New FDP controlling procedures under dependence

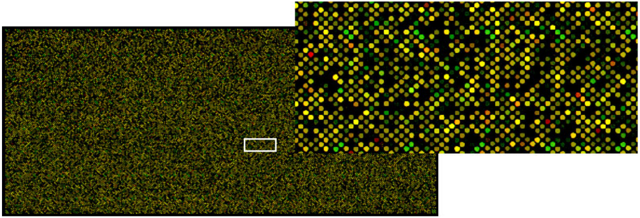
Etienne Roquain[1]
Joint work with S. Delattre[2]

[1]Laboratory LPMA, Université Pierre et Marie Curie (Paris 6), France
[2]Laboratory LPMA, Université Paris Diderot (Paris 7), France

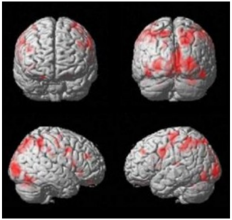
MCP 2017, 21-th June 2017, Riverside

Massive and complex data

- Microarray
  interesting genes?



- Neuroimaging (FMRI)
  activated regions?



- Astronomy
  directions with stars?

Adaptative procedures

- Model: $P = (H, \Gamma)$ with $H \in \{0,1\}^m$ and Γ some "parameter"
- Type I error rates: Crit = FWER, FDR, $\mathbf{P}(FDP > \gamma)$

- Build classical procedure R:

$$\sup_P\{\text{Crit}(R,P)\} \leq \text{ and } \approx \alpha$$

- Build adaptive procedure $R = R(\hat{\Gamma})$:

$$\forall P, \;\; \text{Crit}(R,P) \leq \text{ and } \approx \alpha$$

  - ⋆ $\Gamma =$ proportion of signal: [Holm (1979)], [Storey (2002)], [Benjamini et al (2006)], [Sarkar (2008)], [Blanchard and R. (2009)], [Finner et al (2009)], [Li and Barber (2017)], ...
  - ⋆ $\Gamma =$ effect sizes of signal: [Wasserman and Roeder (2006)], [R. and van de Wiel (2009)],...
  - ⋆ $\Gamma =$ discreteness of the tests: [Tarone (1990)], [Heyse (2011)], [Heller (2012)], ...
  - ⋆ $\Gamma =$ dependence

Adaptative procedures

- Model: $P = (H, \Gamma)$ with $H \in \{0,1\}^m$ and Γ some “parameter"
- Type I error rates: Crit = FWER, FDR, $\mathbf{P}(FDP > \gamma)$
- Build classical procedure R:

$$\sup_P \{\text{Crit}(R, P)\} \leq \text{ and } \approx \alpha$$

- Build adaptive procedure $R = R(\hat{\Gamma})$:

$$\forall P, \ \text{Crit}(R, P) \leq \text{ and } \approx \alpha$$

- ⋆ Γ = proportion of signal: [Holm (1979)], [Storey (2002)], [Benjamini et al (2006)], [Sarkar (2008)], [Blanchard and R. (2009)], [Finner et al (2009)], [Li and Barber (2017)], ...
- ⋆ Γ = effect sizes of signal: [Wasserman and Roeder (2006)], [R. and van de Wiel (2009)],...
- ⋆ Γ = discreteness of the tests: [Tarone (1990)], [Heyse (2011)], [Heller (2012)], ...
- ⋆ Γ = dependence

Adaptative procedures

- Model: $P = (H, \Gamma)$ with $H \in \{0,1\}^m$ and Γ some “parameter"
- Type I error rates: Crit = FWER, FDR, $\mathbf{P}(FDP > \gamma)$

- Build classical procedure R:

$$\sup_P\{\text{Crit}(R,P)\} \leq \text{ and } \approx \alpha$$

- Build adaptive procedure $R = R(\hat{\Gamma})$:

$$\forall P, \;\; \text{Crit}(R,P) \leq \text{ and } \approx \alpha$$

⋆ $\Gamma =$ proportion of signal: [Holm (1979)], [Storey (2002)], [Benjamini et al (2006)], [Sarkar (2008)], [Blanchard and R. (2009)], [Finner et al (2009)], [Li and Barber (2017)], ...

⋆ $\Gamma =$ effect sizes of signal: [Wasserman and Roeder (2006)], [R. and van de Wiel (2009)],...

⋆ $\Gamma =$ discreteness of the tests: [Tarone (1990)], [Heyse (2011)], [Heller (2012)], ...

⋆ $\Gamma =$ dependence

Adaptative procedures

- Model: $P = (H, \Gamma)$ with $H \in \{0,1\}^m$ and Γ some "parameter"
- Type I error rates: Crit = FWER, FDR, $\mathbf{P}(FDP > \gamma)$

- Build classical procedure R:

$$\sup_P \{\text{Crit}(R, P)\} \leq \text{ and } \approx \alpha$$

- Build adaptive procedure $R = R(\hat{\Gamma})$:

$$\forall P, \ \ \text{Crit}(R, P) \leq \text{ and } \approx \alpha$$

- ⋆ $\Gamma =$ proportion of signal: [Holm (1979)], [Storey (2002)], [Benjamini et al (2006)], [Sarkar (2008)], [Blanchard and R. (2009)], [Finner et al (2009)], [Li and Barber (2017)], ...
- ⋆ $\Gamma =$ effect sizes of signal: [Wasserman and Roeder (2006)], [R. and van de Wiel (2009)],...
- ⋆ $\Gamma =$ discreteness of the tests: [Tarone (1990)], [Heyse (2011)], [Heller (2012)], ...
- ⋆ $\Gamma =$ dependence

Dependence and multiple testing

- non-adaptive: survive to dependence
- adaptive: learn dependence

- FWER:
  - non-adaptive: [Bonferroni (1935)], [Benditkis et al (2015)], ...
  - adaptive: [Westfall and Young (1993)], [Romano and Wolf (2005)], ...
- FDR:
  - non-adaptive: [Benjamini and Yekutieli (2001)], [Sarkar (2008)], [Blanchard and R. (2008)],...
  - adaptive: [Romano et al (2008)], [Barber and Candès (2015)],...
- FDP:
  - non-adaptive: [Lehmann and Romano (2005)], [Romano and Shaikh (2006)]...
  - adaptive: [Korn et al (2004)], [Romano and Wolf (2007)], [Guo et al (2014)], [Hemerik and Goeman (2017)]...

Delattre and R. (2015). New procedures controlling the FDP via Romano-Wolf's heuristic. AoS.

Dependence and multiple testing

- non-adaptive: survive to dependence
- adaptive: learn dependence

- FWER:
  - non-adaptive: [Bonferroni (1935)], [Benditkis et al (2015)], ...
  - adaptive: [Westfall and Young (1993)], [Romano and Wolf (2005)], ...
- FDR:
  - non-adaptive: [Benjamini and Yekutieli (2001)], [Sarkar (2008)], [Blanchard and R. (2008)],...
  - adaptive: [Romano et al (2008)], [Barber and Candès (2015)],...
- FDP:
  - non-adaptive: [Lehmann and Romano (2005)], [Romano and Shaikh (2006)]...
  - adaptive: [Korn et al (2004)], [Romano and Wolf (2007)], [Guo et al (2014)], [Hemerik and Goeman (2017)]...

Delattre and R. (2015). New procedures controlling the FDP via Romano-Wolf's heuristic. AoS.

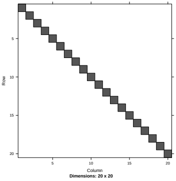
Setting

We observe

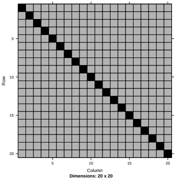
$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix} + \mathcal{N}(0, \Gamma)$$

- Mean : some $\mu_i = 0$, some $\mu_i \neq 0$ ($\mu_i > 0$)
- Noise : various covariance Γ here known !

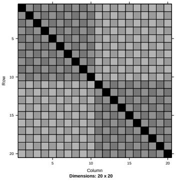
Indep.



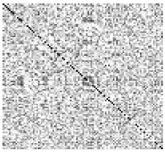
Equicorrelated



3-factor



Real Life



FDP approach

Procedure:

Step-up or step-down procedure with some critical values τ_ℓ, $\ell = 1, \ldots, m$.
Crossing point $\hat\ell$, thresholding at $\tau_{\hat\ell}$

$$\mathrm{FDP} = \frac{\#\text{ false rejections}}{1 \vee \#\text{ rejections}} = \frac{V(\tau_{\hat\ell})}{1 \vee R(\tau_{\hat\ell})}$$

Aim

- Control the FDP : derive critical values τ_ℓ, $1 \leq \ell \leq m$, s.t.

$$\mathbf{P}(\mathrm{FDP}(\tau_{\hat\ell}) > \alpha) \leq \zeta \text{ (and } \approx \zeta) \quad \text{for all } (\mu, \Gamma)$$

- Choose $\tau_\ell = \tau_\ell(\Gamma)$, here Γ known

FDP approach

Procedure:

Step-up or step-down procedure with some critical values τ_ℓ, $\ell = 1, \dots, m$.
Crossing point $\hat{\ell}$, thresholding at $\tau_{\hat{\ell}}$

$$\text{FDP} = \frac{\#\text{ false rejections}}{1 \vee \#\text{ rejections}} = \frac{V(\tau_{\hat{\ell}})}{1 \vee R(\tau_{\hat{\ell}})}$$

Aim

- Control the FDP : derive critical values τ_ℓ, $1 \leq \ell \leq m$, s.t.

$$\mathbf{P}(\text{FDP}(\tau_{\hat{\ell}}) > \alpha) \leq \zeta \text{ (and } \approx \zeta) \quad \text{for all } (\mu, \Gamma)$$

- Choose $\tau_\ell = \tau_\ell(\Gamma)$, here Γ known

FDP approach

Procedure:

Step-up or step-down procedure with some critical values τ_ℓ, $\ell = 1, \ldots, m$.
Crossing point $\hat{\ell}$, thresholding at $\tau_{\hat{\ell}}$

$$\mathrm{FDP} = \frac{\#\text{ false rejections}}{1 \vee \#\text{ rejections}} = \frac{V(\tau_{\hat{\ell}})}{1 \vee R(\tau_{\hat{\ell}})}$$

Aim

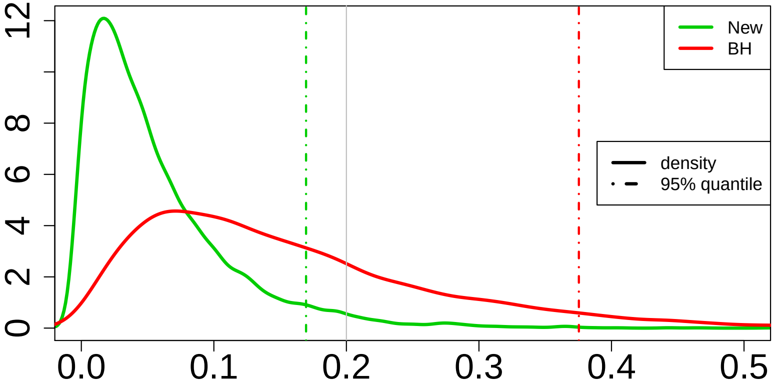
- Control the FDP : derive critical values τ_ℓ, $1 \leq \ell \leq m$, s.t.

$$\mathbf{P}(\mathrm{FDP}(\tau_{\hat{\ell}}) > \alpha) \leq \zeta \text{ (and } \approx \zeta) \quad \text{for all } (\mu, \Gamma)$$

- Choose $\tau_\ell = \tau_\ell(\Gamma)$, here Γ known

Limitation of BH procedure

$m = 1000$, $\alpha = 0.2$, $\zeta = 0.05$, Equi-correlated $\rho = 0.1$



[Korn et al. (2004)], among others

Romano-Wolf's heuristic revisited

Calibrate critical values τ_ℓ's :

Crossing point $R(\tau_{\hat\ell}) = \hat\ell$ assume $\hat\ell = \ell_0$

$$\mathbf{P}(\mathrm{FDP}(\tau_{\ell_0}) > \alpha) = \mathbf{P}(V(\tau_{\ell_0}) > \alpha R(\tau_{\ell_0})) = \mathbf{P}(V(\tau_{\ell_0}) \geq \lfloor \alpha\ell_0 \rfloor + 1) \leq \zeta$$

e.g. using Markov's inequality

$$\mathbf{P}(V(\tau_{\ell_0}) \geq \lfloor \alpha\ell_0 \rfloor + 1) \leq \frac{m_0 \tau_{\ell_0}}{\lfloor \alpha\ell_0 \rfloor + 1} \leq \frac{m \tau_{\ell_0}}{\lfloor \alpha\ell_0 \rfloor + 1} \text{ gives } \tau_\ell = \frac{\zeta(\lfloor \alpha\ell \rfloor + 1)}{m}$$

Several other "bounding devices" adapting to dependence

- K-Markov
- Exact: for $Z \sim \mathcal{N}(0, \Gamma)$,

$$\mathbf{P}(V(t) \geq k) \leq \mathbf{P}(V'(t) \geq k) \text{ for } V'(t) = \sum_{i=1}^m \mathbf{1}\{\overline{\Phi}(Z_i) \leq t\},$$

$\Rightarrow$ Many new critical values !

Romano-Wolf's heuristic revisited

Calibrate critical values τ_ℓ's :

Crossing point $R(\tau_{\hat{\ell}}) = \hat{\ell}$ assume $\hat{\ell} = \ell_0$

$$\mathbf{P}(\mathrm{FDP}(\tau_{\ell_0}) > \alpha) = \mathbf{P}(V(\tau_{\ell_0}) > \alpha R(\tau_{\ell_0})) = \mathbf{P}(V(\tau_{\ell_0}) \geq \lfloor \alpha \ell_0 \rfloor + 1) \leq \zeta$$

e.g. using Markov's inequality

$$\mathbf{P}(V(\tau_{\ell_0}) \geq \lfloor \alpha \ell_0 \rfloor + 1) \leq \frac{m_0 \tau_{\ell_0}}{\lfloor \alpha \ell_0 \rfloor + 1} \leq \frac{m \tau_{\ell_0}}{\lfloor \alpha \ell_0 \rfloor + 1} \text{ gives } \tau_\ell = \frac{\zeta(\lfloor \alpha \ell \rfloor + 1)}{m}$$

Several other "bounding devices" adapting to dependence

- K-Markov
- Exact: for $Z \sim \mathcal{N}(0, \Gamma)$,

$$\mathbf{P}(V(t) \geq k) \leq \mathbf{P}(V'(t) \geq k) \text{ for } V'(t) = \sum_{i=1}^{m} \mathbf{1}\{\overline{\Phi}(Z_i) \leq t\},$$

$\Rightarrow$ Many new critical values !

Romano-Wolf's heuristic revisited

Calibrate critical values τ_ℓ's :

Crossing point $R(\tau_{\hat{\ell}}) = \hat{\ell}$ assume $\hat{\ell} = \ell_0$

$$\mathbf{P}(\mathrm{FDP}(\tau_{\ell_0}) > \alpha) = \mathbf{P}(V(\tau_{\ell_0}) > \alpha R(\tau_{\ell_0})) = \mathbf{P}(V(\tau_{\ell_0}) \geq \lfloor \alpha \ell_0 \rfloor + 1) \leq \zeta$$

e.g. using Markov's inequality

$$\mathbf{P}(V(\tau_{\ell_0}) \geq \lfloor \alpha \ell_0 \rfloor + 1) \leq \frac{m_0 \tau_{\ell_0}}{\lfloor \alpha \ell_0 \rfloor + 1} \leq \frac{m \tau_{\ell_0}}{\lfloor \alpha \ell_0 \rfloor + 1} \text{ gives } \tau_\ell = \frac{\zeta(\lfloor \alpha \ell \rfloor + 1)}{m}$$

Several other "bounding devices" adapting to dependence

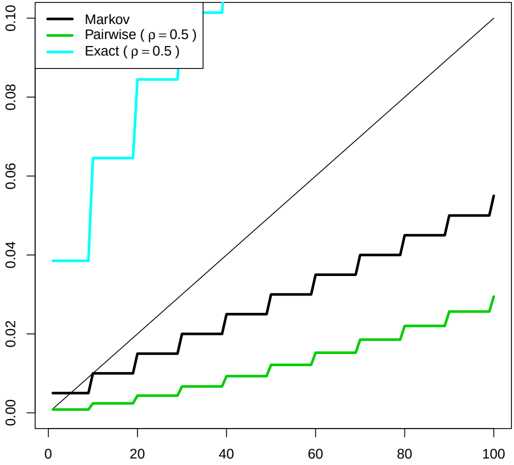
- K-Markov
- Exact: for $Z \sim \mathcal{N}(0, \Gamma)$,

$$\mathbf{P}(V(t) \geq k) \leq \mathbf{P}(V'(t) \geq k) \text{ for } V'(t) = \sum_{i=1}^m \mathbf{1}\{\overline{\Phi}(Z_i) \leq t\},$$

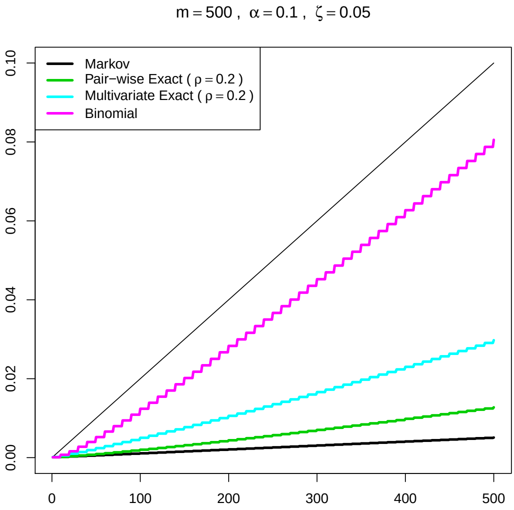
$\Rightarrow$ Many new critical values !

RW heuristic critical values

$m = 100\,,\ \alpha = 0.1\,,\ \zeta = 0.5$



RW heuristic critical values



Theoretical study of RW's heuristic

Finite sample:

$$\mathbf{P}(\text{FDP} > \alpha) \leq \zeta$$

- validation under indep. between null *p*-values and alternative *p*-values [Guo et al (2014)], [Guo and Romano (2007)], [Romano and Wolf (2007)]

- counter-example under equi-correlation
- two theoretically valid modifications of RW's heuristic

Asymptotic:

$$\limsup_m \{\mathbf{P}(\text{FDP} > \alpha)\} \leq \zeta$$

- validation under weak-dependence (CLT for *p*-value process)

- validation under strong, positive dependence in a factor model

Theoretical study of RW's heuristic

Finite sample:

$$\mathbf{P}(\text{FDP} > \alpha) \leq \zeta$$

- validation under indep. between null p-values and alternative p-values [Guo et al (2014)], [Guo and Romano (2007)], [Romano and Wolf (2007)]
- counter-example under equi-correlation
- two theoretically valid modifications of RW's heuristic

Asymptotic:

$$\limsup_m \{\mathbf{P}(\text{FDP} > \alpha)\} \leq \zeta$$

- validation under weak-dependence (CLT for p-value process)
- validation under strong, positive dependence in a factor model

Theoretical study of RW's heuristic

Finite sample:

$$\mathbf{P}(\text{FDP} > \alpha) \leq \zeta$$

- validation under indep. between null p-values and alternative p-values [Guo et al (2014)], [Guo and Romano (2007)], [Romano and Wolf (2007)]
- counter-example under equi-correlation
- two theoretically valid modifications of RW's heuristic

Asymptotic:

$$\limsup_m \{\mathbf{P}(\text{FDP} > \alpha)\} \leq \zeta$$

- validation under weak-dependence (CLT for p-value process)
- validation under strong, positive dependence in a factor model

A finite sample valid correction

Let $I_\ell = \mathbf{1}\{V'(\tau_\ell) \geq \lfloor \alpha\ell \rfloor + 1\}$ and

$$V'(t) = \sum_{i=1}^{m} \mathbf{1}\{\overline{\Phi}(Z_i) \leq t\}, \;\; Z \sim \mathcal{N}(0, \Gamma)$$

Now bound

$$\begin{aligned}
\mathbf{P}(\mathrm{FDP}(\tau_{\hat\ell}) > \alpha) &= \mathbf{P}(V(\tau_{\hat\ell}) \geq \lfloor \alpha\hat\ell \rfloor + 1) \\
&\leq \sum_{\ell=1}^{m} \mathbf{E}(I_\ell \mathbf{1}\{\hat\ell = \ell\}) \\
&= \sum_{\ell=1}^{m} \mathbf{E}(I_\ell \mathbf{1}\{\hat\ell \geq \ell\}) - \sum_{\ell=1}^{m-1} \mathbf{E}(I_\ell \mathbf{1}\{\hat\ell \geq \ell+1\}) \\
&= \sum_{\ell=1}^{m} \mathbf{E}((I_\ell - I_{\ell-1})\mathbf{1}\{\hat\ell \geq \ell\}) \\
&\leq \sum_{\ell=1}^{m} \mathbf{E}((I_\ell - I_{\ell-1})_+) := B(\tau, m, \Gamma)
\end{aligned}$$

Then choose τ to get $B(\tau, m, \Gamma) \leq \zeta$

A finite sample valid correction

Let $I_\ell = \mathbf{1}\{V'(\tau_\ell) \geq \lfloor \alpha\ell \rfloor + 1\}$ and

$$V'(t) = \sum_{i=1}^{m} \mathbf{1}\{\overline{\Phi}(Z_i) \leq t\}, \;\; Z \sim \mathcal{N}(0, \Gamma)$$

Now bound

$$\begin{aligned}
\mathbf{P}(\mathrm{FDP}(\tau_{\hat\ell}) > \alpha) &= \mathbf{P}(V(\tau_{\hat\ell}) \geq \lfloor \alpha\hat\ell \rfloor + 1) \\
&\leq \sum_{\ell=1}^{m} \mathbf{E}(I_\ell \mathbf{1}\{\hat\ell = \ell\}) \\
&= \sum_{\ell=1}^{m} \mathbf{E}(I_\ell \mathbf{1}\{\hat\ell \geq \ell\}) - \sum_{\ell=1}^{m-1} \mathbf{E}(I_\ell \mathbf{1}\{\hat\ell \geq \ell + 1\}) \\
&= \sum_{\ell=1}^{m} \mathbf{E}((I_\ell - I_{\ell-1})\mathbf{1}\{\hat\ell \geq \ell\}) \\
&\leq \sum_{\ell=1}^{m} \mathbf{E}((I_\ell - I_{\ell-1})_+) := B(\tau, m, \Gamma)
\end{aligned}$$

Then choose τ to get $B(\tau, m, \Gamma) \leq \zeta$

A finite sample valid correction

Let $I_\ell = \mathbf{1}\{V'(\tau_\ell) \geq \lfloor \alpha\ell \rfloor + 1\}$ and

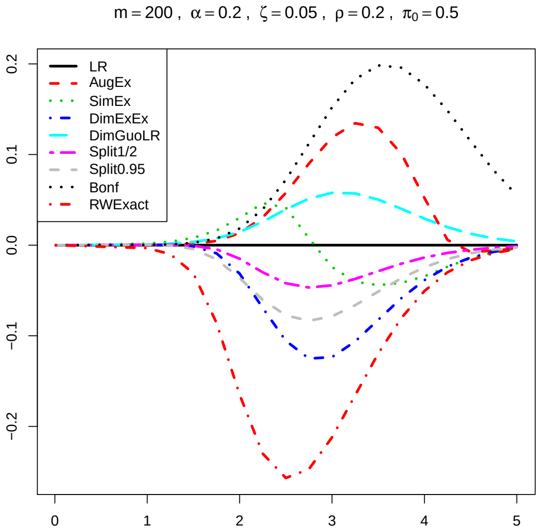
$$V'(t) = \sum_{i=1}^{m} \mathbf{1}\{\overline{\Phi}(Z_i) \leq t\},\ \ Z \sim \mathcal{N}(0, \Gamma)$$

Now bound

$$\begin{aligned}
\mathbf{P}(\mathrm{FDP}(\tau_{\hat\ell}) > \alpha) &= \mathbf{P}(V(\tau_{\hat\ell}) \geq \lfloor \alpha\hat\ell \rfloor + 1) \\
&\leq \sum_{\ell=1}^{m} \mathbf{E}(I_\ell \mathbf{1}\{\hat\ell = \ell\}) \\
&= \sum_{\ell=1}^{m} \mathbf{E}(I_\ell \mathbf{1}\{\hat\ell \geq \ell\}) - \sum_{\ell=1}^{m-1} \mathbf{E}(I_\ell \mathbf{1}\{\hat\ell \geq \ell + 1\}) \\
&= \sum_{\ell=1}^{m} \mathbf{E}((I_\ell - I_{\ell-1}) \mathbf{1}\{\hat\ell \geq \ell\}) \\
&\leq \sum_{\ell=1}^{m} \mathbf{E}((I_\ell - I_{\ell-1})_+) := B(\tau, m, \Gamma)
\end{aligned}$$

Then choose τ to get $B(\tau, m, \Gamma) \leq \zeta$

Illustration : FNR relative to LR



RW heuristic made asymptotical

Equi-correlated case

$$X_i = \mu_i + \rho^{1/2} W + (1-\rho)^{1/2}\zeta_i,\ 1 \le i \le m.$$

Then,

$$\mathbf{P}(V(\tau_\ell) \ge \lfloor \alpha\ell \rfloor + 1) \le \mathbf{P}(V'(\tau_\ell) \ge \lfloor \alpha\ell \rfloor + 1) \approx \mathbf{P}(F(\tau_\ell, W) \ge \alpha\ell/m)$$

for

$$F(t, W) = \overline{\Phi}\left((\overline{\Phi}^{-1}(t) - \rho^{1/2}\, W)/(1-\rho)^{1/2} \right) \text{ increasing w.r.t. } W$$

Choose $F(\tau_\ell, \overline{\Phi}^{-1}(\zeta)) = \alpha\ell/m$, so that,

$$\mathbf{P}(F(\tau_\ell, W) \ge \alpha\ell/m) = \mathbf{P}(F(\tau_\ell, W) \ge F(\tau_\ell, \overline{\Phi}^{-1}(\zeta))) = \mathbf{P}(W \ge \overline{\Phi}^{-1}(\zeta)) = \zeta$$

New critical values

$$\tau_\ell = \overline{\Phi}\left(\rho^{1/2}\overline{\Phi}^{-1}(\zeta) + (1-\rho)^{1/2}\overline{\Phi}^{-1}(\alpha\ell/m) \right)$$

RW heuristic made asymptotical

Equi-correlated case

$$X_i = \mu_i + \rho^{1/2} W + (1-\rho)^{1/2} \zeta_i,\ 1 \le i \le m.$$

Then,

$$\mathbf{P}(V(\tau_\ell) \ge \lfloor \alpha\ell \rfloor + 1) \le \mathbf{P}(V'(\tau_\ell) \ge \lfloor \alpha\ell \rfloor + 1) \approx \mathbf{P}(F(\tau_\ell, W) \ge \alpha\ell/m)$$

for

$$F(t, W) = \overline{\Phi}\left((\overline{\Phi}^{-1}(t) - \rho^{1/2}\, W)/(1-\rho)^{1/2}\right) \text{ increasing w.r.t. } W$$

Choose $F(\tau_\ell, \overline{\Phi}^{-1}(\zeta)) = \alpha\ell/m$, so that,

$$\mathbf{P}(F(\tau_\ell, W) \ge \alpha\ell/m) = \mathbf{P}(F(\tau_\ell, W) \ge F(\tau_\ell, \overline{\Phi}^{-1}(\zeta))) = \mathbf{P}(W \ge \overline{\Phi}^{-1}(\zeta)) = \zeta$$

New critical values

$$\tau_\ell = \overline{\Phi}\left(\rho^{1/2}\overline{\Phi}^{-1}(\zeta) + (1-\rho)^{1/2}\overline{\Phi}^{-1}(\alpha\ell/m)\right)$$

RW heuristic made asymptotical

Equi-correlated case

$$X_i = \mu_i + \rho^{1/2} W + (1-\rho)^{1/2}\zeta_i,\ 1 \le i \le m.$$

Then,

$$\mathbf{P}(V(\tau_\ell) \ge \lfloor \alpha\ell \rfloor + 1) \le \mathbf{P}(V'(\tau_\ell) \ge \lfloor \alpha\ell \rfloor + 1) \approx \mathbf{P}(F(\tau_\ell, W) \ge \alpha\ell/m)$$

for

$$F(t, W) = \overline{\Phi}\left((\overline{\Phi}^{-1}(t) - \rho^{1/2}\, W)/(1-\rho)^{1/2}\right) \text{ increasing w.r.t. } W$$

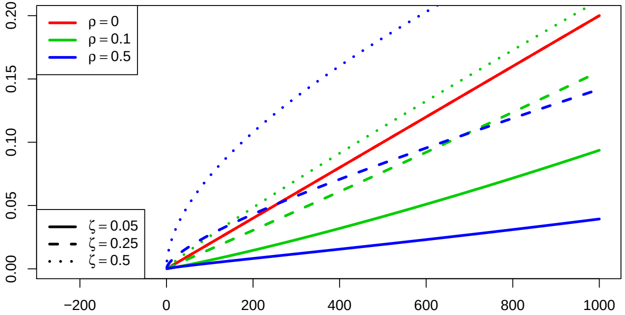
Choose $F(\tau_\ell, \overline{\Phi}^{-1}(\zeta)) = \alpha\ell/m$, so that,

$$\mathbf{P}(F(\tau_\ell, W) \ge \alpha\ell/m) = \mathbf{P}(F(\tau_\ell, W) \ge F(\tau_\ell, \overline{\Phi}^{-1}(\zeta))) = \mathbf{P}(W \ge \overline{\Phi}^{-1}(\zeta)) = \zeta$$

New critical values

$$\tau_\ell = \overline{\Phi}\left(\rho^{1/2}\overline{\Phi}^{-1}(\zeta) + (1-\rho)^{1/2}\overline{\Phi}^{-1}(\alpha\ell/m)\right)$$

Result



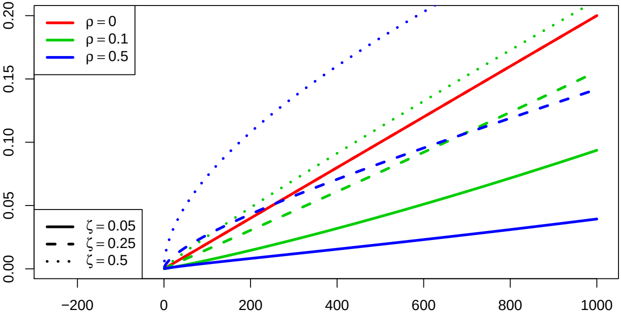
Theorem [Delattre and R. (2015)]

One-sided Gaussian ρ-equi-correlation; $\rho \in [0, 1)$, $m_0/m \to \pi_0 \in (0, 1)$, then

$$\limsup_m \mathbf{P}(\mathrm{FDP}(\mathit{New}) > \alpha) \leq \zeta.$$

Can be generalized to $Y_i = c_i W + \xi_i$, $c_i \geq 0$

Result



Theorem [Delattre and R. (2015)]

One-sided Gaussian ρ-equi-correlation; $\rho \in [0, 1)$, $m_0/m \to \pi_0 \in (0, 1)$, then

$$\limsup_m \mathbf{P}(\mathrm{FDP}(\mathit{New}) > \alpha) \le \zeta.$$

Can be generalized to $Y_i = c_i W + \xi_i$, $c_i \ge 0$

Outlook

Conclusions : FDP control

- incorporating known dependence is possible
- RW's heuristic can be justified in some cases
- and corrected otherwise

Perspectives

- more justifications of RW's heuristic?
- unknown dependence (permutation?)
- *p*-value correction under equi-correlation

Advertising

2-year postdoc position in France about post hoc inference

Outlook

Conclusions : FDP control

- incorporating known dependence is possible
- RW's heuristic can be justified in some cases
- and corrected otherwise

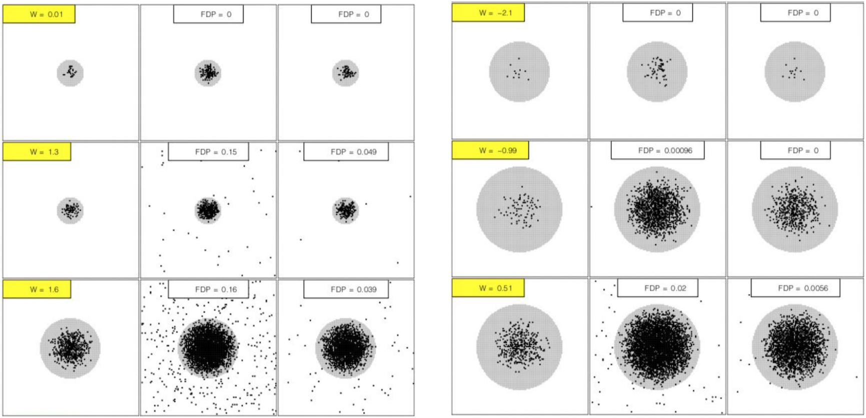
Perspectives

- more justifications of RW's heuristic?
- unknown dependence (permutation?)
- *p*-value correction under equi-correlation

Advertising

2-year postdoc position in France about post hoc inference

Illustration



On going work: p-value correction

Under equi-correlation

$$X_i = \mu_i + \rho^{1/2} W + (1-\rho)^{1/2} \zeta_i,\ 1 \leq i \leq m.$$

How to “kill" W?

- $X_i^\star = X_i - \bar{X}$ (add bias)
- $X_i^\star = X_i - \rho^{1/2} \hat{W}$

Semi-parametric estimation of W:

- $\mu \in \mathbb{R}^m$ unknown sparse signal $m_0/m \geq 1 - c\, m^{-\beta}$, $\beta \in [0, 1)$
- minimax estimation rate of W is

$$m^{-\beta} (\log m)^{-\gamma}$$

with $\gamma = 1/2$ (two-sided) $\gamma = 3/2$ (one-sided).

Todo: properties of BH based on $X_i^\star = X_i - \rho^{1/2} \hat{W}$?

[Leek and Storey (2008)], [Friguet et al. (2009)], [Fan et al. (2012)]

On going work: p-value correction

Under equi-correlation

$$X_i = \mu_i + \rho^{1/2} W + (1-\rho)^{1/2}\zeta_i,\ 1 \leq i \leq m.$$

How to "kill" W?

- $X_i^\star = X_i - \bar{X}$ (add bias)
- $X_i^\star = X_i - \rho^{1/2}\hat{W}$

Semi-parametric estimation of W:

- $\mu \in \mathbb{R}^m$ unknown sparse signal $m_0/m \geq 1 - c\, m^{-\beta}$, $\beta \in [0,1)$
- minimax estimation rate of W is

$$m^{-\beta}(\log m)^{-\gamma}$$

with $\gamma = 1/2$ (two-sided) $\gamma = 3/2$ (one-sided).

Todo: properties of BH based on $X_i^\star = X_i - \rho^{1/2}\hat{W}$?

[Leek and Storey (2008)], [Friguet et al. (2009)], [Fan et al. (2012)]

On going work: p-value correction

Under equi-correlation

$$X_i = \mu_i + \rho^{1/2} W + (1-\rho)^{1/2} \zeta_i, \ 1 \leq i \leq m.$$

How to “kill" W?

- $X_i^\star = X_i - \bar{X}$ (add bias)
- $X_i^\star = X_i - \rho^{1/2} \hat{W}$

Semi-parametric estimation of W:

- $\mu \in \mathbb{R}^m$ unknown sparse signal $m_0/m \geq 1 - c\, m^{-\beta}$, $\beta \in [0, 1)$
- minimax estimation rate of W is

$$m^{-\beta} (\log m)^{-\gamma}$$

  with $\gamma = 1/2$ (two-sided) $\gamma = 3/2$ (one-sided).

Todo: properties of BH based on $X_i^\star = X_i - \rho^{1/2} \hat{W}$?

[Leek and Storey (2008)], [Friguet et al. (2009)], [Fan et al. (2012)]