

Spike-and-slab Bayesian multiple testing procedures

Ismaël Castillo

Sorbonne Université

joint work with Étienne Roquain

Aad's conference 2019





1. Introduction

Frequentist analysis of posteriors

Experiment. $X = X^{(n)}$, $\mathcal{P} = \{P_{\theta}^{(n)}, \theta \in \Theta\}$,

Bayesian setting.

- a) $\theta \sim \Pi$ prior distribution
- b) $X|\theta \sim P_{\theta}^{(n)}$
- c) $\theta|X \sim: \Pi[\cdot|X]$ posterior

Frequentist analysis of $\Pi[\cdot|X]$.

Suppose there exists $\theta_0 \in \Theta$ 'true' value of θ

→ study the properties of $\Pi[\cdot|X]$ in probability under $X \sim P_{\theta_0}^{(n)}$

Posterior distribution for frequentist inference

The posterior is a distribution $\Pi[\cdot | X]$, which one may use for inference

- estimation

- ▶ convergence rate of $\Pi[\cdot | X]$
[Ghosal, Ghosh, van der Vaart 00] [Ghosal&van der Vaart's book]
- ▶ in general not the same as cv. of aspects

- uncertainty quantification

- ▶ credible sets = high probability sets under $\Pi[\cdot | X]$
[Szabó, van der Vaart, van Zanten 15]
- ▶ limiting shape of $\Pi[\cdot | X]$: Bernstein–von Mises

- multiple testing (?)

- ▶ comparing $\Pi[H_0 | X]$ and $\Pi[H_1 | X]$
- ▶ broadly used in practice, but not much so far on frequentist analysis

Sparsity: Gaussian sequence model

Sparse Gaussian sequence model

$$X_i = \theta_i + \varepsilon_i, \quad 1 \leq i \leq n$$

- Unknown $\theta = (\theta_1, \dots, \theta_n)$ is a sparse vector

$$\theta \in \ell_0[s_n] = \{\theta \in \mathbb{R}^n, |\{i : \theta_i \neq 0\}| \leq s_n\}$$

- $\varepsilon_1, \dots, \varepsilon_n$ independent $\mathcal{N}(0, 1)$

Sparsity, spike-and-slab prior

$$X_i = \theta_i + \varepsilon_i, \quad \theta \in \ell_0[s_n]$$

Spike and slab prior on θ

$$\Pi_\alpha \sim \bigotimes_{i=1}^n (1 - \alpha)\delta_0 + \alpha\mathcal{G}$$

$\alpha \in [0, 1]$, $\mathcal{G}$ distribution with density γ on $\mathbb{R}$

Sparsity, spike-and-slab prior

$$X_i = \theta_i + \varepsilon_i, \quad \theta \in \ell_0[s_n]$$

Spike and slab prior on θ

$$\Pi_\alpha \sim \bigotimes_{i=1}^n (1 - \alpha)\delta_0 + \alpha\mathcal{G}$$

$\alpha \in [0, 1]$, $\mathcal{G}$ distribution with density γ on $\mathbb{R}$

The associated posterior distribution given α is, with $g = \gamma * \phi$,

$$\Pi_\alpha[\cdot | X] \sim \bigotimes_{i=1}^n (1 - a_\alpha(X_i))\delta_0 + a_\alpha(X_i)\mathcal{G}_{X_i}$$
$$1 - a_\alpha(X_i) = \frac{(1 - \alpha)\phi(x)}{(1 - \alpha)\phi(x) + \alpha g(x)} = \Pi_\alpha[\theta_i = 0 | X]$$

Choice of α : we take $\alpha = \hat{\alpha}(X) = \hat{\alpha}$ as discussed later

2. Multiple testing, FDR, and posteriors

Multiple testing framework

Suppose one wants to simultaneously test

$$H_{0i} : \theta_i = 0 \quad \text{against} \quad H_{1i} : \theta_i \neq 0, \quad \text{for all } 1 \leq i \leq n$$

Multiple testing procedure $\varphi = (\varphi_i)_{1 \leq i \leq n} \in \{0, 1\}^n$

$\varphi_i = 1$ if and only if the null hypothesis H_{0i} is rejected

The frequentist False Discovery Rate (FDR) of φ at point θ_0 is defined as

$$\text{FDR}(\theta_0, \varphi) = E_{\theta_0} \left[\frac{\sum_{i=1}^n \mathbf{1}\{\theta_{0,i} = 0\} \varphi_i}{1 \vee \sum_{i=1}^n \varphi_i} \right]$$

FDR-specific state-of-the-art method: the BH procedure

For each i , let P_i denote the p -value for testing H_{0i} using a test of the form $\mathbb{1}_{|x_i| \geq \lambda}$

[The BH procedure] Order the p -values

$$P_{(1)} \leq P_{(2)} \leq \dots \leq P_{(n)}$$

Fix desired level $t \in (0, 1)$ and set

$$\hat{k} = \max \left\{ k : P_{(k)} \leq \frac{k}{n} t \right\}$$
$$\varphi_i^{BH} = \mathbb{1}_{P_i \leq P_{(\hat{k})}}, \quad 1 \leq i \leq n$$

Benjamini and Hochberg (95) prove in particular that

$$FDR_{\theta_0}[\varphi^{BH}] = \frac{n - s_n}{n} t \quad \forall \theta_0 \in \ell_0[s_n]$$
$$\leq t$$

Bayesian multiple testing: ℓ -values procedure

Prior distribution Π on θ and data X from the model, define the ℓ -value as

$$\ell_i(X) = \Pi[\theta_i = 0 | X]$$

Let $t \in (0, 1)$ be a desired control level e.g. 5%

Procedure 1 Reject H_{0i} for any coordinate i with

$$\ell_i(X) \leq t$$

That is, set $\varphi_i^\ell = (\varphi_t^\ell)_i = \mathbb{1}_{\ell_i(X) \leq t}$

Bayesian multiple testing: ℓ -values procedure

Prior distribution Π on θ and data X from the model, define the ℓ -value as

$$\ell_i(X) = \Pi[\theta_i = 0 | X]$$

Let $t \in (0, 1)$ be a desired control level e.g. 5%

Procedure 1 Reject H_{0i} for any coordinate i with

$$\ell_i(X) \leq t$$

That is, set $\varphi_i^\ell = (\varphi_t^\ell)_i = \mathbb{1}_{\ell_i(X) \leq t}$

- Efron [07, 08] advocated the use of such procedure with empirical Bayes choice of Π
- For $\Pi = \Pi_\alpha$ spike and slab prior and fixed α , we have $\ell_i(x) = \frac{(1-\alpha)\phi(x)}{(1-\alpha)\phi(x) + \alpha g(x)}$

Bayesian multiple testing: Bayesian FDR

Proposition. For any fixed prior Π , for $\varphi = \varphi^\ell$ (Procedure 1) with cut-off $t \in (0, 1)$,

$$\text{FDR}_B(\Pi, \varphi) := \int \text{FDR}(\theta, \varphi) d\Pi(\theta) \leq t.$$

$$\begin{aligned} \text{FDR}_B(\Pi, \varphi^\ell) &= E \left[E \left[\frac{\sum_{i=1}^n \mathbb{1}_{\theta_i=0} \mathbb{1}_{\varphi_i^\ell(X)=1}}{1 \vee \sum_{i=1}^n \mathbb{1}_{\varphi_i^\ell(X)=1}} \mid X \right] \right] = E_X \left[\frac{\sum_{i=1}^n \ell_i(X) \mathbb{1}_{\{\ell_i(X) \leq t\}}}{1 \vee \sum_{i=1}^n \mathbb{1}_{\{\ell_i(X) \leq t\}}} \right] \\ &\leq t P(\exists i : \ell_i(X) \leq t) \leq t \end{aligned}$$

Bayesian multiple testing: Bayesian FDR

Proposition. For any fixed prior Π , for $\varphi = \varphi^\ell$ (Procedure 1) with cut-off $t \in (0, 1)$,

$$\text{FDR}_B(\Pi, \varphi) := \int \text{FDR}(\theta, \varphi) d\Pi(\theta) \leq t.$$

$$\begin{aligned} \text{FDR}_B(\Pi, \varphi^\ell) &= E \left[E \left[\frac{\sum_{i=1}^n \mathbb{1}_{\theta_i=0} \mathbb{1}_{\varphi_i^\ell(X)=1}}{1 \vee \sum_{i=1}^n \mathbb{1}_{\varphi_i^\ell(X)=1}} \mid X \right] \right] = E_X \left[\frac{\sum_{i=1}^n \ell_i(X) \mathbb{1}_{\{\ell_i(X) \leq t\}}}{1 \vee \sum_{i=1}^n \mathbb{1}_{\{\ell_i(X) \leq t\}}} \right] \\ &\leq t P(\exists i : \ell_i(X) \leq t) \leq t \end{aligned}$$

Frequentist analysis: what about the behaviour of $\text{FDR}(\theta_0, \varphi^\ell)$?

Given nominal level t , do we have $\text{FDR}(\theta_0, \varphi^\ell) \leq t$ for well-chosen Π and cut-off t ?

Relation to [Abramovich, Benjamini, Donoho, Johnstone (2006)]

- [ABDJ 06] define an estimator $\hat{\theta}$ based on the BH idea

$$\hat{\theta}_i^{BH} := X_i \mathbb{1}_{P_i \leq P_{(\hat{k})}}$$

and show that, if $s_n \in [\log^5 n, n^\delta]$ for $0 < \delta < 1$, then for large n

$$\sup_{\theta \in \ell_0[s_n]} E_\theta [\|\hat{\theta}^{BH} - \theta\|_2^2] \leq C_\delta s_n \log(n/s_n)$$

“good risk control $\longleftarrow$ FDR-adapted method”

- Could one go the other way around

“optimal posterior rate $\longrightarrow$ good FDR control?”

Goals for our Bayesian multiple testing method

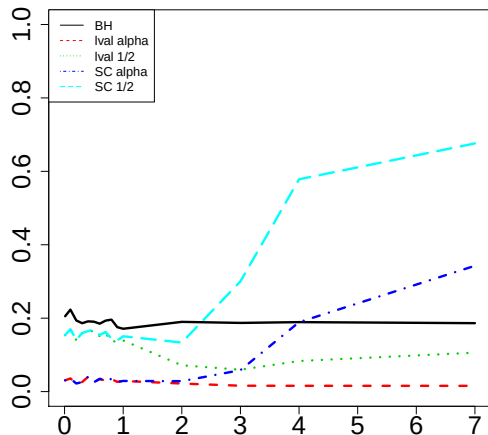
Consider a posterior distribution $\Pi[\cdot | X]$

- it should be adaptive, rate optimal estimation rate
- it should enable uncertainty quantification confidence sets
- from $\Pi[\cdot | X]$, and fixed level t , want to derive [how?] a procedure $\varphi = \varphi(t, X)$ s.t.

$$\sup_{\theta_0 \in \ell_0[s_n]} FDR(\theta_0, \varphi) \leq t \quad [\text{or close to } t] \quad \textcolor{blue}{FDR \text{ control}}$$

Is it at all possible?

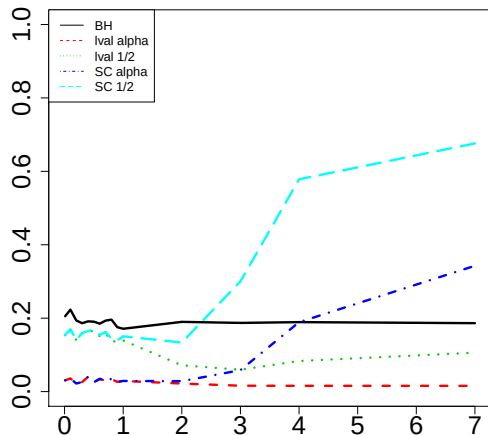
Bayesian multiple testing: frequentist analysis (empirical)



$\Pi = \Pi_{\hat{\alpha}}$ EB-spike and slab

- target level $\alpha = 0.2$
- $n = 1000$, $s_n = 50$
- signal $\theta_{0,i} = M \mathbb{1}_{i \leq s_n}$
- x-axis: grid of signal strength M
- y-axis: empirical $\text{FDR}(\theta_0, \varphi^\ell)$

Bayesian multiple testing: frequentist analysis (empirical)



$\Pi = \Pi_{\hat{\alpha}}$ EB-spike and slab

- target level $\alpha = 0.2$
- $n = 1000$, $s_n = 50$
- signal $\theta_{0,i} = M \mathbb{1}_{i \leq s_n}$
- x-axis: grid of signal strength M
- y-axis: empirical $\text{FDR}(\theta_0, \varphi^\ell)$

Remark. For a horseshoe-based procedure [van der Pas, Szabo, van der Vaart 17] observed similar empirical evidence as green and red lines

3. Empirical Bayes Spike and Slab:

the method and some optimality properties

Empirical Bayes Spike and Slab

$$\begin{aligned}X|\theta &\sim \otimes_{i=1}^n \mathcal{N}(\theta_i, 1) \\ \theta|\alpha &\sim \otimes_{i=1}^n (1-\alpha)\delta_0 + \alpha\mathcal{G} =: \Pi_\alpha\end{aligned}$$

Idea: estimate α by maximising the marginal likelihood of $X|\alpha$

$$p(X|\alpha) = \prod_{i=1}^n \int \phi(X_i - \theta_i) d\Pi(\theta_i|\alpha)$$

Setting $g = \phi * \gamma$, this leads to [with $a_n \approx 1/n$]

$$\hat{\alpha} = \arg \max_{\alpha \in [a_n, 1]} \prod_{i=1}^n (1-\alpha)\phi(X_i) + \alpha g(X_i)$$

Next: make inference using plug-in posterior $\Pi_{\hat{\alpha}}[\cdot|X]$

Empirical Bayes Spike and Slab: aspects of posterior

$$\Pi_{\hat{\alpha}}[\cdot | X] = \otimes_{i=1}^n (1 - a_{\hat{\alpha}}(X_i))\delta_0 + a_{\hat{\alpha}}(X_i)\mathcal{G}_{X_i}$$

Theorem (convergence for posterior mean and median) [Johnstone Silverman 04]

Suppose γ has 'heavy tails' [at least Laplace, at most Cauchy].

Let $\hat{\theta} = \bar{\theta} = \int \theta d\Pi_{\hat{\alpha}}(\theta | X)$ or the posterior median of $\Pi_{\hat{\alpha}}[\cdot | X]$. Then for large M

$$\sup_{\theta_0 \in \ell_0[s_n]} E_{\theta_0}[\|\hat{\theta} - \theta_0\|^2] \leq Ms_n \log(n/s_n).$$

Empirical Bayes SAS: optimal minimax rate

Theorem [C. and Misner 18] Let Π_α be the SAS prior with γ the **Cauchy** density. Chose $\hat{\alpha}$ by empirical Bayes as before. Then for large enough M

$$\sup_{\theta_0 \in \ell_0[s_n]} E_{\theta_0} \int \|\theta - \theta_0\|^2 d\Pi_{\hat{\alpha}}(\theta | X) \leq Ms_n \log(n/s_n)$$

Empirical Bayes SAS: optimal minimax rate

Theorem [C. and Misner 18] Let Π_α be the SAS prior with γ the **Cauchy** density. Chose $\hat{\alpha}$ by empirical Bayes as before. Then for large enough M

$$\sup_{\theta_0 \in \ell_0[s_n]} E_{\theta_0} \int \|\theta - \theta_0\|^2 d\Pi_{\hat{\alpha}}(\theta | X) \leq Ms_n \log(n/s_n)$$

- Complete EB posterior is optimal if γ heavy tailed
[tails inbetween x^{-3} and $x^{-1-\delta}$ work]
- does **not** work if slab γ is Laplace [lower bound rate $e^{\sqrt{c \log n/s_n} s_n}$]
→ fundamental difference between posterior and aspects of it
- empirical Bayes estimate $\hat{\alpha}$ more sensible to tails of slab than full Bayes
→ difference between empirical Bayes and full Bayes

Empirical Bayes SAS: sharp minimax constant

Theorem [C. and Misner 19+] Let Π_α be the SAS prior with γ a 'super-heavy' tailed density [$\propto 1/(x \log^2(x))$]. Chose $\hat{\alpha}$ by empirical Bayes as before. Then

$$\sup_{\theta_0 \in \ell_0[s_n]} E_{\theta_0} \int \|\theta - \theta_0\|^2 d\Pi_{\hat{\alpha}}(\theta | X) \sim 2s_n \log(n/s_n).$$

Empirical Bayes SAS: confidence sets

Let $\gamma \in (0, 1)$ be a fixed level

From the posterior $\Pi[\cdot | X]$ one can build $\mathcal{A}$ such that

$$\Pi[\mathcal{A} | X] \geq 1 - \gamma \quad \text{credible set}$$

Requirements: for some large subset of sparse vectors

$$\lim_{n \rightarrow \infty} P_{\theta_0}(\theta_0 \in \mathcal{A}) \geq 1 - \gamma \quad \text{confidence set}$$

$$E_{\theta_0} \text{Diam}_d(\mathcal{A}) \sim r_n^* \quad \text{diameter}$$

where r_n^* optimal (minimax) diameter

→ SAS empirical Bayes confidence sets investigated in [C. and Szabó 19]

4. Main results on multiple testing

Bayesian multiple testing: ℓ -values result

Theorem 1. [C. and Roquain 19 (simplified)]

Let $\Pi_{\hat{\alpha}}$ SAS prior with Empirical Bayes choice of α , with slab Laplace or Cauchy.

$$\ell_i(X) = \Pi_{\hat{\alpha}}[\theta_i = 0 | X]$$

$$\varphi_i^\ell(X) = \mathbb{1}_{\ell_i(X) \leq t} \quad \text{for fixed } t \in (0, 1).$$

Suppose $s_n = O(n^\delta)$ for some $0 < \delta < 1$. Then as $n \rightarrow \infty$,

$$\lim_{n \rightarrow \infty} \sup_{\theta_0 \in \ell_0[s_n]} FDR(\theta_0, \varphi^\ell) = 0.$$

Bayesian multiple testing: ℓ -values result

Theorem 1. [C. and Roquain 19 (simplified)]

Let $\Pi_{\hat{\alpha}}$ SAS prior with Empirical Bayes choice of α , with slab Laplace or Cauchy.

$$\ell_i(X) = \Pi_{\hat{\alpha}}[\theta_i = 0 | X]$$

$$\varphi_i^\ell(X) = \mathbb{1}_{\ell_i(X) \leq t} \quad \text{for fixed } t \in (0, 1).$$

Suppose $s_n = O(n^\delta)$ for some $0 < \delta < 1$. Then as $n \rightarrow \infty$,

$$\lim_{n \rightarrow \infty} \sup_{\theta_0 \in \ell_0[s_n]} FDR(\theta_0, \varphi^\ell) = 0.$$

Remark. The same holds for ‘marginal credible intervals’ multiple testing method

Bayesian multiple testing: ℓ -values comments

- A more precise result: for large enough n , uniformly over $\theta_0 \in \ell_0[s_n]$,

$$\sup_{\theta_0 \in \ell_0[s_n]} FDR(\theta_0, \varphi^\ell) \leq C \cdot \begin{cases} \frac{\log \log n}{\sqrt{\log n}} & \text{for Laplace slab} \\ \frac{\log \log n}{\log n} & \text{for Cauchy slab} \end{cases}$$

- A main difficulty for the proof is to get a **uniform** control i.e. over **all** $\theta_0 \in \ell_0[s_n]$
 - ▶ special case of strong signals is easiest
 - ▶ otherwise one has to control the EB estimate $\hat{\alpha}$ in a fairly sharp way
 - ▶ even in the absence of self-similarity-type conditions
- Can one find Bayes procedure such that, for given level t ,
→ FDR controlled at level “close to” t ?

Bayesian multiple testing: q -values

[Storey 2003] defines q -values as

$$q(x) = \Pi(\theta_i = 0 \mid |X_i| \geq |x|)$$

Bayesian multiple testing: q -values

[Storey 2003] defines q -values as

$$q(x) = \Pi(\theta_i = 0 \mid |X_i| \geq |x|)$$

- Intuition behind q -values mimics 'FDR' from the Bayesian perspective

$$q(x) = \frac{\Pi[\theta_i = 0, |X_i| \geq |x|]}{\Pi[|X_i| \geq |x|]} = \frac{E[\sum_{i=1}^n \mathbb{1}_{\theta_i=0} \mathbb{1}_{|X_i| \geq |x|}]}{E[\sum_{i=1}^n \mathbb{1}_{|X_i| \geq |x|}]}$$

Bayesian multiple testing: q -values

[Storey 2003] defines q -values as

$$q(x) = \Pi(\theta_i = 0 \mid |X_i| \geq |x|)$$

- Intuition behind q -values mimics 'FDR' from the Bayesian perspective

$$q(x) = \frac{\Pi[\theta_i = 0, |X_i| \geq |x|]}{\Pi[|X_i| \geq |x|]} = \frac{E[\sum_{i=1}^n \mathbb{1}_{\theta_i=0} \mathbb{1}_{|X_i| \geq |x|}]}{E[\sum_{i=1}^n \mathbb{1}_{|X_i| \geq |x|}]}$$

- For spike and slab prior Π_α and setting $\overline{F}(s) = \int_s^\infty f(x)dx$ for a density f ,

$$q_{\text{spike and slab}}(\alpha; x) = q(\alpha; x) = \frac{(1 - \alpha)\overline{\Phi}(|x|)}{(1 - \alpha)\overline{\Phi}(|x|) + \alpha \overline{G}(|x|)}$$

Bayesian multiple testing: q-values

Procedure 2 Reject H_{0i} for any coordinate i with

$$q_i(X) \leq t$$

That is, set $\varphi_i^q = (\varphi_t^q)_i = \mathbb{1}_{q_i(X) \leq t}$

- Comparison with ℓ -values for spike and slab prior

$$q(x) \leq \ell(x)$$

- For a fixed prior Π and $\mathbb{1}_{q(X_i) \geq t}$, one has the ‘decision-theoretic’ statement,

$$FDR_B(\Pi, \varphi_t^\ell) \leq FDR_B(\Pi, \varphi_t^q) \leq t.$$

Bayesian multiple testing: q -values

Theorem 2. [C. and Roquain 19]

Let $\Pi_{\hat{\alpha}}$ SAS prior with Empirical Bayes choice of α , with slab **Laplace** or **Cauchy**. Consider the corresponding procedure based on q -values, for $t \in (0, 1)$

$$\varphi_i^q(X) = \mathbf{1}\{ q(\hat{\alpha}; X_i) \leq t \}$$

Suppose $s_n = O(n^\delta)$ for some $0 < \delta < 1$. Then for n large enough,

$$\sup_{\theta_0 \in \ell_0[s_n]} FDR(\theta_0, \varphi^q) \leq Ct \log(1/t).$$

- the upper-bound is less conservative than ℓ -values
- q -values scale well with respect to the desired control level t (up to $\log(1/t)$ term)

Bayesian multiple testing: Q-values

Define a BMT procedure φ^{Q-val} as

$$\varphi_i^Q = \mathbf{1}\{q(\hat{\alpha}; X_i) \leq t\} \cdot \mathbf{1}\{\hat{\alpha} > a_n\}$$

$$a_n = \frac{\sqrt{\log n}}{n\overline{G}(\sqrt{2.1 \log n})} \asymp \frac{1}{n} \cdot \begin{cases} \log n & \text{Cauchy case} \\ \sqrt{\log n} e^{\sqrt{\log n}} & \text{Laplace case} \end{cases}$$

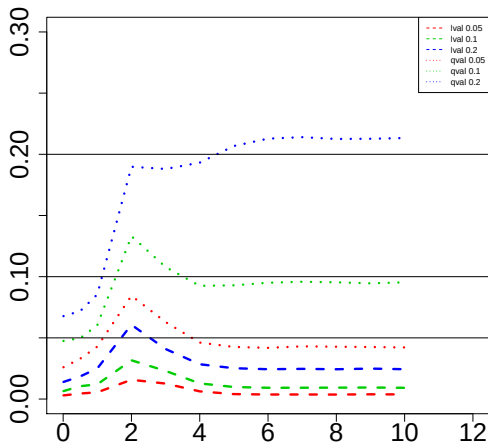
Theorem 2. (bis) [C. and Roquain 19]

Under the same conditions as before, consider the Q-values procedure. Suppose $s_n = O(n^\delta)$ for some $0 < \delta < 1$. Then for n large enough,

$$\sup_{\theta_0 \in \ell_0[s_n]} FDR(\theta_0, \varphi^Q) \leq Ct$$

- *Remark.* Constant C is explicit in terms of δ [and is close to 1 in simulations]

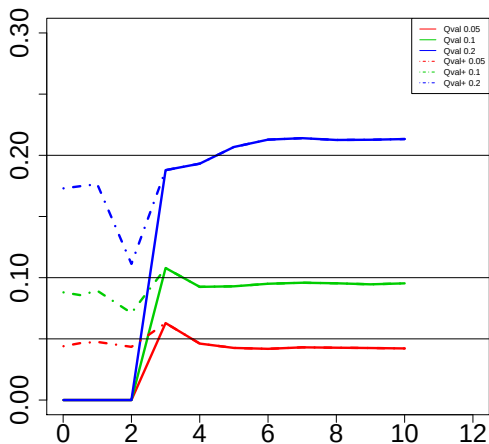
Bayesian multiple testing: q-values



$\Pi = \Pi_{\hat{\alpha}}$ EB-spike and slab

- $n = 10000$, $s_n/n = 0.01$
- signal $\theta_{0,i} = M\mathbb{1}_{i \leq s_n}$
- x-axis: grid of signal strength M
- y-axis: empirical $\text{FDR}(\theta_0, \varphi^q)$

Bayesian multiple testing: Q-values



$\Pi = \Pi_{\hat{\alpha}}$ EB-spike and slab

- $n = 10000$, $s_n/n = 0.01$
- signal $\theta_{0,i} = M\mathbb{1}_{i \leq s_n}$
- x-axis: grid of signal strength M
- y-axis: empirical $\text{FDR}(\theta_0, \varphi^Q)$
 $\text{FDR}(\theta_0, \varphi^{Q+})$

Sharp FDR-control for q/Q -values under strong signal

Is the bound for q/Q -values **sharp**? Indeed, we already had for large enough n

$$\sup_{\theta_0 \in \ell_0[s_n]} FDR(\theta_0, \varphi^\ell) \leq C \frac{\log \log n}{(\log n)^{\kappa/2}} \leq Ct$$

Is $FDR(\theta_0, \varphi^q)$ close to Ct at least for a subset of θ_0 s?

Definition. Let $\mathcal{L}_0[s_n] = \mathcal{L}_0[s_n; a]$ be a set of signals, for arbitrary $a > 1$,

$$\mathcal{L}_0[s_n] = \left\{ \theta \in \ell_0[s_n] : |\theta_i| \geq a \sqrt{2 \log \frac{n}{s_n}} \text{ for all } i \text{ such that } \theta_i \neq 0 \right\}$$

Sharp FDR-control for q -values under strong signal

Theorem 3. [C. and Roquain 19]

Let $\Pi_{\hat{\alpha}}$ SAS prior with Empirical Bayes choice of α , with slab Laplace or Cauchy. Consider the corresponding procedure based on q -values [or Q -values], for $t \in (0, 1)$

$$\varphi_i^q(X) = \varphi_i^q(X, t) = \mathbf{1}\{q(\hat{\alpha}; X_i) \leq t\}$$

Suppose $s_n = O(n^\delta)$ for some $0 < \delta < 1$. Then uniformly for $\theta_0 \in \mathcal{L}_0[s_n; a]$ ($a > 1$),

$$\text{FDR}(\theta_0, \varphi^q) \rightarrow t, \quad (n \rightarrow \infty)$$

FDR+FNR control under sharp signal condition

False Negative Rate [“type II” multiple testing error]

$$\text{FNR}(\theta_0, \varphi) = E_{\theta_0} \left[\frac{\sum_{i=1}^n \mathbf{1}\{\theta_{0,i} \neq 0\}(1 - \varphi_i(X))}{1 \vee \sum_{i=1}^n \mathbf{1}\{\theta_{0,i} \neq 0\}} \right].$$

FDR+FNR control under sharp signal condition

False Negative Rate ["type II" multiple testing error]

$$\text{FNR}(\theta_0, \varphi) = E_{\theta_0} \left[\frac{\sum_{i=1}^n \mathbf{1}\{\theta_{0,i} \neq 0\}(1 - \varphi_i(X))}{1 \vee \sum_{i=1}^n \mathbf{1}\{\theta_{0,i} \neq 0\}} \right].$$

Theorem 4. [C. and Roquain 19]

Let φ^ℓ, φ^q as before for $t \in (0, 1)$. Then uniformly for $\theta_0 \in \mathcal{L}_0[s_n; a]$ ($a > 1$),

$$\text{FNR}(\theta_0, \varphi^\ell) \rightarrow 0, \quad \text{FNR}(\theta_0, \varphi^q) \rightarrow 0.$$

Corollary. Let $\mathfrak{R}(\theta_0, \varphi) = \text{FDR}(\theta_0, \varphi) + \text{FNR}(\theta_0, \varphi)$. Then unif. over $\mathcal{L}_0[s_n; a]$ ($a > 1$),

$$\mathfrak{R}(\theta_0, \varphi^\ell) \rightarrow 0, \quad \mathfrak{R}(\theta_0, \varphi^q) \rightarrow t.$$

Remark [sharp boundary] For $a < 1$, $\liminf_n \inf_{\varphi \in \mathcal{C}} \sup_{\theta_0 \in \mathcal{L}_0[s_n; a]} \mathfrak{R}(\theta_0, \varphi) \geq 1$.

Conclusion

- Posteriors based on spike and slab priors, here calibrated using Empirical Bayes, enable optimal inference for sparse classes in terms of
 - ▶ optimal contraction rates, uncertainty quantification
 - ▶ FDR control
- FDR control of Bayes procedures
 - ▶ can be guaranteed from a frequentist perspective for adaptive rate procedures
 - ▶ spike and slab EB priors provide such a control

= theoretical validation of practical use of posteriors for multiple testing

- Perspectives/work in progress
 - ▶ ℓ/q -values can be deployed in many settings → theory?
 - ▶ multiple testing → ANR grant project **BASICS**
 - sparse models: regression and dependence, sparse random graphs
 - different priors
 - other procedures



Postdoc supervisors

Front: Aad (Ismaël)

Right: Mark van de Wiel (Étienne)

[Picture taken at Aad's 50's party @ VU]