{P}_t\left(\{p_i \leq t\}\right) =$
  $\mathbb{P}_t\left(\{p_i \leq t'\}\right)$ □

Storey-BH

Proof of the Third Lemma, due to Romain Périer

- If $t > \lambda$, $\{\hat{t}_\alpha^{\text{St-BH}} \geq t\} = \varnothing \in \mathcal{F}_t$
- Let $t \leq \lambda$

$$\begin{aligned}
\{\hat{t}_\alpha^{\text{St-BH}} \geq t\} &= \{\tau_{\hat{k}^{\text{St-BH}}} \geq t\} \\
&= \left\{\alpha \frac{\hat{k}^{\text{St-BH}} \vee 1}{\hat{m}_0} \wedge \lambda \geq t\right\} \\
&= \left\{\alpha \frac{\hat{k}^{\text{St-BH}} \vee 1}{\hat{m}_0} \geq t\right\} \text{ because } t \leq \lambda \\
&= \left\{\hat{k}^{\text{St-BH}} \vee 1 \geq \frac{\hat{m}_0 t}{\alpha}\right\} = \left\{\hat{k}^{\text{St-BH}} \vee 1 \geq \left\lceil \frac{\hat{m}_0 t}{\alpha} \right\rceil\right\} \\
&= \left\{1 \geq \left\lceil \frac{\hat{m}_0 t}{\alpha} \right\rceil\right\} \cup \left\{\hat{k}^{\text{St-BH}} \geq \left\lceil \frac{\hat{m}_0 t}{\alpha} \right\rceil\right\}
\end{aligned}$$

- with $\left\{1 \geq \left\lceil \frac{\hat{m}_0 t}{\alpha} \right\rceil\right\} \in \mathcal{F}_\lambda \subseteq \mathcal{F}_t$

Storey-BH

Proof of the Third Lemma, due to Romain Périer

- Let $\mathcal{M} = (1-\lambda)^{-1}[\![1, m+1]\!]$ the finite set $\hat{m}_0$ belongs to

$$
\begin{aligned}
\left\{\hat{k}^{\text{St-BH}} \geq \left\lceil \frac{\hat{m}_0 t}{\alpha} \right\rceil \right\} &= \left\{\exists k \geq \left\lceil \frac{\hat{m}_0 t}{\alpha} \right\rceil, p_{(k)} \leq \tau_k \right\} \\
&= \left\{\exists k \geq \left\lceil \frac{\hat{m}_0 t}{\alpha} \right\rceil, \sum_{i=1}^{m} \mathbb{1}_{\{p_i \leq \tau_k\}} \geq k \right\} \\
&= \left\{\exists k \geq \left\lceil \frac{\hat{m}_0 t}{\alpha} \right\rceil, \sum_{i=1}^{m} \mathbb{1}_{\left\{p_i \leq \alpha \frac{k \vee 1}{\hat{m}_0} \wedge \lambda \right\}} \geq k \right\} \\
&= \bigcup_{m_0 \in \mathcal{M}} \{\hat{m}_0 = m_0\} \cap \left\{\exists k \geq \left\lceil \frac{m_0 t}{\alpha} \right\rceil, \sum_{i=1}^{m} \mathbb{1}_{\left\{p_i \leq \alpha \frac{k \vee 1}{m_0} \wedge \lambda \right\}} \right. \\
&= \bigcup_{m_0 \in \mathcal{M}} \{\hat{m}_0 = m_0\} \cap \bigcup_{k \geq \left\lceil \frac{m_0 t}{\alpha} \right\rceil} \left\{ \sum_{i=1}^{m} \mathbb{1}_{\left\{p_i \leq \alpha \frac{k \vee 1}{m_0} \wedge \lambda \right\}} \geq k \right\}
\end{aligned}
$$

Storey-BH

Proof of the Third Lemma, due to Romain Périer

- $\{\hat{m}_0 = m_0\} \in \mathcal{F}_\lambda \subseteq \mathcal{F}_t$ for all $m_0 \in \mathcal{M}$
- $\left\{\sum_{i=1}^{m} \mathbb{1}_{\left\{p_i \le \alpha \frac{k \vee 1}{m_0} \wedge \lambda\right\}} \ge k\right\} \in \mathcal{F}_{\alpha \frac{k\vee 1}{m_0} \wedge \lambda}$ for all $m_0 \in \mathcal{M}$, $k \ge \left\lceil \frac{m_0 t}{\alpha} \right\rceil$, but:

$$\begin{aligned}
\alpha \frac{k \vee 1}{m_0} \wedge \lambda &\ge \alpha \frac{k}{m_0} \wedge \lambda \\
&\ge \frac{\alpha}{m_0} \left\lceil \frac{m_0 t}{\alpha} \right\rceil \wedge \lambda \\
&\ge \frac{\alpha}{m_0} \frac{m_0 t}{\alpha} \wedge \lambda \\
&\ge t \wedge \lambda = t
\end{aligned}$$

- So $\mathcal{F}_{\alpha \frac{k \vee 1}{m_0} \wedge \lambda} \subseteq \mathcal{F}_t$ too

□

Storey-BH

Proof of the Theorem

- $\hat{t}_\alpha^{\text{St-BH}} = \widehat{G}_\lambda\left(\hat{t}_\alpha^{\text{St-BH}}\right) = \frac{\alpha}{\hat{m}_0}\left(\left|R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right| \vee 1\right) \wedge \lambda$
- If $\widehat{\text{FDP}}^{\text{St-BH}}(\lambda) \geq \alpha$, $\frac{\hat{m}_0\lambda}{|R(\lambda)|\vee 1} \geq \alpha$ so $\hat{t}_\alpha^{\text{St-BH}} = \frac{\alpha}{\hat{m}_0}\left(\left|R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right| \vee 1\right)$
- And so $\text{FDP}\left(R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right) = \frac{V\left(R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right)}{\left|R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right|\vee 1} = \frac{\alpha V\left(R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right)}{\hat{m}_0\hat{t}_\alpha^{\text{St-BH}}} = \alpha\frac{1-\lambda}{m-|R(\lambda)|+1}\frac{V\left(R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right)}{\hat{t}_\alpha^{\text{St-BH}}}$
- If $\widehat{\text{FDP}}^{\text{St-BH}}(\lambda) < \alpha$ then $\hat{t}_\alpha^{\text{St-BH}} = \lambda$ and $\frac{1}{|R(\lambda)|\vee 1} < \alpha\frac{1-\lambda}{m-|R(\lambda)|+1}\frac{1}{\lambda}$
- And so $\text{FDP}\left(R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right) = \text{FDP}\left(R(\lambda)\right) < \alpha\frac{1-\lambda}{m-|R(\lambda)|+1}\frac{V(R(\lambda))}{\lambda}$

Storey-BH

Proof of the Theorem

$$
\begin{aligned}
\mathrm{FDR}\left(R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right) &= \mathbb{E}\left[\mathrm{FDP}\left(R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right)\mathbb{1}_{\left\{\widehat{\mathrm{FDP}}^{\text{St-BH}}(\lambda)\geq\alpha\right\}}\right] \\
&\quad + \mathbb{E}\left[\mathrm{FDP}\left(R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right)\mathbb{1}_{\left\{\widehat{\mathrm{FDP}}^{\text{St-BH}}(\lambda)<\alpha\right\}}\right] \\
&\leq \mathbb{E}\left[\alpha\frac{1-\lambda}{m-|R(\lambda)|+1}\frac{V\left(R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right)}{\hat{t}_\alpha^{\text{St-BH}}}\mathbb{1}_{\left\{\widehat{\mathrm{FDP}}^{\text{St-BH}}(\lambda)\geq\alpha\right\}}\right] \\
&\quad \mathbb{E}\left[\alpha\frac{1-\lambda}{m-|R(\lambda)|+1}\frac{V(R(\lambda))}{\lambda}\mathbb{1}_{\left\{\widehat{\mathrm{FDP}}^{\text{St-BH}}(\lambda)<\alpha\right\}}\right]
\end{aligned}
$$

Storey-BH

Proof of the Theorem

$$\mathbb{E}\left[\alpha \frac{1-\lambda}{m-|R(\lambda)|+1} \frac{V\left(R\left(\hat{t}_{\alpha}^{\text{St-BH}}\right)\right)}{\hat{t}_{\alpha}^{\text{St-BH}}} \mathbb{1}_{\left\{\widehat{\mathrm{FDP}}^{\text{St-BH}}(\lambda)\geq\alpha\right\}}\right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[\alpha \frac{1-\lambda}{m-|R(\lambda)|+1} \frac{V\left(R\left(\hat{t}_{\alpha}^{\text{St-BH}}\right)\right)}{\hat{t}_{\alpha}^{\text{St-BH}}} \mathbb{1}_{\left\{\widehat{\mathrm{FDP}}^{\text{St-BH}}(\lambda)\geq\alpha\right\}} \middle| \mathcal{F}_{\lambda}\right]\right]$$

$$= \mathbb{E}\left[\alpha \frac{1-\lambda}{m-|R(\lambda)|+1} \mathbb{E}\left[\frac{V\left(R\left(\hat{t}_{\alpha}^{\text{St-BH}}\right)\right)}{\hat{t}_{\alpha}^{\text{St-BH}}} \middle| \mathcal{F}_{\lambda}\right] \mathbb{1}_{\left\{\widehat{\mathrm{FDP}}^{\text{St-BH}}(\lambda)\geq\alpha\right\}}\right]$$

Storey-BH
Proof of the Theorem

- By-product of optional stopping theorem: $\frac{V\left(R\left(t \vee \hat{t}_\alpha^{\text{St-BH}}\right)\right)}{t \vee \hat{t}_\alpha^{\text{St-BH}}}$ is also a reverse-time martingale w.r.t. $(\mathcal{F}_t)_{t \in]0,1]}$
- Also note that $\lambda \geq \hat{t}_\alpha^{\text{St-BH}} \geq \tau_1 = \frac{\alpha}{\hat{m}_0} \wedge \lambda \geq \frac{\alpha(1-\lambda)}{m+1} \wedge \lambda$ a.s.

$$\mathbb{E}\left[\frac{V\left(R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right)}{\hat{t}_\alpha^{\text{St-BH}}} \middle| \mathcal{F}_\lambda \right] = \mathbb{E}\left[\frac{V\left(R\left(\left(\frac{\alpha(1-\lambda)}{m+1} \wedge \lambda\right) \vee \hat{t}_\alpha^{\text{St-BH}}\right)\right)}{\left(\frac{\alpha(1-\lambda)}{m+1} \wedge \lambda\right) \vee \hat{t}_\alpha^{\text{St-BH}}} \middle| \mathcal{F}_\lambda \right]$$

$$= \frac{V\left(R\left(\lambda \vee \hat{t}_\alpha^{\text{St-BH}}\right)\right)}{\lambda \vee \hat{t}_\alpha^{\text{St-BH}}}$$

$$= \frac{V\left(R\left(\lambda\right)\right)}{\lambda}$$

Storey-BH

Proof of the Theorem

$$
\begin{aligned}
\mathrm{FDR}\left(R\left(\hat{t}_\alpha^{\text{St-BH}}\right)\right) &\leq \alpha\mathbb{E}\left[\frac{1-\lambda}{m-|R(\lambda)|+1}\frac{V(R(\lambda))}{\lambda}\right] \\
&\leq \alpha\mathbb{E}\left[\frac{1-\lambda}{m_0-V(R(\lambda))+1}\frac{V(R(\lambda))}{\lambda}\right] \\
&\leq \alpha\sum_{k=1}^{m_0}\frac{1-\lambda}{\lambda}\frac{k}{m_0-k+1}\binom{m_0}{k}\lambda^k(1-\lambda)^{m_0-k} \\
&\leq \alpha\sum_{k=1}^{m_0}\binom{m_0}{k-1}\lambda^{k-1}(1-\lambda)^{m_0-k+1} \\
&\leq \alpha\sum_{k=0}^{m_0-1}\binom{m_0}{k}\lambda^{k}(1-\lambda)^{m_0-k} \\
&\leq \alpha(1-\lambda^{m_0}) \leq \alpha \quad \square
\end{aligned}
$$

- Proof can be adapted to prove BH a 3rd time, but requires uniformity

Adaptivity to signal strength and location

Introduction to hypothesis weighting

- SU and SD procedures are implicitly adaptive to signal strength: strong signal $\Rightarrow$ small p_i's $\Rightarrow$ large $\hat{k}^{SU}/\hat{k}^{SD}$
- What if we have prior knowledge about the hypotheses likely to be (strong) signal?
- We can encode that into weights and plug them into the procedure:
  - Compare p_i to $w_i \tau_k$ instead of τ_k, $w_i \geq 0$, with a bounding condition on the w_i's
  - If i likely to be (strong) signal: small w_i, which makes larger w_j's affordable for other hypotheses
- Weights can be random

weighted-Benjamini-Hochberg procedure (wBH)

[Genovese, Roeder, and Wasserman (2006)]

- Let $w_1, \dots, w_m$ nonnegative random variables and consider the weighted FDP estimator $\widehat{\mathrm{FDP}}^{\mathrm{wBH}}(t) = \frac{mt}{\sum_{i=1}^m \mathbb{1}_{\{p_i \leq w_i t\}} \vee 1}$
- Let $\hat{t}_\alpha^{w\text{-}heur} = \sup\left\{t \in [0,1] : \frac{\alpha}{m}\left(\sum_{i=1}^m \mathbb{1}_{\{p_i \leq w_i t\}} \vee 1\right) \geq t\right\}$
- Alternatively, let

$$q_i = \begin{cases} 0 & \text{if} \quad p_i = 0, w_i = 0 \\ 2 & \text{if} \quad p_i \neq 0, w_i = 0 \\ \frac{p_i}{w_i} & \text{if} \quad w_i \neq 0 \end{cases}$$

- Remarks:
  - $q_i \leq t$ if and only if $p_i \leq w_i t$
  - Not the same ordering for the q_i's than the p_i's: denote it $q_{\langle 1 \rangle} \leq \ldots \leq q_{\langle m \rangle}$
  - The weighted p-values q_i's are not valid p-values because not necessarily super-uniform under the null
  - All previous deterministic results on SU procedures hold nonetheless

weighted-Benjamini-Hochberg procedure (wBH)

- wBH can be defined as BH applied to the q_i's: $\hat{k}^{\mathrm{wBH}} = \max\left\{k \in [\![0, m]\!], q_{\langle k\rangle} \leq \alpha \frac{k \vee 1}{m}\right\}$, $\hat{t}_\alpha^{\mathrm{wBH}} = \alpha \frac{\hat{k}^{\mathrm{wBH}} \vee 1}{m}$ and $R^{\mathrm{wBH}} = \left\{i : q_i \leq \hat{t}_\alpha^{\mathrm{wBH}}\right\} = \left\{i : p_i \leq w_i \hat{t}_\alpha^{\mathrm{wBH}}\right\}$
- As before, $\hat{t}_\alpha^{w\text{-}heur} = \hat{t}_\alpha^{\mathrm{wBH}}$, because $\hat{t}_\alpha^{w\text{-}heur} = \sup\left\{t \in [0, 1] : \frac{\alpha}{m}\left(\sum_{i=1}^m \mathbb{1}_{\{q_i \leq t\}} \vee 1\right) \geq t\right\}$, same proof but with q_i instead of p_i

weighted-Benjamini-Hochberg procedure (wBH)

FDR control

Theorem [Genovese, Roeder, and Wasserman (2006)]

Assume that for all $P \in \mathfrak{F}$, the $(p_i)_{i \in \mathcal{H}_0}$ are independent, that they are independent from the $(p_i)_{i \in \mathcal{H}_1}$, that the $(w_i)_{i \in \mathcal{H}_0}$ are independent, that they are independent from the $(w_i)_{i \in \mathcal{H}_1}$, that (p_i) and (w_i) are independent, and finally that the w_i's are integrable with $\sum_{i=1}^m \mathbb{E}[w_i] \leq m$. Then for all $P \in \mathfrak{F}$,

$$\mathrm{FDR}\left(R^{\mathrm{wBH}}\right) \leq \alpha \frac{\sum_{i \in \mathcal{H}_0} \mathbb{E}[w_i]}{m} \leq \alpha.$$

- Includes the case of deterministic (and possibly grouped) weights based on prior knowledge

Lemma

Under the same conditions, for all $i \in \mathcal{H}_0$, $\mathbb{P}(q_i \leq t) \leq t\mathbb{E}[w_i]$.

weighted-Benjamini-Hochberg procedure (wBH)

Proofs

$$\begin{aligned}\mathbb{P}(q_i \le t) &= \mathbb{E}\left[\mathbb{P}(q_i \le t | w_i)\right] \\ &= \mathbb{E}\left[\mathbb{P}(p_i \le t w_i | w_i)\right] \\ &\le \mathbb{E}[t w_i] \text{ by independence and super-uniformity} \\ &\le t\mathbb{E}[w_i] \quad \square\end{aligned}$$

- For FDR control, same proof as BH for independent case, thanks to all deterministic Lemmas on SU procedures:

$$\begin{aligned}\mathrm{FDR}\left(R^{\mathrm{wBH}}\right) &= \sum_{i \in \mathcal{H}_0} \sum_{k=1}^{m} \frac{1}{k} \mathbb{P}\left(q_i \le \alpha \frac{k}{m}\right) \mathbb{P}\left(\hat{k}^{-i} = k-1\right) \\ &\le \frac{\alpha}{m} \sum_{i \in \mathcal{H}_0} \mathbb{E}[w_i] \sum_{k=1}^{m} \mathbb{P}\left(\hat{k}^{-i} = k-1\right) \\ &\le \alpha \frac{\sum_{i \in \mathcal{H}_0} \mathbb{E}[w_i]}{m} \le \alpha \quad \square\end{aligned}$$

- Weights are independent of the data here, for adaptive weights see e.g. [Roquain and Wiel (2009)], [Durand (2019)]

Table of contents

An example of discrete test

Binomial test

- The simplest example: $X_1, \dots, X_n$ i.i.d $\sim \mathcal{B}(p)$, $p \in [0,1]$
- $X = \sum_{i=1}^n X_i$
- $\mathfrak{F} = \{\mathcal{B}(n,p), p \in [0,1]\}$, discrete distributions
- Is the coin rigged? $\Leftrightarrow H_0 = \left\{\mathcal{B}\left(n, \frac{1}{2}\right)\right\}$
- $\hat{p}_i(X), \bar{p}_i(X), \breve{p}_i(X)$ also discrete, but $\breve{p}_i(X)$ not the best suited for bilateral tests

Another example of discrete test

Fisher's exact test

- Testing association between an allele and a phenotype of interest

	Phenotype 1	Phenotype 2	Total
Allele A	$n_{1,\mathrm{A}}$	$n_{2,\mathrm{A}}$	n_A
Allele a	$n_{1,\mathrm{a}}$	$n_{2,\mathrm{a}}$	n_a
Total	n_1	n_2	N

- For large samples, χ^2 approximation: $\frac{\left(n_{1,\mathrm{A}}-\frac{n_1 n_\mathrm{A}}{N}\right)^2}{\frac{n_1 n_\mathrm{A}}{N}} + \frac{\left(n_{1,\mathrm{a}}-\frac{n_1 n_\mathrm{a}}{N}\right)^2}{\frac{n_1 n_\mathrm{a}}{N}} + \frac{\left(n_{2,\mathrm{A}}-\frac{n_2 n_\mathrm{A}}{N}\right)^2}{\frac{n_2 n_\mathrm{A}}{N}} + \frac{\left(n_{2,\mathrm{a}}-\frac{n_2 n_\mathrm{a}}{N}\right)^2}{\frac{n_2 n_\mathrm{a}}{N}}$ follows $\chi^2(1)$ distribution under H_0
- What if we want an exact test ?
- Under H_0, conditionally to n_1 and n_A, $n_{1,\mathrm{A}} \sim \mathcal{H}(N, n_1, n_\mathrm{A}) = \mathcal{H}(N, n_\mathrm{A}, n_1)$, hypergeometric hence discrete
- $\hat{p}_i(X), \bar{p}_i(X), \breve{p}_i(X)$ also discrete, but $\breve{p}_i(X)$ not the best suited for bilateral tests

Generic construction of p-values with discreteness

Following the idea of "the probability of an event at least as extreme as"

- Assume we have at hand a test statistic $T_i : \mathcal{X} \longrightarrow \mathbb{R}$ such that $\forall P \in H_{0,i}$, $\exists \mathcal{A}_{i,P}$ countable or finite such that $T_i(X) \in \mathcal{A}_{i,P}$ a.s.
- Then let

$$\begin{aligned}
\check{p}_i(X) &= \sup_{P \in H_{0,i}} \sum_{\substack{k \in \mathcal{A}_{i,P} \\ \mathbb{P}_{Z\sim P}(T_i(Z)=k) \leq \mathbb{P}_{\substack{Z\sim P \\ Z \perp X}}(T_i(Z)=T_i(X)|X)}} \mathbb{P}_{\substack{Z\sim P \\ Z \perp X}}(T_i(Z)=k) \\
&= \sup_{P \in H_{0,i}} \mathbb{P}_{Z\sim P}\left(T_i(Z) \in \{k \in \mathcal{A}_{i,P} : (T_i)_{\#P}(\{k\}) \leq (T_i)_{\#P}(\{T_i(X)\})\}\right) \\
&= \sup_{P \in H_{0,i}} \sum_{\substack{k \in \mathcal{A}_{i,P} \\ (T_i)_{\#P}(\{k\}) \leq (T_i)_{\#P}(\{T_i(X)\})}} (T_i)_{\#P}(\{k\}) \\
&= \sup_{P \in H_{0,i}} (T_i)_{\#P}\left(\{k \in \mathcal{A}_{i,P} : (T_i)_{\#P}(\{k\}) \leq (T_i)_{\#P}(\{T_i(X)\})\}\right)
\end{aligned}$$

- $\check{p}_i = \mathbb{P}($ to realize a value of the support lesser or as common as $T_i(X))$

Generic construction of p-values with discreteness

Following the idea of "the probability of an event at least as extreme as"

Theorem

$\check{p}_i$ is an appropriate p-value, that is, it is super-uniform under the null:
Let $Q \in H_{0,i}$, $X \sim Q$, then

$$\forall x \in \mathbb{R}, \mathbb{P}\left(\check{p}_i(X) \leq x\right) \leq u(x). \quad (4)$$

- This is actually more general than with discrete support given that discrete and $\subseteq \mathbb{R} \Rightarrow$ countable but not the reverse
- Proof:
- As before, $\check{p}_i(X) \geq \check{p}_{i,Q}(X) = \sum_{\substack{k \in \mathcal{A}_{i,Q} \\ (T_i)_{\#Q}(\{k\}) \leq (T_i)_{\#Q}(\{T_i(X)\})}} (T_i)_{\#Q}(\{k\})$ so proving $\check{p}_{i,Q}(X) \succeq \mathcal{U}([0,1])$ is sufficient
- As before, $\check{p}_{i,Q}(X) \in [0,1]$ a.s. and right-continuity of the c.d.f so we only need to check (4) for $x \in]0,1[$

Generic construction of p-values with discreteness
Proof

- Note that
$$\check{p}_{i,Q}(X) \in \mathcal{S}_{i,Q} = \left\{ \sum_{\substack{k \in \mathcal{A}_{i,Q} \\ (T_i)_{\#Q}(\{k\}) \leq (T_i)_{\#Q}(\{\ell\})}} (T_i)_{\#Q}(\{k\}) : \ell \in \mathcal{A}_{i,Q} \right\}$$
a.s., and $\mathcal{S}_{i,Q}$ is countable or finite

Generic construction of p-values with discreteness

Proof

- Also note that for $x \in \mathcal{S}_{i,Q}$, $x = \sum_{\substack{k\in\mathcal{A}_{i,Q}\\ (T_i)_{\#Q}(\{k\})\leq (T_i)_{\#Q}(\{\ell\})}} (T_i)_{\#Q}(\{k\})$, $\ell \in \mathcal{A}_{i,Q}$, the c.d.f. of $s_{i,Q}(X)$ in x is

$$\mathbb{P}\left(\check{p}_{i,Q}(X) \leq x\right) = \mathbb{P}\left(\sum_{\substack{k\in\mathcal{A}_{i,Q}\\ (T_i)_{\#Q}(\{k\})\leq (T_i)_{\#Q}(\{T_i(X)\})}} (T_i)_{\#Q}(\{k\}) \leq \sum_{\substack{k\in\mathcal{A}_{i,Q}\\ (T_i)_{\#Q}(\{k\})\leq (T}} \right.$$

$$= \mathbb{P}\left((T_i)_{\#Q}(\{T_i(X)\}) \leq (T_i)_{\#Q}(\{\ell\})\right)$$

$$= \mathbb{P}\left(T_i(X) \in \{k \in \mathcal{A}_{i,Q} : (T_i)_{\#Q}(\{k\}) \leq (T_i)_{\#Q}(\{\ell\})\}\right)$$

$$= (T_i)_{\#Q}\left(\{k \in \mathcal{A}_{i,Q} : (T_i)_{\#Q}(\{k\}) \leq (T_i)_{\#Q}(\{\ell\})\}\right)$$

$$= \sum_{\substack{k\in\mathcal{A}_{i,Q}\\ (T_i)_{\#Q}(\{k\})\leq (T_i)_{\#Q}(\{\ell\})}} (T_i)_{\#Q}(\{k\}) = x$$

- $\Rightarrow$ The c.d.f. of $\check{p}_{i,Q}(X)$ is the identity on the support of $\check{p}_{i,Q}(X)$

Generic construction of p-values with discreteness

Proof

- Let $x \in]0,1[$, if $x < x'$ for all $x' \in \mathcal{S}_{i,Q}$ then $\mathbb{P}\left(\check{p}_{i,Q}(X) \leq x\right) = 0 \leq x$
- Else let $\underline{x} = \sup\{x' \in \mathcal{S}_{i,Q}, x' \leq x\}$ and note that $\mathbb{P}\left(\check{p}_{i,Q}(X) \leq x\right) = \mathbb{P}\left(\check{p}_{i,Q}(X) \leq \underline{x}\right)$
- If $\underline{x} \in \mathcal{S}_{i,Q}$ (i.e. it's a max, e.g. if $\mathcal{A}_{i,Q}$ is finite), then $\mathbb{P}\left(\check{p}_{i,Q}(X) \leq \underline{x}\right) = \underline{x} \leq x$
- Else, $\mathbb{P}\left(\check{p}_{i,Q}(X) \leq \underline{x}\right) = \mathbb{P}\left(\check{p}_{i,Q}(X) < \underline{x}\right) = \lim_{\substack{t \to \underline{x} \\ t < \underline{x}}} \mathbb{P}\left(\check{p}_{i,Q}(X) < t\right)$ by left-continuity
- Let $t_n \in \{x' \in \mathcal{S}_{i,Q}, x' \leq x\}, t_n \to \underline{x}$, $\mathbb{P}\left(\check{p}_{i,Q}(X) < t_n\right) \leq \mathbb{P}\left(\check{p}_{i,Q}(X) \leq t_n\right) = t_n \leq x$ $\square$

Generic construction of p-values with discreteness

Corollary

If, for all $P \in H_{0,i}$, $(T_i)_{\#P}$ does not depend on P, and if $\mathcal{A}_i = \mathcal{A}_{i,P}$ is finite, then $\mathcal{S}_i = \mathcal{S}_{i,P}$ is finite too and we can order its elements $x_1 < \cdots < x_N = 1$ for some N and describe the c.d.f. of $\check{p}_i(X)$ really simply:

$$\forall P \in H_{0,i}, X \sim P, \forall x \in \mathbb{R},$$

$$\mathbb{P}\left(\check{p}_i(X) \le x\right) = \begin{cases} 0 & \text{if} \quad x < x_1 \\ x_n & \text{if} \quad x_n \le x < x_{n+1}, n < N \\ 1 & \text{if} \quad x \ge 1 \end{cases} \tag{5}$$

- Denote $k_1, \ldots, k_D$ the distinct elements of $\mathcal{A}_i$ and order $(T_i)_{\#P}\left(\{k_\cdot\}\right)_{(1)} \le \ldots \le (T_i)_{\#P}\left(\{k_\cdot\}\right)_{(D)}$
- Assuming all are $\neq 0$ (so take the smallest $\mathcal{A}_i$ possible) and no ties, then $N = D$ and $x_n = \sum_{\nu=1}^{n} (T_i)_{\#P}\left(\{k_\cdot\}\right)_{(\nu)}$, in particular $\mathbb{P}\left(\check{p}_i(X) = x_n\right) = (T_i)_{\#P}\left(\{k_\cdot\}\right)_{(n)}$
- $x_N = (T_i)_{\#P}\left(\mathcal{A}_i\right) = 1$, always
- If ties, $N < D$

Generic construction of p-values with discreteness

Remark

- If, for all $P \in H_{0,i}$, $\mathcal{A}_i = \mathcal{A}_{i,P}$ does not depend on P (here, $(T_i)_{\#P}$ can) and is finite, then the set of possible values for $\hat{p}_i(X)$, $\bar{p}_i(X)$, $\breve{p}_i(X)$, $\check{p}_i(X)$ are also finite and do not depend on P
- $\hat{p}_i(X) \in \left\{ \sup_{P \in H_{0,i}} \sum_{\substack{k \in \mathcal{A}_i \\ k \geq \ell}} (T_i)_{\#P}(\{k\}) : \ell \in \mathcal{A}_i \right\} =$
  $\left\{ \sup_{P \in H_{0,i}} (T_i)_{\#P}([\ell, \infty[) : \ell \in \mathcal{A}_i \right\}$ a.s.
- $\check{p}_i(X) \in \left\{ \sup_{P \in H_{0,i}} \sum_{\substack{k \in \mathcal{A}_i \\ (T_i)_{\#P}(\{k\}) \leq (T_i)_{\#P}(\{\ell\})}} (T_i)_{\#P}(\{k\}) : \ell \in \mathcal{A}_i \right\}$ a.s.

Generic construction of p-values with discreteness

Back to Fisher's test

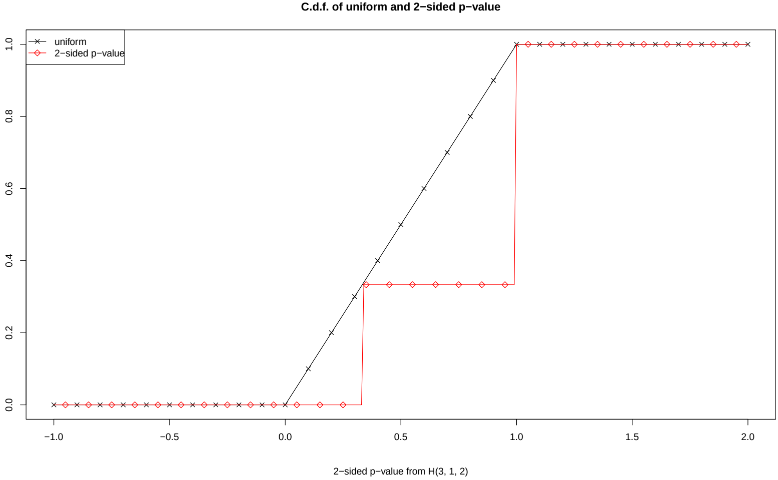
	Phenotype 1	Phenotype 2	Total
Allele A	0	1	$n_{\mathrm{A}} = 1$
Allele a	2	0	$n_{\mathrm{a}} = 2$
Total	$n_1 = 2$	$n_2 = 1$	$N = 3$

- Conditionnally to $n_{\mathrm{A}} = 1, n_1 = 2$, without association, $n_{1,\mathrm{A}} \sim \mathcal{H}(3, 1, 2) = P_0$
- $P_0(\{0\}) = \frac{\binom{2}{0}\binom{1}{1}}{\binom{3}{1}} = \frac{1}{3}$, $P_0(\{1\}) = \frac{\binom{2}{1}\binom{1}{0}}{\binom{3}{1}} = \frac{2}{3}$
- Then $\breve{p}_i(n_{1,\mathrm{A}}) = 2\min(P_0(]-\infty, n_{1,\mathrm{A}}]), P_0([n_{1,\mathrm{A}}, \infty[)) = \frac{2}{3}\mathbb{1}_{\{n_{1,\mathrm{A}}=0\}} + \frac{4}{3}\mathbb{1}_{\{n_{1,\mathrm{A}}=1\}}$
- Whereas
$$\breve{p}_i(n_{1,\mathrm{A}}) = \sum_{\substack{k\in\{0,1\} \\ P_0(\{k\})\leq P_0(\{n_{1,\mathrm{A}}\})}} P_0(\{k\}) = \frac{1}{3}\mathbb{1}_{\{n_{1,\mathrm{A}}=0\}} + \frac{2}{3}\mathbb{1}_{\{n_{1,\mathrm{A}}=1\}}$$
- Clearly $\breve{p}_i(n_{1,\mathrm{A}})$ is less conservative than $\breve{p}_i(n_{1,\mathrm{A}})$, furthermore $\breve{p}_i(X)$ can be > 1 (as soon as X has an atom of $\mathbb{P} > \frac{1}{2}$, and opening one interval makes it invalid)

Generic construction of p-values with discreteness

Back to Fisher's test

- C.d.f. of $\check{p}_i(n_{1,\mathrm{A}})$ under H_0, that is if if $n_{1,\mathrm{A}} \sim \mathcal{H}(3,1,2)$:



The issue with discrete p-values

Strict super-uniformity

$$\forall P \in H_{0,i}, X \sim P, \mathbb{P}\left(p_i(X) \leq x\right) \leq u(x) \text{ and } \exists x, \mathbb{P}\left(p_i(X) \leq x\right) < u(x)$$

i.e. under the null, our p-values are larger than uniforms

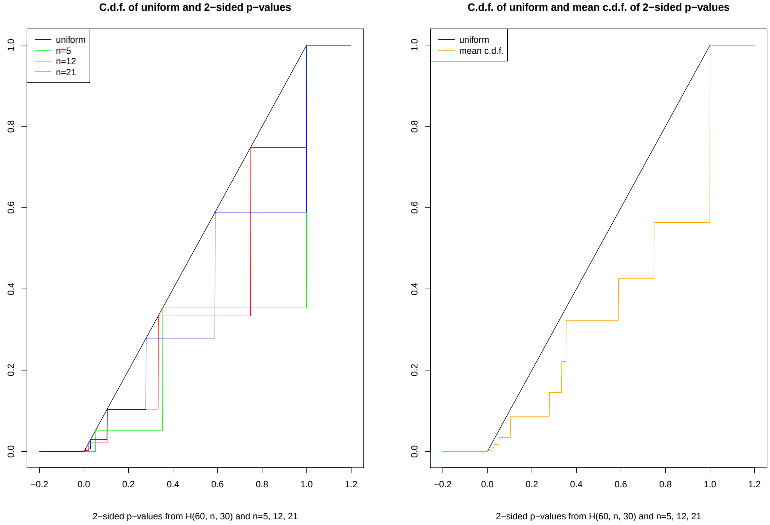
Problem

Usual MT procedures designed for uniform p-values (seen as the worst case)

- As discrete p-values are larger than uniforms, classic thresholds are too low, too conservative $\Longrightarrow$ loss of power
- Goal: use the knowledge of the discrete c.d.f. under the null to improve power

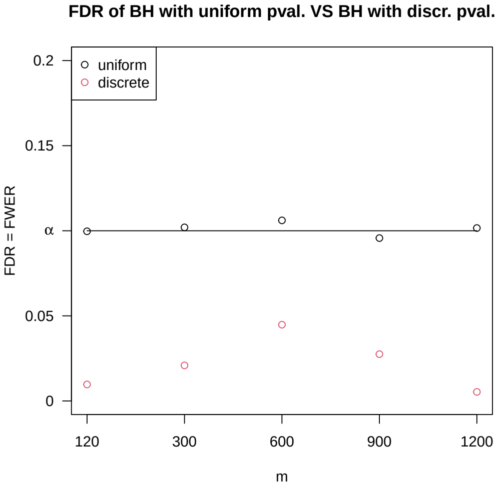
The issue with discrete p-values

- C.d.f. plots of 2-sided p-values associated with $\mathcal{H}(60, 5, 30)$, $\mathcal{H}(60, 12, 30)$ and $\mathcal{H}(60, 21, 30)$



The issue with discrete p-values

- BH under full null, $m/3$ 2-sided p-values derived from $\mathcal{H}(60, 5, 30)$, $m/3$ from $\mathcal{H}(60, 12, 30)$, $m/3$ from $\mathcal{H}(60, 21, 30)$
- MC estimation of the FDR with 10^4 replications



Assumption for the remainder of the section

- $\exists$ a finite set $\mathcal{S}_i$ such that $\forall P \in H_{0,i}, X \sim P, \mathbb{P}(p_i(X) \in \mathcal{S}_i) = 1$
- See previous remark for a sufficient condition and construction
- Also let $\underline{s}_i = \min \mathcal{S}_i$

Tarone-Bonferroni procedures

Increasing power for discrete tests [Tarone (1990)]

- A simple idea: if $\underline{s}_i > \alpha$, $H_{0,i}$ can never be wrongly rejected so might not count it when adjusting for multiplicity
- Let $R_1 = \{i \in [\![1, m]\!] : \underline{s}_i \leq \alpha\}$, $m(1) = |R_1|$, $\hat{t}_\alpha^{\text{TB}} = \frac{\alpha}{m(1)}$ and $R^{\text{TB}} = R\left(\hat{t}_\alpha^{\text{TB}}\right)$
- $\hat{t}_\alpha^{\text{TB}} \geq \hat{t}_\alpha^{\text{Bonf}}$: less conservative than Bonferroni
- $\text{FWER}\left(R^{\text{TB}}\right) = \sum_{i \in \mathcal{H}_0} \mathbb{P}\left(p_i \leq \frac{\alpha}{m(1)}\right) = \sum_{i \in \mathcal{H}_0 \cap R_1} \mathbb{P}\left(p_i \leq \frac{\alpha}{m(1)}\right) \leq \alpha \frac{|\mathcal{H}_0 \cap R_1|}{m(1)} \leq \alpha$ □
- We can do better : $\forall k \in [\![1, m]\!]$, let $R_k = \left\{i \in [\![1, m]\!] : \underline{s}_i \leq \frac{\alpha}{k}\right\}$, $m(k) = |R_k|$, actually $\text{FWER}\left(R^{\text{TB}}\right)$ is bounded by $\alpha \frac{|\mathcal{H}_0 \cap R_{m(1)}|}{m(1)}$ which is even smaller ⇒ "fixed point" research

Tarone-Bonferroni procedures

Increasing power for discrete tests

- Let $K^* = \min\{k \in [\![1, m]\!] : m(k) \leq k\}$, non-empty set because $m(m(1)) \leq m(1)$, $\hat{t}_\alpha^{\text{TB-ref}} = \frac{\alpha}{K^*}$ and $R^{\text{TB-ref}} = R\left(\hat{t}_\alpha^{\text{TB-ref}}\right)$
- For any fixed k,

$$\forall P \in \mathfrak{F}, \mathbb{P}\left(\exists i \in \mathcal{H}_0 : p_i \leq \frac{\alpha}{k}\right) \leq \sum_{i \in \mathcal{H}_0 \cap R_k} \mathbb{P}\left(p_i \leq \frac{\alpha}{k}\right) \leq \alpha \frac{m(k)}{k},$$

  which shows that $\text{FWER}\left(R^{\text{TB}}\right), \text{FWER}\left(R^{\text{TB-ref}}\right) \leq \alpha$ □
- K^* is the optimal choice, TB-refined is even less conservative

FDR control with discrete p-values

Heyse procedure

- Recall the previous plot of the mean c.d.f. of the discrete p-values
- $\Rightarrow$ idea: "invert" this mean c.d.f. at $\alpha\frac{k}{m}$ and apply a SU procedure [Heyse (2011)]
- Let $F_i : t \mapsto \sup_{P\in H_{0,i}} \mathbb{P}_{X\sim P}\left(p_i(X) \leq t\right)$: worst-case c.d.f., and $\overline{F}(t) = \frac{1}{m}\sum_{i=1}^{m} F_i(t)$
- Let $\mathcal{S} = \bigcup_{i=1}^{m} \mathcal{S}_i$, $\tau_k = \max\left\{t \in \mathcal{S} : \overline{F}(t) \leq \alpha\frac{k}{m}\right\}$
- $R^{\text{Heyse}} = R^{\text{SU}}\left(\tau\right)$
- BH is also the SU procedure with $\xi_k = \max\left\{t \in \mathcal{S} : t \leq \alpha\frac{k}{m}\right\}$ (effective critical values), $\overline{F}\left(\xi_k\right) \leq \xi_k \leq \alpha\frac{k}{m}$ so $\tau_k \geq \xi_k$: Heyse less conservative than BH, only with heterogeneity though: if $F_i = F_j = \overline{F}$ and the assumption $F_i(t) = t$, $\forall t \in \mathcal{S}_i = \mathcal{S}$ then $\overline{F}(t) = t$ for all $t \in \mathcal{S}$ and $\tau_k = \xi_k$
- Problem: R^{Heyse} doesn't control the FDR! [Döhler, Durand, and Roquain (2018)]

FDR control with discrete p-values

- Heyse *almost* works though, it works up to a small rescaling factor
- Let $\tau_m = \max\left\{t \in \mathcal{S} : \frac{1}{m}\sum_{i=1}^{m} \frac{F_i(t)}{1-F_i(t)} \leq \alpha\right\}$
- For $k < m$, let $\tau_k = \max\left\{t \in \mathcal{S} : t \leq \tau_m, \sum_{i=1}^{m} \frac{F_i(t)}{1-F_i(\tau_m)} \leq \alpha k\right\}$
- Let $R^{\text{HSU}} = R^{\text{SU}}(\tau)$
- Can be more conservative than BH but not that much, and in practice isn't

Theorem [Döhler, Durand, and Roquain (2018)]

Assume that for all $P \in \mathfrak{F}$, the $(p_i)_{i\in\mathcal{H}_0}$ are independent, and they are independent from the $(p_i)_{i\in\mathcal{H}_1}$.
Then for all $P \in \mathfrak{F}$,

$$\text{FDR}\left(R^{\text{HSU}}\right) \leq \alpha$$

FDR control with discrete p-values

- We can do even better by implicit adaptivity to m_0:
- Let τ_m the same
- For $k < m$, let $\tau_k =$ $\max\left\{t \in \mathcal{S} : t \leq \tau_m, \left(\left(\frac{F_i(t)}{1-F_i(\tau_m)}\right)_{(1)} + \cdots + \left(\frac{F_i(t)}{1-F_i(\tau_m)}\right)_{(m-k+1)}\right) \leq \alpha k\right\}$
- Idea: if k "good" rejections, $m_0 \leq m - k + 1$ so only control needed for the worst case with $m - k + 1$ kept null hypotheses
- Let $R^{\text{AHSU}} = R^{\text{SU}}(\tau)$
- Less conservative than HSU because of larger critical values

Theorem [Döhler, Durand, and Roquain (2018)]

Assume that for all $P \in \mathfrak{F}$, the $(p_i)_{i \in \mathcal{H}_0}$ are independent, and they are independent from the $(p_i)_{i \in \mathcal{H}_1}$.
Then for all $P \in \mathfrak{F}$,

$$\text{FDR}\left(R^{\text{AHSU}}\right) \leq \alpha$$

FDR control with discrete p-values

- Both FDR controls come from the same bound

Theorem [Döhler, Durand, and Roquain (2018)]

Assume that for all $P \in \mathfrak{F}$, the $(p_i)_{i\in\mathcal{H}_0}$ are independent, and they are independent from the $(p_i)_{i\in\mathcal{H}_1}$. Let any critical value sequence τ with $F_i(\tau_m) < 1$ for all $i \in [\![1, m]\!]$.
Then for all $P \in \mathfrak{F}$,

$$\mathrm{FDR}\left(R^{\mathrm{SU}}(\tau)\right) \leq \max_{1\leq k\leq m} \max_{\substack{A\subseteq[\![1,m]\!]\\ |A|=m-k+1}} \frac{1}{k}\sum_{i\in A}\frac{F_i(\tau_k)}{1-F_i(\tau_m)}$$

FDR control with discrete p-values

Proof of the Theorem

- Recall the Lemma on SU procedures:
  $\{p_i \leq \tau_{\hat{k}^{\mathrm{SU}}},\ \hat{k}^{\mathrm{SU}} = k\} = \{p_i \leq \tau_k,\ \hat{k}^{-i} = k-1\}$
- Another one: let $(\sigma_1, \dots, \sigma_m) = (\tau_2, \dots, \tau_m, \tau_m)$ and
  $\hat{k}^{\#} = \max\{k : p_{(k)} \leq \sigma_k\} = \left|R^{\mathrm{SU}}(\sigma)\right|$. Then $p_i > \tau_m \Rightarrow \hat{k}^{-i} = \hat{k}^{\#}$
- Proof: $p_{\left(\hat{k}^{-i}\right)} \leq p^{-i}_{\left(\hat{k}^{-i}\right)} \leq \tau^{-i}_{\hat{k}^{-i}} = \sigma_{\hat{k}^{-i}}$ so $\hat{k}^{-i} \leq \hat{k}^{\#}$, always
- Let $p_i = p_{(k_i)}$, note that $p^{-i}_{(k)} = p_{(k)}$ for all $k < k_i$ and $p^{-i}_{(k)} = p_{(k+1)}$ for all $m-1 \geq k \geq k_i$
- $p_i > \tau_m$ entails $p_{(k_i)} = p_i > \tau_m \geq \sigma_{\hat{k}^{\#}} \geq p_{(\hat{k}^{\#})}$ so $k_i > \hat{k}^{\#}$ (also entails $m > \hat{k}^{\#}$) so $p^{-i}_{(\hat{k}^{\#})} = p_{(\hat{k}^{\#})}$
- Finally $p^{-i}_{(\hat{k}^{\#})} = p_{(\hat{k}^{\#})} \leq \sigma_{\hat{k}^{\#}} = \tau^{-i}_{\hat{k}^{\#}}$ and $\hat{k}^{\#} \leq \hat{k}^{-i}$ $\square$

- Starts like the proof of BH:

$$\begin{aligned}
\mathrm{FDR}\left(R^{\mathrm{SU}}(\tau)\right) &= \sum_{i\in\mathcal{H}_0}\sum_{k=1}^{m}\frac{1}{k}\mathbb{P}\left(p_i \le \tau_k\right)\mathbb{P}\left(\hat{k}^{-i}=k-1\right)\\
&\le \sum_{i\in\mathcal{H}_0}\sum_{k=1}^{m}\frac{1}{k}F_i(\tau_k)\mathbb{P}\left(\hat{k}^{-i}=k-1\right)\\
&\le \sum_{i\in\mathcal{H}_0}\mathbb{E}\left[\frac{F_i\left(\tau_{\hat{k}^{-i}+1}\right)}{\hat{k}^{-i}+1}\right]
\end{aligned}$$

FDR control with discrete p-values

Proof of the Theorem

- Hide 1: $1 - F_i(\tau_m) \leq 1 - \mathbb{P}(p_i \leq \tau_m)$ so

$$
\begin{aligned}
1 &\leq \frac{\mathbb{P}(p_i > \tau_m)}{1 - F_i(\tau_m)} \\
&= \mathbb{E}\left[\frac{\mathbb{1}_{\{p_i > \tau_m\}}}{1 - F_i(\tau_m)}\right] \\
&= \mathbb{E}\left[\frac{\mathbb{1}_{\{p_i > \tau_m\}}}{1 - F_i(\tau_m)} \middle| \hat{k}^{-i}\right] \text{ by independence}
\end{aligned}
$$

- Hence

$$\begin{aligned}
\mathrm{FDR}\left(R^{\mathrm{SU}}(\tau)\right) &\leq \sum_{i\in\mathcal{H}_0} \mathbb{E}\left[\frac{F_i\left(\tau_{\hat{k}^{-i}+1}\right)}{\hat{k}^{-i}+1} \times 1\right] \\
&\leq \sum_{i\in\mathcal{H}_0} \mathbb{E}\left[\frac{F_i\left(\tau_{\hat{k}^{-i}+1}\right)}{\hat{k}^{-i}+1} \times \mathbb{E}\left[\frac{\mathbb{1}_{\{p_i>\tau_m\}}}{1-F_i(\tau_m)}\middle|\hat{k}^{-i}\right]\right] \\
&= \sum_{i\in\mathcal{H}_0} \mathbb{E}\left[\frac{F_i\left(\tau_{\hat{k}^{-i}+1}\right)}{\hat{k}^{-i}+1} \frac{\mathbb{1}_{\{p_i>\tau_m\}}}{1-F_i(\tau_m)}\right] \\
&\leq \sum_{i\in\mathcal{H}_0} \mathbb{E}\left[\frac{F_i\left(\tau_{\hat{k}^{\#}+1}\right)}{1-F_i(\tau_m)} \frac{\mathbb{1}_{\{p_i>\tau_m\}}}{\hat{k}^{\#}+1} \mathbb{1}_{\{\hat{k}^{\#}<m\}}\right] \quad \text{by the new Lemma}
\end{aligned}$$

FDR control with discrete p-values

Proof of the Theorem

- Hence

$$\mathrm{FDR}\left(R^{\mathrm{SU}}(\tau)\right) \leq \sum_{i\in\mathcal{H}_0} \mathbb{E}\left[\frac{F_i\left(\tau_{\hat{k}^\#+1}\right)}{1-F_i(\tau_m)}\frac{\mathbb{1}_{\{p_i>\sigma_{\hat{k}^\#}\}}}{\hat{k}^\#+1}\mathbb{1}_{\{\hat{k}^\#<m\}}\right] \text{ because } \tau_m \geq \sigma_{\hat{k}}$$

$$\leq \mathbb{E}\left[\sum_{i\in\mathcal{H}_0}\frac{F_i\left(\tau_{\hat{k}^\#+1}\right)}{1-F_i(\tau_m)}\frac{\mathbb{1}_{\{p_i>\sigma_{\hat{k}^\#}\}}}{\hat{k}^\#+1}\mathbb{1}_{\{\hat{k}^\#<m\}}\right]$$

- $A = \{i : p_i > \sigma_{\hat{k}^\#}\} = [\![1, m]\!] \setminus R^{\mathrm{SU}}(\sigma)$ so $|A| = m - \hat{k}^\#$ by property of SU

$$\mathrm{FDR}\left(R^{\mathrm{SU}}(\tau)\right) \leq \mathbb{E}\left[\max_{\substack{A\subseteq[\![1,m]\!]\\|A|=m-\hat{k}^\#}}\sum_{i\in\mathcal{H}_0\cap A}\frac{F_i\left(\tau_{\hat{k}^\#+1}\right)}{1-F_i(\tau_m)}\frac{1}{\hat{k}^\#+1}\mathbb{1}_{\{\hat{k}^\#<m\}}\right]$$

$$\leq \max_{0\leq k\leq m-1}\max_{\substack{A\subseteq[\![1,m]\!]\\|A|=m-k}}\sum_{i\in\mathcal{H}_0\cap A}\frac{F_i\left(\tau_{k+1}\right)}{1-F_i(\tau_m)}\frac{1}{k+1} \qquad \square$$

FDR control with discrete p-values

- Analog Lemmas for FDR bound and procedures SD
- HSD: SD with $\tau_k = \max\left\{t \in \mathcal{S} : \sum_{i=1}^m \frac{F_i(t)}{1-F_i(t)} \leq \alpha k\right\}$
- AHSD: SD with
  $\tau_k = \max\left\{t \in \mathcal{S} : \left(\left(\frac{F_i(t)}{1-F_i(t)}\right)_{(1)} + \cdots + \left(\frac{F_i(t)}{1-F_i(t)}\right)_{(m-k+1)}\right) \leq \alpha k\right\}$
- Higher critical values than HSU and AHSU, but SD: no one generally better than the other

Theorem [Döhler, Durand, and Roquain (2018)]

Assume that for all $P \in \mathfrak{F}$, the $(p_i)_{i\in\mathcal{H}_0}$ are independent, and they are independent from the $(p_i)_{i\in\mathcal{H}_1}$. Let any critical value sequence τ with $F_i(\tau_m) < 1$ for all $i \in [\![1, m]\!]$.
Then for all $P \in \mathfrak{F}$,

$$\mathrm{FDR}\left(R^{\mathrm{SD}}(\tau)\right) \leq \max_{1\leq k\leq m} \max_{\substack{A\subseteq[\![1,m]\!] \\ |A|=m-k+1}} \frac{1}{k} \sum_{i\in A} \frac{F_i(\tau_k)}{1-F_i(\tau_k)}$$

Table of contents

Exploratory analysis in multiple testing

Exploratory analysis: searching interesting hypotheses that will be cautiously investigated after.

Desired properties [Goeman and Solari (2011)]:

- Mildness: allows some false positives
- Flexibility: the procedure does not prescribe, but advise
- Post hoc: take decisions on the procedure after seing the data

> [Goeman and Solari (2011)]
>
> This **reverses the traditional roles** of the user and procedure in multiple testing. Rather than [...] to let the user choose the quality criterion, and to let the procedure return the collection of rejected hypotheses, the **user chooses the collection of rejected hypotheses freely**, and the multiple testing procedure returns the **associated quality criterion**.

FWER is somewhat flexible, FDR is somewhat mild

Post hoc and replication crisis

Post hoc done wrong: p-hacking

- Pre-selecting variables that seem significant, exclude others
- Theoretical results no longer hold because the selection step is random
- Example: selecting the 1000 smallest p-values in a genetic study with 10^6 variants

- p-hacking may be one of the causes of the replication crisis (many published results non reproducible)
- $\Rightarrow$ need for exploratory analysis MT procedures with the above properties
- Larger field: selective inference

Post hoc inference

a.k.a. simultaneous inference

Confidence bounds on any set of selected variables

A confidence bound is a (random: depends on X) function $\widehat{V}$ such that

$$\forall P \in \mathfrak{F}, \forall \alpha \in]0,1[, \mathbb{P}\left(\forall S \subset [\![1,m]\!], V(S) \leq \widehat{V}(S)\right) \geq 1-\alpha$$

- Hence for any selected $\widehat{S} = \widehat{S}(X)$, $\mathbb{P}\left(V(\widehat{S}) \leq \widehat{V}(\widehat{S})\right) \geq 1-\alpha$ holds
- Also FDP control: $\mathbb{P}\left(\forall S \subset [\![1,m]\!], \text{FDP}(S) \leq \frac{\widehat{V}(S)}{|S| \vee 1}\right) \geq 1-\alpha$, hence (far) better than FDR control
- Originates from [Genovese and Wasserman (2006)],[Meinshausen (2006)]
- A guarantee over any selected set instead of a rejected set, advise some $\widehat{S}$ instead of prescribe one R : the MT paradigm is reversed

Post hoc inference

Some first, trivial bounds

- $\widehat{V}(S) = |S|$
- Let a procedure R controlling the FWER, then $\widehat{V}(S) = |S \setminus R|$ is a valid post hoc bound

$$\mathbb{P}(\exists S : |S \cap \mathcal{H}_0| > |S \setminus R|) \leq \mathbb{P}(\exists S : |S \cap \mathcal{H}_0 \cap R^c| + |S \cap \mathcal{H}_0 \cap R| > |S \cap R$$
$$\leq \mathbb{P}(\exists S : |S \cap \mathcal{H}_0 \cap R| > 0)$$
$$\leq \mathbb{P}(|\mathcal{H}_0 \cap R| > 0) \leq \alpha \quad \square$$

- Let a procedure R controlling the k-FWER, then $\widehat{V}(S) = |S \setminus R| + k - 1$ is a valid post hoc bound

$$\mathbb{P}(\exists S : |S \cap \mathcal{H}_0| > |S \setminus R| + k - 1) \leq \mathbb{P}(\exists S : |S \cap \mathcal{H}_0 \cap R| > k - 1)$$
$$\leq \mathbb{P}(|\mathcal{H}_0 \cap R| > k - 1) \leq \alpha \quad \square$$

BNR technology

[Blanchard, Neuvial, and Roquain (2020)]

Key concept: reference family

- $\mathfrak{R} = (R_k, \zeta_k)_{k \in \mathcal{K}}$ with $R_k \subseteq [\![1, m]\!]$, $\zeta_k \in [\![0, |R_k|]\!]$ (everything can depend on X) such that the Joint Error Rate (JER):

$$\mathrm{JER}(\mathfrak{R}) = \mathbb{P}\left(\exists k, |R_k \cap \mathcal{H}_0| > \zeta_k\right)$$

  is controlled at level α for all $P \in \mathfrak{F}$
- Conversely, $\forall P \in \mathfrak{F}, \mathbb{P}_{X \sim P}\left(\forall k, |R_k \cap \mathcal{H}_0| \leq \zeta_k\right) \geq 1 - \alpha$
- Confidence bound only on the $K = |\mathcal{K}|$ members of $\mathfrak{R}$
- $\Longrightarrow$ Derivation of a global confidence bound by interpolation

BNR technology

[Blanchard, Neuvial, and Roquain (2020)]

- Idea: we get the following info on $\mathcal{H}_0$:
  $\mathcal{H}_0 \in \mathcal{A}(\mathfrak{R}) = \{A \subseteq [\![1, m]\!], \forall k, |R_k \cap A| \leq \zeta_k\}$

Two different bounds

- $V^*_{\mathfrak{R}}(S) = \max_{A \in \mathcal{A}(\mathfrak{R})} |S \cap A|$ optimal but hard to compute
- $\overline{V}_{\mathfrak{R}}(S) = \min_{k \in \mathcal{K}} (\zeta_k + |S \setminus R_k|) \wedge |S|$ easier to compute

- $\overline{V}_{\mathfrak{R}}$ is worse than $V^*_{\mathfrak{R}}$, proof: let $A \in \mathcal{A}(\mathfrak{R})$
- $|S \cap A| = |S \cap A \cap R_k| + |S \cap A \cap R_k{}^c| \leq |A \cap R_k| + |S \cap R_k{}^c| \leq \zeta_k + |S \setminus R_k|$
- True for all k: $|S \cap A| \leq \overline{V}_{\mathfrak{R}}(S)$, true for all A: $V^*_{\mathfrak{R}}(S) \leq \overline{V}_{\mathfrak{R}}(S)$ □

BNR technology

Proposition

Assume that the R_k's are nested, that is $R_k \subseteq R_{k'}$ or $R_{k'} \subseteq R_k$ for $k, k' \in \mathcal{K}$. Then $V^*_{\mathfrak{R}}(S) = \overline{V}_{\mathfrak{R}}(S)$ for all $S \subseteq [\![1, m]\!]$.

- In the following we identify $\mathcal{K}$ and $[\![1, K]\!]$ such that $R_k \subseteq R_{k'}$ for $k \leq k'$
- $\overline{V}_{\mathfrak{R}}(S) = \min_{k \leq K} (\zeta_k + |S \setminus R_k|) \wedge |S| = \min_{k \leq K} (\zeta_k + |S \setminus (R_k \cap S)|) \wedge |S| = \overline{V}_{\mathfrak{R}_{\wedge S}}(S)$ with $\mathfrak{R}_{\wedge S} = (R_k \cap S, \zeta_k)_{k \leq K}$
- Let $\tilde{\zeta}_k = \overline{V}_{\mathfrak{R}_{\wedge S}}(R_k \cap S) = \min_{j \leq K}(\zeta_j + |(R_k \cap S) \setminus (R_j \cap S)|) \wedge |R_k \cap S|$ and consider $\tilde{\mathfrak{R}} = (R_k \cap S, \tilde{\zeta}_k)_{k \leq K}$
- By taking $j = k$, $\tilde{\zeta}_k \leq (\zeta_k + |(R_k \cap S) \setminus (R_k \cap S)|) \wedge |R_k \cap S| \leq \zeta_k$ so $\overline{V}_{\tilde{\mathfrak{R}}}(S) \leq \overline{V}_{\mathfrak{R}_{\wedge S}}(S) = \overline{V}_{\mathfrak{R}}(S)$

BNR technology

Proof of the Proposition

- Useful set property : $|E \setminus G| \leq |E \setminus F| + |F \setminus G|$

$$
\begin{aligned}
\overline{V}_{\tilde{\mathfrak{R}}}(S) &= \min_{k \leq K} \left(\min_{j \leq K} (\zeta_j + |(R_k \cap S) \setminus (R_j \cap S)|) \wedge |R_k \cap S| + |S \setminus (R_k \cap S)| \right) \\
&= \min_{k \leq K} \left(\min_{j \leq K} (\zeta_j + |(R_k \cap S) \setminus (R_j \cap S)|) + |S \setminus (R_k \cap S)| \right) \wedge |S| \\
&= \min_{k \leq K} \min_{j \leq K} (\zeta_j + |(R_k \cap S) \setminus (R_j \cap S)| + |S \setminus (R_k \cap S)|) \wedge |S| \\
&\geq \min_{j \leq K} (\zeta_j + |S \setminus (R_j \cap S)|) \wedge |S| = \overline{V}_{\mathfrak{R}_{\wedge S}}(S) = \overline{V}_{\mathfrak{R}}(S)
\end{aligned}
$$

- So $\overline{V}_{\tilde{\mathfrak{R}}}(S) = \overline{V}_{\mathfrak{R}}(S)$ (self-consistency result)
- Remark: this intermediate result does not use the nestedness and is true in general

BNR technology

Proof of the Proposition

- Let's construct $A \subseteq S$, $A \in \mathcal{A}(\mathfrak{R})$ such that $|A| \geq \overline{V}_{\mathfrak{R}}(S)$, will imply $V^*_{\mathfrak{R}}(S) \geq \overline{V}_{\mathfrak{R}}(S)$
- By nestedness,

$$
\begin{aligned}
\tilde{\zeta}_k &= \min_{j \leq K}(\zeta_j + |(R_k \cap S) \setminus (R_j \cap S)|) \wedge |R_k \cap S| \\
&\leq \min_{j \leq K}(\zeta_j + |(R_{k+1} \cap S) \setminus (R_j \cap S)|) \wedge |R_{k+1} \cap S| \\
&= \tilde{\zeta}_{k+1}
\end{aligned}
$$

BNR technology

Proof of the Proposition

- Furthermore,

$$\begin{aligned}
\tilde{\zeta}_{k+1} &= \min_{j \le K}(\zeta_j + |(R_{k+1} \cap S) \setminus (R_j \cap S)|) \wedge |R_{k+1} \cap S| \\
&\le \min_{j \le K}(\zeta_j + |(R_{k+1} \cap S) \setminus (R_k \cap S)| + |(R_k \cap S) \setminus (R_j \cap S)|) \wedge |R_{k+1} \cap S \\
&= (|(R_{k+1} \cap S) \setminus (R_k \cap S)| + \min_{j \le K}(\zeta_j + |(R_k \cap S) \setminus (R_j \cap S)|)) \wedge |R_{k+1} \cap \\
&= |(R_{k+1} \cap S) \setminus (R_k \cap S)| + \min_{j \le K}(\zeta_j + |(R_k \cap S) \setminus (R_j \cap S)|) \wedge |R_k \cap S| \\
&\text{(nestedness: } |R_{k+1} \cap S| = |(R_{k+1} \cap S) \setminus (R_k \cap S)| + |R_k \cap S|) \\
&= |(R_{k+1} \cap S) \setminus (R_k \cap S)| + \tilde{\zeta}_k
\end{aligned}$$

- So $0 \le \tilde{\zeta}_{k+1} - \tilde{\zeta}_k \le |(R_{k+1} \cap S) \setminus (R_k \cap S)|$

BNR technology

Proof of the Proposition

- Let $B_k = \{\tilde{\zeta}_k - \tilde{\zeta}_{k-1}$ elements of $(R_k \cap S) \setminus (R_{k-1} \cap S)\}$, $1 \leq k \leq K$, with $R_0 = \varnothing$ and $\tilde{\zeta}_0 = 0$
- Let $A = \bigcup_{k=1}^{K} B_k \cup (S \setminus (R_K \cap S))$, disjoint union because of nestedness, $A \subseteq S$
- $|R_k \cap A| = \left|\bigcup_{\ell=1}^{k} B_\ell\right| = \sum_{\ell=1}^{k} |B_\ell| = \sum_{\ell=1}^{k} (\tilde{\zeta}_\ell - \tilde{\zeta}_{\ell-1}) = \tilde{\zeta}_k \leq \zeta_k$ so $A \in \mathcal{A}(\mathfrak{R})$
- $|A| = \sum_{\ell=1}^{K} |B_\ell| + |S \setminus (R_K \cap S)| = \tilde{\zeta}_K + |S \setminus (R_K \cap S)| = (\tilde{\zeta}_K + |S \setminus (R_K \cap S)|) \wedge |S|$ because $A \subseteq S$
- Finally $\overline{V}_{\mathfrak{R}}(S) = \overline{V}_{\tilde{\mathfrak{R}}}(S) \leq (\tilde{\zeta}_K + |S \setminus (R_K \cap S)|) \wedge |S| = |A|$ □

BNR technology

- How to construct effectively a reference family $(R_k, \zeta_k)_{k \in \mathcal{K}}$ with JER control?
- One approach: constrain $\zeta_k = k - 1$, $R_k = \{i \in [\![1, m]\!], p_i \leq t_k\}$, $k \in [\![1, m]\!]$, $t_k \nearrow$ and search for valid $(t_k)_{1 \leq k \leq m}$
- In this case, $R_k \subseteq R_{k+1}$: nestedness hence $\overline{V}_{\mathfrak{R}}$ optimal
- In this case, $\text{JER}(\mathfrak{R}) = \mathbb{P}(\exists k, |R_k \cap \mathcal{H}_0| \geq k)$: k-FWER but simultaneous over all k
- $(t_k)_{1 \leq k \leq m}$ can be constructed with probabilistic inequalities

Simes and Hommel inequalities

[Hommel (1983)], [Simes (1986)]

- Let $U_1, \dots U_{m_0}$ m_0 super-uniform random variables
- Then $\mathbb{P}\left(\exists i \leq m_0, U_{(i)} \leq \frac{\alpha i}{m_0 H_{m_0}}\right) \leq \alpha$ (Hommel inequality)
- If, furthermore, they are wPRDS on $[\![1, m_0]\!]$, $\mathbb{P}\left(\exists i \leq m_0, U_{(i)} \leq \frac{\alpha i}{m_0}\right) \leq \alpha$ (Simes inequality)

Simes and Hommel inequalities

Proofs

- Consider the model $\mathfrak{F}_U = \{\mathbb{P}_{U_1,\dots,U_{m_0}}\}$ with $H_{0,i} = \mathfrak{F}_U$ for all $i \in [\![1, m_0]\!]$, the U_i's are valid p-values
- Note that FWER = FDR when all null hypotheses are true, which is the case here

$$\mathbb{P}\left(\exists i \leq m_0, U_{(i)} \leq \frac{\alpha i}{m_0 H_{m_0}}\right) = \text{FWER}(R^{\text{BY}}) \\ = \text{FDR}(R^{\text{BY}}) \\ \leq \alpha \quad \square$$

- Same proof for Simes and wPRDS using the FDR control of BH $\square$

BNR technology

- Consequence: $\forall P \in \mathfrak{F}$,
  $\mathbb{P}\left(\exists i \leq m_0, p_{(i:\mathcal{H}_0)} \leq \frac{\alpha i}{mH_m}\right) \leq \mathbb{P}\left(\exists i \leq m_0, p_{(i:\mathcal{H}_0)} \leq \frac{\alpha i}{m_0 H_{m_0}}\right) \leq \alpha$ and similarly with wPRDS on $\mathcal{H}_0$, $\mathbb{P}\left(\exists i \leq m_0, p_{(i:\mathcal{H}_0)} \leq \frac{\alpha i}{m}\right) \leq \alpha$
- $t_k = \frac{\alpha k}{mH_m}$ induces JER control, and if wPRDS on $\mathcal{H}_0$ $\forall P \in \mathfrak{F}$, $t_k = \frac{\alpha k}{m}$ too
- Proof: let $c_m = H_m$ or 1 depending on the case (Hommel or Simes)

$$\begin{aligned}\exists k \leq K : |R_k \cap \mathcal{H}_0| \geq k &\Leftrightarrow \exists k \leq m_0 : \left|\left\{i \in [\![1, m]\!] : p_i \leq \frac{\alpha k}{mc_m}\right\} \cap \mathcal{H}_0\right| \geq k \\ &\Leftrightarrow \exists k \leq m_0 : \left|\left\{i \in \mathcal{H}_0 : p_i \leq \frac{\alpha k}{mc_m}\right\}\right| \geq k \\ &\Leftrightarrow \exists k \leq m_0 : p_{(k:\mathcal{H}_0)} \leq \frac{\alpha k}{mc_m} \quad \square\end{aligned}$$

BNR technology

Theorem [Blanchard, Neuvial, and Roquain (2020)]

The bound $V^*_{\mathfrak{R}_{Hommel}} : S \mapsto \min_{1\leq k\leq m}\left(k-1+\sum_{i\in S}\mathbb{1}_{\left\{p_i>\frac{\alpha k}{mH_m}\right\}}\right)\wedge|S|$ is a valid confidence bound, associated to the reference family $\mathfrak{R}_{Hommel}=\left(\left\{i : p_i\leq\frac{\alpha k}{mH_m}\right\}, k-1\right)_{k\in[\![1,m]\!]}$.

If, for all $P\in\mathfrak{F}$, the (p_i) are wPRDS with $\mathcal{H}_0$ as the subset, the bound $V^*_{\mathfrak{R}_{Simes}} : S \mapsto \min_{1\leq k\leq m}\left(k-1+\sum_{i\in S}\mathbb{1}_{\left\{p_i>\frac{\alpha k}{m}\right\}}\right)\wedge|S|$ is a valid confidence bound, associated to the reference family $\mathfrak{R}_{Simes}=\left(\left\{i : p_i\leq\frac{\alpha k}{m}\right\}, k-1\right)_{k\in[\![1,m]\!]}$.

Closed testing

[Marcus, Peritz, and Gabriel (1976)]

- Designed for FWER control
- Form $H_{0,I} = \bigcap_{i \in I} H_{0,i}$ all intersection hypotheses
- Have a collection of α-level local intersection tests ϕ_I:
  $\forall P \in H_{0,I}, \mathbb{P}_{X \sim P}(\phi_I(X) = 1) \leq \alpha$
- Examples:
  - Bonferroni local test $\phi_I = 1$ if $\exists i \in I : p_i \leq \frac{\alpha}{|I|}$
  - Hommel local test $\phi_I = 1$ if $\exists i \in I : p_{(i:I)} \leq \frac{\alpha i}{|I| H_{|I|}}$
  - Simes local test $\phi_I = 1$ if $\exists i \in I : p_{(i:I)} \leq \frac{\alpha i}{|I|}$ (under wPRDS on $\mathcal{H}_0$)
  - Proofs: if $P \in H_{0,I}$, $P \in H_{0,i}$ for all $i \in I$ so $\mathcal{L}(p_i) \succeq \mathcal{U}([0,1])$ for all $i \in I$ □

Closed testing

- Closed testing: iteratively test $H_{0,I}$ only if all $H_{0,J}$, $J \supsetneq I$, are rejected, then reject the individual hypotheses $H_{0,i}$ such that $H_{0,\{i\}}$ has been rejected: $R^{Closed} = \{i \in [\![1,m]\!] : \forall I \subseteq [\![1,m]\!] \text{ with } i \in I, \phi_I = 1\}$
- $\forall P \in \mathfrak{F}, \text{FWER}\left(R^{Closed}\right) \leq \alpha$
- $P \in H_{0,\mathcal{H}_0}$ (tautological), so

$$\begin{aligned} \text{FWER}\left(R^{Closed}\right) &= \mathbb{P}\left(\exists i \in \mathcal{H}_0 : \forall I \subseteq [\![1,m]\!], i \in I, \phi_I = 1\right) \\ &\leq \mathbb{P}\left(\phi_{\mathcal{H}_0} = 1\right) \\ &\leq \alpha \quad \square \end{aligned}$$

- Remark: each intersection test at level α, no multiplicity adjustment to the number of intersection hypotheses tested (only $\phi_{\mathcal{H}_0}$ matters)

Closed testing

A fun result

Proposition

Assume that closed testing is conducted with the Bonferroni intersection test $\phi_I = \mathbb{1}_{\left\{\exists i \in I : p_i \leq \frac{\alpha}{|I|}\right\}}$. Then $R^{Closed} = R^{\text{HB}}$ a.s.

- First note that $\forall k \in [\![1, m]\!], \forall I$ such that $|I| = m - k + 1$, $\exists k' \leq k$ such that $p_{(k')} \in \{p_i : i \in I\}$, because if $p_i > p_{(k)}$ for all $i \in I$ then $|I| \leq m - k$
- If $p_i \leq \frac{\alpha}{m - \hat{k}^{\text{HB}} + 1}$ (implies $\hat{k}^{\text{HB}} \geq 1$), let I such that $i \in I$, we want $\phi_I = 1$. 2 cases.
- If $|I| \leq m - \hat{k}^{\text{HB}} + 1$ then $p_i \leq \frac{\alpha}{m - \hat{k}^{\text{HB}} + 1} \leq \frac{\alpha}{|I|}$ so $\phi_I = 1$
- If $|I| > m - \hat{k}^{\text{HB}} + 1$ (implies $\hat{k}^{\text{HB}} \geq 2$), $|I| = m - k + 1$ with $k \in [\![1, \hat{k}^{\text{HB}}[\![$. Let $k' \leq k$ such that $p_{(k')} \in \{p_i : i \in I\}$. $k' \leq \hat{k}^{\text{HB}}$ so by definition of SD procedures $p_{(k')} \leq \frac{\alpha}{m - k' + 1} \leq \frac{\alpha}{m - k + 1} = \frac{\alpha}{|I|}$ so $\phi_I = 1$
- Hence $R^{\text{HB}} \subseteq R^{Closed}$

Closed testing

Proof of the Proposition

- Let $i \in R^{Closed}$: $\phi_I = 1$ for all I such that $i \in I$
- Let $\tilde{k} = \min\left\{k \in [\![1, m]\!] : p_i \leq \frac{\alpha}{m-k+1}\right\}$, well-defined because $\phi_{\{i\}} = 1$ so $p_i \leq \alpha$
- Goal : show that $\tilde{k} \leq \hat{k}^{\text{HB}}$, will imply $i \in R^{\text{HB}}$
- By recursion, $p_{(k')} \leq \frac{\alpha}{m-k'+1}$ for all $k' \in [\![1, \tilde{k}]\!]$, imply $\tilde{k} \leq \hat{k}^{\text{HB}}$ by definition
- $k' = 1$: $\phi_{[\![1,m]\!]} = 1$ so $p_{(1)} \leq \frac{\alpha}{m}$
- Let $k' < \tilde{k}$, by definition of $\tilde{k}$, $p_i > \frac{\alpha}{m-k'+1} \geq p_{(k')} \geq \cdots \geq p_{(1)}$
- So $i \in I = [\![1, m]\!] \setminus \{(1), \ldots, (k')\}$ with $|I| = m - k'$. $i \in R^{Closed}$ so $\phi_I = 1$, hence $\exists j \in I : p_j \leq \frac{\alpha}{|I|} = \frac{\alpha}{m-(k'+1)+1}$ hence $p_{(k'+1)} = \min_{j \in I} p_i \leq \frac{\alpha}{m-(k'+1)+1}$ □

Closed testing for post hoc inference

[Goeman and Solari (2011)]

Main idea

The closed testing provides more information than just the individual rejections:

- Let $\mathcal{X}$ the (random) set of all I such that we rejected $H_{0,I}$
- Simultaneous guarantee over all $H_{0,I}$, $I \in \mathcal{X}$:

$$\forall P \in \mathfrak{F}, \mathbb{P}\left(\exists I \in \mathcal{X}, P \in H_{0,I}\right) \leq \alpha$$

- Proof: as before, if $P \in H_{0,I}$, $I \subseteq \mathcal{H}_0$, so $\mathcal{H}_0 \in \mathcal{X}$, so $\phi_{\mathcal{H}_0} = 1$ □

Closed testing for post hoc inference

- A simple example where the closed testing is more informative than the resulting FWER procedure:
- $p_1 = \frac{2\alpha}{3}$, $p_2 = \frac{2\alpha}{3}$, $p_3 = 1$, and Simes intersection test
- $p_{(k)} \leq \alpha \frac{k}{3}$ for $k = 1$ and 2 so $H_{0,\{1,2,3\}}$ rejected
- $p_{(2)} \leq \alpha \frac{2}{2}$ so $H_{0,\{1,2\}}$ rejected
- But $p_{(1)} > \frac{\alpha}{2}$, $p_{(2)} > \frac{\alpha}{2}$ and $p_{(3)} > \alpha$ so $H_{0,\{1,3\}}$ and $H_{0,\{2,3\}}$ conserved
- Hence $H_{0,\{1\}}$, $H_{0,\{2\}}$ and $H_{0,\{3\}}$ all conserved and $R^{Closed} = \varnothing$, but we learned that there is signal in $H_{0,\{1,2,3\}}$ and $H_{0,\{1,2\}}$!

Closed testing for post hoc inference

Confidence bound derivation

- The proposed confidence bound is $V_{\mathrm{GS}}(S) = \max_{\substack{J \subseteq S \\ J \notin \mathcal{X}}} |J|$
- Uses all information in $\mathcal{X}$, not just singletons
- First note that $V_{\mathrm{GS}}(S) = \max_{J \notin \mathcal{X}} |S \cap J|$, $\leq$ obvious, and if $J \notin \mathcal{X}$, $S \cap J \in \mathcal{X}$ would imply $J \in \mathcal{X}$ by closure, so $S \cap J \notin \mathcal{X}$ and $\geq$ achieved
- $V_{\mathrm{GS}}(S) = \max_{\substack{J \subseteq S \\ J \notin \mathcal{X}}} |J|$ is a valid confidence bound because

$$\begin{aligned}
\mathbb{P}\left(\exists S, |S \cap \mathcal{H}_0| > V_{\mathrm{GS}}(S)\right) &\leq \mathbb{P}\left(\exists S, |S \cap \mathcal{H}_0| > \max_{J \notin \mathcal{X}} |S \cap J|\right) \\
&\leq \mathbb{P}\left(\mathcal{H}_0 \in \mathcal{X}\right) \\
&\leq \mathbb{P}\left(\phi_{\mathcal{H}_0} = 1\right) \leq \alpha \quad \square
\end{aligned}$$

Closed testing for post hoc inference

JER equivalence

Proposition

$\mathfrak{R} = (I, |I| - 1)_{I \in \mathcal{X}}$ controls the JER and $V_{\mathrm{GS}}(S) = V^*_{\mathfrak{R}}(S)$.

$$
\begin{aligned}
\mathbb{P}\left(\exists I \in \mathcal{X} : |I \cap \mathcal{H}_0| > |I| - 1\right) &\leq \mathbb{P}\left(\exists I \in \mathcal{X} : |I \cap \mathcal{H}_0| = |I|\right) \\
&\leq \mathbb{P}\left(\exists I \in \mathcal{X} : I \subseteq \mathcal{H}_0\right) \\
&\leq \mathbb{P}\left(\mathcal{H}_0 \in \mathcal{X}\right) \text{ by closure} \\
&\leq \mathbb{P}\left(\phi_{\mathcal{H}_0} = 1\right) \leq \alpha
\end{aligned}
$$

- Recall $V_{\mathrm{GS}}(S) = \max_{J \in \mathcal{X}^c} |S \cap J|$
- $\mathcal{A}(\mathfrak{R})^c = \{A : \exists I \in \mathcal{X}, |I \cap A| = |I|\} = \{A : \exists I \in \mathcal{X}, I \cap A = I\} = \{A : \exists I \in \mathcal{X}, I \subseteq A\} = \mathcal{X}$ by closure
- So $V_{\mathrm{GS}}(S) = \max_{J \in \mathcal{A}(\mathfrak{R})} |S \cap J| = V^*_{\mathfrak{R}}(S)$ □

Closed testing for post hoc inference

JER equivalence

Proposition

Reciprocally, let $\mathfrak{R}$ that controls the JER, then there exists a collection of intersection tests for which $V_{\mathrm{GS}}(S) = V^*_{\mathfrak{R}}(S)$.

- Let $\phi_I = \mathbb{1}_{\{I \notin \mathcal{A}(\mathfrak{R})\}}$, valid test : let $P \in H_{0,I}$, so $I \subseteq \mathcal{H}_0$, then

$$\begin{aligned}\mathbb{P}\left(I \notin \mathcal{A}(\mathfrak{R})\right) &= \mathbb{P}\left(\exists k \in \mathcal{K} : |I \cap R_k| > \zeta_k\right) \\ &\leq \mathbb{P}\left(\exists k \in \mathcal{K} : |\mathcal{H}_0 \cap R_k| > \zeta_k\right) \leq \alpha\end{aligned}$$

- By definition, $\mathcal{A}(\mathfrak{R})$ = exactly the conserved intersection hypotheses, so trivially $\mathcal{A}(\mathfrak{R}) \subseteq \mathcal{X}^{\mathrm{c}}$ and $V^*_{\mathfrak{R}}(S) \leq V_{\mathrm{GS}}(S)$
- Conversely, if $J \in \mathcal{X}^{\mathrm{c}}$, there is $B \in \mathcal{A}(\mathfrak{R})$ such that $J \subseteq B$ so $|S \cap J| \leq |S \cap B|$ and so $V_{\mathrm{GS}}(S) \leq V^*_{\mathfrak{R}}(S)$ □

Back to BNR technology

[Durand et al. (2020)]

- How to construct effectively a reference family $(R_k, \zeta_k)_{k\in\mathcal{K}}$ with JER control?
- Another approach: constrain R_k to some deterministic regions (using prior knowledge like gene ontologies) and (super-)estimate $|R_k \cap \mathcal{H}_0|$ to get a ζ_k

Proposition

If the R_k form a partition of $[\![1, m]\!]$, then $V^*_{\mathfrak{R}}(S) = \sum_{k\in\mathcal{K}} \zeta_k \wedge |S \cap R_k|$.

- Let any $A \in \mathcal{A}(\mathfrak{R})$, $|R_k \cap A| \leq \zeta_k$ so $|A \cap S| = \sum_{k\in\mathcal{K}} |A \cap S \cap R_k|$ with $|A \cap S \cap R_k| \leq |R_k \cap A| \leq \zeta_k$ and $|A \cap S \cap R_k| \leq |S \cap R_k|$ so by taking the max, $V^*_{\mathfrak{R}}(S) \leq \sum_{k\in\mathcal{K}} \zeta_k \wedge |S \cap R_k|$
- Construct $A = \bigcup_{k\in\mathcal{K}} \{\zeta_k \wedge |S \cap R_k| \text{ elements of } S \cap R_k\}$, $A \in \mathcal{A}(\mathfrak{R})$ so $\sum_{k\in\mathcal{K}} \zeta_k \wedge |S \cap R_k| = |A| \leq V^*_{\mathfrak{R}}(S)$ □

Back to BNR technology

ζ_k computation

Theorem

Assume that for all $P \in \mathfrak{F}$, the $(p_i)_{i\in\mathcal{H}_0}$ are independent, and they are independent from the $(p_i)_{i\in\mathcal{H}_1}$.

Assume that $\mathcal{K}$ and the R_k are deterministic. Let $C_\lambda = \sqrt{\frac{1}{2}\log\left(\frac{1}{\lambda}\right)}$ for all $\lambda \in]0,1[$. Let

$$\zeta_k = |R_k| \wedge \min_{t\in[0,1[} \left\lfloor \frac{C_{\frac{\alpha}{K}}}{2(1-t)} + \left(\frac{C^2_{\frac{\alpha}{K}}}{4(1-t)^2} + \frac{\sum_{i\in R_k} \mathbb{1}_{\{p_i>t\}}}{1-t} \right)^{1/2} \right\rfloor^2$$

Then, if $\frac{\alpha}{K} < \frac{1}{2}$, $\mathfrak{R}$ controls the JER at level α.

Back to BNR technology

ζ_k computation

- In practice,

$$\zeta_k = |R_k| \wedge \min_{0 \le \ell \le |R_k|} \left\lfloor \frac{C_{\frac{\alpha}{K}}}{2(1-p_{(\ell:R_k)})} + \left(\frac{C^2_{\frac{\alpha}{K}}}{4(1-p_{(\ell:R_k)})^2} + \frac{|R_k|-\ell}{1-p_{(\ell:R_k)}} \right)^{1/2} \right\rfloor^2$$

- Entry cost: $\zeta_k \ge \left\lfloor C^2_{\frac{\alpha}{K}} \right\rfloor = \left\lfloor \log\left(\frac{K}{\alpha}\right) \right\rfloor \ge 1$ as soon as $\alpha \le e^{-2}K$: impossible to detect regions made of pure signal
- $\frac{\alpha}{K}$: union bound correction w.r.t. the number of regions
- Dependency on α and K are only through a log

Back to BNR technology

Proof of the Theorem

- Dvoretzky-Kiefer-Wolfowitz-Massart inequality [Massart (1990)]: let any $S \subseteq [\![1, m]\!]$, $S_0 = S \cap \mathcal{H}_0$ $\nu = |S_0|$ and $U_1, \dots, U_m$ i.i.d. r.v. with $\mathbb{P}_{U_1} = \mathcal{U}([0,1])$. For all $\varepsilon \geq \sqrt{\frac{1}{2\nu} \log 2}$,

$$\mathbb{P}\left(\sup_{t\in\mathbb{R}}\left(\frac{1}{\nu}\sum_{i\in S_0}\mathbb{1}_{\{U_i\leq t\}} - u(t)\right) > \varepsilon\right) = \mathbb{P}\left(\sup_{t\in[0,1[}\left(\frac{1}{\nu}\sum_{i\in S_0}\mathbb{1}_{\{U_i\leq t\}} - t\right) > \varepsilon\right.$$
$$\leq e^{-2\nu\varepsilon^2}$$

- Let any $\lambda < \frac{1}{2}$ and $\varepsilon = \sqrt{\frac{1}{2\nu}\log\left(\frac{1}{\lambda}\right)} = \frac{1}{\sqrt{\nu}} C_\lambda$, $\varepsilon \geq \sqrt{\frac{1}{2\nu}\log 2}$

Back to BNR technology

Proof of the Theorem

- So, $\mathbb{P}\left(\sup_{t\in[0,1[}\left(\frac{1}{\nu}\sum_{i\in S_0}\mathbb{1}_{\{U_i\leq t\}}-t\right)>\varepsilon\right)\leq\lambda$
- $\mathbb{P}\left(\sup_{t\in[0,1[}\left(\frac{1}{\nu}\sum_{i\in S_0}\mathbb{1}_{\{U_i\leq t\}}-t\right)\leq\varepsilon\right)\geq 1-\lambda$
- $\mathbb{P}\left(\inf_{t\in[0,1[}\left(t-\frac{1}{\nu}\sum_{i\in S_0}\mathbb{1}_{\{U_i\leq t\}}\right)\geq -\varepsilon\right)\geq 1-\lambda$
- $\mathbb{P}\left(\forall t\in[0,1[, t-\frac{1}{\nu}\sum_{i\in S_0}\mathbb{1}_{\{U_i\leq t\}}\geq -\varepsilon\right)\geq 1-\lambda$
- $\mathbb{P}\left(\forall t\in[0,1[, \left(1-\frac{1}{\nu}\sum_{i\in S_0}\mathbb{1}_{\{U_i\leq t\}}\right)-(1-t)\geq -\varepsilon\right)\geq 1-\lambda$
- $\mathbb{P}\left(\forall t\in[0,1[, \frac{1}{\nu}\sum_{i\in S_0}\mathbb{1}_{\{U_i> t\}}-(1-t)\geq -\varepsilon\right)\geq 1-\lambda$

Back to BNR technology

Proof of the Theorem

- With $\mathbb{P} \geq 1 - \lambda$, for all $t \in [0, 1[$, $\sum_{i \in S_0} \mathbb{1}_{\{U_i > t\}} - \nu(1 - t) + \sqrt{\nu} C_\lambda \geq 0$, let $x = \sqrt{\nu}$ and solve this second degree polynom in x
- $\Delta = C_\lambda^2 + 4(1 - t) \sum_{i \in S_0} \mathbb{1}_{\{U_i > t\}} > 0$, the polynom is ≥ 0 inside of its two real roots $\frac{C_\lambda \pm \sqrt{C_\lambda^2 + 4(1-t) \sum_{i \in S_0} \mathbb{1}_{\{U_i > t\}}}}{2(1-t)}$, one is ≤ 0 and the other ≥ 0, and $x = \sqrt{\nu} \geq 0$, so

$$
\begin{aligned}
x &\leq \frac{C_\lambda + \sqrt{C_\lambda^2 + 4(1 - t) \sum_{i \in S_0} \mathbb{1}_{\{U_i > t\}}}}{2(1 - t)} \\
&= \frac{C_\lambda}{2(1 - t)} + \left(\frac{C_\lambda^2}{4(1 - t)^2} + \frac{\sum_{i \in S_0} \mathbb{1}_{\{U_i > t\}}}{1 - t} \right)^{1/2}
\end{aligned}
$$

Back to BNR technology

Proof of the Theorem

- With $\mathbb{P} \geq 1 - \lambda$, for all $t \in [0, 1[$,

$$\nu \leq \left(\frac{C_\lambda}{2(1-t)} + \left(\frac{C_\lambda^2}{4(1-t)^2} + \frac{\sum_{i \in S_0} \mathbb{1}_{\{U_i > t\}}}{1-t} \right)^{1/2} \right)^2$$

- Let $b_{it} = \mathbb{P}(p_i \leq t)$, for $i \in \mathcal{H}_0$, $b_{it} \leq t$ so $\mathbb{1}_{\{U_i > t\}} \leq \mathbb{1}_{\{U_i > b_{it}\}}$
- Then,

$$\mathbb{P}\left(\nu \leq \min_{t \in [0,1[} \left(\frac{C_\lambda}{2(1-t)} + \left(\frac{C_\lambda^2}{4(1-t)^2} + \frac{\sum_{i \in S_0} \mathbb{1}_{\{U_i > b_{it}\}}}{1-t} \right)^{1/2} \right)^2 \right) \geq 1 - \lambda$$

Back to BNR technology

Proof of the Theorem

- Lemma: $\left(\sum_{i\in S_0} \mathbb{1}_{\{U_i > b_{it}\}}\right)_{t\in[0,1[} \overset{\mathcal{L}}{=} \left(\sum_{i\in S_0} \mathbb{1}_{\{p_i > t\}}\right)_{t\in[0,1[}$
- So

$$\mathbb{P}\left(\nu \leq \min_{t\in[0,1[}\left(\frac{C_\lambda}{2(1-t)} + \left(\frac{C_\lambda^2}{4(1-t)^2} + \frac{\sum_{i\in S_0} \mathbb{1}_{\{p_i>t\}}}{1-t}\right)^{1/2}\right)^2\right) \geq 1-\lambda$$

- And finally

$$\mathbb{P}\left(|S\cap\mathcal{H}_0| \leq \min_{t\in[0,1[}\left(\frac{C_\lambda}{2(1-t)} + \left(\frac{C_\lambda^2}{4(1-t)^2} + \frac{\sum_{i\in S} \mathbb{1}_{\{p_i>t\}}}{1-t}\right)^{1/2}\right)^2\right) \geq 1$$

- Apply this to $S = R_k$ and $\lambda = \frac{\alpha}{K}$, add the $\lfloor\cdot\rfloor$ and $|R_k|\wedge$ freely, and use a union bound to conclude □

Back to BNR technology

Proof of the Lemma

- We show that the marginals finite-dimensional are equal, only with two marginals w.l.o.g.: let $t_1 < t_2 \in [0, 1[$
- We show $\left(\sum_{i\in S_0} \mathbb{1}_{\{U_i > b_{it_1}\}}, \sum_{i\in S_0} \mathbb{1}_{\{U_i > b_{it_2}\}}\right) \overset{\mathcal{L}}{=} \left(\sum_{i\in S_0} \mathbb{1}_{\{p_i > t_1\}}, \sum_{i\in S_0} \mathbb{1}_{\{p_i > t_2\}}\right)$ with the equality of the characteristic functions

$$\begin{aligned}\phi(s,u) &= \mathbb{E}\left[\exp\left(\imath s \sum_{i\in S_0} \mathbb{1}_{\{U_i > b_{it_1}\}} + \imath u \sum_{i\in S_0} \mathbb{1}_{\{U_i > b_{it_2}\}}\right)\right] \\ &= \prod_{i\in S_0} \mathbb{E}\left[\exp\left(\imath s \mathbb{1}_{\{U_i > b_{it_1}\}} + \imath u \mathbb{1}_{\{U_i > b_{it_2}\}}\right)\right]\end{aligned}$$

  by independence
- Same for $\left(\sum_{i\in S_0} \mathbb{1}_{\{p_i > t_1\}}, \sum_{i\in S_0} \mathbb{1}_{\{p_i > t_2\}}\right)$: showing equality inside the product is enough

Back to BNR technology
Proof of the Lemma

- $(b_{it})_t$ nondecreasing so

$$
\begin{aligned}
\phi_i(s,u) &= \mathbb{E}\left[\exp\left(\imath s \mathbb{1}_{\{U_i > b_{it_1}\}} + \imath u \mathbb{1}_{\{U_i > b_{it_2}\}}\right)\right] \\
&= \mathbb{E}\left[\exp\left(\imath (s+u) \mathbb{1}_{\{U_i > b_{it_2}\}} + \imath s \mathbb{1}_{\{b_{it_2} \geq U_i > b_{it_1}\}}\right)\right] \\
&= \int_{[0,1]} e^{\imath (s+u) \mathbb{1}_{\{x > b_{it_2}\}} + \imath s \mathbb{1}_{\{b_{it_2} \geq x > b_{it_1}\}}} \mathrm{d}x \\
&= \int_{]b_{it_2},1]} e^{\imath (s+u)} \mathrm{d}x + \int_{]b_{it_1}, b_{it_2}]} e^{\imath s} \mathrm{d}x + \int_{[0,b_{it_1}]} \mathrm{d}x \\
&= e^{\imath (s+u)}(1 - b_{it_2}) + e^{\imath s}(b_{it_2} - b_{it_1}) + b_{it_1}
\end{aligned}
$$

Back to BNR technology

Proof of the Lemma

- Similarly,

$$\begin{aligned}
\psi_i(s,u) &= \mathbb{E}\left[\exp\left(\imath s \mathbb{1}_{\{p_i>t_1\}} + \imath u \mathbb{1}_{\{p_i>t_2\}}\right)\right] \\
&= \mathbb{E}\left[\exp\left(\imath (s+u) \mathbb{1}_{\{p_i>t_2\}} + \imath s \mathbb{1}_{\{t_2\geq p_i>t_1\}}\right)\right] \\
&= \int_{[0,1]} e^{\imath(s+u)\mathbb{1}_{\{x>t_2\}}+\imath s\mathbb{1}_{\{t_2\geq x>t_1\}}} \mathbb{P}_{p_i}(\mathrm{d}x) \\
&= \int_{]t_2,1]} e^{\imath(s+u)} \mathbb{P}_{p_i}(\mathrm{d}x) + \int_{]t_1,t_2]} e^{\imath s} \mathbb{P}_{p_i}(\mathrm{d}x) + \int_{[0,t_1]} \mathbb{P}_{p_i}(\mathrm{d}x) \\
&= e^{\imath(s+u)} \mathbb{P}_{p_i}(]t_2,1]) + e^{\imath s} \mathbb{P}_{p_i}(]t_1,t_2]) + \mathbb{P}_{p_i}([0,t_1]) \\
&= e^{\imath(s+u)}(1-b_{it_2}) + e^{\imath s}(b_{it_2}-b_{it_1}) + b_{it_1} \quad \square
\end{aligned}$$

Topics that were not covered

- Bayesian multiple testing, ℓ-values, q-values
- Knock-offs
- Permutation p-values, conformal p-values
- Sequential/online multiple testing
- Multiple testing with e-values
- And many more

Bibliography I

Aickin, M and H Gensler (1996). "Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods.". In: *American Journal of Public Health* 86.5. PMID: 8629727, pp. 726–728. DOI: 10.2105/AJPH.86.5.726. eprint: https://doi.org/10.2105/AJPH.86.5.726. URL: https://doi.org/10.2105/AJPH.86.5.726.

Benjamini, Yoav and Yosef Hochberg (1995). "Controlling the false discovery rate: a practical and powerful approach to multiple testing". In: *J. Roy. Statist. Soc. Ser. B* 57.1, pp. 289–300. ISSN: 0035-9246. URL: http://links.jstor.org/sici?sici=0035-9246(1995)57:1<289:CTFDRA>2.0.CO;2-E&origin=MSN.

Benjamini, Yoav and Daniel Yekutieli (2001). "The control of the false discovery rate in multiple testing under dependency". In: *Ann. Statist.* 29.4, pp. 1165–1188. ISSN: 0090-5364. DOI: 10.1214/aos/1013699998. URL: https://doi.org/10.1214/aos/1013699998.

Blanchard, Gilles, Pierre Neuvial, and Etienne Roquain (2020). "Post hoc confidence bounds on false positives using reference families". In: *Ann. Statist.* 48.3, pp. 1281–1303. ISSN: 0090-5364. DOI: 10.1214/19-AOS1847. URL: https://doi.org/10.1214/19-AOS1847.

Blanchard, Gilles and Etienne Roquain (2008). "Two simple sufficient conditions for FDR control". In: *Electron. J. Stat.* 2, pp. 963–992. ISSN: 1935-7524. DOI: 10.1214/08-EJS180. URL: https://doi.org/10.1214/08-EJS180.

Bibliography II

- Bonferroni, Carlo (1936). "Teoria statistica delle classi e calcolo delle probabilita". In: *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze* 8, pp. 3–62.
- Döhler, Sebastian, Guillermo Durand, and Etienne Roquain (2018). "New FDR bounds for discrete and heterogeneous tests". In: *Electron. J. Stat.* 12.1, pp. 1867–1900. DOI: 10.1214/18-EJS1441. URL: https://doi.org/10.1214/18-EJS1441.
- Durand, Guillermo (2019). "Adaptive *p*-value weighting with power optimality". In: *Electron. J. Stat.* 13.2, pp. 3336–3385. DOI: 10.1214/19-ejs1578. URL: https://doi.org/10.1214/19-ejs1578.
- Durand, Guillermo et al. (2020). "Post hoc false positive control for structured hypotheses". In: *Scand. J. Stat.* 47.4, pp. 1114–1148. ISSN: 0303-6898. DOI: 10.1111/sjos.12453. URL: https://doi.org/10.1111/sjos.12453.
- Genovese, Christopher R., Kathryn Roeder, and Larry Wasserman (2006). "False discovery control with *p*-value weighting". In: *Biometrika* 93.3, pp. 509–524. ISSN: 0006-3444,1464-3510. DOI: 10.1093/biomet/93.3.509. URL: https://doi.org/10.1093/biomet/93.3.509.
- Genovese, Christopher R. and Larry Wasserman (2006). "Exceedance control of the false discovery proportion". In: *J. Amer. Statist. Assoc.* 101.476, pp. 1408–1417. ISSN: 0162-1459. DOI: 10.1198/016214506000000339. URL: https://doi.org/10.1198/016214506000000339.

Bibliography III

Giraud, Christophe (2021). *Introduction to High-Dimensional Statistics Second Edition*. Chapman and Hall/CRC.

Goeman, Jelle J. and Aldo Solari (2011). "Multiple testing for exploratory research". In: *Statist. Sci.* 26.4, pp. 584–597. ISSN: 0883-4237. DOI: 10.1214/11-STS356. URL: https://doi.org/10.1214/11-STS356.

Heyse, Joseph F (2011). "A false discovery rate procedure for categorical data". In: *Recent advances in biostatistics: False discovery rates, survival analysis, and related topics*. World Scientific, pp. 43–58.

Holm, Sture (1979). "A simple sequentially rejective multiple test procedure". In: *Scand. J. Statist.* 6.2, pp. 65–70. ISSN: 0303-6898.

Hommel, G. (1983). "Tests of the overall hypothesis for arbitrary dependence structures". In: *Biometrical J.* 25.5, pp. 423–430. ISSN: 0323-3847.

Lehmann, E. L. and Joseph P. Romano (2005). "Generalizations of the familywise error rate". In: *Ann. Statist.* 33.3, pp. 1138–1154. ISSN: 0090-5364,2168-8966. DOI: 10.1214/009053605000000084. URL: https://doi.org/10.1214/009053605000000084.

Marcus, Ruth, Eric Peritz, and K. R. Gabriel (1976). "On closed testing procedures with special reference to ordered analysis of variance". In: *Biometrika* 63.3, pp. 655–660. ISSN: 0006-3444. DOI: 10.1093/biomet/63.3.655. URL: https://doi.org/10.1093/biomet/63.3.655.

Bibliography IV

Massart, P. (1990). "The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality". In: *Ann. Probab.* 18.3, pp. 1269–1283. ISSN: 0091-1798,2168-894X. URL: http://links.jstor.org/sici?sici=0091-1798(199007)18:3<1269:TTCITD>2.0.CO;2-Q&origin=MSN.

Meinshausen, Nicolai (2006). "False discovery control for multiple tests of association under general dependence". In: *Scand. J. Statist.* 33.2, pp. 227–237. ISSN: 0303-6898. DOI: 10.1111/j.1467-9469.2005.00488.x. URL: https://doi.org/10.1111/j.1467-9469.2005.00488.x.

Roquain, Etienne (2015). "Contributions to multiple testing theory for high-dimensional data". PhD thesis. Université Pierre et Marie Curie.

Roquain, Etienne and Mark A. van de Wiel (2009). "Optimal weighting for false discovery rate control". In: *Electron. J. Stat.* 3, pp. 678–711. DOI: 10.1214/09-EJS430. URL: https://doi.org/10.1214/09-EJS430.

Simes, R. J. (1986). "An improved Bonferroni procedure for multiple tests of significance". In: *Biometrika* 73.3, pp. 751–754. ISSN: 0006-3444. DOI: 10.1093/biomet/73.3.751. URL: https://doi.org/10.1093/biomet/73.3.751.

Storey, John D., Jonathan E. Taylor, and David Siegmund (2004). "Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach". In: *J. R. Stat. Soc. Ser. B Stat. Methodol.* 66.1, pp. 187–205. ISSN: 1369-7412. DOI: 10.1111/j.1467-9868.2004.00439.x. URL: https://doi.org/10.1111/j.1467-9868.2004.00439.x.

Bibliography V

Tarone, Robert E (1990). "A modified Bonferroni method for discrete data". In: *Biometrics*, pp. 515–522.