Step-up procedure with data-driven optimal weights for grouped hypotheses

Guillermo Durand
LPMA

MCP 2017

22/06/2017

Motivation
Grouped hypotheses

Context

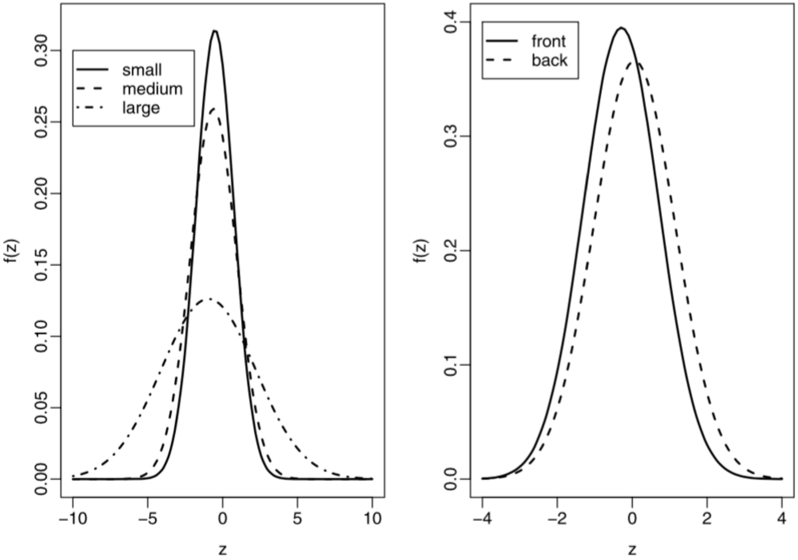
The hypotheses have a group structure :
same distribution under $\mathcal{H}_1$ in each group

Examples :

- The Adequate Yearly Progress data set where grouping schools by size avoids a preference for large schools (Cai and Sun 2009).
- An fMRI study where voxels from the front and the back of the brain have different distribution (Cai and Sun 2009).
- Grouping genes by pathway.

Motivation

Examples from Cai and Sun 2009



Topic of this talk :

Handle the group structure by weighting p-values (Holm 1979, Genovese, Roeder, and Wasserman 2006, Blanchard and Roquain 2008, Hu, Zhao, and Zhou 2010, Zhao and Zhang 2014, ...)

One procedure : IHW (Ignatiadis et al. 2016), first suggested in the discussion section of Roquain and Van De Wiel 2009.

Here : philosophy of IHW and relation to previous work, and my contributions on it, including a variant.

Table of contents

Table of contents

The well-known BH procedure

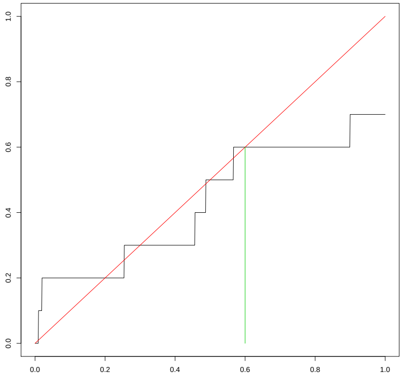
- Order p-values : $p_{(1)} \leq \cdots \leq p_{(m)}$
- Compute $\hat{k} = \max\{k : p_{(k)} \leq \alpha k/m\}$
- Reject all $p_i \leq \alpha \frac{\hat{k}}{m}$
- FDR control at level $\pi_0 \alpha$ when wPRDS

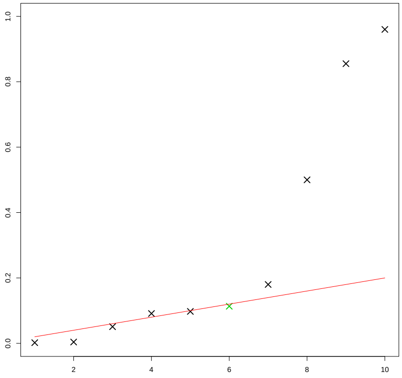
Useful other formulation

$\frac{\hat{k}}{m} = \max\{u : \widehat{G}(u) \geq u\} := \mathcal{I}\left(\widehat{G}\right)$ where

$$\widehat{G} : u \mapsto m^{-1} \sum_{i=1}^{m} \mathbb{1}_{\{p_i \leq \alpha u\}}, u \in [0, 1]$$

An illustration of $\mathcal{I}(\widehat{G})$





Weighted-BH (WBH)

Take weights $(w_i)_{1\leq i\leq m}$ s.t. $\sum_i w_i = m$ (weight vector), form

$$\widehat{G}_w : u \mapsto m^{-1} \sum_{i=1}^{m} \mathbb{1}_{\{p_i \leq \alpha u w_i\}}$$

and reject all $p_i \leq \alpha \hat{u} w_i$ with $\hat{u} = \mathcal{I}\left(\widehat{G}_w\right)$.

BH is a WBH procedure with $w_i = 1\ \forall i$.

Weighted-Step-Up (WSU)

A generalization : non-linear weight functions

From Roquain and Van De Wiel 2009 :

Take W s.t. $(W_i(u))_i$ is a weight vector $\forall u$ and

$$\widehat{G}_W : u \mapsto m^{-1} \sum_{i=1}^{m} \mathbb{1}_{\{p_i \leq \alpha u W_i(u)\}}$$

is nondecreasing, then $\text{WSU}\,(W) = \{i : p_i \leq \alpha \hat{u} W_i\,(\hat{u})\}$ with $\hat{u} = \mathcal{I}\left(\widehat{G}_W\right)$.

Weighted-Step-Up (WSU)

A practical way to compute $\mathcal{I}\left(\widehat{G}_W\right)$

No need to compute $W(u)$ for each u !

$\forall k \in [\![1, m]\!]$, compute all $\frac{p_i}{W_i(\frac{k}{m})}$ and take q_k the k-th smallest.
Let $q_0 = 0$.
Then $\mathcal{I}\left(\widehat{G}_W\right) = m^{-1} \max\{k \in [\![0, m]\!] : q_k \leq \alpha \frac{k}{m}\}$.

Now : what W to choose ?

Optimal weighting

- Unconditional model : $\forall i, \mathbb{P}(i \in \mathcal{H}_0) = \pi_0$.
- C.d.f. under the alternative : F_i.
- Consider the procedure $R_{u,w}$ rejecting p_i if $p_i \leq \alpha u w_i$.
- Its power is $P_w^{(m)}(u) = (1-\pi_0) m^{-1} \sum_{i=1}^m F_i(\alpha u w_i)$
- Maximize it for all u :

Definition of optimal weights :

$$W^{*(m)}(u) = \underset{w : \sum_i^m w_i = m}{\operatorname{argmax}} P_w^{(m)}(u)$$

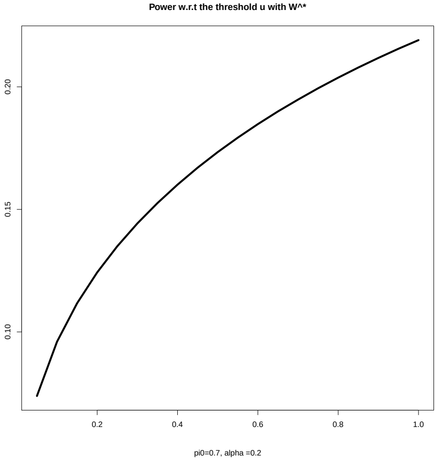
Optimal weighting
Existence and uniqueness

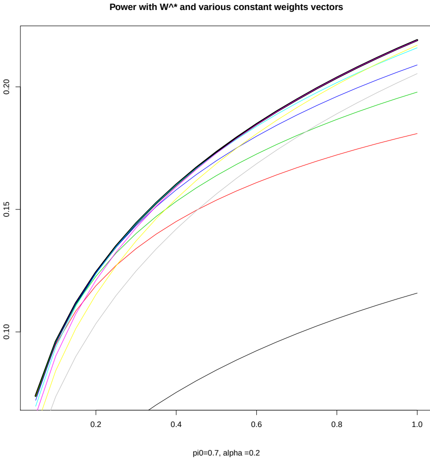
Assume some regularity properties of the F_i (concavity, differentiability, fulfilled in the gaussian 1-sided framework).

Theorem (Roquain and Van De Wiel 2009)

Then we have existence, uniqueness and continuity of $W^{*(m)}$, and $u \mapsto uW_i^{*(m)}(u)$ is non-decreasing.
Moreover, WSU $\left(W^{*(m)}\right)$ asymptotically enjoys FDR control at level $\pi_0\alpha$ and power optimality among all WBH procedures.

Illustration of $W^*(u)$ as an argmax





Optimal weighting

Main problem and resulting motivation

- F_i unknown under the alternative ! So is W^*.
- Goal : estimate W^*, obtain asymptotical results on FDR control and power optimality.
- Leads to data-driven optimal weighting and IHW.

Table of contents

Data-driven optimal weighting

- Assume that the p-values have uniform distribution under the null.

The trick :

$W^*(u)$ is also the unique maximizer of

$$G_w(u) = \mathbb{E}\left[\widehat{G}_w(u)\right] = \frac{\pi_0}{m}\sum_i^m (\alpha u w_i \wedge 1) + \mathrm{Pow}_w(u)$$

the mean proportion of rejections done by $R_{u,w}$.

Data-driven optimal weighting

$\Longrightarrow$ estimate W^* by maximizing G_w's empiric counterpart $\widehat{G}_w$.

Define $\widehat{W}^*(u)$ as :

$$\widehat{W}^*(u) \in \underset{w \geq 0 : \sum_i w_i = m}{\operatorname{argmax}} \widehat{G}_w(u) = \operatorname{argmax} \frac{1}{m} \sum_{i=1}^{m} \mathbb{1}_{p_i \leq \alpha u w_i}$$

Optimal data-driven step-up procedure :

$$\mathrm{WSU}\left(\widehat{W}^*\right)$$

(same as IHW)

Data-driven optimal weighting

Assumptions

- Assumptions of Roquain and Van De Wiel 2009.
- G groups of sizes $(m_g)_{g \leq G}$, where p-values have the same distribution.
- p-values are independent.
- $\alpha > \alpha^* = \left(\pi_0 + (1 - \pi_0) \sum_g f_g(0^+)\right)^{-1}$ (Chi 2007).
- $\frac{m_g}{m} \underset{m \to \infty}{\longrightarrow} \pi_g > 0$.

Proofs of the following results inspired by (Roquain and Van De Wiel 2009), (Zhao and Zhang 2014) and (Hu, Zhao, and Zhou 2010).

First main result

Theorem (FDR control)

$$\mathrm{FDP}\left(\mathrm{WSU}\left(\widehat{W}^*\right)\right) \xrightarrow{a.s.} \pi_0\alpha$$

$$\mathrm{FDR}\left(\mathrm{WSU}\left(\widehat{W}^*\right)\right) \longrightarrow \pi_0\alpha$$

Only at level α in Ignatiadis et al. 2016 but with less assumptions.

Second main result

Theorem (power optimality)

$$\lim_{m\to\infty} \mathrm{Pow}\left(\mathrm{WSU}\left(\widehat{W}^*\right)\right) = \lim_{m\to\infty} \mathrm{Pow}\left(\mathrm{WSU}\left(W^{*(m)}\right)\right)$$

and (corollary)

$$\lim_{m\to\infty} \mathrm{Pow}\left(\mathrm{WSU}\left(\widehat{W}^*\right)\right) \geq \limsup_{m\to\infty} \mathrm{Pow}\left(\mathrm{WSU}\left(w^{(m)}\right)\right)$$

for all sequences of determinisitic weights $(w^{(m)})$.

Stabilization variant

- WSU $\left(\widehat{W}^*\right)$ overfits so with low signal we loose the FDR control in finite sample.
- We should prefer BH then.
- $\Longrightarrow$ test if there is signal before choosing the procedure

Stabilized WSU$_\beta$ $\left(\widehat{W}^*\right)$

$$\mathrm{sWSU}_\beta\left(\widehat{W}^*\right) := \begin{cases} \mathrm{WSU}\left(\widehat{W}^*\right) & \text{if } \phi_\beta = 1 \\ \mathrm{BH} & \text{if } \phi_\beta = 0 \end{cases}$$

with $\phi_\beta = \mathbb{1}_{\{Z_m > q_{\beta,m}\}}$, $Z_m = \sqrt{m}\sup_{u\in[0,1]}\left(\widehat{G}_{\widehat{W}^*}(u) - \alpha u\right)$ and $q_{\beta,m}$ the $(1-\beta)$ quantile of Z_{0m}.

Stabilization variant

Main idea

Small signal $\Longrightarrow$ Z_m close to Z_{0m} in distribution, and

$$\begin{aligned}
\mathrm{FDR}\left(\mathrm{sWSU}_\beta\left(\widehat{W}^*\right)\right) &= \mathbb{E}\left[\phi_\beta \,\mathrm{FDP}\left(\mathrm{WSU}\left(\widehat{W}^*\right)\right)\right. \\
&\quad \left. + (1-\phi_\beta)\,\mathrm{FDP}\,(\mathrm{BH})\right] \\
&\leq \mathbb{E}\left[\phi_\beta + \mathrm{FDP}\,(\mathrm{BH})\right] \\
&\leq \mathbb{P}\left(Z_m > q_{\beta,m}\right) + \mathrm{FDR}\,(\mathrm{BH}) \\
&\lesssim \mathbb{P}\left(Z_{0m} > q_{\beta,m}\right) + \mathrm{FDR}\,(\mathrm{BH}) \\
&\leq \beta + \mathrm{FDR}\,(\mathrm{BH})
\end{aligned}$$

$\beta \to 0$ for asymptotic control ?

Stabilization variant
Result

Theorem

$\text{sWSU}_\beta\left(\widehat{W}^*\right)$ is asymptotically equivalent to $\text{WSU}\left(\widehat{W}^*\right)$ because $\phi_\beta \xrightarrow{a.s.} 1$ when $m \to \infty$, even if $\beta = \beta_m \to 0$ not too slowly $\left(\beta_m \geq \exp\left(-m^{1-\nu}\right)\right)$.

Proof relies on the DKWM inequality (Massart 1990).

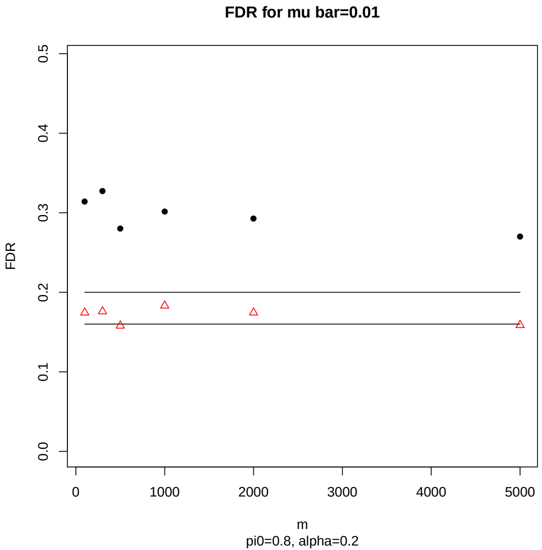
Table of contents

About the computation of $\widehat{W}^*$

Key ideas

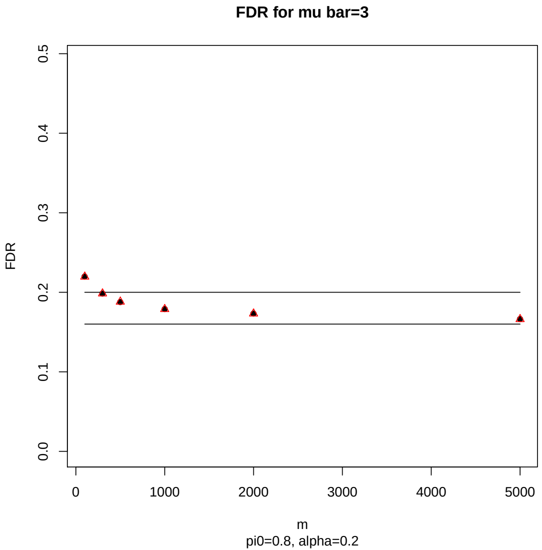
- Compute only $\widehat{W}^*(u)$ for $u = \frac{1}{m}, \frac{2}{m}, \dots, \frac{m-1}{m}, 1$.
- Max over $w : \sum m_g w_g = m \Leftrightarrow$ max over $w : \sum m_g w_g \leq m$.
- Fixing u, $w \mapsto \widehat{G}_w(u)$ discrete, only jumps at the $\frac{p_{g,i}}{\alpha u} \Longrightarrow$ search $\widehat{W}^*_g(u)$ as a $\frac{p_{g,i_g}}{\alpha u}$ such that $\sum m_g \frac{p_{g,i_g}}{\alpha u} \leq m$.
- $\widehat{G}_w(u)$ nondecreasing in u AND w : try to reject 1 hyp, then 2, then 3... for $u = \frac{1}{m}$, when fail at k hyp, try to reject k hyp for $u = \frac{2}{m}$, ...

FDR plot



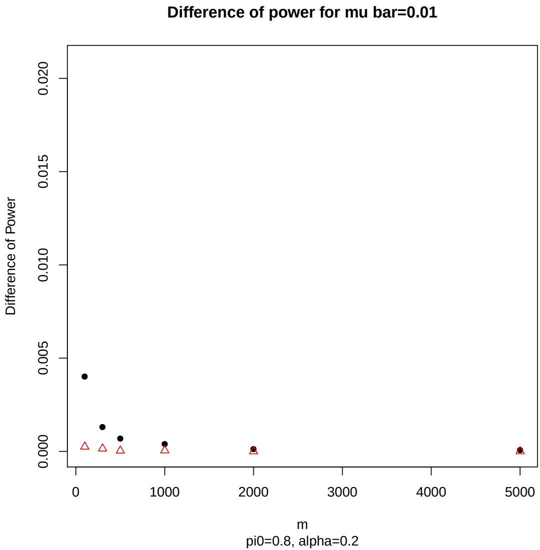
- $\pi_1 = \pi_2 = 0.5$
- $\mu_1 = \bar{\mu}$ and $\mu_2 = 2\bar{\mu}$.
- x axis : m.
- y axis : the FDR of our procedure over 1000 replications.

FDR plot



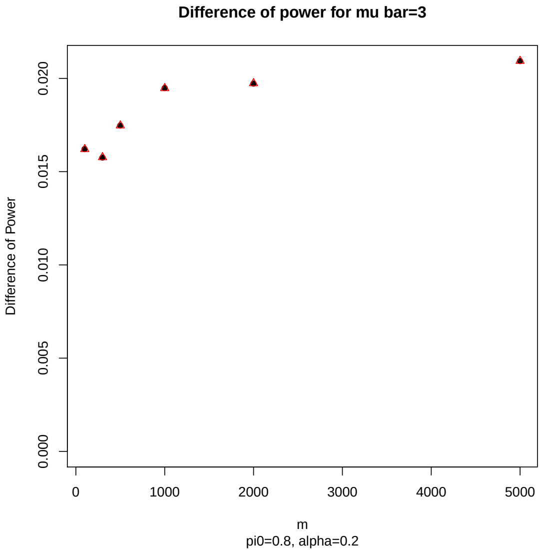
- $\pi_1 = \pi_2 = 0.5$
- $\mu_1 = \bar{\mu}$ and $\mu_2 = 2\bar{\mu}$.
- $\beta = 0.05$.
- x axis : m.
- y axis : the FDR of our procedure over 1000 replications.

Difference of power with BH



- $\pi_1 = \pi_2 = 0.5$
- $\mu_1 = \bar{\mu}$ and $\mu_2 = 2\bar{\mu}$.
- $\beta = 0.05$.
- x axis : m.
- y axis : the power of our procedure over 1000 replications minus the power of BH.

Difference of power with BH



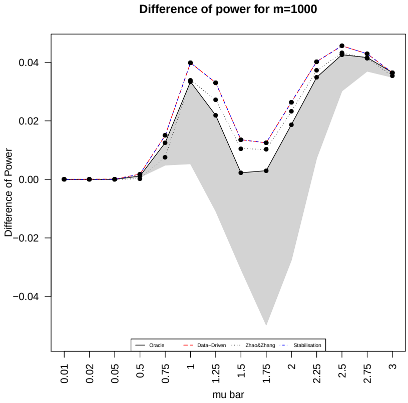
- $\pi_1 = \pi_2 = 0.5$
- $\mu_1 = \bar{\mu}$ and $\mu_2 = 2\bar{\mu}$.
- x axis : m.
- y axis : the power of our procedure over 1000 replications minus the power of BH.

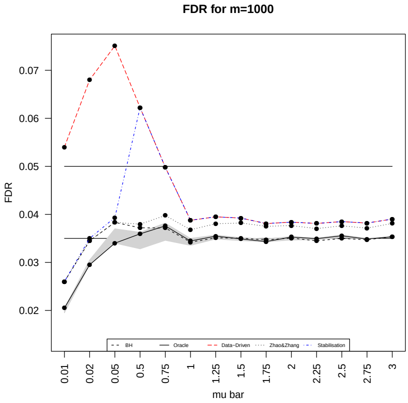
Comparison with other methods

- WSU $\left(\widehat{W}^*\right)$ and sWSU $\left(\widehat{W}^*\right)$ plotted.
- Varying signal $\bar{\mu}$.
- Oracle is the oracle method of Roquain and Van De Wiel 2009.
- Zhao&Zhang is a an adapation of Zhao and Zhang 2014 without $\hat{\pi}_0$.
- Many weighted-BH procedures plotted $\Longrightarrow$ grey area in the figures.

Comparison with other methods

$\alpha = 0.05$, 70% true null, $m_1 = m_2 = 500$, $\beta = 0.01$





Comparison with other methods

$\alpha = 0.05$, 70% true null, $m_1 = m_2 = 500$, $\beta = 0.01$

- $\mu_1 = \bar{\mu}$ and $\mu_2 = 2\bar{\mu}$.
- 1000 replications.
- Overfitting in our method.
- Good control for small and large $\bar{\mu}$ but not medium.
- Choose smaller β ? (0.01 here)

Current work
Go further than Ignatiadis et al. 2016

- Good theoretical properties but with strong assumptions.
- Notably, in real life π_0 is a π_{0g} depending on g.
- Estimate G_w with a better estimator than $\widehat{G}_w$?
- Finite sample guarantees ?

Bibliography I

- Blanchard, Gilles and Etienne Roquain (2008). "Two simple sufficient conditions for FDR control". In: *Electronic journal of Statistics* 2, pp. 963–992.
- Cai, T. Tony and Wenguang Sun (2009). "Simultaneous testing of grouped hypotheses: Finding needles in multiple haystacks". In: *Journal of the American Statistical Association* 104.488, pp. 1467–1481.
- Chi, Zhiyi (2007). "On the performance of FDR control: constraints and a partial solution". In: *The Annals of Statistics* 35.4, pp. 1409–1431.
- Genovese, Christopher R., Kathryn Roeder, and Larry Wasserman (2006). "False discovery control with p-value weighting". In: *Biometrika*, pp. 509–524.
- Holm, Sture (1979). "A simple sequentially rejective multiple test procedure". In: *Scandinavian journal of statistics*, pp. 65–70.

Bibliography II

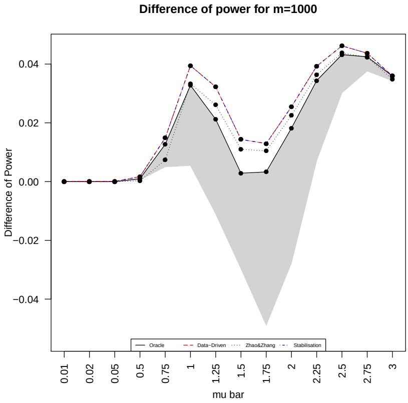
- Hu, James X., Hongyu Zhao, and Harrison H. Zhou (2010). "False discovery rate control with groups". In: *Journal of the American Statistical Association* 105.491.
- Ignatiadis, Nikolaos et al. (2016). "Data-driven hypothesis weighting increases detection power in genome-scale multiple testing". In: *Nature methods* 13.7, pp. 577–580.
- Massart, Pascal (1990). "The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality". In: *The Annals of Probability*, pp. 1269–1283.
- Roquain, Etienne and Mark A. Van De Wiel (2009). "Optimal weighting for false discovery rate control". In: *Electronic Journal of Statistics* 3, pp. 678–711.
- Zhao, Haibing and Jiajia Zhang (2014). "Weighted p-value procedures for controlling FDR of grouped hypotheses". In: *Journal of Statistical Planning and Inference* 151, pp. 90–106.

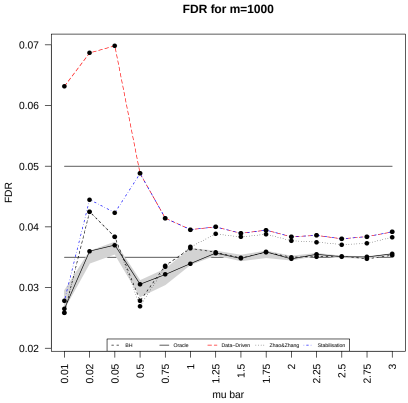
The end

Thank you for your attention !

Comparison with other methods

$\alpha = 0.05$, 70% true null, $m_1 = m_2 = 500$, $\beta = 0.05$





Existence and uniqueness of oracle optimal weights
Assumptions

From Roquain and Van De Wiel 2009 :

- F_i is strictly concave and continuous on $[0, 1]$
- F_i has a derivative f_i on $(0, 1)$
- $f_i(0^+)$ is constant for all i, same for $f_i(1^-)$
- $\lim_{y \to f_i(0^+)} \frac{f_j^{-1}(y)}{f_i^{-1}(y)}$ exists in $[0, \infty]$ for all i, j

These hypotheses are fulfilled in the gaussian 1-sided framework.

Optimal weighting
Existence and uniqueness

Proof ideas

Compute an explicit formula using the Lagrange multiplier method :

$$L(\lambda, w) = m^{-1} \sum_{i=1}^{m} F_i(\alpha u w_i) - \lambda \left(\sum_{i=1}^{m} w_g - m \right)$$

gives us

$$W_i^*(u) = \frac{1}{\alpha u} f_i^{-1} \left(\Psi^{-1}(\alpha u) \right)$$

where $\Psi(x) = m^{-1} \sum_{i=1}^{m} f_i^{-1}(x)$.

Some notations

- From now W^* is the asymptotic optimal weight when the F_g are known :

$$W^*(u) = \underset{w:\sum \pi_g w_g = 1}{\operatorname{argmax}} G_w^\infty(u) = \underset{w:\sum \pi_g w_g = 1}{\operatorname{argmax}} \sum_g \pi_g D_g(\alpha u w_g)$$

with $D_g(\cdot) = \pi_0 \max(\cdot, 1) + (1 - \pi_0) F_g(\cdot)$.

- $P_W^\infty(u) = (1 - \pi_0) \sum_g \pi_g F_g(\alpha u W_g(u))$.
- $\hat{u} = \mathcal{I}\left(\widehat{G}_{\widehat{W}^*}\right)$ and $u^* = \mathcal{I}(G_{W^*}^\infty)$.

A chain of technical results

A first lemma

$$\sup_{u\in[0,1]} \sup_{w\in(\mathbb{R}^+)^G} \left| \widehat{G}_w(u) - G_w^\infty(u) \right| \xrightarrow{a.s.} 0$$

by Glivenko-Cantelli theorem and $\frac{m_g}{m} \to \pi_g$.

The main technical proposition

Proposition

$$\sup_{u\in[0,1]} \left| \widehat{G}_{\widehat{W}^*}(u) - G^{\infty}_{W^*}(u) \right| \xrightarrow{a.s.} 0$$

or, equivalently,

$$\sup_{u\in[0,1]} \left| G^{\infty}_{\widehat{W}^*}(u) - G^{\infty}_{W^*}(u) \right| \xrightarrow{a.s.} 0.$$

The main technical proposition
Proof ideas

- Manipulate with the triangular inequality and remove the absolute values when able by using the maximality of $\widehat{G}_{\widehat{W}^*}(u)$ and $G^{\infty}_{W^*}(u)$

Problem

They are not maxima on the same sets :
$K^m = \{w : m^{-1} \sum m_g w_g = 1\}$ versus $K^\infty = \{w : \sum \pi_g w_g = 1\}$

The main technical proposition

Proof ideas

- We introduce two shifts $\delta(u) = \sum \pi_g \widehat{W}_g^*(u) - 1$ and $\delta'(u) = \sum \frac{m_g}{m} W_g^*(u) - 1$.
- Then we form shifted weights $\widehat{W}^\sim(u) = \widehat{W}^*(u) - \boldsymbol{\delta}(u) \in K^\infty$ and $W^\sim(u) = W^*(u) - \boldsymbol{\delta}'(u) \in K^m$.

The main technical proposition
Final ideas

- Use that $\left| G^{\infty}_{\widehat{W}^{\sim}}(u) - G^{\infty}_{W^*}(u) \right| = G^{\infty}_{W^*}(u) - G^{\infty}_{\widehat{W}^{\sim}}(u)$.
- End up with

$$\sup_u \left| G^{\infty}_{\widehat{W}^*}(u) - G^{\infty}_{W^*}(u) \right| \leq \sup_u \left(\widehat{G}_{W^{\sim}}(u) - \widehat{G}_{\widehat{W}^*}(u) \right) + \zeta_m$$

for some $\zeta_m \xrightarrow{a.s.} 0$.

- Use that $\widehat{G}_{W^{\sim}}(u) - \widehat{G}_{\widehat{W}^*}(u) \leq 0$. $\square$

The second important proposition

Proposition

$$\hat{u} \xrightarrow[m\to\infty]{a.s.} u^*$$

from which $\widehat{G}_{\widehat{W}^*}(\hat{u}) \xrightarrow{a.s.} G^{\infty}_{W^*}(u^*)$ because $\widehat{G}_{\widehat{W}^*}(\hat{u}) = \hat{u}$ and $G^{\infty}_{W^*}(u^*) = u^*$.

Note $X_m = \sup_{u\in[0,1]} \left| \widehat{G}_{\widehat{W}^*}(u) - G^{\infty}_{W^*}(u) \right| \overset{a.s.}{\to} 0$, take a δ in $(0, u^*)$, note $u^0 = u^* - \delta$ and for all $\delta' \geq \delta$, $u' = u^* + \delta'$.

The second important proposition
Proof

- $s_\delta = \max_{\delta' \geq \delta} (G^{\infty}_{W^*}(u') - u') < 0$ because if $s_\delta = 0$ it would contradict u^* maximality.
- $sup_{\delta' \geq \delta} \left(\widehat{G}_{\widehat{W}^*}(u') - u' \right) \leq s_\delta + X_m \to s_\delta < 0$
- So when $m \to \infty$ we must have $\hat{u} < u^* + \delta$.

The second important proposition
Proof

- $G^{\infty}_{W^*}(u^0) \geq G^{\infty}_{w}(u^0)$ with $w = W^*(u^*)$ by maximality.
- $G^{\infty}_{w}(u^0) = \frac{G^{\infty}_{w}(u^0)}{u^0} u^0 > \frac{G^{\infty}_{w}(u^*)}{u^*} u^0 = u^0$ by strict concavity.
- $\widehat{G}_{\widehat{W}^*}(u^0) - u^0 \geq G^{\infty}_{W^*}(u^0) - u^0 - X_m \to G^{\infty}_{W^*}(u^0) - u^0 > 0$.
- So when $m \to \infty$ we must have $\hat{u} > u^* - \delta$. □

Third and last proposition

We have shown that $\widehat{G}_{\widehat{W}^*}(\hat{u}) \xrightarrow{a.s.} u^*$, that is for the denominator of the FDP. Showing that the numerator converges to $\pi_0 \alpha u^*$ is straightforward after this :

Proposition

$$\widehat{W}^*(\hat{u}) \xrightarrow{a.s.} W^*(u^*),$$

or, equivalently,

$$\widehat{W}^\sim(\hat{u}) \xrightarrow{a.s.} W^*(u^*).$$

Third and last proposition

Proof ideas

- One can show with the previous results and the triagular inequality that $\left| G^{\infty}_{\widehat{W}^{\sim}(\hat{u})}(u^*) - G^{\infty}_{W^*}(u^*) \right| \xrightarrow{a.s.} 0$.
- By contradiction, if $\widehat{W}^{\sim}(\hat{u}) \overset{a.s.}{\nrightarrow} W^*(u^*)$ then we find a $w^I \neq W^*(u^*)$ maximizing $G^{\infty}_{w}(u^*)$ but $W^*(u^*)$ is unique. $\square$

Optimality in power
Proof ideas

- First, $\text{Pow}\left(\widehat{W}^*\right) = \mathbb{E}\left[\widehat{P}_{\widehat{W}^*}(\hat{u})\right]$ where $\widehat{P}_W(u)$ is m^{-1} times the number of true alternative rejected.
- $\widehat{P}_{\widehat{W}^*}(\hat{u}) \xrightarrow{a.s.} P^{\infty}_{W^*}(u^*)$ and $\widehat{P}_{W^{*(m)}}(\hat{u}) \xrightarrow{a.s.} P^{\infty}_{W^*}(u^*)$.
- For each limit point for $\text{Pow}(w^{(m)})$ there is a limit point w for $w^{(m)}$.
- $\hat{u}^{(m'')} \xrightarrow{a.s.} \mathcal{I}\left(G^{\infty}_w\right)$ and then
- $\widehat{P}_{w^{(m'')}}\left(\hat{u}^{(m'')}\right) \xrightarrow{a.s.} P^{\infty}_w\left(\mathcal{I}\left(G^{\infty}_w\right)\right) \leq P^{\infty}_{W^*}\left(\mathcal{I}\left(G^{\infty}_w\right)\right) \leq P^{\infty}_{W^*}(u^*)$. $\square$

More about the computation of $\widehat{W}^*$

Start of the algorithm

- Fix $u = \frac{1}{m}$, form $\tilde{p}_{gi} = \frac{p_{gi}}{\alpha u}$ and order the $\tilde{p}_{gi}$ in each group :

$$\tilde{p}_{g,1} \leq \cdots \leq \tilde{p}_{g,m_g}.$$

  Also note $\tilde{p}_{g,0} = 0$.

- If $\forall g, \tilde{p}_{g,1} > m$, no rejection and move to $u = \frac{2}{m}$. If $\exists g, \tilde{p}_{g,1} \leq m$, continue and at least 1 rejection.

More about the computation of $\widehat{W}^*$

Start of the algorithm

- Form all G-tuples $\boldsymbol{j} : \sum j_g = 2$ and check if there is one $\boldsymbol{j}$ such that $\sum m_g \tilde{p}_{g,j_g} \leq m$
  - If there is one, at least 2 rejections and continue with G-tuples of sum equal to 3.
  - If not, 1 rejection and use a $w_g = \tilde{p}_{g,j_g}$ suitable for 1 rejection, and move to $u = \frac{2}{m}$.

More about the computation of $\widehat{W}^*$
At rejection level k

- Form all G-tuples $\boldsymbol{j} : \sum j_g = k$ and check if there is one $\boldsymbol{j}$ such that $\sum m_g \tilde{p}_{g,j_g} \leq m$
  - If there is one, at least k rejections and continue with G-tuples of sum equal to $k+1$.
  - If not, $k-1$ rejections and use a $w_g = \tilde{p}_{g,j_g}$ suitable for $k-1$ rejections, and move to $u = \frac{2}{m}$.

Illustration of $\widehat{W}^*(u)$

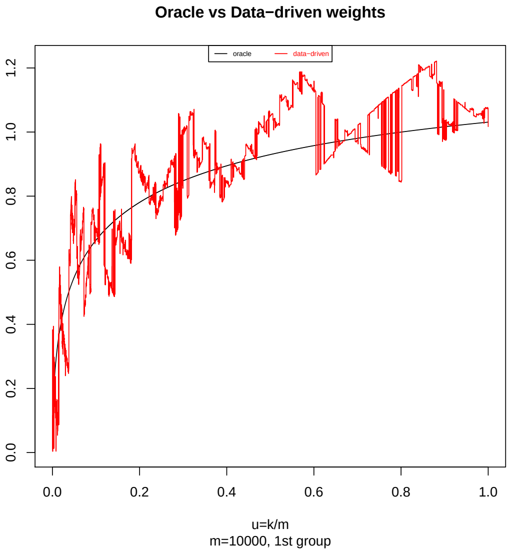
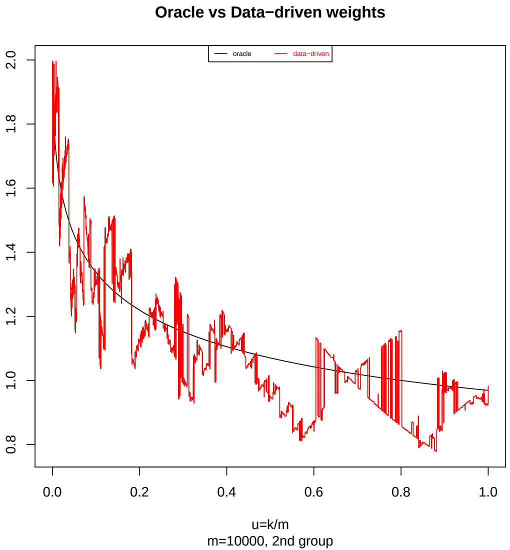
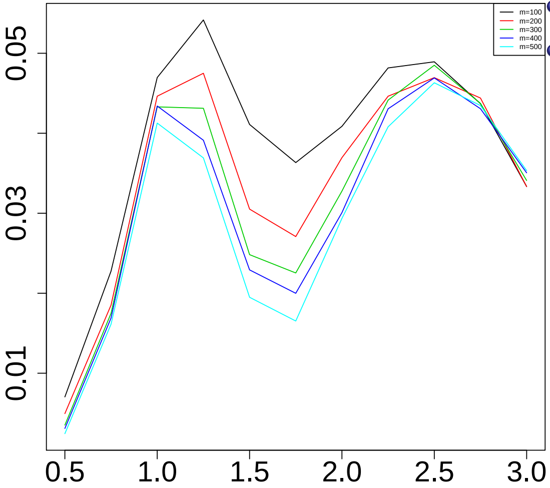


Illustration of $\widehat{W}^*(u)$



The overfitting decreases in m

$\alpha = 0.05$, 70% true null, $\pi_1 = \pi_2 = 0.5$



- $\mu_1 = \bar{\mu}$ and $\mu_2 = 2\bar{\mu}$.
- x axis : $\bar{\mu}$.
- y axis : the power of our procedure over 1000 replications minus the power of BH.