Resampling-based confidence regions and multiple tests

Sylvain Arlot[1,2,3]
joint work with Gilles Blanchard[4] Étienne Roquain[5]

[1]CNRS

[2]École Normale Supérieure de Paris

[3]Willow, Inria Paris-Rocquencourt

[4]Universität Potsdam, Institut für Mathematik

[5]Université Pierre et Marie Curie, LPMA

JSTAR 2010, Statistiques en grande dimension, Rennes
October, 21st 2010

Model

Observations: $\mathbf{Y} = (Y^1, \dots, Y^n) = \begin{pmatrix} Y_1^1 & \dots & Y_1^n \\ \vdots & & \vdots \\ \vdots & & \vdots \\ Y_K^1 & \dots & Y_K^n \end{pmatrix}$

$Y^1, \dots, Y^n \in \mathbb{R}^K$ i.i.d. symmetric

- Unknown mean $\mu = (\mu_k)_k$
- Unknown covariance matrix Σ

- $n \ll K$

Aims: Find a confidence region for μ or $\{k \text{ s.t. } \mu_k \neq 0\}$

Model

Observations: $\mathbf{Y} = (Y^1, \dots, Y^n) = \begin{pmatrix} Y_1^1 & \dots & Y_1^n \\ \vdots & & \vdots \\ \vdots & & \vdots \\ Y_K^1 & \dots & Y_K^n \end{pmatrix}$

$Y^1, \dots, Y^n \in \mathbb{R}^K$ i.i.d. symmetric

- Unknown mean $\mu = (\mu_k)_k$
- Unknown covariance matrix Σ
- $n \ll K$

Aims: Find a confidence region for μ or $\{k \text{ s.t. } \mu_k \neq 0\}$

Model

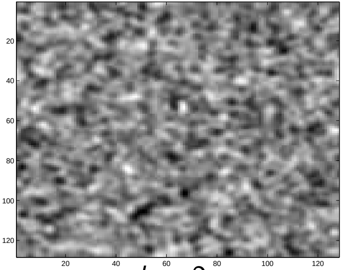
Observations: $\mathbf{Y} = (Y^1, \ldots, Y^n) = \begin{pmatrix} Y_1^1 & \ldots & Y_1^n \\ \vdots & & \vdots \\ \vdots & & \vdots \\ Y_K^1 & \ldots & Y_K^n \end{pmatrix}$

$Y^1, \ldots, Y^n \in \mathbb{R}^K$ i.i.d. symmetric

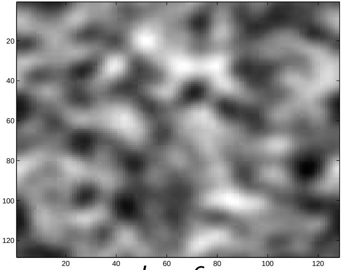
- Unknown mean $\mu = (\mu_k)_k$
- Unknown covariance matrix Σ
- $n \ll K$

Aims: Find a confidence region for μ or $\{k \text{ s.t. } \mu_k \neq 0\}$

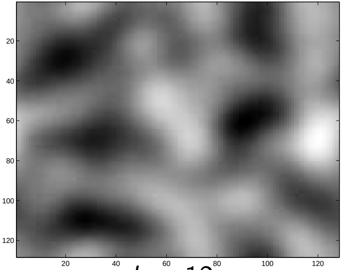
Illustration (1): $K = 16384 \gg n$, spatially correlated noise



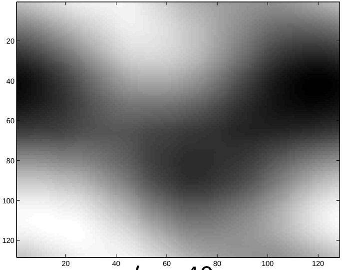
$b = 2$



$b = 6$



$b = 12$



$b = 40$

Illustration (2): $K = 16384 \gg n$, textured noise

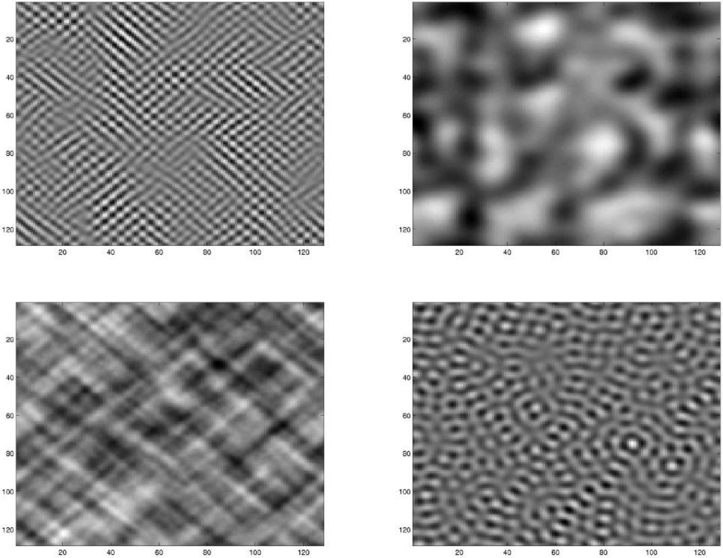
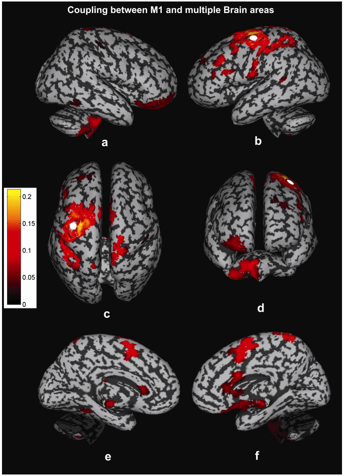
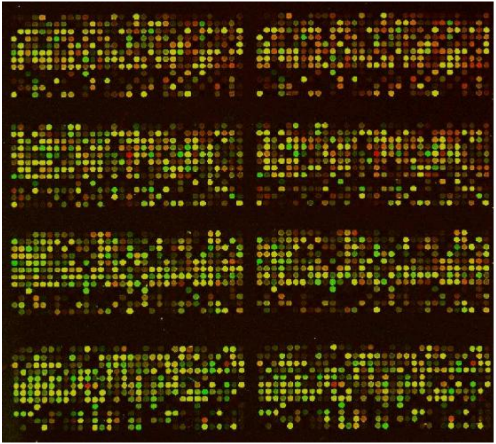


Illustration (3): neuroimaging and microarrays



(neural activity)



(gene expression levels)

Multiple simultaneous hypothesis testing

For every k we test: $H_{0,k}$: "$\mu_k = 0$" against $H_{1,k}$: "$\mu_k \neq 0$".

A *multiple testing procedure* rejects:

$$R(\mathbf{Y}) \subset \{1, \dots, K\}.$$

Type I errors measured by the Family Wise Error Rate:

$$\mathrm{FWER}(R) = \mathbb{P}(\exists k \in R(\mathbf{Y}) \text{ s.t. } \mu_k = 0).$$

$\Rightarrow$ build a procedure R such that $\mathrm{FWER}(R) \leq \alpha$?

- strong control of the FWER: $\forall \mu \in \mathbb{R}^K$
- power: $\mathrm{Card}(R(\mathbf{Y}))$ as large as possible

Multiple simultaneous hypothesis testing

For every k we test: $H_{0,k}$: "$\mu_k = 0$" against $H_{1,k}$: "$\mu_k \neq 0$".

A *multiple testing procedure* rejects:

$$R(\mathbf{Y}) \subset \{1, \ldots, K\} .$$

Type I errors measured by the Family Wise Error Rate:

$$\mathrm{FWER}(R) = \mathbb{P}\left(\exists k \in R(\mathbf{Y}) \text{ s.t. } \mu_k = 0\right) .$$

$\Rightarrow$ build a procedure R such that $\mathrm{FWER}(R) \leq \alpha$?

- strong control of the FWER: $\forall \mu \in \mathbb{R}^K$
- power: $\mathrm{Card}(R(\mathbf{Y}))$ as large as possible

Multiple simultaneous hypothesis testing

For every k we test: $H_{0,k}$: "$\mu_k = 0$" against $H_{1,k}$: "$\mu_k \neq 0$".

A *multiple testing procedure* rejects:

$$R(\mathbf{Y}) \subset \{1, \dots, K\} .$$

Type I errors measured by the Family Wise Error Rate:

$$\mathrm{FWER}(R) = \mathbb{P}\left(\exists k \in R(\mathbf{Y}) \text{ s.t. } \mu_k = 0\right) .$$

$\Rightarrow$ build a procedure R such that $\mathrm{FWER}(R) \leq \alpha$?

- strong control of the FWER: $\forall \mu \in \mathbb{R}^K$
- power: $\mathrm{Card}(R(\mathbf{Y}))$ as large as possible

Multiple simultaneous hypothesis testing

For every k we test: $H_{0,k}$: "$\mu_k = 0$" against $H_{1,k}$: "$\mu_k \neq 0$".

A *multiple testing procedure* rejects:

$$R(\mathbf{Y}) \subset \{1, \ldots, K\} .$$

Type I errors measured by the Family Wise Error Rate:

$$\mathrm{FWER}(R) = \mathbb{P}\left(\exists k \in R(\mathbf{Y}) \,\text{s.t.}\, \mu_k = 0\right) .$$

$\Rightarrow$ build a procedure R such that $\mathrm{FWER}(R) \leq \alpha$?

- strong control of the FWER: $\forall \mu \in \mathbb{R}^K$
- power: $\mathrm{Card}(R(\mathbf{Y}))$ as large as possible

Multiple simultaneous hypothesis testing

For every k we test: $H_{0,k}$: "$\mu_k = 0$" against $H_{1,k}$: "$\mu_k \neq 0$".

A *multiple testing procedure* rejects:

$$R(\mathbf{Y}) \subset \{1, \ldots, K\}.$$

Type I errors measured by the Family Wise Error Rate:

$$\mathrm{FWER}(R) = \mathbb{P}\left(\exists k \in R(\mathbf{Y}) \, \text{s.t.} \, \mu_k = 0\right).$$

$\Rightarrow$ build a procedure R such that $\mathrm{FWER}(R) \leq \alpha$?

- strong control of the FWER: $\forall \mu \in \mathbb{R}^K$
- power: $\mathrm{Card}(R(\mathbf{Y}))$ as large as possible

Thresholding

Reject $\quad R(\mathbf{Y}) = \{k \text{ s.t. } \sqrt{n}|\overline{\mathbf{Y}}_k| > t\}$

where

- $\overline{\mathbf{Y}} = \frac{1}{n}\sum_{i=1}^{n} Y^i$ empirical mean
- $t = t_\alpha(\mathbf{Y})$ threshold (independent of $k \in \{1, \ldots, K\}$)

$$\begin{aligned} \mathrm{FWER}(R) &= \mathbb{P}(\exists k \quad \text{s.t.} \quad \mu_k = 0 \text{ and } \sqrt{n}|\overline{\mathbf{Y}}_k| > t) \\ &= \mathbb{P}(\sqrt{n} \sup_{k \text{ s.t. } \mu_k = 0} |\overline{\mathbf{Y}}_k| > t) \\ &\leq \mathbb{P}(\sqrt{n} \sup_{k} |\overline{\mathbf{Y}}_k| > t) \\ &= \mathbb{P}\left(\|\overline{\mathbf{Y}} - \mu\|_\infty > tn^{-1/2}\right) \end{aligned}$$

So, L^∞ confidence region $\Rightarrow$ control of the FWER

Thresholding

Reject $\quad R(\mathbf{Y}) = \{k \text{ s.t. } \sqrt{n}|\overline{\mathbf{Y}}_k| > t\}$

where

- $\overline{\mathbf{Y}} = \frac{1}{n}\sum_{i=1}^{n} Y^i$ empirical mean
- $t = t_\alpha(\mathbf{Y})$ threshold (independent of $k \in \{1, \ldots, K\}$)

$$
\begin{aligned}
\mathrm{FWER}(R) &= \mathbb{P}(\exists k \quad \text{s.t.} \quad \mu_k = 0 \text{ and } \sqrt{n}|\overline{\mathbf{Y}}_k| > t) \\
&= \mathbb{P}(\sqrt{n} \sup_{k \text{ s.t. } \mu_k = 0} |\overline{\mathbf{Y}}_k| > t) \\
&\leq \mathbb{P}(\sqrt{n} \sup_k |\overline{\mathbf{Y}}_k| > t) \\
&= \mathbb{P}\left(\|\overline{\mathbf{Y}} - \mu\|_\infty > tn^{-1/2}\right)
\end{aligned}
$$

So, L^∞ confidence region $\quad \Rightarrow \quad$ control of the FWER

Bonferroni threshold, Gaussian case

Union bound:

$$\begin{aligned} \mathrm{FWER}(R) &\leq \mathbb{P}\left(\exists k \quad \text{s.t.} \quad \sqrt{n}\left|\overline{\mathbf{Y}}_k - \mu_k\right| > t\right) \\ &\leq K \sup_k \mathbb{P}(\sqrt{n}|\overline{\mathbf{Y}}_k - \mu_k| > t) \\ &\leq 2K\overline{\Phi}(t/\sigma) \ , \end{aligned}$$

where $\overline{\Phi}$ is the standard Gaussian upper tail function.

Bonferroni's threshold: $t_\alpha^{\mathrm{Bonf}} = \sigma \overline{\Phi}^{-1}(\alpha/(2K))$.

- deterministic threshold
- too conservative if there are strong correlations between the coordinates Y_k

$\Rightarrow$ how to do better ?

Bonferroni threshold, Gaussian case

Union bound:

$$\begin{aligned} \mathrm{FWER}(R) &\leq \mathbb{P}\left(\exists k \quad \text{s.t.} \quad \sqrt{n}\left|\overline{\mathbf{Y}}_k - \mu_k\right| > t\right) \\ &\leq K \sup_k \mathbb{P}(\sqrt{n}|\overline{\mathbf{Y}}_k - \mu_k| > t) \\ &\leq 2K\overline{\Phi}(t/\sigma) \quad , \end{aligned}$$

where $\overline{\Phi}$ is the standard Gaussian upper tail function.

Bonferroni's threshold: $t_\alpha^{\mathrm{Bonf}} = \sigma\overline{\Phi}^{-1}(\alpha/(2K))$.

- deterministic threshold
- too conservative if there are strong correlations between the coordinates Y_k

$\Rightarrow$ how to do better ?

Goal

For every $p \in [1, +\infty]$, find a threshold $t_\alpha(\mathbf{Y})$ such that

$$\forall \mu \in \mathbb{R}^K \qquad \mathbb{P}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p > t_\alpha(\mathbf{Y})\right) \leq \alpha \ .$$

$\Rightarrow$ L^p confidence ball for μ at level α (FWER$(R) \leq \alpha$ if $p = +\infty$)

- Non-asymptotic: $\forall K, n$
- General correlations
- Assumptions:
  (**Gauss**) $Y^i \sim \mathcal{N}(\mu, \Sigma)$ with $\sigma = (\Sigma_{k,k})_{k=1\dots K}$ known
  (**SB**) $Y^i - \mu \sim \mu - Y^i$ (symmetry) and $\forall k, \left|Y_k^i\right| \leq M$ a.s.

Ideal threshold: $t = q_\alpha^\star$, $(1-\alpha)$-quantile of $\mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p\right)$

$q_\alpha^\star$ depends on $\mathcal{D}(\mathbf{Y})$ unknown $\Rightarrow$ estimated by **resampling**.

Goal

For every $p \in [1, +\infty]$, find a threshold $t_\alpha(\mathbf{Y})$ such that

$$\forall \mu \in \mathbb{R}^K \qquad \mathbb{P}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p > t_\alpha\left(\mathbf{Y}\right)\right) \leq \alpha \ .$$

$\Rightarrow$ L^p confidence ball for μ at level α (FWER(R) $\leq \alpha$ if $p = +\infty$)

- Non-asymptotic: $\forall K, n$
- General correlations
- Assumptions:
  (**Gauss**) $Y^i \sim \mathcal{N}(\mu, \Sigma)$ with $\sigma = (\Sigma_{k,k})_{k=1\dots K}$ known
  (**SB**) $Y^i - \mu \sim \mu - Y^i$ (symmetry) and $\forall k$, $\left|Y_k^i\right| \leq M$ a.s.

Ideal threshold: $t = q_\alpha^\star$, $(1-\alpha)$-quantile of $\mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p\right)$

$q_\alpha^\star$ depends on $\mathcal{D}(\mathbf{Y})$ unknown $\Rightarrow$ estimated by **resampling**.

Goal

For every $p \in [1, +\infty]$, find a threshold $t_\alpha(\mathbf{Y})$ such that

$$\forall \mu \in \mathbb{R}^K \qquad \mathbb{P}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p > t_\alpha\left(\mathbf{Y}\right)\right) \leq \alpha \ .$$

$\Rightarrow$ L^p confidence ball for μ at level α ($\mathrm{FWER}(R) \leq \alpha$ if $p = +\infty$)

- Non-asymptotic: $\forall K, n$
- General correlations
- Assumptions:
  (**Gauss**) $Y^i \sim \mathcal{N}(\mu, \Sigma)$ with $\sigma = (\Sigma_{k,k})_{k=1\dots K}$ known
  (**SB**) $Y^i - \mu \sim \mu - Y^i$ (symmetry) and $\forall k$, $\left|Y^i_k\right| \leq M$ a.s.

Ideal threshold: $t = q^\star_\alpha$, $(1-\alpha)$-quantile of $\mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p\right)$

$q^\star_\alpha$ depends on $\mathcal{D}(\mathbf{Y})$ unknown $\Rightarrow$ estimated by **resampling**.

Classical procedures and known results

- Parametric statistics: no, because $K^2 \gg nK$.
- Asymptotic results (e.g. [van der Vaart and Wellner 1996]): not valid, because $K \gg n$.
- Holm's multiple testing procedure: better than Bonferroni, but unefficient with strong correlations
- (Multiple) tests by symmetrization:
  do not work if $H_{0,k} = \{\mu_k \leq 0\}$.
  do not give confidence regions.
  not translation-invariant $\Rightarrow$ less powerful if $\|\mu\|_\infty$ is large

Classical procedures and known results

- Parametric statistics: no, because $K^2 \gg nK$.
- Asymptotic results (e.g. [van der Vaart and Wellner 1996]): not valid, because $K \gg n$.
- Holm's multiple testing procedure: better than Bonferroni, but unefficient with strong correlations
- (Multiple) tests by symmetrization:
  do not work if $H_{0,k} = \{\mu_k \leq 0\}$.
  do not give confidence regions.
  not translation-invariant $\Rightarrow$ less powerful if $\|\mu\|_\infty$ is large

Classical procedures and known results

- Parametric statistics: no, because $K^2 \gg nK$.
- Asymptotic results (e.g. [van der Vaart and Wellner 1996]): not valid, because $K \gg n$.

- Holm's multiple testing procedure: better than Bonferroni, but unefficient with strong correlations

- (Multiple) tests by symmetrization:
  do not work if $H_{0,k} = \{\mu_k \leq 0\}$.
  do not give confidence regions.
  not translation-invariant $\Rightarrow$ less powerful if $\|\mu\|_\infty$ is large

Resampling principle [Efron 1979 ; ...]

Sample $\quad Y^1, \ldots, Y^n \overset{resampling}{\longrightarrow} (W_1, Y^1), \ldots, (W_n, Y^n) \quad$ weighted sample

- "Y^i is kept W_i times in the resample"
- Weight vector: $(W_1, \ldots, W_n)$, independent of $\mathbf{Y}$
- Example 1: Efron's bootstrap $\Leftrightarrow$ n-sample with replacement $\Leftrightarrow (W_1, \ldots, W_n) \sim \mathcal{M}(n; n^{-1}, \ldots n^{-1})$
- Example 2: Rademacher weights: W_i i.i.d. $\sim \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1 \Leftrightarrow$ subsampling, with subsample size $\approx n/2$

Heuristics: $\mathcal{D}(\text{sample}|\text{true distribution}) \approx \mathcal{D}(\text{resample}|\text{sample})$

Resampling principle [Efron 1979 ; ...]

Sample $Y^1, \ldots, Y^n \overset{resampling}{\longrightarrow} (W_1, Y^1), \ldots, (W_n, Y^n)$ weighted sample

- "Y^i is kept W_i times in the resample"
- Weight vector: $(W_1, \ldots, W_n)$, independent of $\mathbf{Y}$
- Example 1: Efron's bootstrap $\Leftrightarrow$ n-sample with replacement $\Leftrightarrow (W_1, \ldots, W_n) \sim \mathcal{M}(n; n^{-1}, \ldots n^{-1})$
- Example 2: Rademacher weights: W_i i.i.d. $\sim \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1 \Leftrightarrow$ subsampling, with subsample size $\approx n/2$

Heuristics: $\mathcal{D}(\text{sample}|\text{true distribution}) \approx \mathcal{D}(\text{resample}|\text{sample})$

Resampling principle [Efron 1979 ; ...]

Sample $\quad Y^1, \ldots, Y^n \overset{\text{resampling}}{\longrightarrow} (W_1, Y^1), \ldots, (W_n, Y^n) \quad$ weighted sample

- "Y^i is kept W_i times in the resample"
- Weight vector: $(W_1, \ldots, W_n)$, independent of $\mathbf{Y}$
- Example 1: Efron's bootstrap $\Leftrightarrow$ n-sample with replacement $\Leftrightarrow (W_1, \ldots, W_n) \sim \mathcal{M}(n; n^{-1}, \ldots n^{-1})$
- Example 2: Rademacher weights: W_i i.i.d. $\sim \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1 \Leftrightarrow$ subsampling, with subsample size $\approx n/2$

Heuristics: $\mathcal{D}(\text{sample}|\text{true distribution}) \approx \mathcal{D}(\text{resample}|\text{sample})$

Resampling principle [Efron 1979 ; ...]

Sample $\quad Y^1, \ldots, Y^n \stackrel{\textit{resampling}}{\longrightarrow} (W_1, Y^1), \ldots, (W_n, Y^n) \quad$ weighted sample

- "Y^i is kept W_i times in the resample"
- Weight vector: $(W_1, \ldots, W_n)$, independent of $\mathbf{Y}$
- Example 1: Efron's bootstrap $\Leftrightarrow$ n-sample with replacement $\Leftrightarrow (W_1, \ldots, W_n) \sim \mathcal{M}(n; n^{-1}, \ldots n^{-1})$
- Example 2: Rademacher weights: W_i i.i.d. $\sim \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1 \Leftrightarrow$ subsampling, with subsample size $\approx n/2$

Heuristics: $\mathcal{D}(\text{sample}|\text{true distribution}) \approx \mathcal{D}\,(\text{resample}|\text{sample})$

Concentration method

- $\|\overline{\mathbf{Y}} - \mu\|_p$ concentrates around its expectation, standard-deviation $\leq \|\sigma\|_p n^{-1/2}$
- Estimate $\mathbb{E}\left[\|\overline{\mathbf{Y}} - \mu\|_p\right]$ by resampling

$$\Rightarrow q_\alpha^{\text{conc}}(\mathbf{Y}) = \text{cst} \times \sqrt{n}\mathbb{E}\left[\|\overline{\mathbf{Y}}_W - \overline{W}\,\overline{\mathbf{Y}}\|_p | \mathbf{Y}\right] + \text{remainder}(\|\sigma\|_p, \alpha, n)$$

Works well if expectations ($\propto \sqrt{\log(K)}$) are larger than fluctuations ($\propto \overline{\Phi}^{-1}(\alpha/2)$)

Concentration method

- $\|\overline{\mathbf{Y}} - \mu\|_p$ concentrates around its expectation, standard-deviation $\leq \|\sigma\|_p n^{-1/2}$
- Estimate $\mathbb{E}\left[\|\overline{\mathbf{Y}} - \mu\|_p\right]$ by resampling

$$\Rightarrow q_\alpha^{\text{conc}}(\mathbf{Y}) = \text{cst} \times \sqrt{n}\mathbb{E}\left[\|\overline{\mathbf{Y}}_W - \overline{W}\,\overline{\mathbf{Y}}\|_p | \mathbf{Y}\right] + \text{remainder}(\|\sigma\|_p, \alpha, n)$$

Works well if expectations ($\propto \sqrt{\log(K)}$) are larger than fluctuations ($\propto \overline{\Phi}^{-1}(\alpha/2)$)

Quantile method

Ideal threshold: $q_\alpha^\star = (1-\alpha)\text{-quantile of } \mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p\right)$

$\Rightarrow$ Resampling estimate of $q_\alpha^\star$:
$q_\alpha^{\text{quant}}(\mathbf{Y}) = (1-\alpha)\text{-quantile of } \mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}}_W - \overline{W}\,\overline{\mathbf{Y}}\|_p \middle| \mathbf{Y}\right)$

$$\text{with} \qquad \overline{W} := \frac{1}{n}\sum_{i=1}^{n} W_i$$

$$\overline{\mathbf{Y}}_W := \frac{1}{n}\sum_{i=1}^{n} W_i Y^i \qquad \text{Resampling empirical mean}$$

$q_\alpha^{\text{quant}}(\mathbf{Y})$ depends only on $\mathbf{Y}$ $\Rightarrow$ can be computed with real data

Quantile method

Ideal threshold: $q_\alpha^\star = (1-\alpha)$-quantile of $\mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p\right)$

$\Rightarrow$ Resampling estimate of $q_\alpha^\star$:
$q_\alpha^{\text{quant}}(\mathbf{Y}) = (1-\alpha)$-quantile of $\mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}}_W - \overline{W}\,\overline{\mathbf{Y}}\|_p | \mathbf{Y}\right)$

$$\text{with} \qquad \overline{W} := \frac{1}{n}\sum_{i=1}^n W_i$$

$$\overline{\mathbf{Y}}_W := \frac{1}{n}\sum_{i=1}^n W_i Y^i \qquad \text{Resampling empirical mean}$$

$q_\alpha^{\text{quant}}(\mathbf{Y})$ depends only on $\mathbf{Y}$ $\Rightarrow$ can be computed with real data

Quantile method

Ideal threshold: $q_\alpha^\star = (1-\alpha)$-quantile of $\mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p\right)$

$\Rightarrow$ Resampling estimate of $q_\alpha^\star$:
$q_\alpha^{\text{quant}}(\mathbf{Y}) = (1-\alpha)$-quantile of $\mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}}_W - \overline{W}\,\overline{\mathbf{Y}}\|_p | \mathbf{Y}\right)$

$$\text{with} \qquad \overline{W} := \frac{1}{n}\sum_{i=1}^n W_i$$

$$\overline{\mathbf{Y}}_W := \frac{1}{n}\sum_{i=1}^n W_i Y^i \qquad \text{Resampling empirical mean}$$

$q_\alpha^{\text{quant}}(\mathbf{Y})$ depends only on $\mathbf{Y}$ $\Rightarrow$ can be computed with real data

Concentration theorem

Theorem

Assume (**Gauss**) *and* W *Rademacher. For every* $\alpha \in (0,1)$ *,*

$$q_\alpha^{conc,1}(\mathbf{Y}) := \frac{\sqrt{n}\mathbb{E}\left[\|\overline{\mathbf{Y}}_W - \overline{W}\,\overline{\mathbf{Y}}\|_p|\mathbf{Y}\right]}{B_W} + \|\sigma\|_p \overline{\Phi}^{-1}(\alpha/2)\left[\frac{C_W}{\sqrt{n}B_W} + 1\right]$$

satisfies

$$\mathbb{P}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p > q_\alpha^{conc,1}(\mathbf{Y})\right) \leq \alpha$$

with $\sigma := \left(\sqrt{\operatorname{var}(Y_k^1)}\right)_{1 \leq k \leq K}$, and

$$B_W := \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2\right)^{1/2} = 1 - \mathcal{O}(n^{-1/2}) \quad \text{and} \quad C_W = 1$$

Main tool: Gaussian concentration theorem [Cirel'son, Ibragimov and Sudakov, 1976]

Remarks

- Valid for quite general weights (with B_W and C_W are independent of K and easy to compute).
- Similar result under assumption (**SB**) (with larger constants).
- $\|\cdot\|_p$ can be replaced by $\sup_k (\cdot)_+ \Rightarrow$ one-sided multiple tests
- Almost deterministic threshold:
  $\Rightarrow$ if $p = \infty$, $q_\alpha^{\text{conc},2}(\mathbf{Y}) \approx \min\left(t_\alpha^{\text{bonf}}, q_\alpha^{\text{conc},1}(\mathbf{Y})\right)$ still has a FWER $\leq \alpha$.

Remarks

- Valid for quite general weights (with B_W and C_W are independent of K and easy to compute).
- Similar result under assumption (**SB**) (with larger constants).
- $\|\cdot\|_p$ can be replaced by $\sup_k (\cdot)_+ \Rightarrow$ one-sided multiple tests
- Almost deterministic threshold:
  $\Rightarrow$ if $p = \infty$, $q_\alpha^{\mathrm{conc},2}(\mathbf{Y}) \approx \min\left(t_\alpha^{\mathrm{bonf}}, q_\alpha^{\mathrm{conc},1}(\mathbf{Y})\right)$ still has a FWER $\leq \alpha$.

Practical computation

- Monte-Carlo approximation with $W^1, \ldots, W^B$ only (with an additionnal term of order $B^{-1/2}$)
- Alternatively, V-fold cross-validation weights
  $\Rightarrow$ computation time $\propto V$, accuracy $\propto C_W B_W^{-1} \approx \sqrt{n/V}$.

- Estimation of σ : under (**Gauss**), if

$$\widehat{\sigma_k} := \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(Y_k^i - \overline{\mathbf{Y}} \right)^2}$$

then for every $\delta \in (0,1)$,

$$\mathbb{P}\left(\|\sigma\|_p \leq \left(C_n - \frac{1}{\sqrt{n}} \overline{\Phi}^{-1}\left(\frac{\delta}{2}\right)\right) \|\widehat{\sigma}\|_p \right) \geq 1 - \delta \quad ,$$

with $C_n = 1 - \mathcal{O}(n^{-1})$.

Practical computation

- Monte-Carlo approximation with $W^1, \ldots, W^B$ only (with an additionnal term of order $B^{-1/2}$)
- Alternatively, V-fold cross-validation weights
  $\Rightarrow$ computation time $\propto V$, accuracy $\propto C_W B_W^{-1} \approx \sqrt{n/V}$.
- Estimation of σ : under (**Gauss**), if

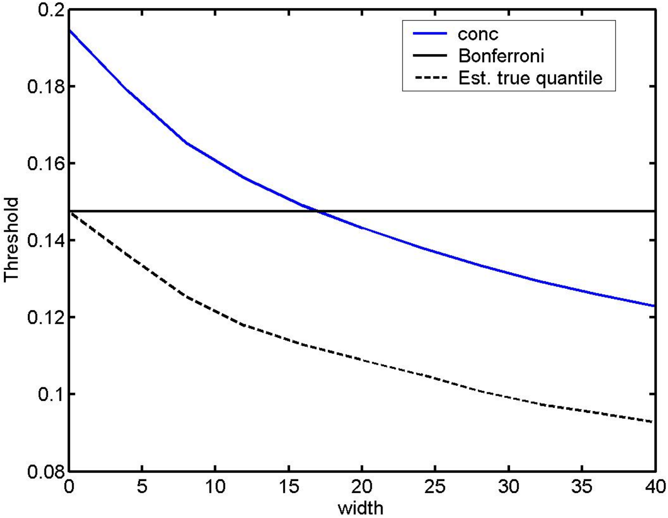
$$\widehat{\sigma_k} := \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(Y_k^i - \overline{\mathbf{Y}} \right)^2}$$

then for every $\delta \in (0,1)$,

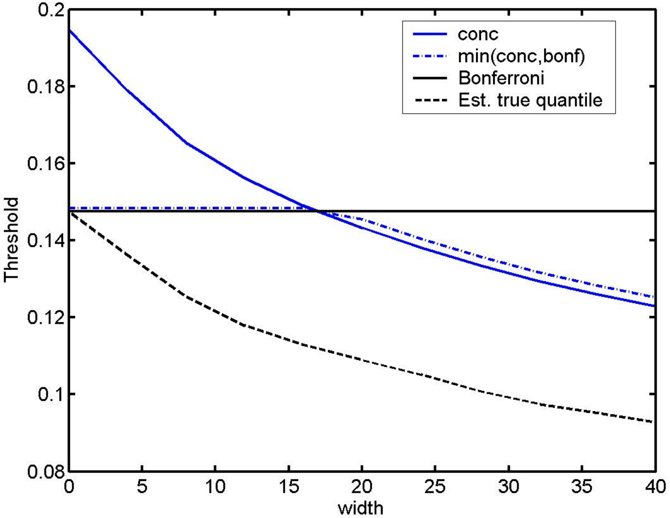
$$\mathbb{P} \left(\|\sigma\|_p \leq \left(C_n - \frac{1}{\sqrt{n}} \overline{\Phi}^{-1} \left(\frac{\delta}{2} \right) \right) \|\widehat{\sigma}\|_p \right) \geq 1 - \delta \quad ,$$

with $C_n = 1 - \mathcal{O}(n^{-1})$.

Simulations: $p = +\infty$, $n = 1000$, $K = 16384$, $\sigma = 1$



Simulations: $p = +\infty$, $n = 1000$, $K = 16384$, $\sigma = 1$



Quantile method

- Rademacher weights only: W_i i.i.d. $\sim \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$
- $q_\alpha^{\text{quant}}(\mathbf{Y}) = (1-\alpha)$-quantile of

$$\mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}}_W - \overline{W}\,\overline{\mathbf{Y}}\|_p \middle| \mathbf{Y}\right)$$

Heuristics $\Rightarrow$ should satisfy $\mathbb{P}\left(\|\overline{\mathbf{Y}} - \mu\|_p > q_\alpha^{\text{quant}}(\mathbf{Y})\right) \leq \alpha$

Quantile theorem: need for a supplementary term

Theorem

Y symmetric. W Rademacher. $\alpha, \delta, \gamma \in (0,1)$.
If f is a non-negative threshold with level bounded by $\alpha\gamma$:

$$\mathbb{P}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p > f(\mathbf{Y})\right) \leq \alpha\gamma$$

Then,

$$q_{\alpha}^{quant+f}(\mathbf{Y}) = q_{\alpha(1-\delta)(1-\gamma)}^{quant}(\mathbf{Y}) + \sqrt{\frac{2\log(2/(\delta\alpha))}{n}} f(\mathbf{Y})$$

yields a level bounded by α :

$$\mathbb{P}\left(\sqrt{n}\|\overline{\mathbf{Y}} - \mu\|_p > q_{\alpha}^{quant+f}(\mathbf{Y})\right) \leq \alpha$$

Remarks

- Uses only the symmetry of Y around its mean
- $\|\cdot\|_p$ can be replaced by $\sup_k(\cdot)_+ \Rightarrow$ one-sided multiple tests
- The supplementary threshold f only appears in a second-order term.

  (**Gauss**) $\Rightarrow$ three thresholds:
  take f among $q_{\alpha\gamma}^{\mathrm{Bonf}}$ (if $p = +\infty$), $q_{\alpha\gamma}^{\mathrm{conc},1}$ and $q_{\alpha\gamma}^{\mathrm{conc},2}$

  (**SB**) $\Rightarrow f = q_{\alpha\gamma}^{\mathrm{conc,SB}}$
- In simulation experiments, f is almost unnecessary.
- $q_{\alpha(1-\delta)(1-\gamma)}^{\mathrm{quant}}(\mathbf{Y})$ can be replaced by a Monte-Carlo estimated quantile (i.e., simulate only $W^1, \ldots, W^B$)
  $\Rightarrow$ loose at most $(B+1)^{-1}$ in the level, nothing if $\alpha(1-\delta)(1-\gamma)(B+1) \in \mathbb{N}$.

Remarks

- Uses only the symmetry of Y around its mean
- $\|\cdot\|_p$ can be replaced by $\sup_k(\cdot)_+ \Rightarrow$ one-sided multiple tests
- The supplementary threshold f only appears in a second-order term.

  (**Gauss**) $\Rightarrow$ three thresholds:
  take f among $q_{\alpha\gamma}^{\mathrm{Bonf}}$ (if $p = +\infty$), $q_{\alpha\gamma}^{\mathrm{conc},1}$ and $q_{\alpha\gamma}^{\mathrm{conc},2}$

  (**SB**) $\Rightarrow f = q_{\alpha\gamma}^{\mathrm{conc,SB}}$
- In simulation experiments, f is almost unnecessary.
- $q_{\alpha(1-\delta)(1-\gamma)}^{\mathrm{quant}}(\mathbf{Y})$ can be replaced by a Monte-Carlo estimated quantile (i.e., simulate only $W^1, \ldots, W^B$)
  $\Rightarrow$ loose at most $(B+1)^{-1}$ in the level, nothing if $\alpha(1-\delta)(1-\gamma)(B+1) \in \mathbb{N}$.

Remarks

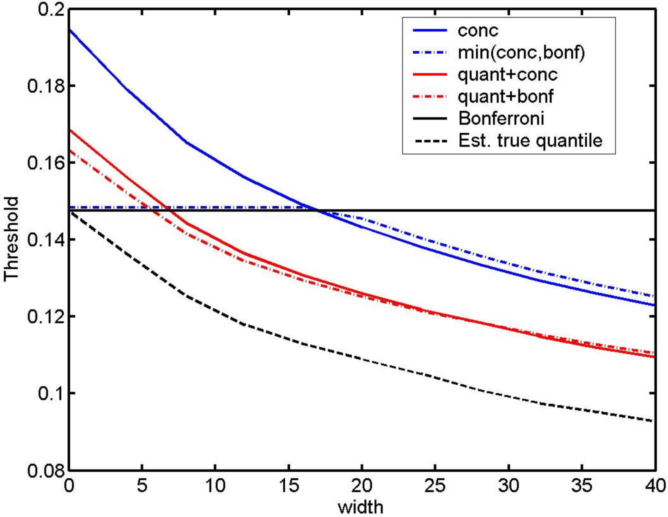
- Uses only the symmetry of Y around its mean
- $\lVert\cdot\rVert_p$ can be replaced by $\sup_k(\cdot)_+ \Rightarrow$ one-sided multiple tests
- The supplementary threshold f only appears in a second-order term.

  (**Gauss**) $\Rightarrow$ three thresholds:
  take f among $q_{\alpha\gamma}^{\mathrm{Bonf}}$ (if $p = +\infty$), $q_{\alpha\gamma}^{\mathrm{conc},1}$ and $q_{\alpha\gamma}^{\mathrm{conc},2}$

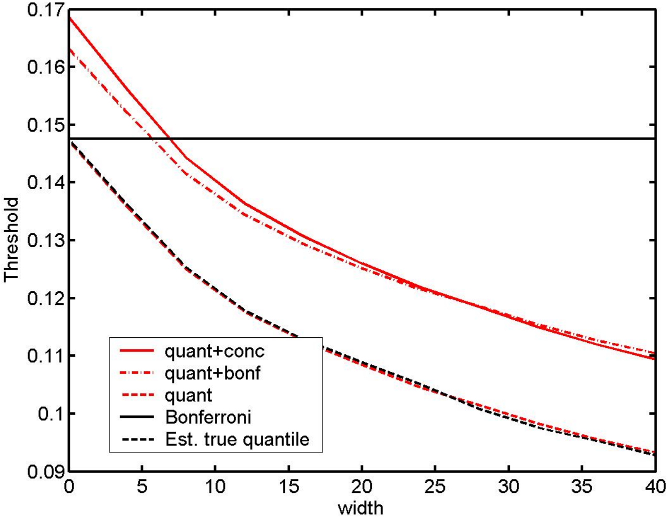
  (**SB**) $\Rightarrow f = q_{\alpha\gamma}^{\mathrm{conc,SB}}$

- In simulation experiments, f is almost unnecessary.
- $q_{\alpha(1-\delta)(1-\gamma)}^{\mathrm{quant}}(\mathbf{Y})$ can be replaced by a Monte-Carlo estimated quantile (i.e., simulate only $W^1, \ldots, W^B$)
  $\Rightarrow$ loose at most $(B+1)^{-1}$ in the level, nothing if $\alpha(1-\delta)(1-\gamma)(B+1) \in \mathbb{N}$.

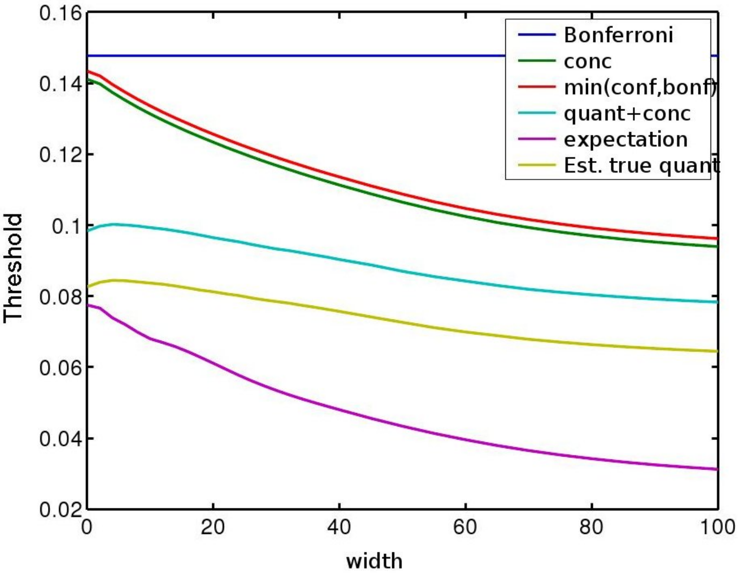
Simulations: $n = 1000$, $K = 16384$, $\sigma = 1$, $p = +\infty$



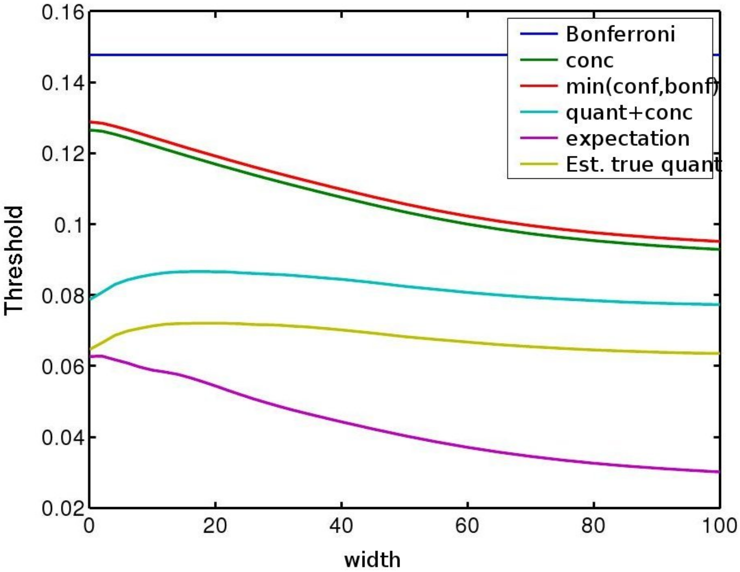
Simulations: without the additive term?



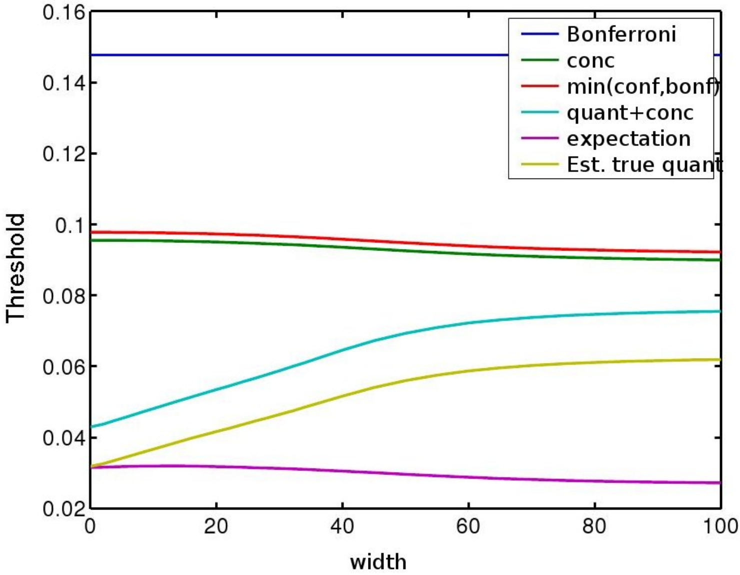
Simulations: role of p ($n = 1000$, $p = 16$)



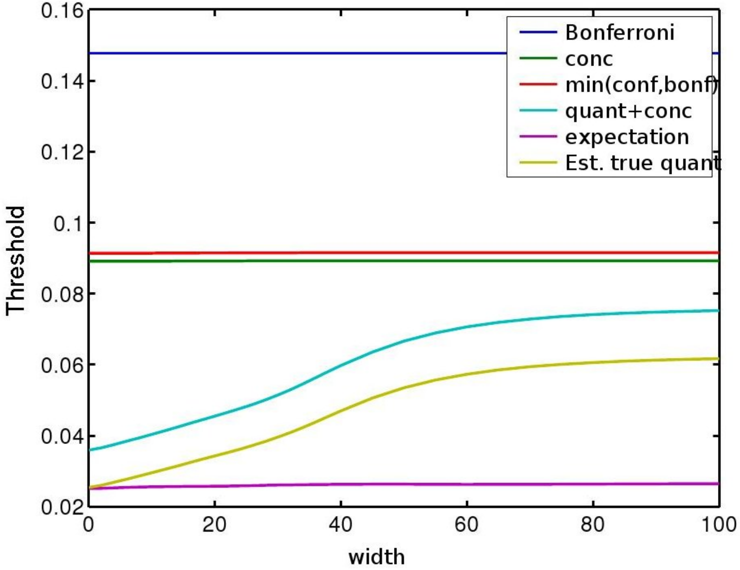
Simulations: role of p ($n = 1000$, $p = 10$)



Simulations: role of p ($n = 1000$, $p = 2$)



Simulations: role of p ($n = 1000$, $p = 1$)



Quantile approach vs. Symmetrization

- Define $q_\alpha^{\text{sym}}(\mathbf{Y}, p)$ as the $(1-\alpha)$-quantile of

$$\mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}}_W\|_p \middle| \mathbf{Y}\right)$$

- Symmetrization argument: if $\mathbf{Y}$ symmetric and W Rademacher, then

$$\mathbb{P}\left(\|\overline{\mathbf{Y}} - \mu\|_p > q_\alpha^{\text{sym}}(\mathbf{Y} - \mu, p)\right) \leq \alpha$$

since

$$\overline{(\mathbf{Y} - \mu)}_W = \frac{1}{n}\sum_{i=1}^n W_i(Y^i - \mu) \sim \frac{1}{n}\sum_{i=1}^n (Y^i - \mu) = \overline{\mathbf{Y}} - \mu \quad .$$

- $q_\alpha^{\text{sym}}(\mathbf{Y} - \mu, p)$ unknown $\Rightarrow$ replacing μ by $\overline{\mathbf{Y}}$ leads to

$$q_\alpha^{\text{quant}}(\mathbf{Y}, p) = q_\alpha^{\text{sym}}(\mathbf{Y} - \overline{\mathbf{Y}}, p)$$

Quantile approach vs. Symmetrization

- Define $q_\alpha^{\mathrm{sym}}(\mathbf{Y}, p)$ as the $(1-\alpha)$-quantile of

$$\mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}}_W\|_p \middle| \mathbf{Y}\right)$$

- Symmetrization argument: if $\mathbf{Y}$ symmetric and W Rademacher, then

$$\mathbb{P}\left(\|\overline{\mathbf{Y}} - \mu\|_p > q_\alpha^{\mathrm{sym}}(\mathbf{Y} - \mu, p)\right) \leq \alpha$$

since

$$\overline{(\mathbf{Y} - \mu)}_W = \frac{1}{n}\sum_{i=1}^{n} W_i (Y^i - \mu) \sim \frac{1}{n}\sum_{i=1}^{n} (Y^i - \mu) = \overline{\mathbf{Y}} - \mu \quad .$$

- $q_\alpha^{\mathrm{sym}}(\mathbf{Y} - \mu, p)$ unknown $\Rightarrow$ replacing μ by $\overline{\mathbf{Y}}$ leads to

$$q_\alpha^{\mathrm{quant}}(\mathbf{Y}, p) = q_\alpha^{\mathrm{sym}}(\mathbf{Y} - \overline{\mathbf{Y}}, p)$$

Quantile approach vs. Symmetrization

- Define $q_\alpha^{\mathrm{sym}}(\mathbf{Y}, p)$ as the $(1-\alpha)$-quantile of

$$\mathcal{D}\left(\sqrt{n}\|\overline{\mathbf{Y}}_W\|_p \middle| \mathbf{Y}\right)$$

- Symmetrization argument: if $\mathbf{Y}$ symmetric and W Rademacher, then

$$\mathbb{P}\left(\|\overline{\mathbf{Y}} - \mu\|_p > q_\alpha^{\mathrm{sym}}(\mathbf{Y} - \mu, p)\right) \le \alpha$$

since

$$\overline{(\mathbf{Y} - \mu)}_W = \frac{1}{n}\sum_{i=1}^{n} W_i (Y^i - \mu) \sim \frac{1}{n}\sum_{i=1}^{n}(Y^i - \mu) = \overline{\mathbf{Y}} - \mu \quad .$$

- $q_\alpha^{\mathrm{sym}}(\mathbf{Y} - \mu, p)$ unknown $\Rightarrow$ replacing μ by $\overline{\mathbf{Y}}$ leads to

$$q_\alpha^{\mathrm{quant}}(\mathbf{Y}, p) = q_\alpha^{\mathrm{sym}}(\mathbf{Y} - \overline{\mathbf{Y}}, p)$$

Multiple testing with the uncentered quantile threshold

If $t \geq q_\alpha^{\mathrm{sym}}((\mathbf{Y}_k)_{k \text{ s.t. } \mu_k=0}, +\infty)$, then

$$\mathrm{FWER}(R) = \mathbb{P}(\exists k \quad \text{s.t.} \quad \mu_k = 0 \text{ and } \sqrt{n}|\overline{\mathbf{Y}}_k| > t)$$
$$= \mathbb{P}(\sqrt{n} \sup_{k \text{ s.t. } \mu_k=0} |\overline{\mathbf{Y}}_k| > t) \leq \alpha$$

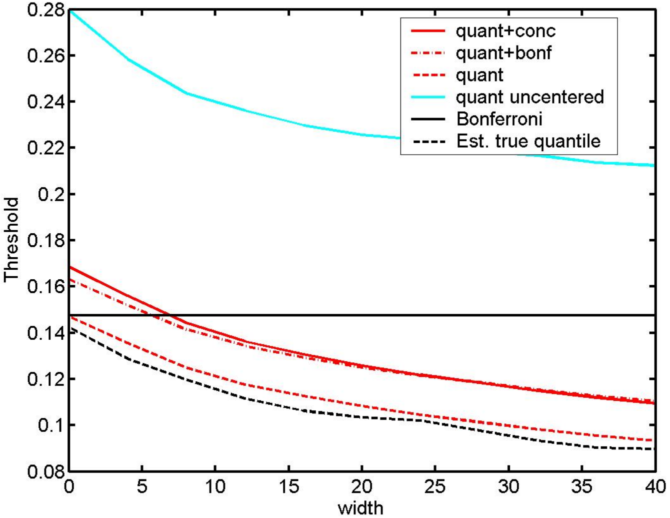
by symmetry of $\mathbf{Y}$.

This holds in particular for

$$q_\alpha^{\text{quant. uncent.}}(\mathbf{Y}) := q_\alpha^{\mathrm{sym}}(\mathbf{Y}, +\infty) \geq q_\alpha^{\mathrm{sym}}((\mathbf{Y})_{k \text{ s.t. } \mu_k=0}, +\infty)$$

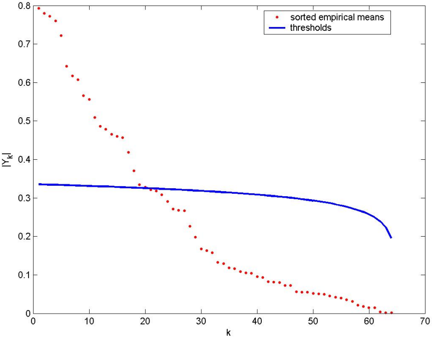
$\Rightarrow$ can be used for multiple testing, but more conservative, especially when the signal μ is strong

Simulations: $n = 1000$, $K = 16384$, $\sigma = 1$, $0 \leq \mu_k \leq 2.9$



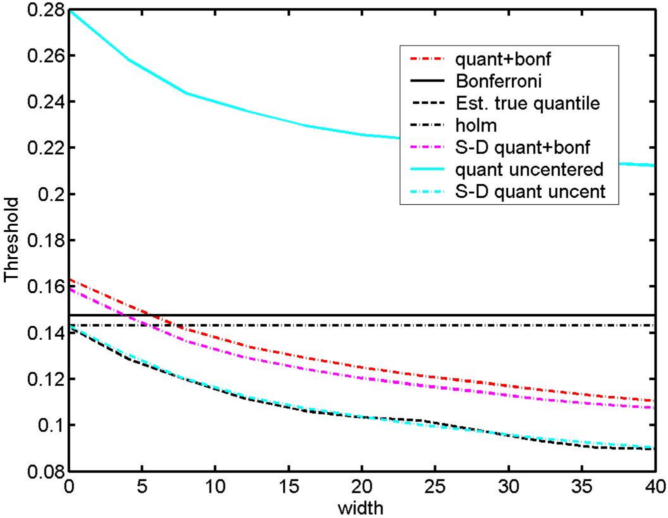
Step-down procedure [Holm 1979 ; Westfall and Young 1993 ; Romano and Wolf 2005]

- Reorder the coordinates: $\left|\overline{\mathbf{Y}}_{\sigma(1)}\right| \geq \cdots \geq \left|\overline{\mathbf{Y}}_{\sigma(K)}\right|$
- Define the thresholds $t_k = t(\mathbf{Y}_{\sigma(k)}, \ldots, \mathbf{Y}_{\sigma(K)})$ for $k = 1, \ldots, K$
- Define $\widehat{k} = \max\left\{k \text{ s.t. } \forall k' \leq k, \overline{\mathbf{Y}}_{\sigma(k')} > t_{k'}(\mathbf{Y})\right\}$
- Reject $H_{0,k}$ for all $k \leq \widehat{k}$

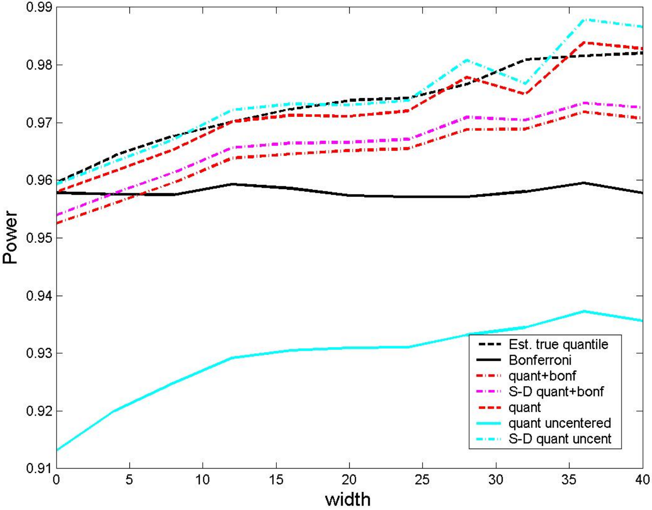


$\Rightarrow$ this procedure has a FWER controlled by α if each t_k has (use that $t_{\mathcal{K}} = t((\mathbf{Y}_k)_{k \in \mathcal{K}})$ is a non-decreasing function of $\mathcal{K}$).

Simulations: $n = 1000$, $K = 16384$, $\sigma = 1$, $0 \leq \mu_k \leq 2.9$



Simulations: power ($0 \leq \mu_k \leq 2.9$)



Hybrid approach, computational cost

- Idea: (centered) quantile better when the signal is strong, uncentered quantile better for weak signals
- First step: use $q_{\alpha}^{\text{quant+Bonf}}(\mathbf{Y})$ to reject a first set of hypotheses.
- Second step: apply the step-down procedure associated with $q_{\alpha(1-\gamma)}^{\text{quant. uncent.}}(\mathbf{Y})$ to the remaining hypotheses.
- Goal: reduce the computational cost / increase the power for a given number of iterations

Hybrid approach, computational cost

- Idea: (centered) quantile better when the signal is strong, uncentered quantile better for weak signals
- First step: use $q_{\alpha}^{\text{quant+Bonf}}(\mathbf{Y})$ to reject a first set of hypotheses.
- Second step: apply the step-down procedure associated with $q_{\alpha(1-\gamma)}^{\text{quant. uncent.}}(\mathbf{Y})$ to the remaining hypotheses.
- Goal: reduce the computational cost / increase the power for a given number of iterations

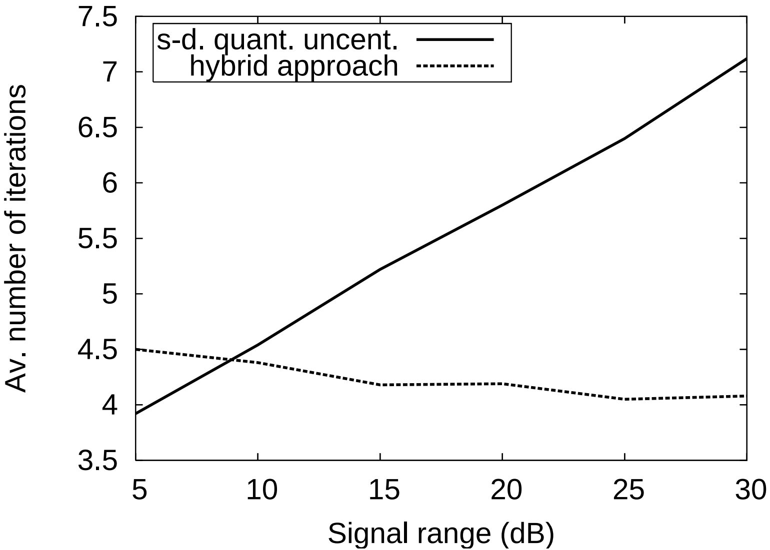
Hybrid approach, computational cost

- Idea: (centered) quantile better when the signal is strong, uncentered quantile better for weak signals

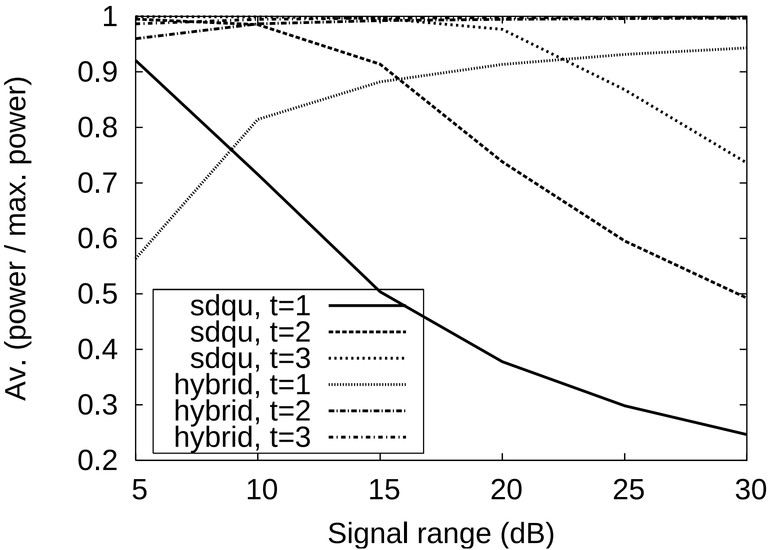
- First step: use $q_{\alpha}^{\text{quant+Bonf}}(\mathbf{Y})$ to reject a first set of hypotheses.
- Second step: apply the step-down procedure associated with $q_{\alpha(1-\gamma)}^{\text{quant. uncent.}}(\mathbf{Y})$ to the remaining hypotheses.

- Goal: reduce the computational cost / increase the power for a given number of iterations

Hybrid approach: average number of iterations



With early stopping: power vs. number of iterations



Conclusion

Two resampling-based procedures:

- concentration (almost deterministic threshold)
- quantile (related to symmetrization techniques)

$\Rightarrow$ Multiple testing + Confidence regions

- FWER / level control (non-asymptotic, K may be $\gg n$)
- very general correlation structures allowed
- Simulations: efficient in presence of correlations
- step-down procedures are possible

- Open problems:
  quantile threshold without f?
  with other weights?
  with non-symmetric $\mathbf{Y}$?

Conclusion

Two resampling-based procedures:

- concentration (almost deterministic threshold)
- quantile (related to symmetrization techniques)

$\Rightarrow$ Multiple testing + Confidence regions

- FWER / level control (non-asymptotic, K may be $\gg n$)
- very general correlation structures allowed
- Simulations: efficient in presence of correlations
- step-down procedures are possible

- Open problems:
  quantile threshold without f?
  with other weights?
  with non-symmetric $\mathbf{Y}$?

Conclusion

Two resampling-based procedures:

- concentration (almost deterministic threshold)
- quantile (related to symmetrization techniques)

$\Rightarrow$ Multiple testing + Confidence regions

- FWER / level control (non-asymptotic, K may be $\gg n$)
- very general correlation structures allowed
- Simulations: efficient in presence of correlations
- step-down procedures are possible

- Open problems:
  quantile threshold without f?
  with other weights?
  with non-symmetric $\mathbf{Y}$?