Hunting for significance with multiple testing

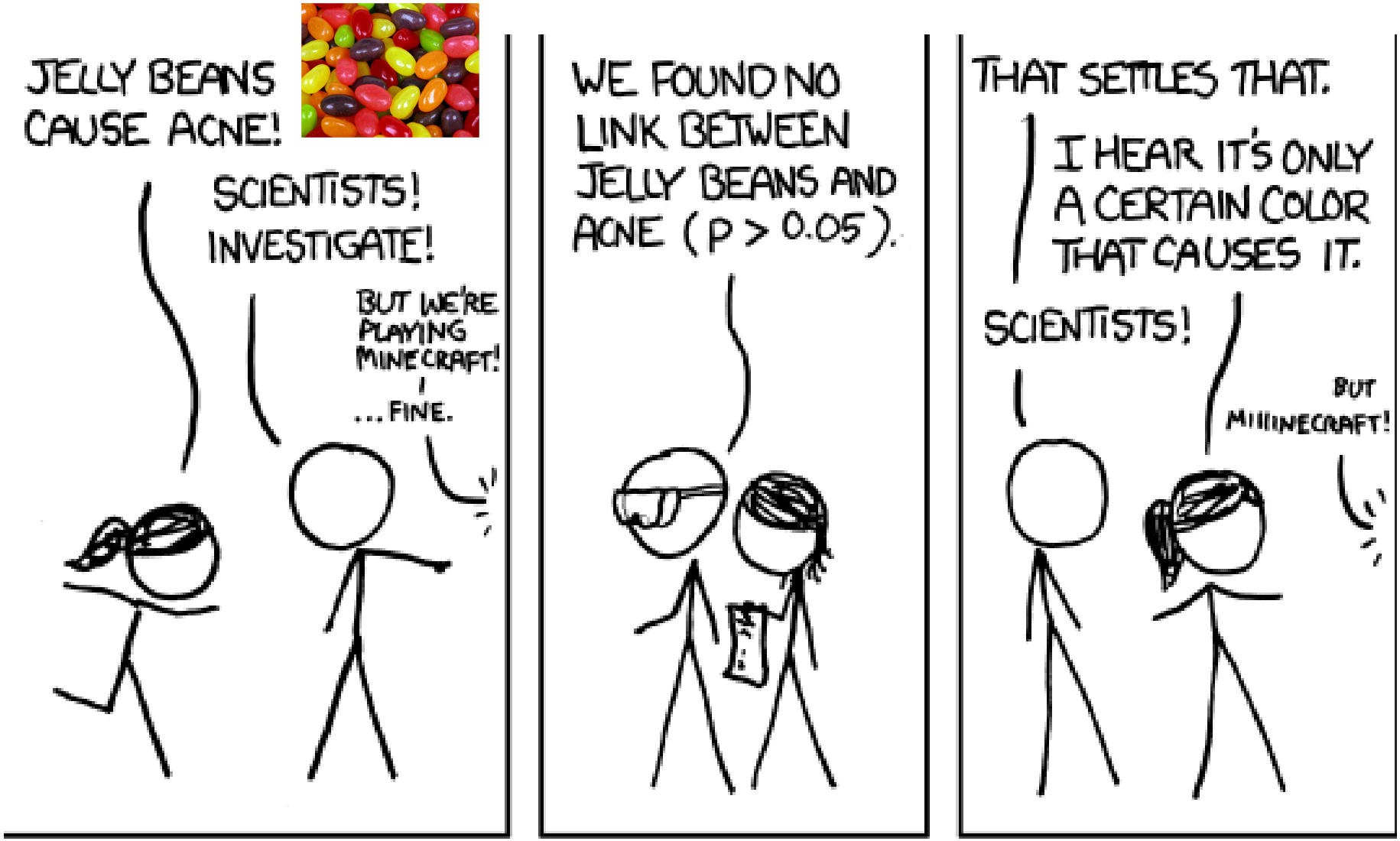
Etienne Roquain[1]

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

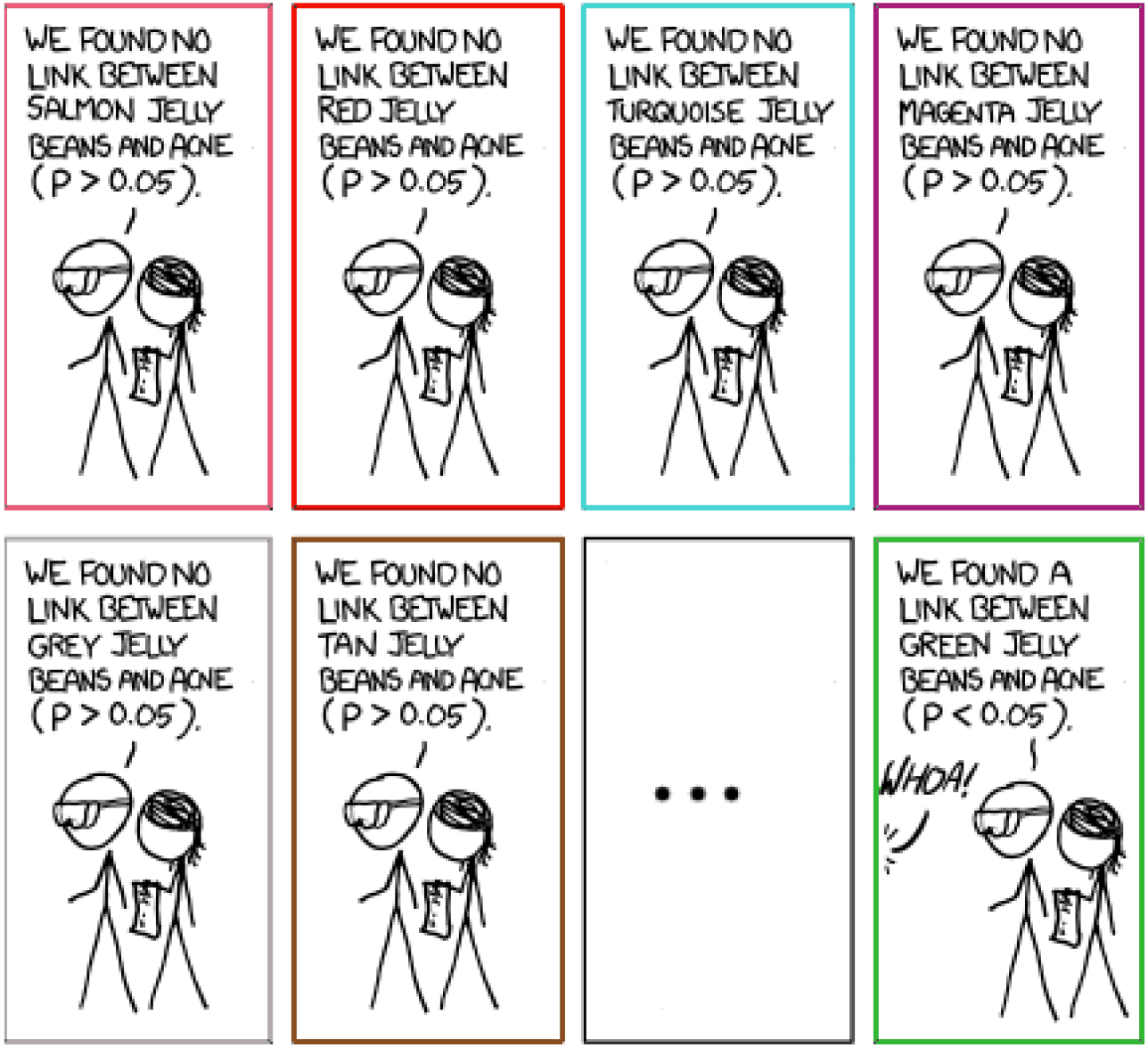
Séminaire MODAL'X, 19 mai 2016

1. Introduction
2. Classical results
3. FDP control
4. Posthoc inference

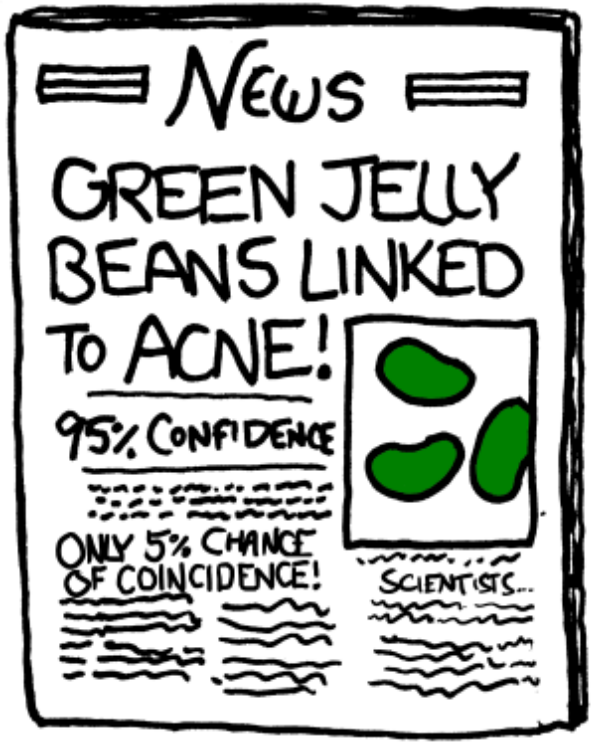
A “multiple testing story" (http://xkcd.com)



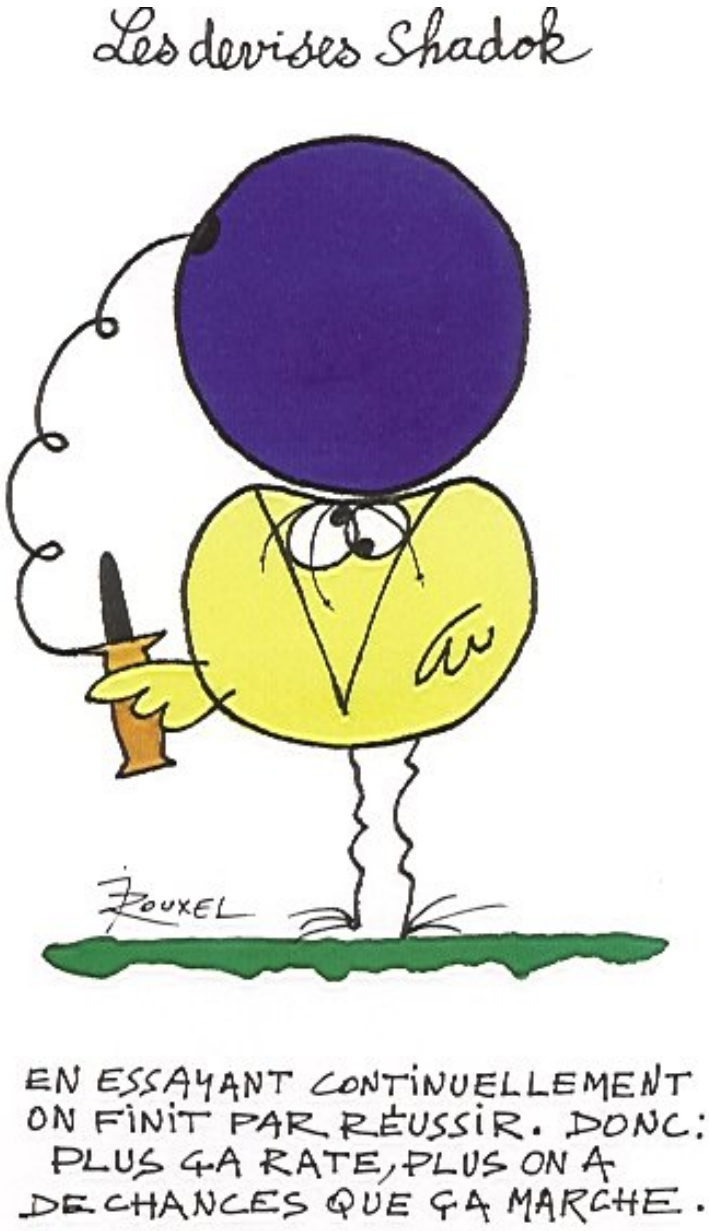
A “multiple testing story" (http://xkcd.com)



A “multiple testing story" (http://xkcd.com)



Shadoks motto



False positive in practice

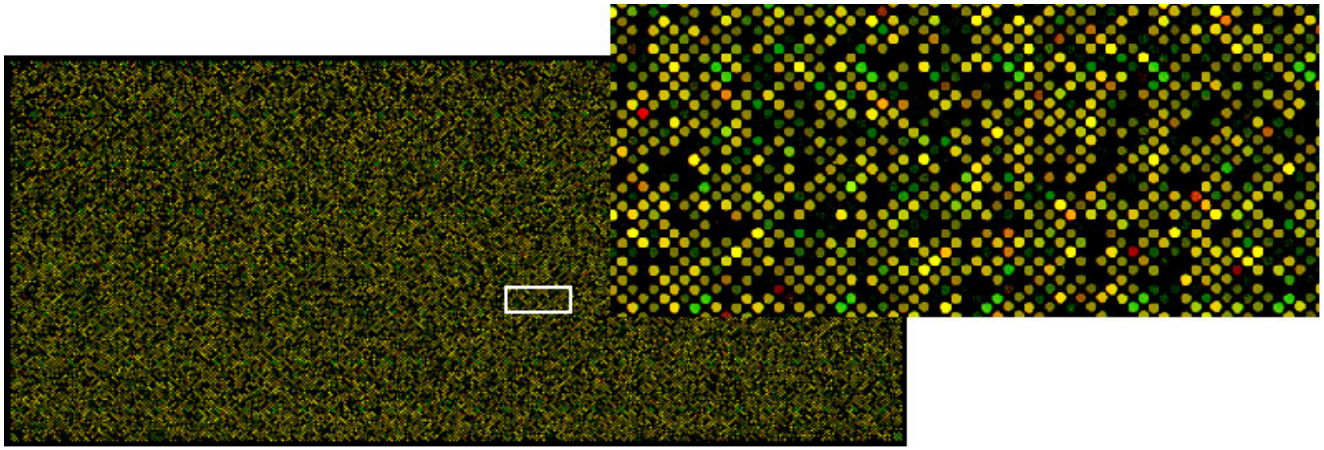
Abstract of [Giovannucci et al. (1995)]

"Between 1986 and 1992, 812 new cases of prostate cancer [...] were documented. Of 46 vegetables and fruits or related products, four were significantly associated with lower prostate cancer risk; of the four–tomato sauce (P for trend = .001), tomatoes (P for trend = .03), and pizza (P for trend = .05), but not strawberries–were primary sources of lycopene. Combined intake of tomatoes, tomato sauce, tomato juice, and pizza (which accounted for 82% of lycopene intake) was inversely associated with risk of prostate cancer (multivariate RR = 0.65; 95% CI = 0.44-0.95, for consumption frequency greater than 10 versus less than 1.5 servings per week; P for trend = .01) and advanced (stages C and D) prostate cancers (multivariate RR = 0.47; 95% CI = 0.22-1.00; P for trend = .03)."

Should we eat pizza to avoid prostate cancer?

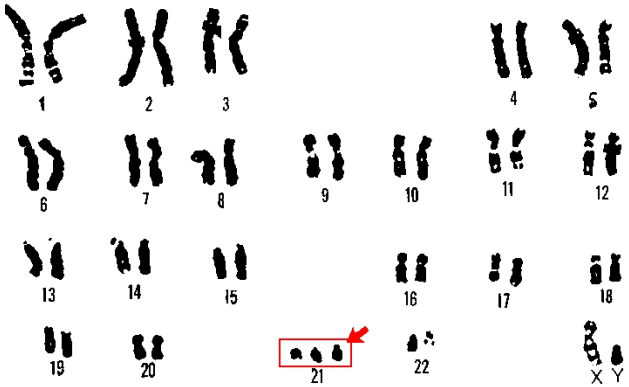
Massive and complex datasets

- Gene expressions
  interesting genes?



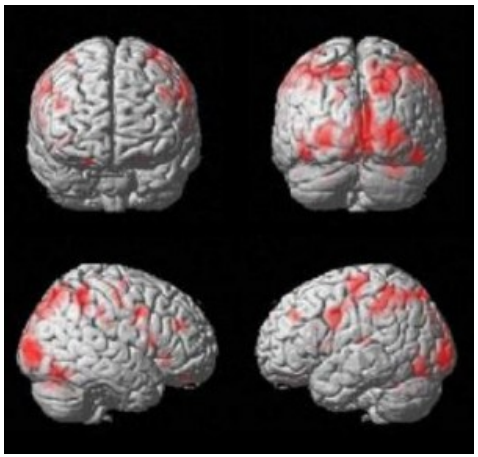
- Genome-wide assoc. study
  interesting SNPs?

- Copy number alterations
  interesting probes?



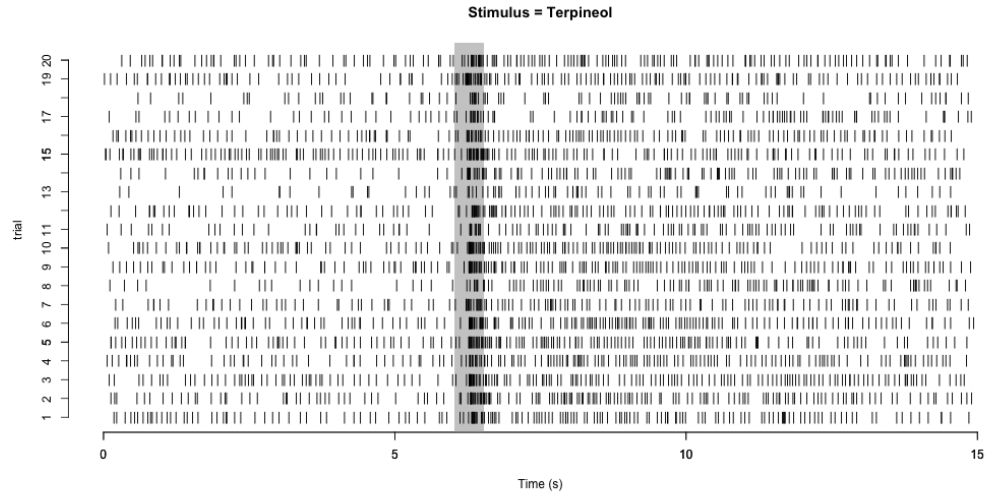
Massive and complex datasets

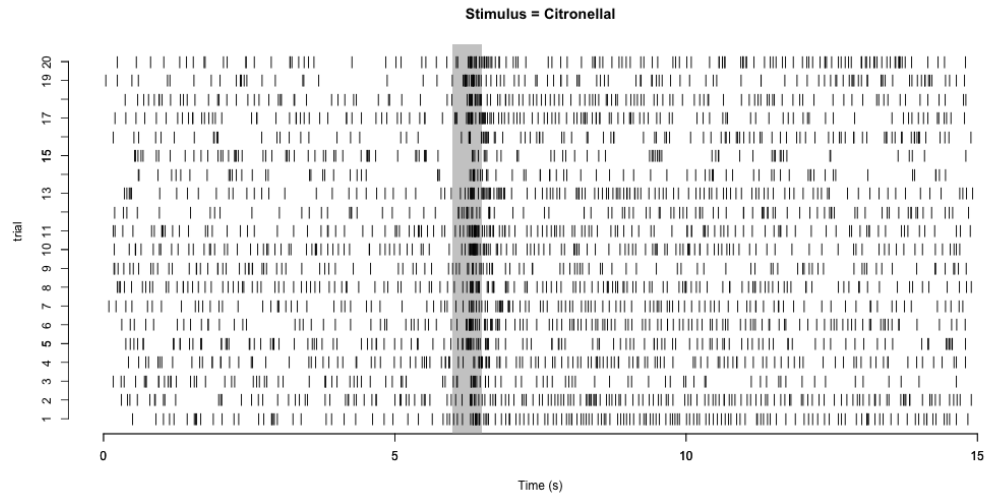
- **Neuroimaging (FMRI)**
  activated regions?



- **Astronomy**
  directions with stars?

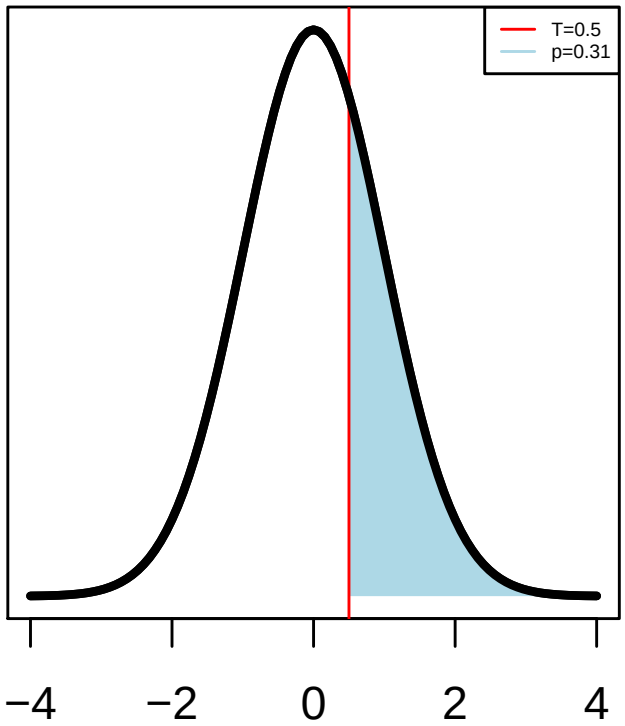
- **Neuroscience**
  periods for neuronal response?

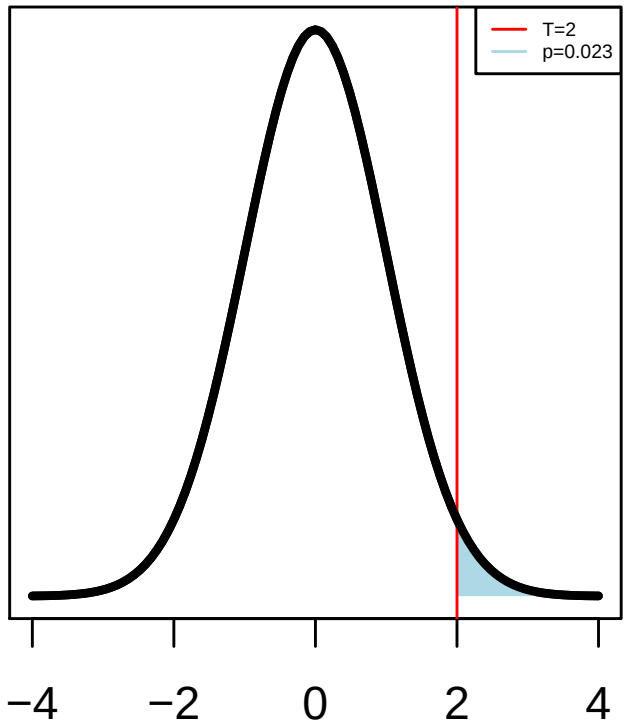




Single testing

- Data: $X = (X_1, \ldots, X_n)$ i.i.d. $\mathcal{N}(\mu, 1)$, $\mu \in \mathbb{R}$ unknown
- Null hypothesis H_0: "$\mu \leq 0$" against alt. H_1: "$\mu > 0$"
- Test statistic: $T(X) = n^{-1/2} \sum_{i=1}^{n} X_i$, p-value $p(X) = \overline{\Phi}(T(X))$



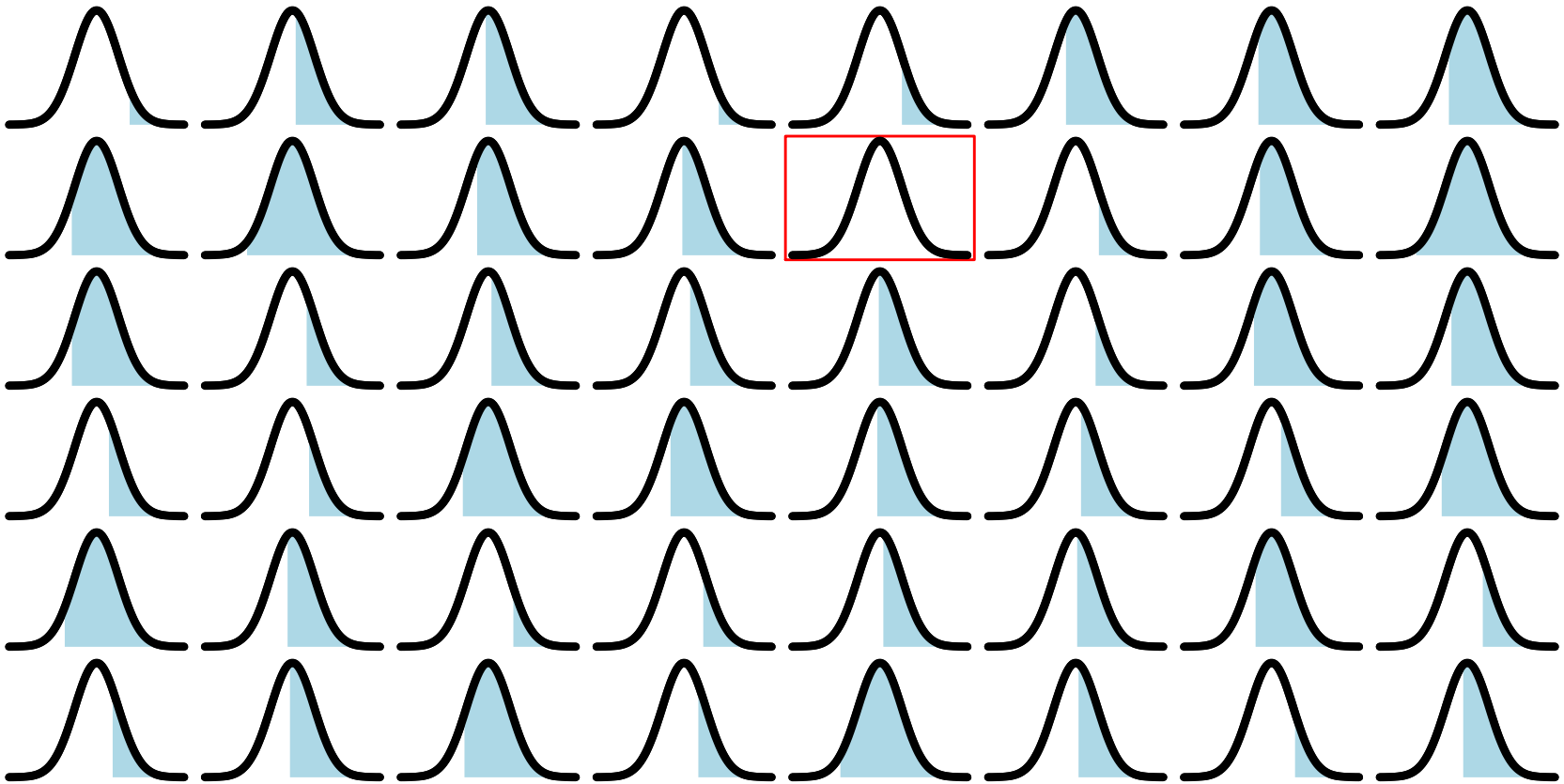


Test of level α: reject H_0 if $p(X) \leq \alpha$

Risk under H_0: $\mathbf{P}(p(X) \leq \alpha) \leq \alpha$

Multiple testing

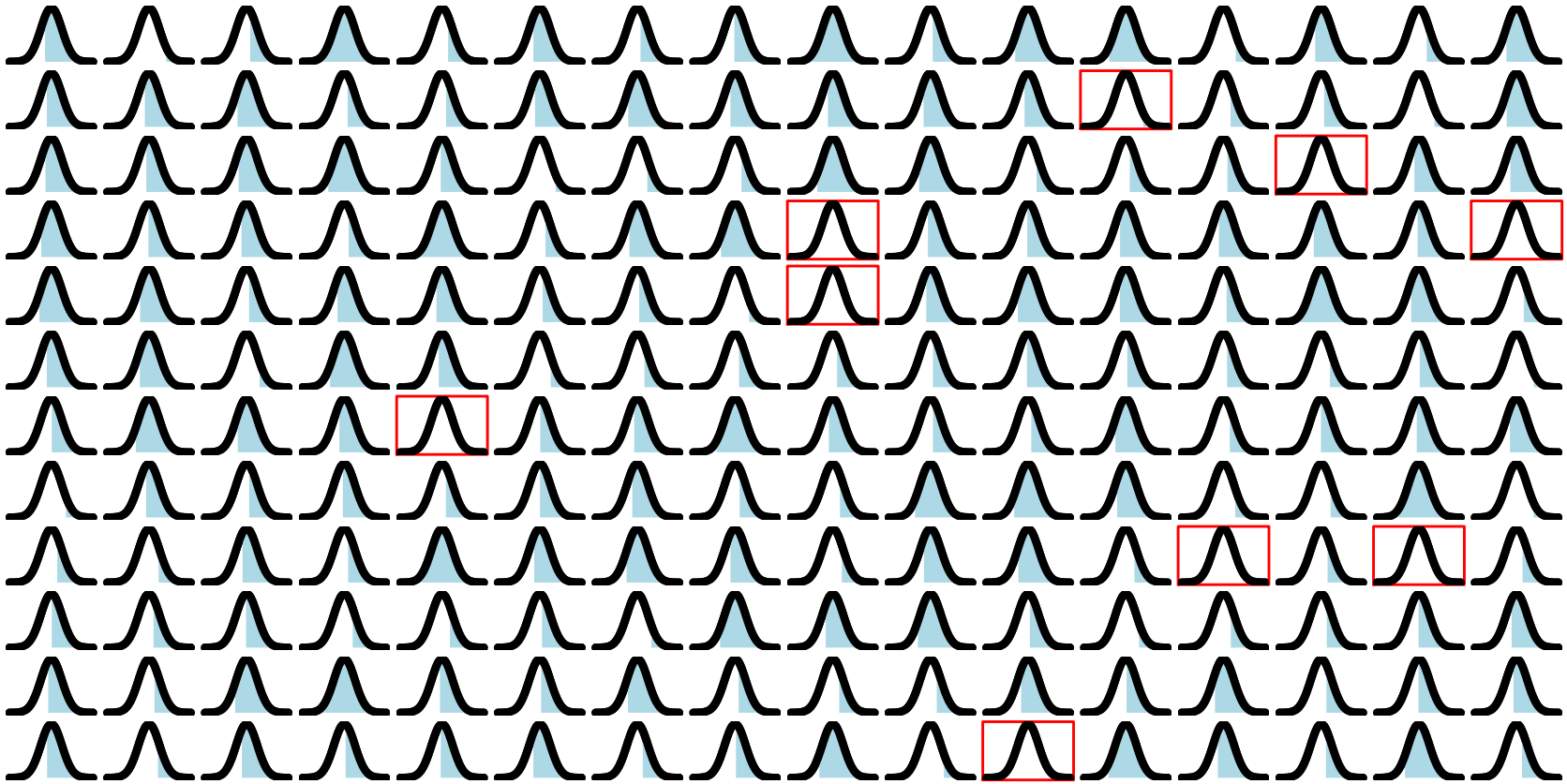
Pure noise, $m = 48$ indep. tests:



At least one false positive: with prob. $1 - (1 - 0.05)^{48} \geq 0.91$

Multiple testing

Pure noise, $m = 192$ indep. tests:

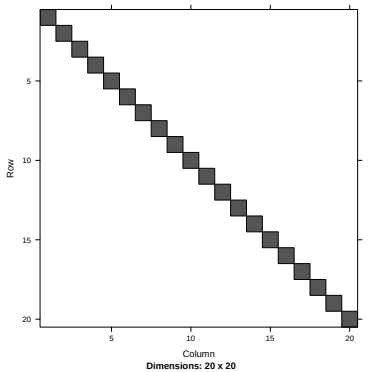


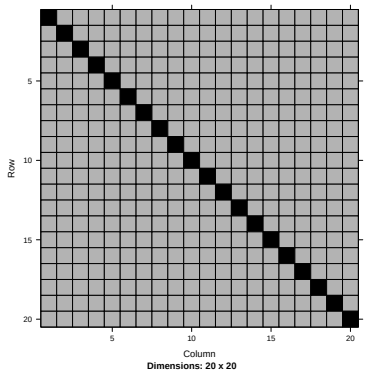
At least one false positive: with prob. $1 - (1 - 0.05)^{192} \geq 1 - 6 \times 10^{-5}$

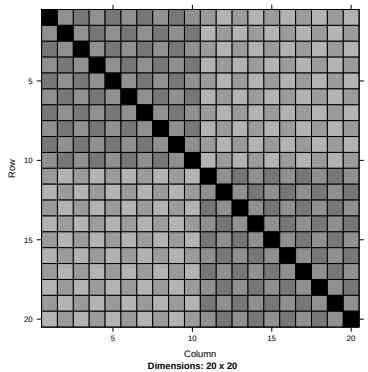
Canonical setting

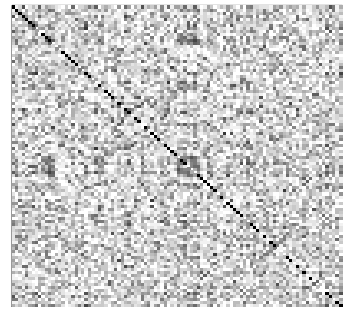
$X_i =$ avg group 2 - avg group 1 (rescaled) for item i, $1 \leq i \leq m$.

- Gaussian model :
$$X \sim \mathcal{N}(\mu, \Gamma) \in \mathbb{R}^m$$
- Hypotheses : $H_{0,i}$: "$\mu_i \leq 0$" vs $H_{1,i}$: "$\mu_i > 0$", $1 \leq i \leq m$.
- Parameter of interest : $\theta_i = \mathbf{1}\{\mu_i > 0\}$, $1 \leq i \leq m$.
- p-values: $p_i = \overline{\Phi}(X_i)$, $1 \leq i \leq m$.
- Dependence: covariance matrix Γ (known or unknown)









Aim : recover θ from $(p_i(X), 1 \leq i \leq m)$

General setting

- Data $X \in (\mathcal{X}, \mathfrak{X})$ with $X \sim P \in \mathcal{P}$ (model)
- $H_{0,i}$: "$P \in \mathcal{P}_{0,i}$", $1 \leq i \leq m$, null hypotheses for P
- true/false label $\theta = \theta(P) \in \{0,1\}^m$ such that

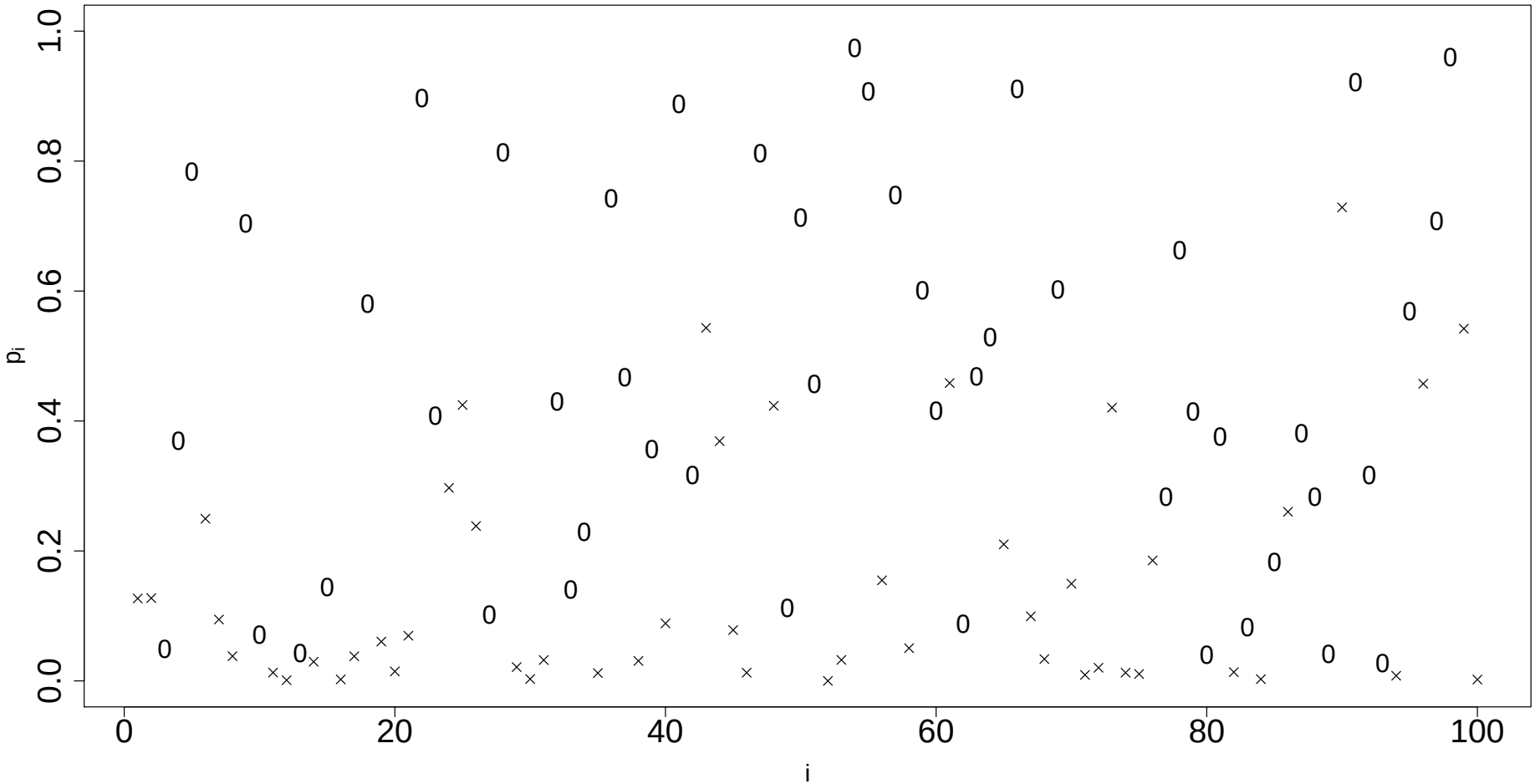
$$\theta_i = 0 \text{ if and only if } H_{0,i} \text{ is true for } P$$

- Assume p-values $(p_i(X), 1 \leq i \leq m)$ such that

$$\text{if } \theta_i = 0, \quad p_i(X) \succeq U(0,1)$$
$$\text{if } \theta_i = 1, \quad p_i(X) \sim \text{let arbitrary (typically "small")}$$

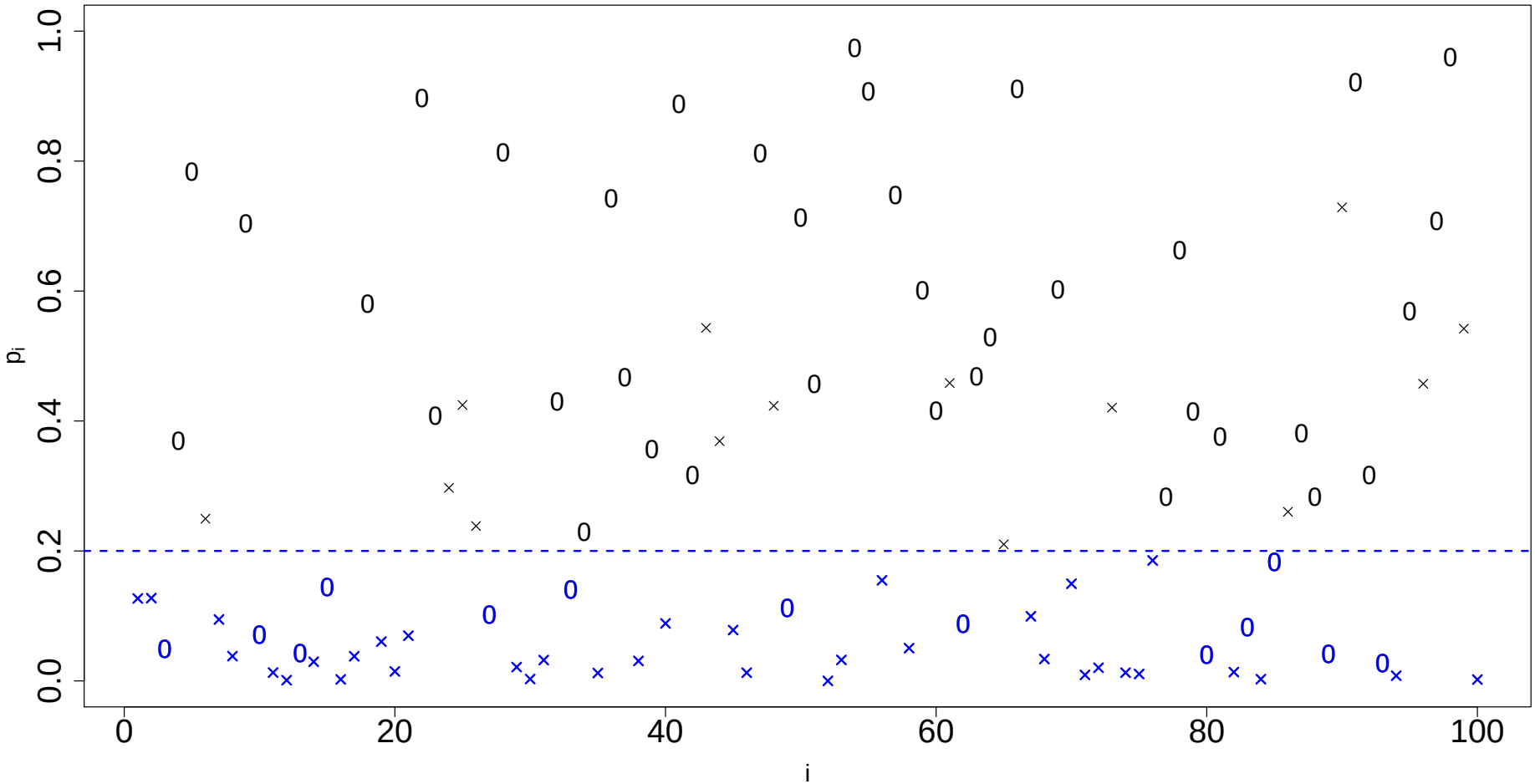
Aim : recover θ from $(p_i(X), 1 \leq i \leq m)$

Thresholding



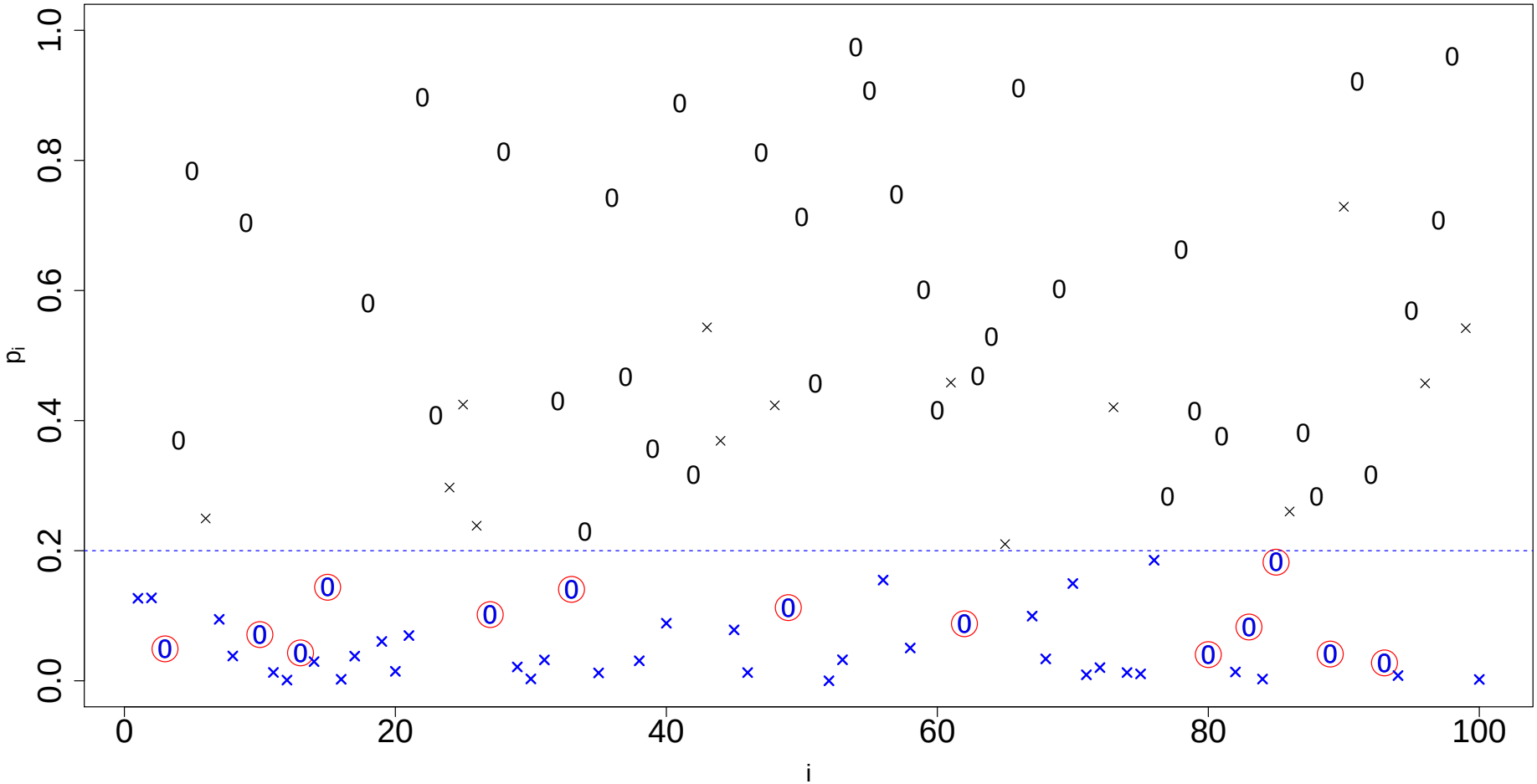
- Rejection set $R = \{1 \leq i \leq m \ : \ p_i(X) \leq \widehat{t}\}$
- Some false positives

Thresholding



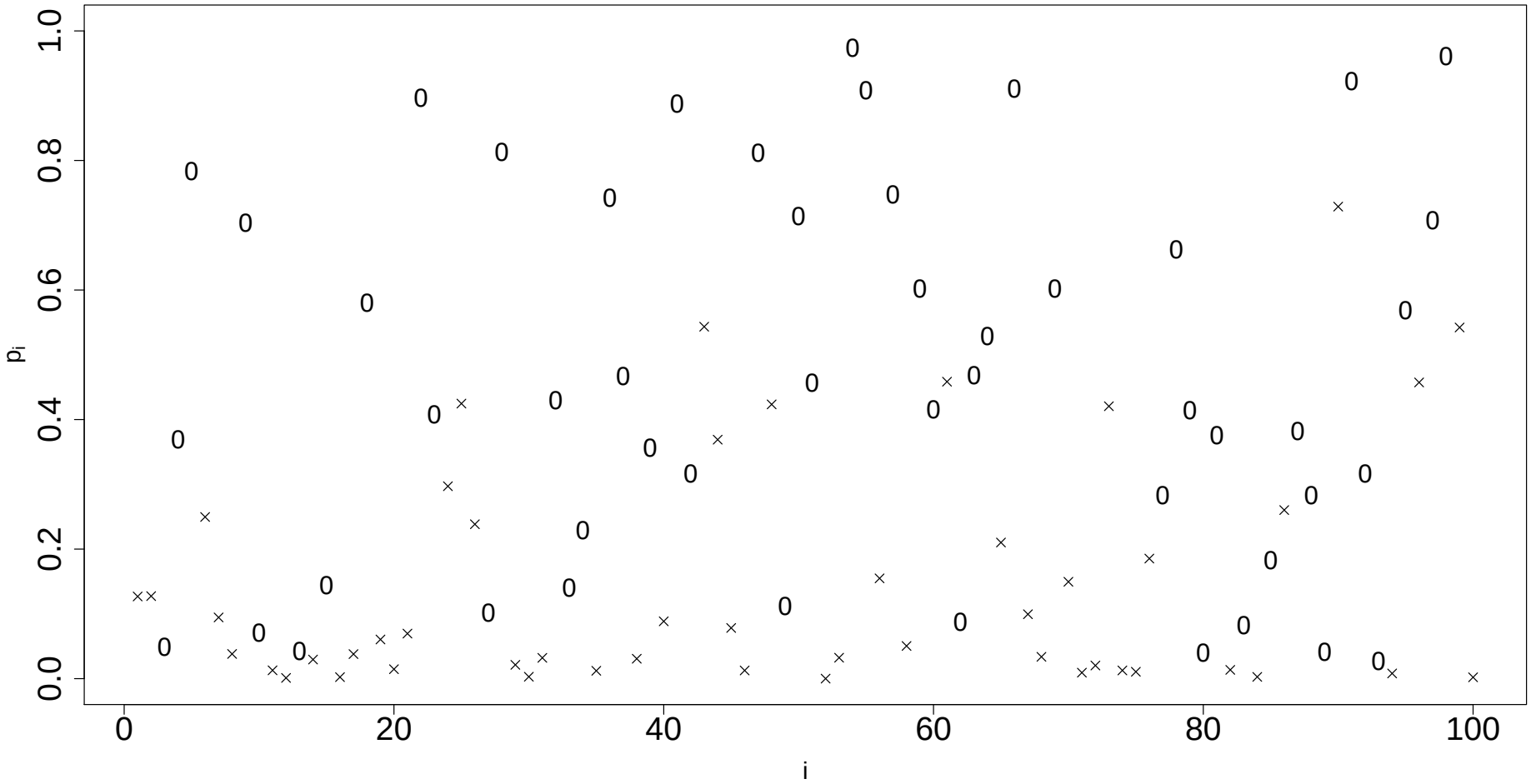
- Rejection set $R = \{1 \leq i \leq m \ : \ p_i(X) \leq \widehat{t}\}$
- Some false positives

Thresholding



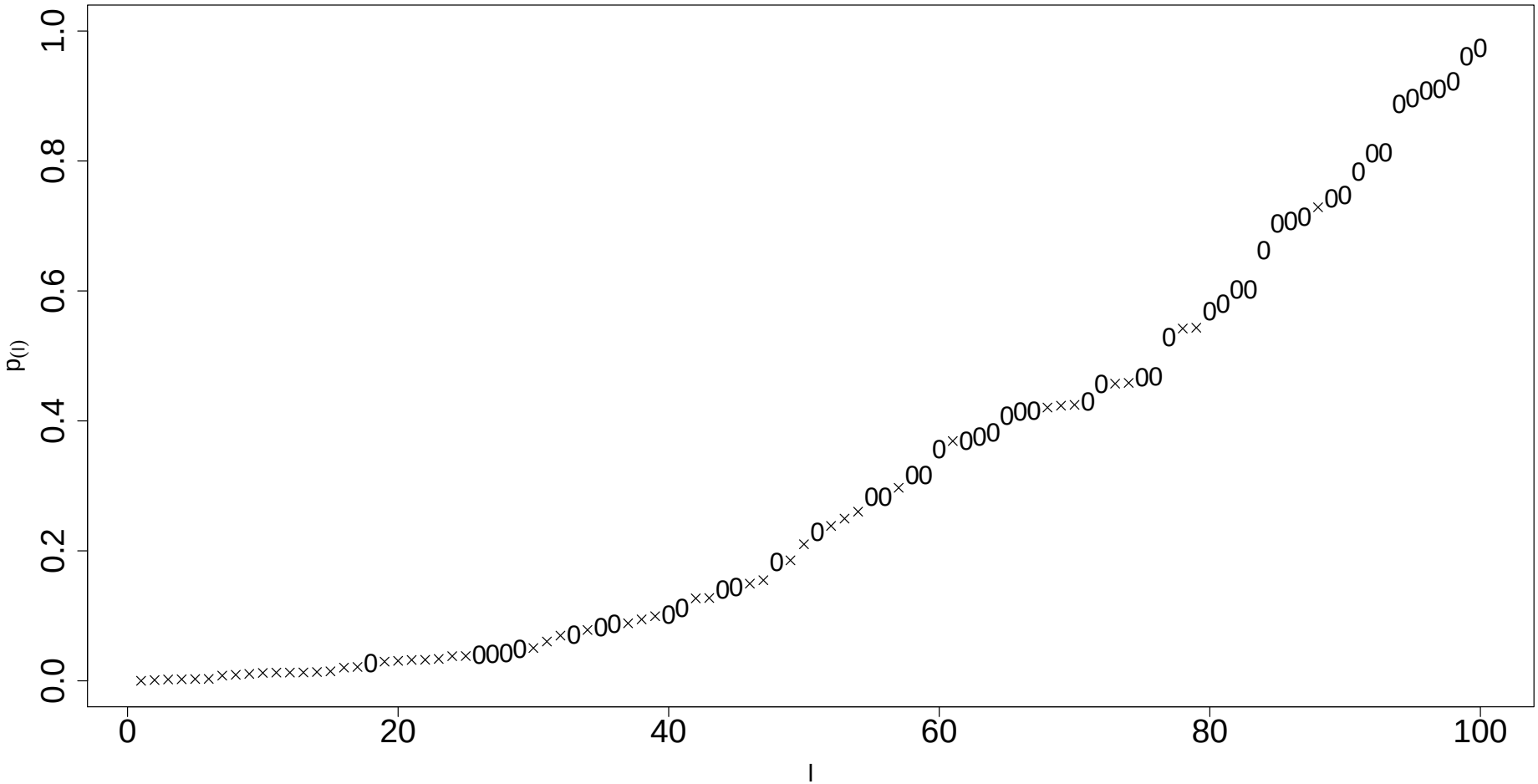
- Rejection set $R = \{1 \leq i \leq m \ : \ p_i(X) \leq \widehat{t}\}$
- Some false positives

Ordering *p*-values



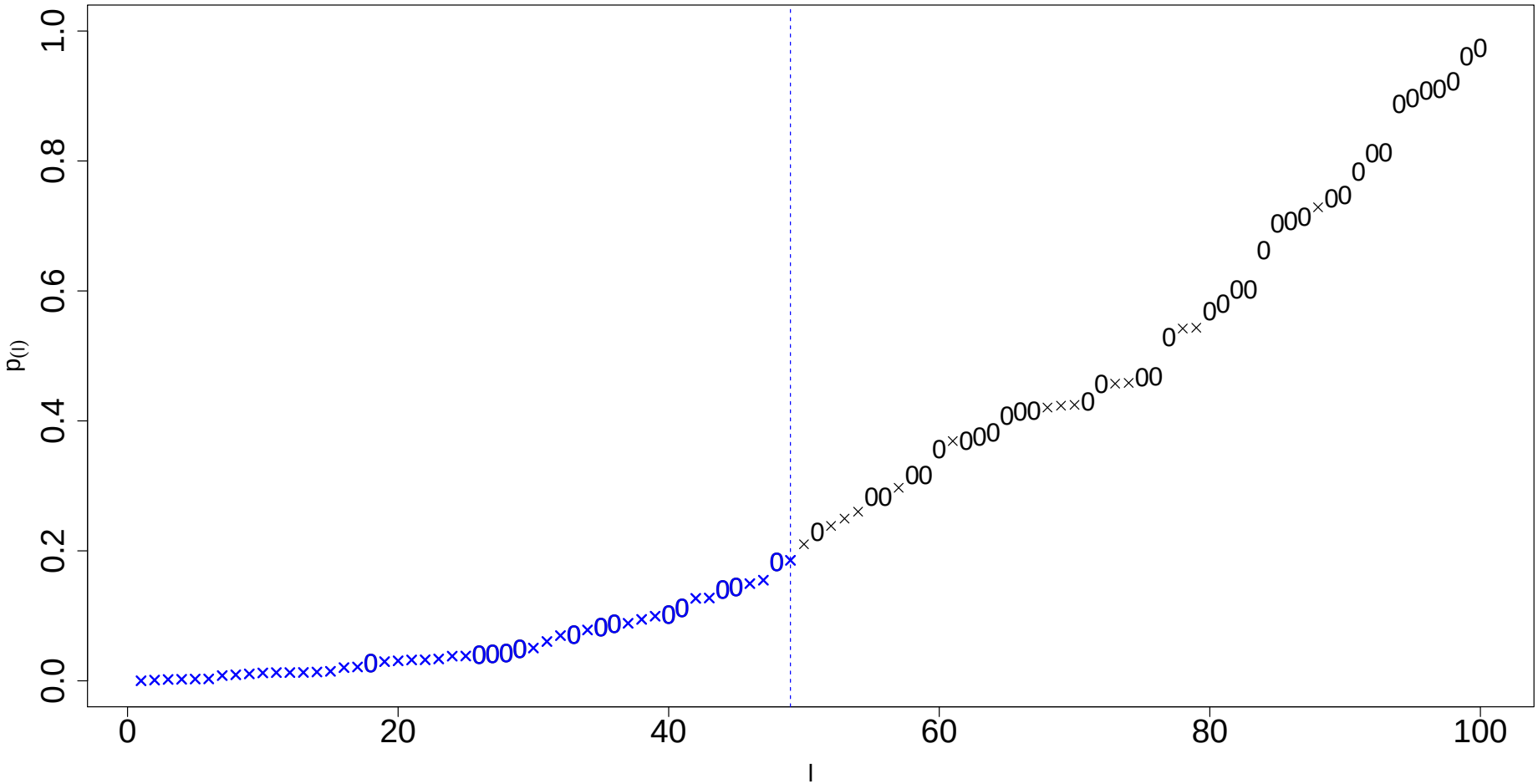
- Stopping rule
- Reject $\widehat{\ell}$ smallest *p*-values

Ordering *p*-values



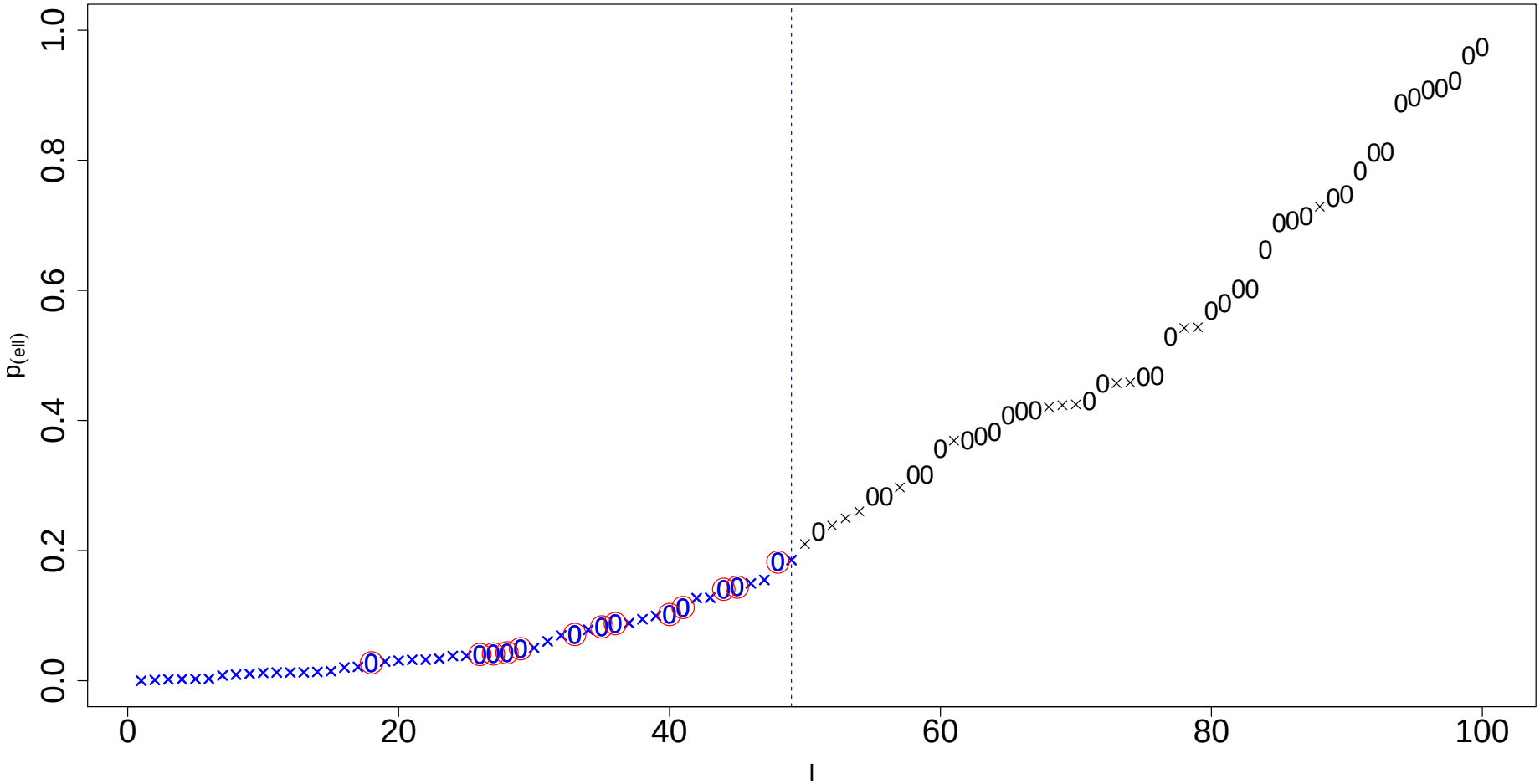
- Stopping rule
- Reject $\widehat{\ell}$ smallest *p*-values

Ordering *p*-values



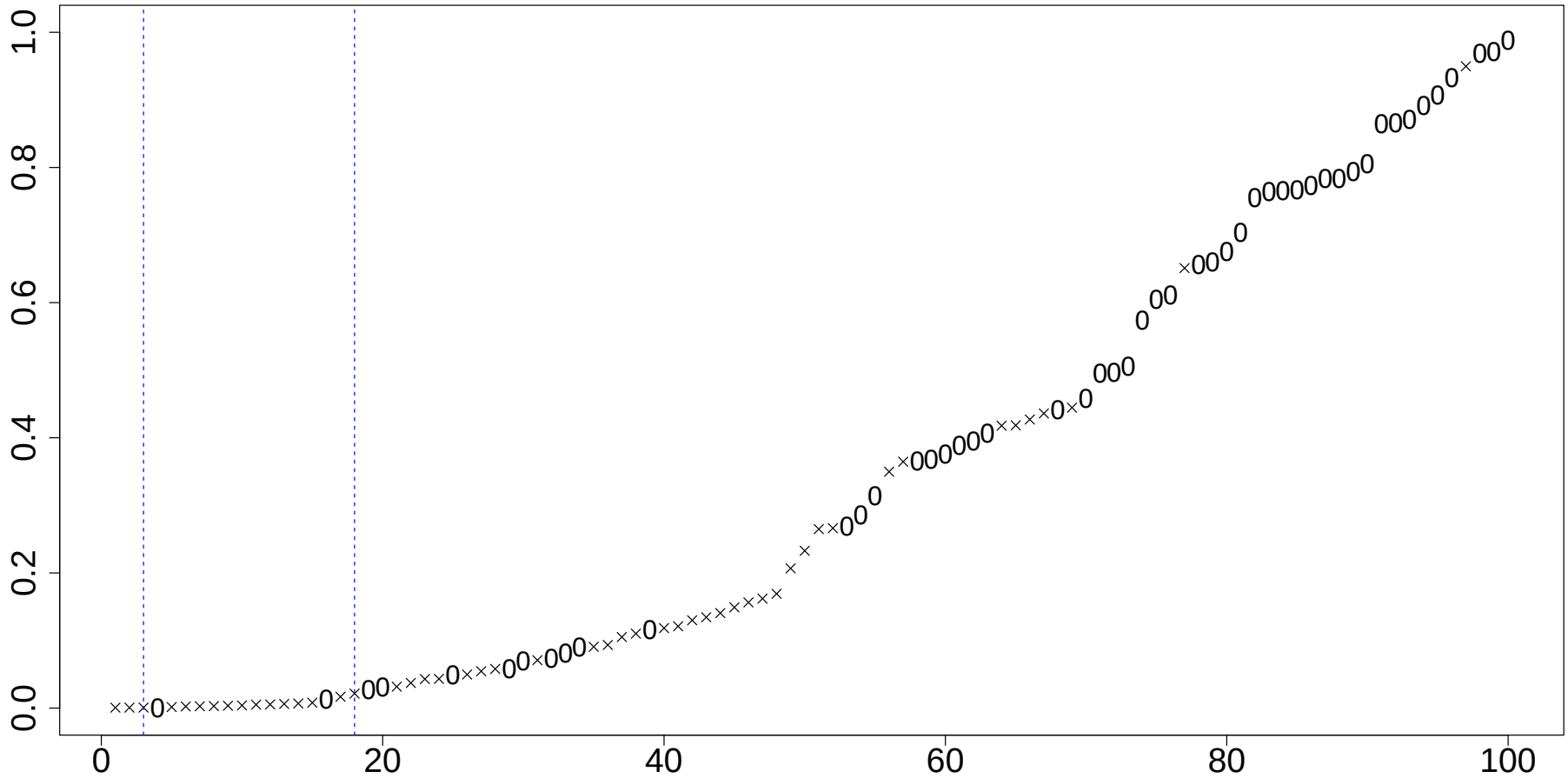
- Stopping rule
- Reject $\widehat{\ell}$ smallest *p*-values

Ordering *p*-values



- Stopping rule
- Reject $\widehat{\ell}$ smallest *p*-values

Choosing procedure?



- "Quality" criterion?

1 Introduction

2 Classical results

3 FDP control

4 Posthoc inference

Family-wise error rate (FWER)

For a threshold t, $V(t) = \sum_{i=1}^{m}(1-\theta_i)\mathbf{1}\{p_i \leq t\}$,

$$\text{FWER}(t,P) = \mathbf{P}(V(t) \geq 1) = \mathbf{P}(p_{(1:\mathcal{H}_0)} \leq t)$$

FWER control

α given, find $t = t_\alpha$ with $\forall P \in \mathcal{P}, \quad \text{FWER}(t,P) \leq \alpha$,

- Clear interpretation
- Bonferroni $t = \alpha/m$ (union bound)
- Possible refinements : $X \sim \mathcal{N}(\mu, \Gamma)$ (one-sided), Γ known

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

  Choose t such that $\mathbf{P}_{Z\sim\mathcal{N}(0,\Gamma)}(\overline{\Phi}(Z)_{(1:m)} \leq t) \leq \alpha$
- Relaxation to k-FWER.

Family-wise error rate (FWER)

For a threshold t, $V(t) = \sum_{i=1}^{m}(1-\theta_i)\mathbf{1}\{p_i \leq t\}$,

$$\mathrm{FWER}(t,P) = \mathbf{P}(V(t) \geq 1) = \mathbf{P}(p_{(1:\mathcal{H}_0)} \leq t)$$

FWER control

α given, find $t = t_\alpha$ with $\forall P \in \mathcal{P}, \quad \mathrm{FWER}(t,P) \leq \alpha$,

- Clear interpretation
- Bonferroni $t = \alpha/m$ (union bound)
- Possible refinements : $X \sim \mathcal{N}(\mu, \Gamma)$ (one-sided), Γ known

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

  Choose t such that $\mathbf{P}_{Z\sim\mathcal{N}(0,\Gamma)}(\overline{\Phi}(Z)_{(1:m)} \leq t) \leq \alpha$
- Relaxation to k-FWER.

Family-wise error rate (FWER)

For a threshold t, $V(t) = \sum_{i=1}^{m}(1-\theta_i)\mathbf{1}\{p_i \leq t\}$,

$$\text{FWER}(t,P) = \mathbf{P}(V(t) \geq 1) = \mathbf{P}(p_{(1:\mathcal{H}_0)} \leq t)$$

FWER control

α given, find $t = t_\alpha$ with $\forall P \in \mathcal{P}, \quad \text{FWER}(t,P) \leq \alpha$,

- Clear interpretation
- Bonferroni $t = \alpha/m$ (union bound)
- Possible refinements : $X \sim \mathcal{N}(\mu, \Gamma)$ (one-sided), Γ known

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

Choose t such that $\mathbf{P}_{Z\sim\mathcal{N}(0,\Gamma)}(\overline{\Phi}(Z)_{(1:m)} \leq t) \leq \alpha$

- Relaxation to k-FWER.

Family-wise error rate (FWER)

For a threshold t, $V(t) = \sum_{i=1}^{m} (1 - \theta_i) \mathbf{1}\{p_i \leq t\}$,

$$\text{FWER}(t, P) = \mathbf{P}(V(t) \geq 1) = \mathbf{P}(p_{(1:\mathcal{H}_0)} \leq t)$$

FWER control

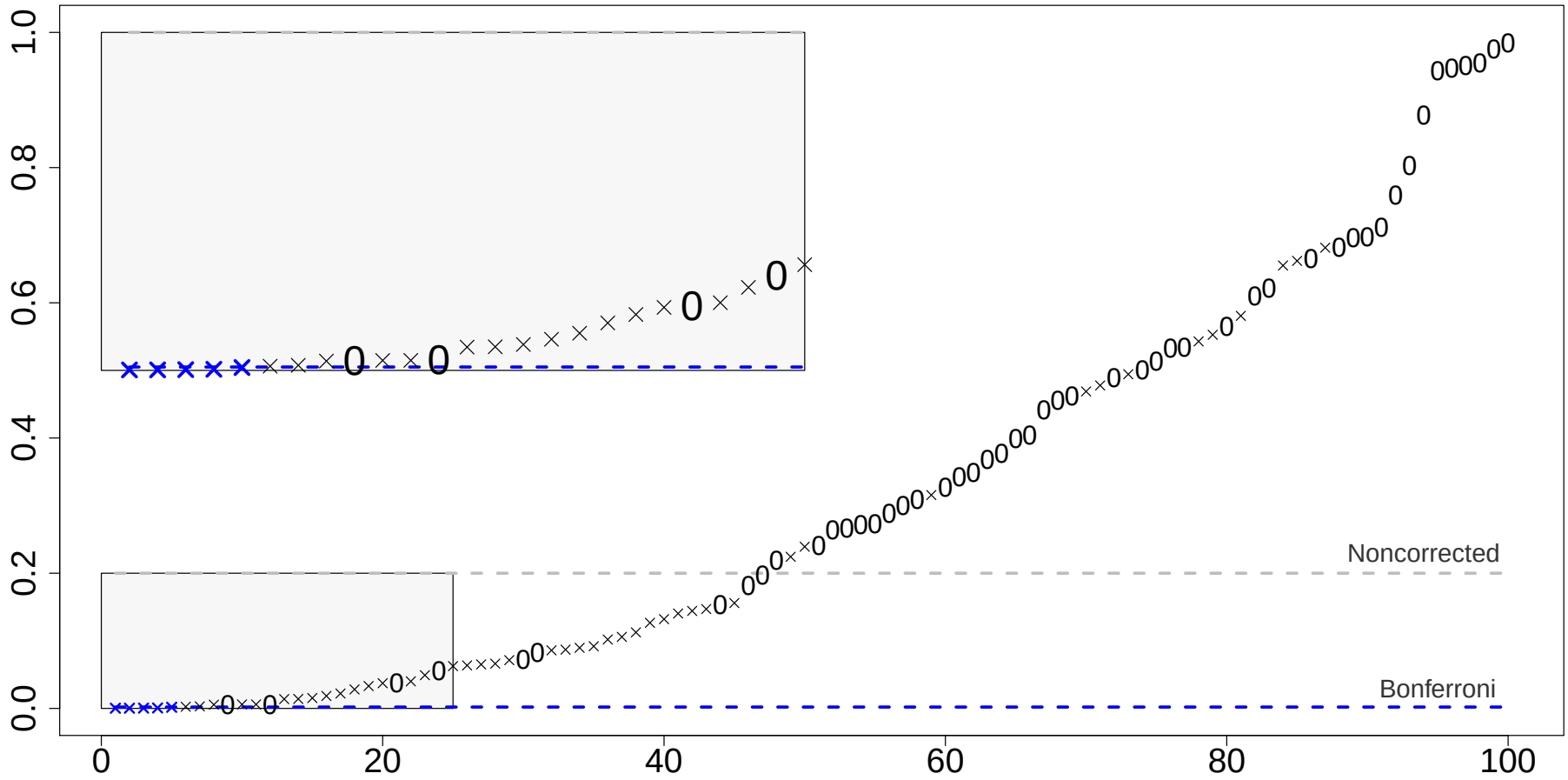
α given, find $t = t_\alpha$ with $\forall P \in \mathcal{P}, \quad \text{FWER}(t, P) \leq \alpha$,

- Clear interpretation
- Bonferroni $t = \alpha/m$ (union bound)
- Possible refinements : $X \sim \mathcal{N}(\mu, \Gamma)$ (one-sided), Γ known

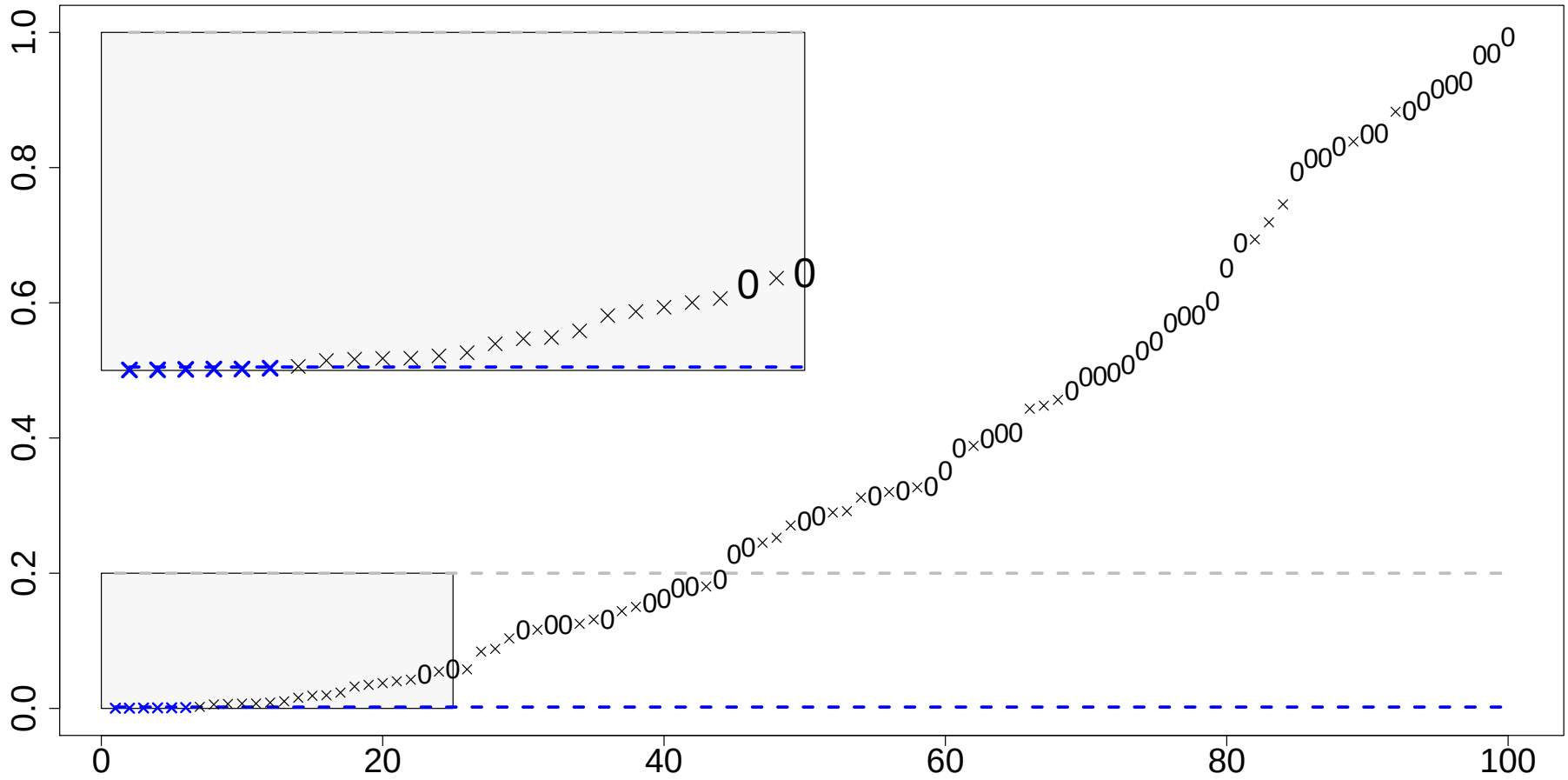
$$V(t) \leq V'(t) = \sum_{i=1}^{m} \mathbf{1}\{\overline{\Phi}(Z_i) \leq t\}, \quad Z \sim \mathcal{N}(0, \Gamma)$$

  Choose t such that $\mathbf{P}_{Z \sim \mathcal{N}(0,\Gamma)}(\overline{\Phi}(Z)_{(1:m)} \leq t) \leq \alpha$
- Relaxation to k-FWER.

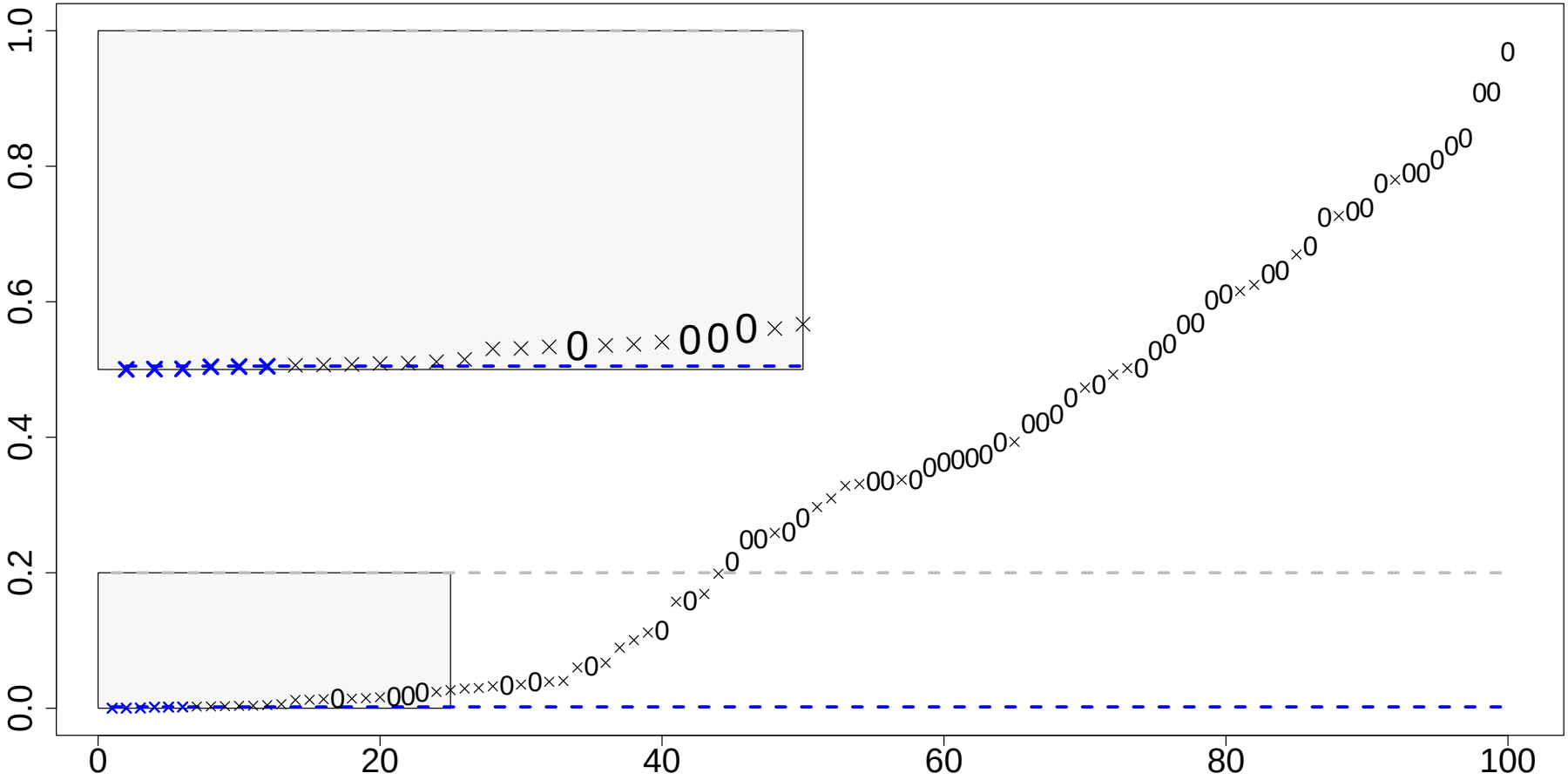
FWER(Bonf.) illustration, $\alpha = 0.2$, $m = 100$



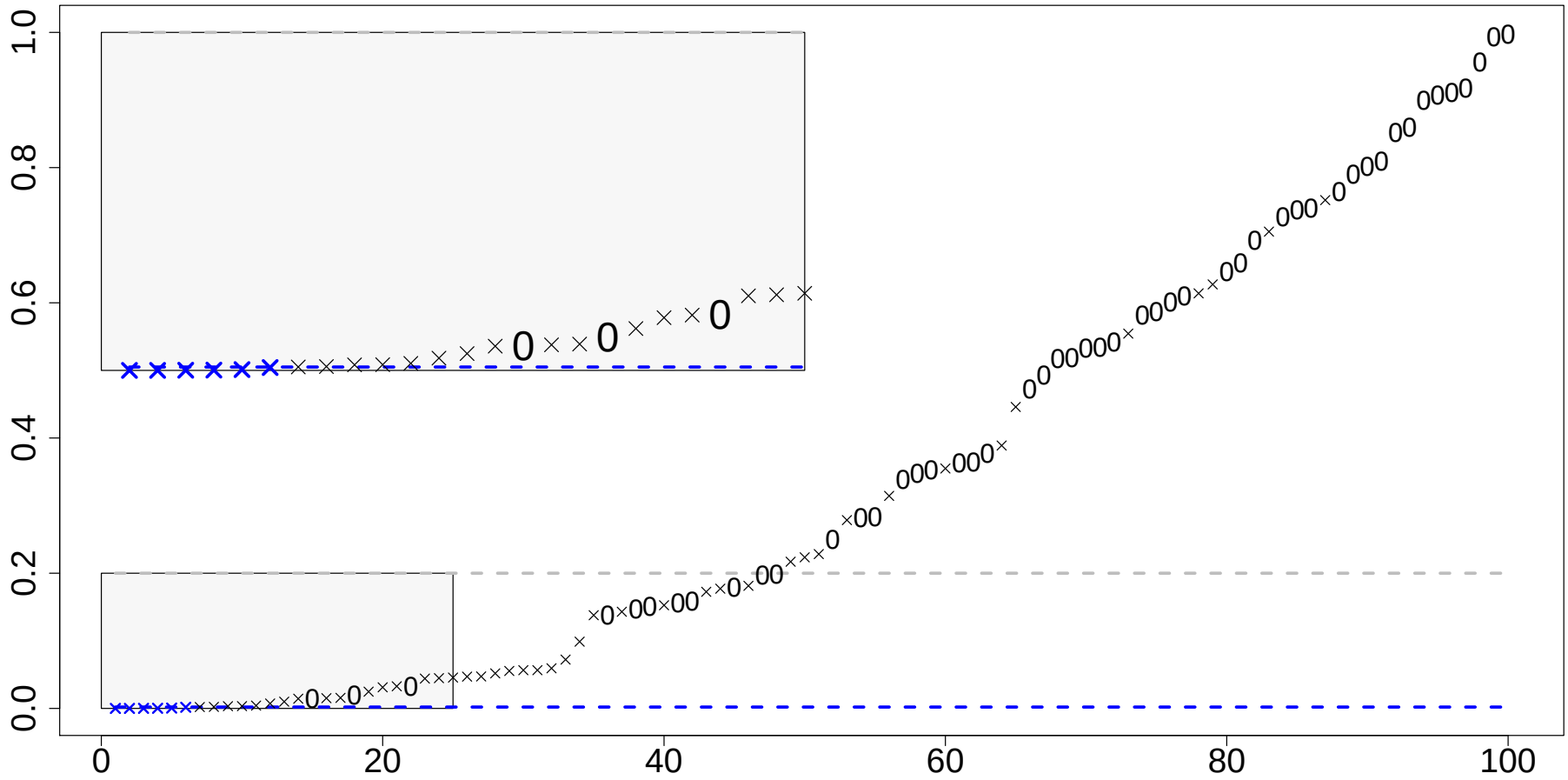
FWER(Bonf.) illustration, $\alpha = 0.2$, $m = 100$



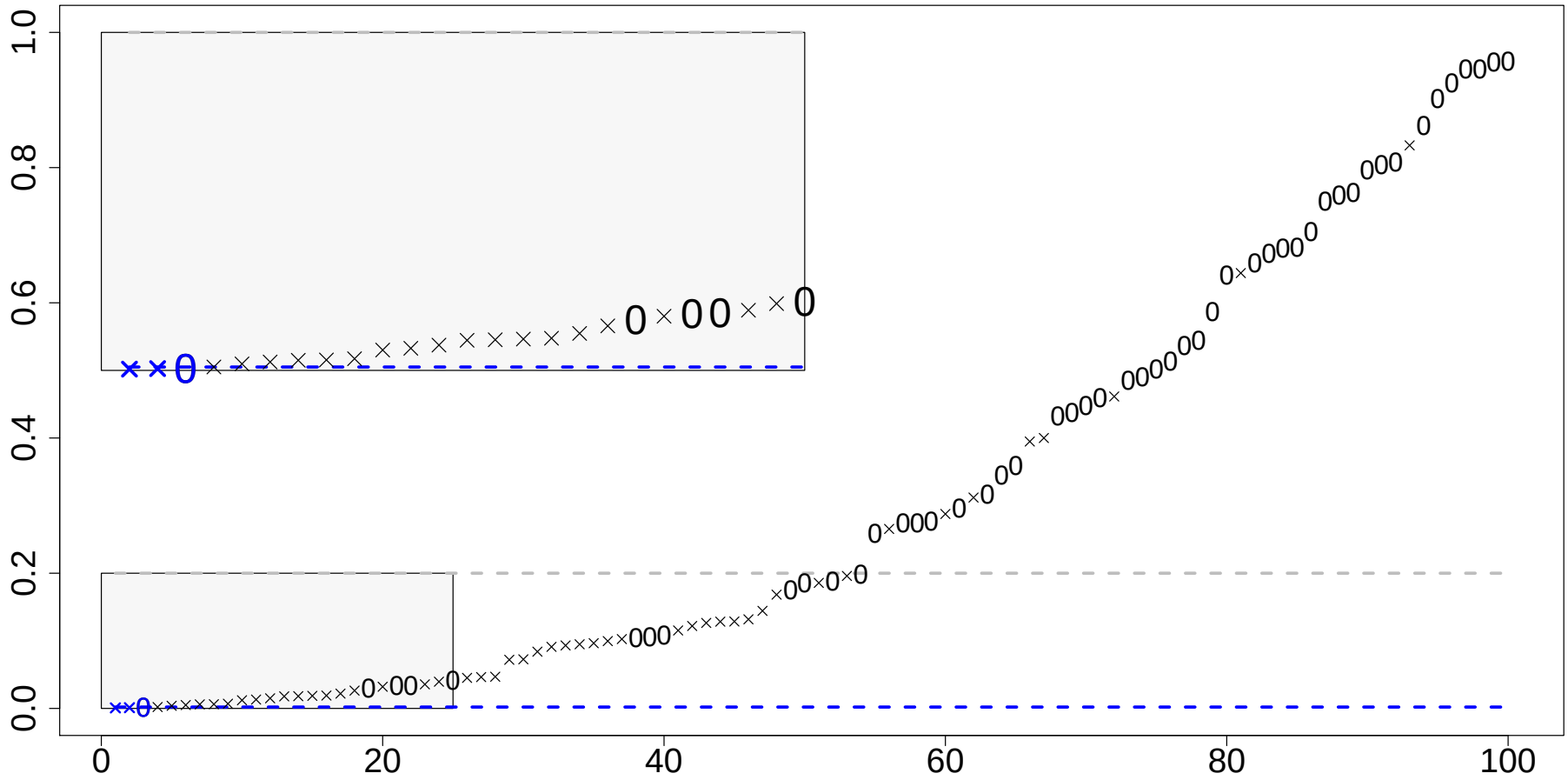
FWER(Bonf.) illustration, $\alpha = 0.2$, $m = 100$



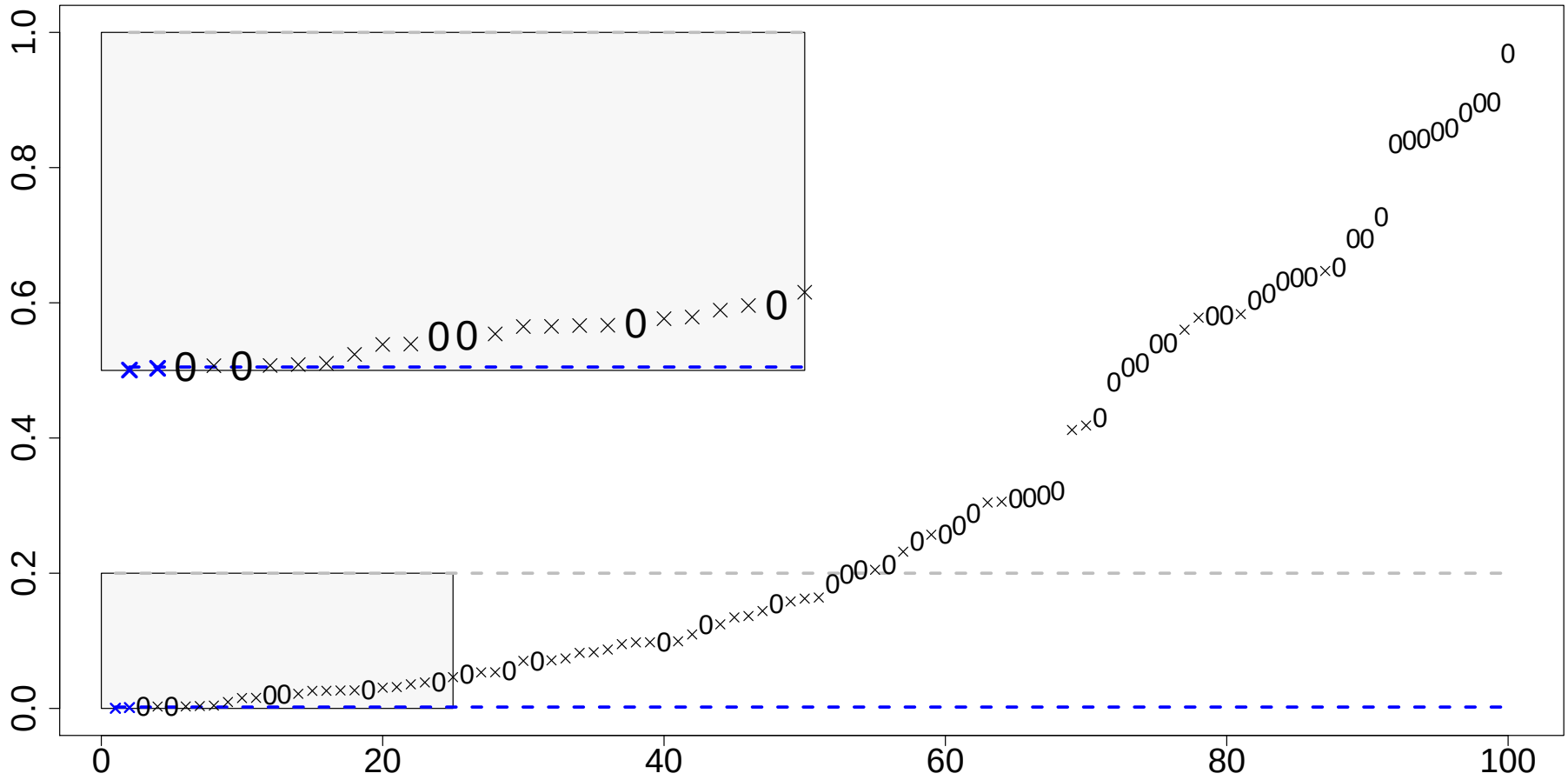
FWER(Bonf.) illustration, $\alpha = 0.2$, $m = 100$



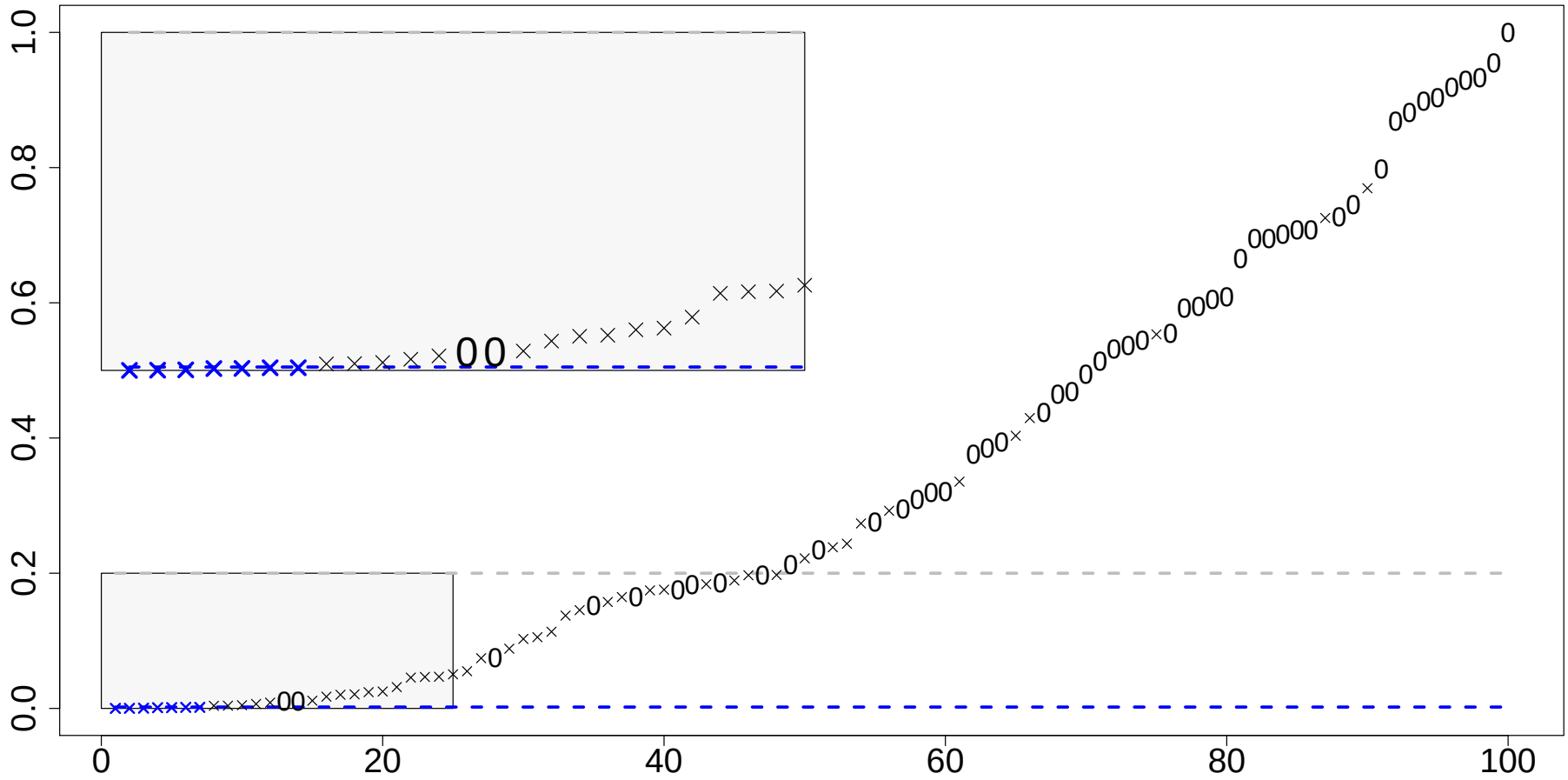
FWER(Bonf.) illustration, $\alpha = 0.2$, $m = 100$



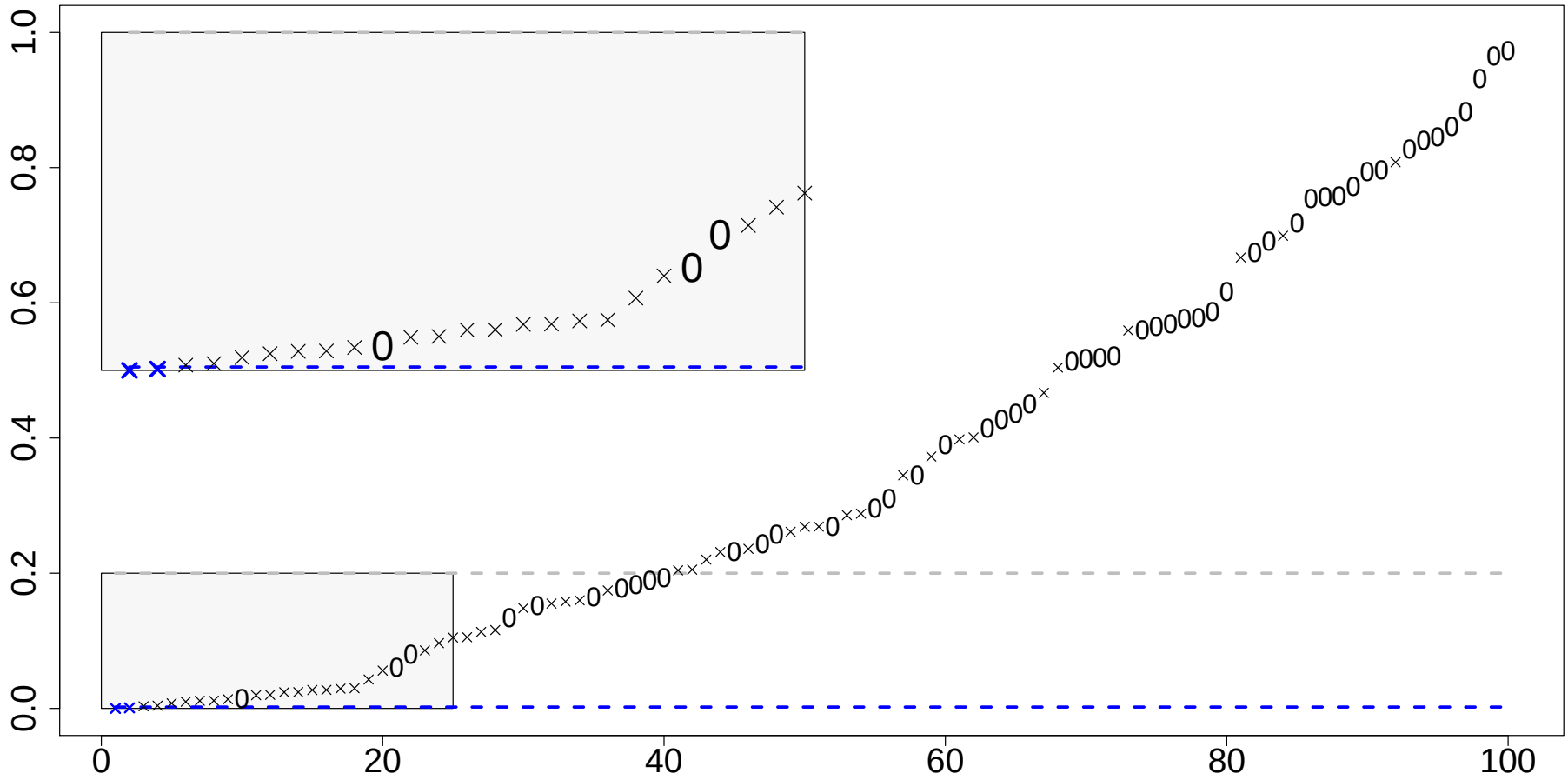
FWER(Bonf.) illustration, $\alpha = 0.2$, $m = 100$



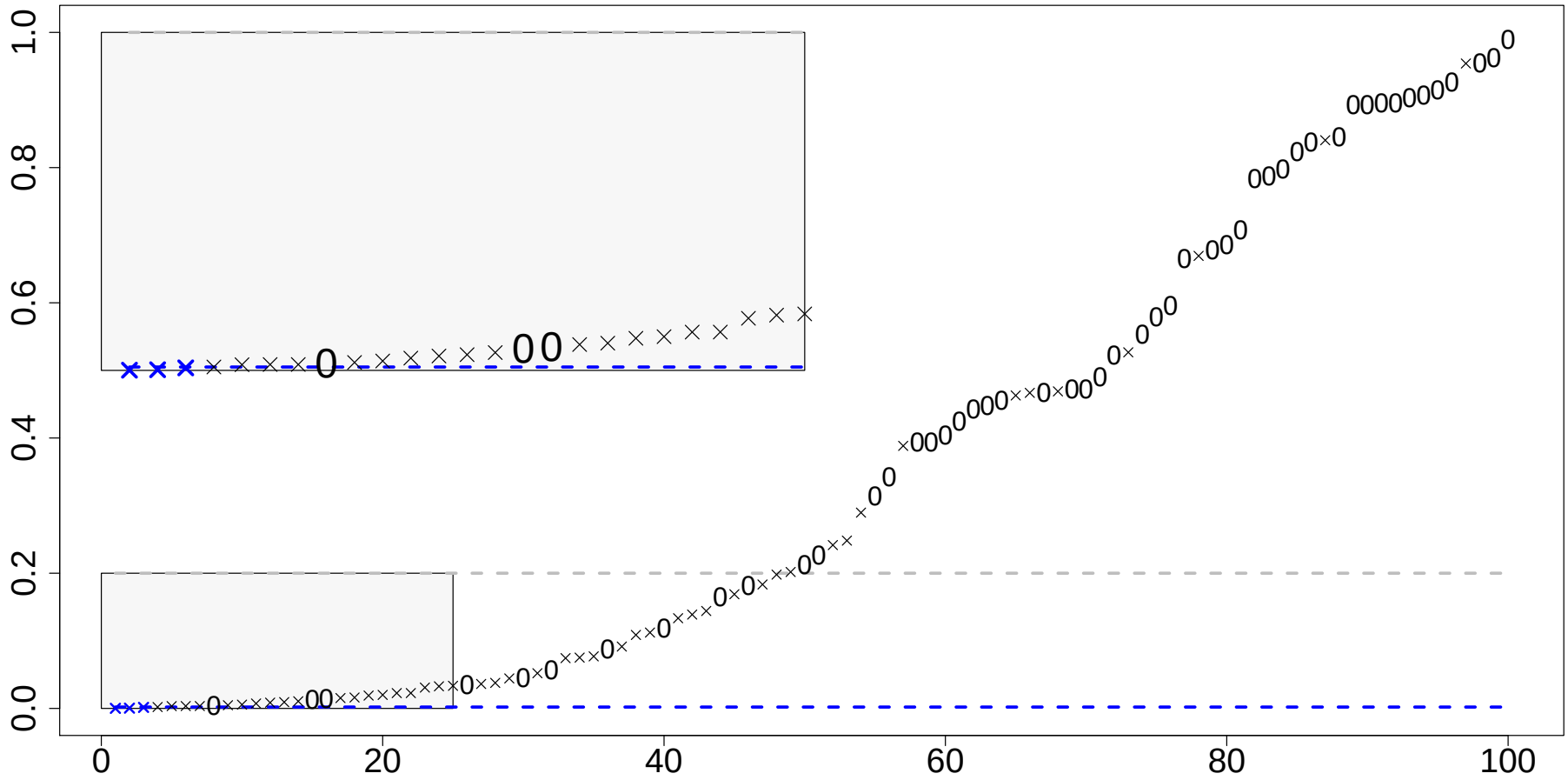
FWER(Bonf.) illustration, $\alpha = 0.2$, $m = 100$



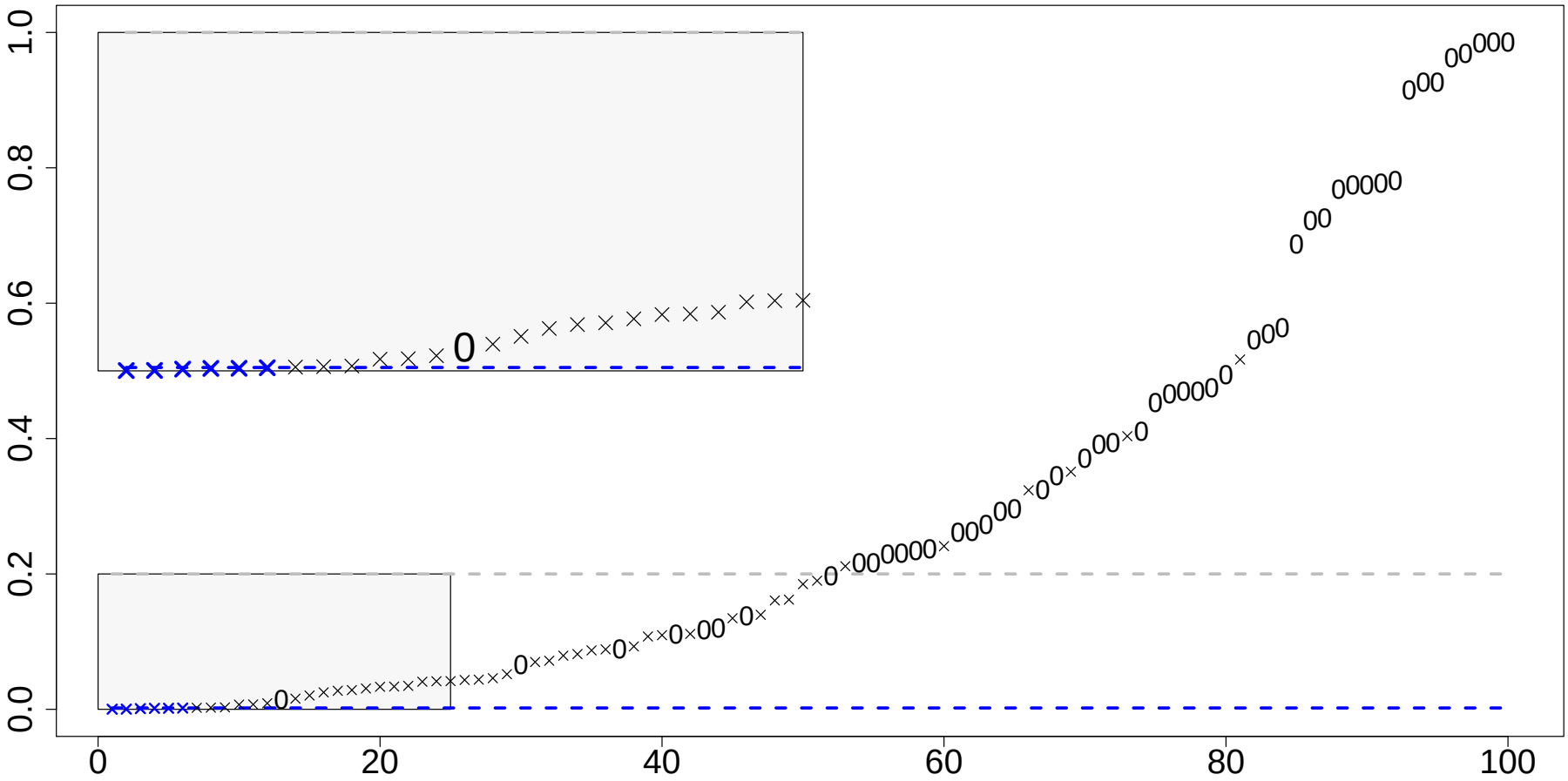
FWER(Bonf.) illustration, $\alpha = 0.2$, $m = 100$



FWER(Bonf.) illustration, $\alpha = 0.2$, $m = 100$



FWER(Bonf.) illustration, $\alpha = 0.2$, $m = 100$



False discovery rate (FDR)

What should we choose for k ?

- 4 false positives out of 10 positives?
- 7 false positives out of 100 positives?

Idea : link k to the number of positives.

For a threshold $\widehat{t}$, $R(\widehat{t}) = \sum_{i=1}^{m} \mathbf{1}\{p_i \leq \widehat{t}\}$,

$$\mathrm{FDR}(\widehat{t}, P) = \mathbf{E}[\mathrm{FDP}(\widehat{t}, P)], \quad \mathrm{FDP}(\widehat{t}, P) = \frac{V(t)}{R(t)} \qquad \left(\frac{0}{0} = 0\right)$$

FDR control

α given, find $\widehat{t} = \widehat{t}_\alpha$ with $\forall P \in \mathcal{P}, \quad \mathrm{FDR}(\widehat{t}, P) \leq \alpha$.

If FDR ≤ 0.05:

- if 20 rejections, allow at most 1 false discovery (on average)
- if 1000 rejections, allow at most 50 false discoveries (on average)

False discovery rate (FDR)

What should we choose for k ?

- 4 false positives out of 10 positives?
- 7 false positives out of 100 positives?

Idea : link k to the number of positives.

For a threshold $\widehat{t}$, $R(\widehat{t}) = \sum_{i=1}^{m} \mathbf{1}\{p_i \leq \widehat{t}\}$,

$$\mathrm{FDR}(\widehat{t}, P) = \mathbf{E}[\mathrm{FDP}(\widehat{t}, P)], \quad \mathrm{FDP}(\widehat{t}, P) = \frac{V(t)}{R(t)} \qquad \left(\frac{0}{0} = 0\right)$$

FDR control

α given, find $\widehat{t} = \widehat{t}_\alpha$ with $\forall P \in \mathcal{P}, \quad \mathrm{FDR}(\widehat{t}, P) \leq \alpha$.

If FDR ≤ 0.05:

- if 20 rejections, allow at most 1 false discovery (on average)
- if 1000 rejections, allow at most 50 false discoveries (on average)

False discovery rate (FDR)

What should we choose for k ?

- 4 false positives out of 10 positives?
- 7 false positives out of 100 positives?

Idea : link k to the number of positives.

For a threshold $\widehat{t}$, $R(\widehat{t}) = \sum_{i=1}^{m} \mathbf{1}\{p_i \leq \widehat{t}\}$,

$$\mathrm{FDR}(\widehat{t}, P) = \mathbf{E}[\mathrm{FDP}(\widehat{t}, P)], \quad \mathrm{FDP}(\widehat{t}, P) = \frac{V(t)}{R(t)} \qquad \left(\frac{0}{0} = 0\right)$$

FDR control

α given, find $\widehat{t} = \widehat{t}_\alpha$ with $\forall P \in \mathcal{P}, \quad \mathrm{FDR}(\widehat{t}, P) \leq \alpha$.

If FDR ≤ 0.05:

- if 20 rejections, allow at most 1 false discovery (on average)
- if 1000 rejections, allow at most 50 false discoveries (on average)

False discovery rate (FDR)

What should we choose for k ?

- 4 false positives out of 10 positives?
- 7 false positives out of 100 positives?

Idea : link k to the number of positives.

For a threshold $\widehat{t}$, $R(\widehat{t}) = \sum_{i=1}^{m} \mathbf{1}\{p_i \leq \widehat{t}\}$,

$$\mathrm{FDR}(\widehat{t}, P) = \mathbf{E}[\mathrm{FDP}(\widehat{t}, P)], \quad \mathrm{FDP}(\widehat{t}, P) = \frac{V(t)}{R(t)} \qquad \left(\frac{0}{0} = 0\right)$$

FDR control

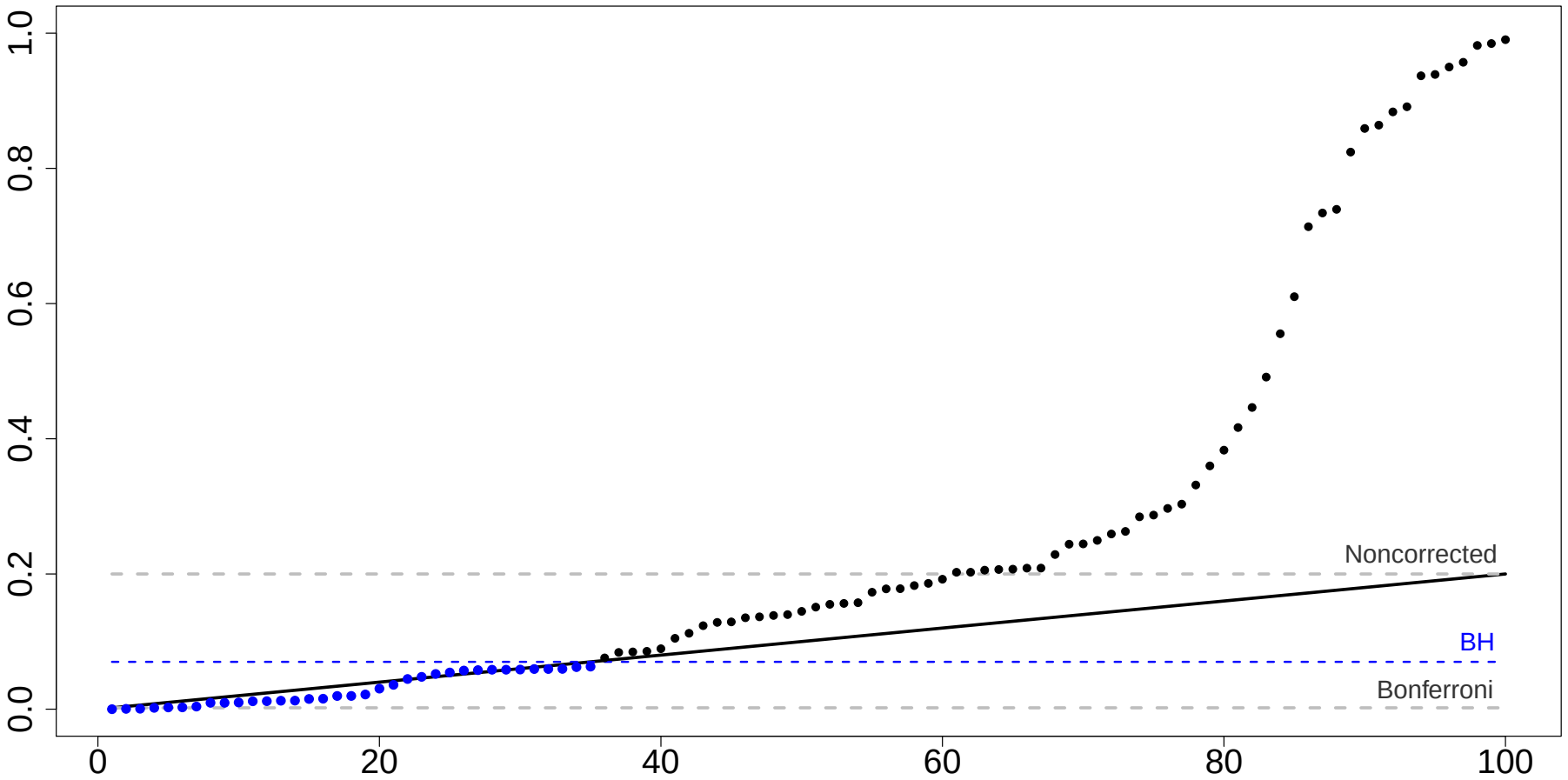
α given, find $\widehat{t} = \widehat{t}_\alpha$ with $\forall P \in \mathcal{P}, \quad \mathrm{FDR}(\widehat{t}, P) \leq \alpha$.

If FDR ≤ 0.05:

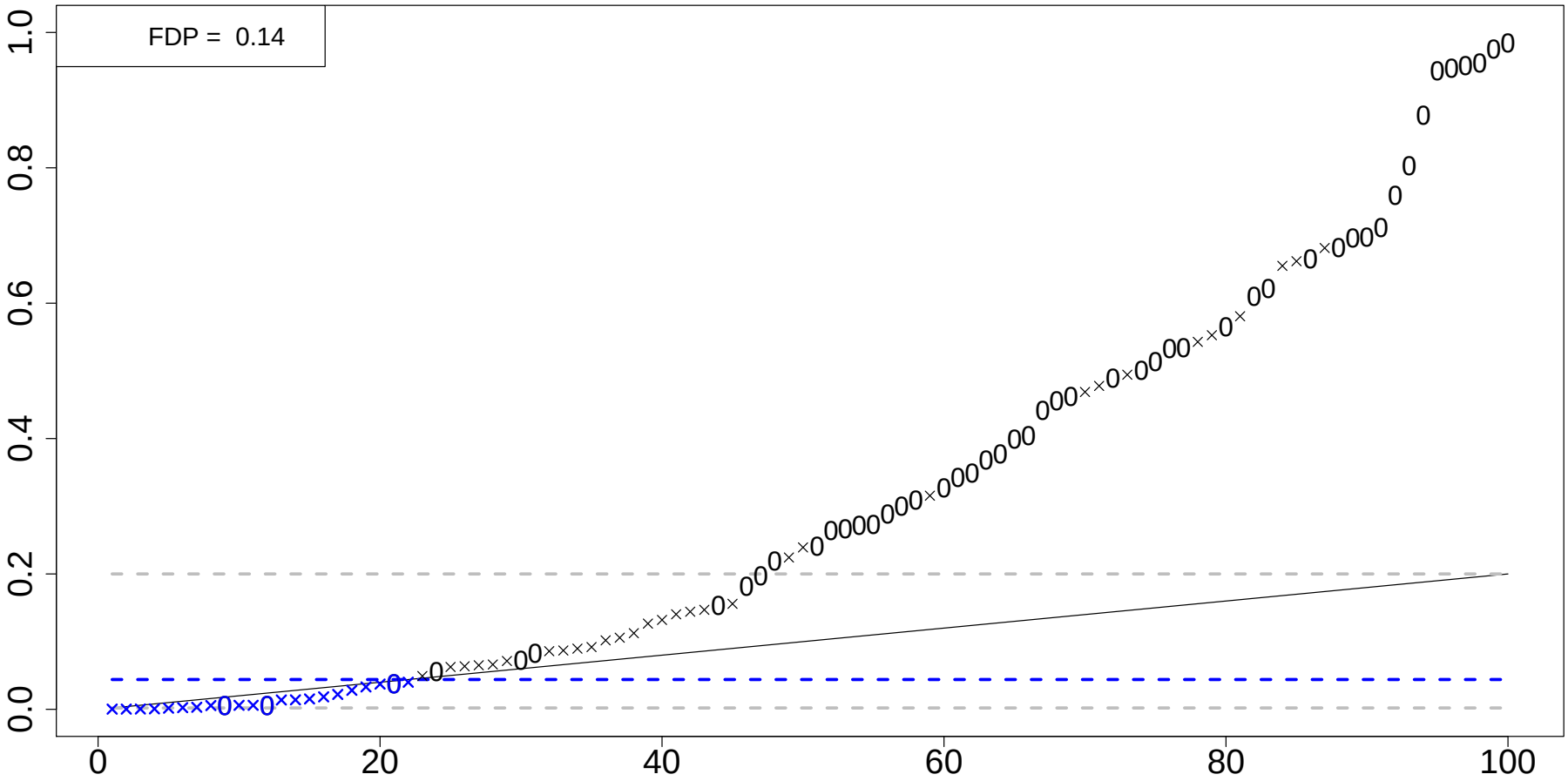
- if 20 rejections, allow at most 1 false discovery (on average)
- if 1000 rejections, allow at most 50 false discoveries (on average)

BH procedure [Benjamini and Hochberg, 1995]

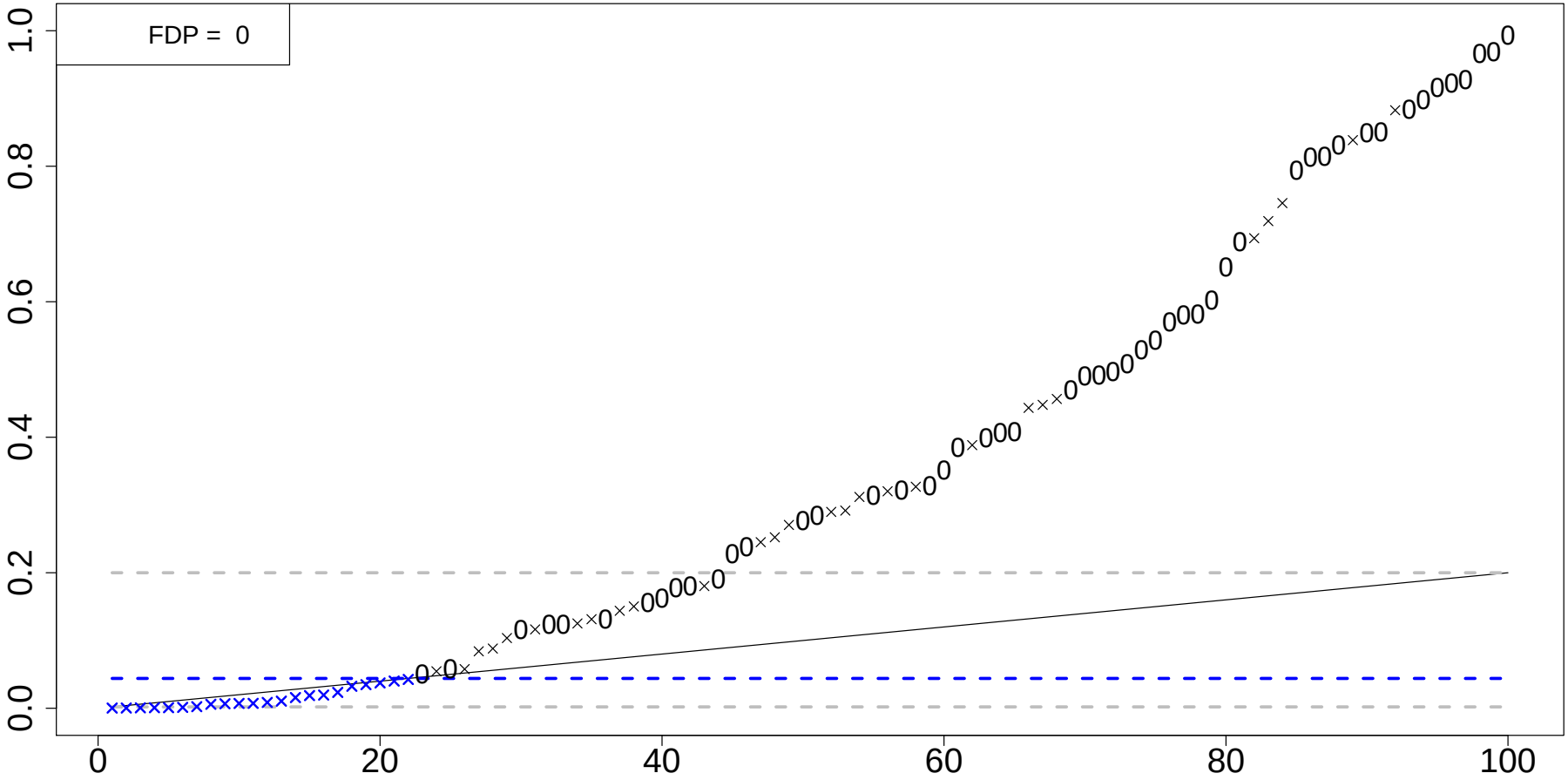
$$\widehat{t} = \alpha(\widehat{\ell} \vee 1)/m \quad \text{with} \quad \widehat{\ell} = \max\{0 \leq \ell \leq m \,:\, p_{(\ell)} \leq \alpha\ell/m\}$$



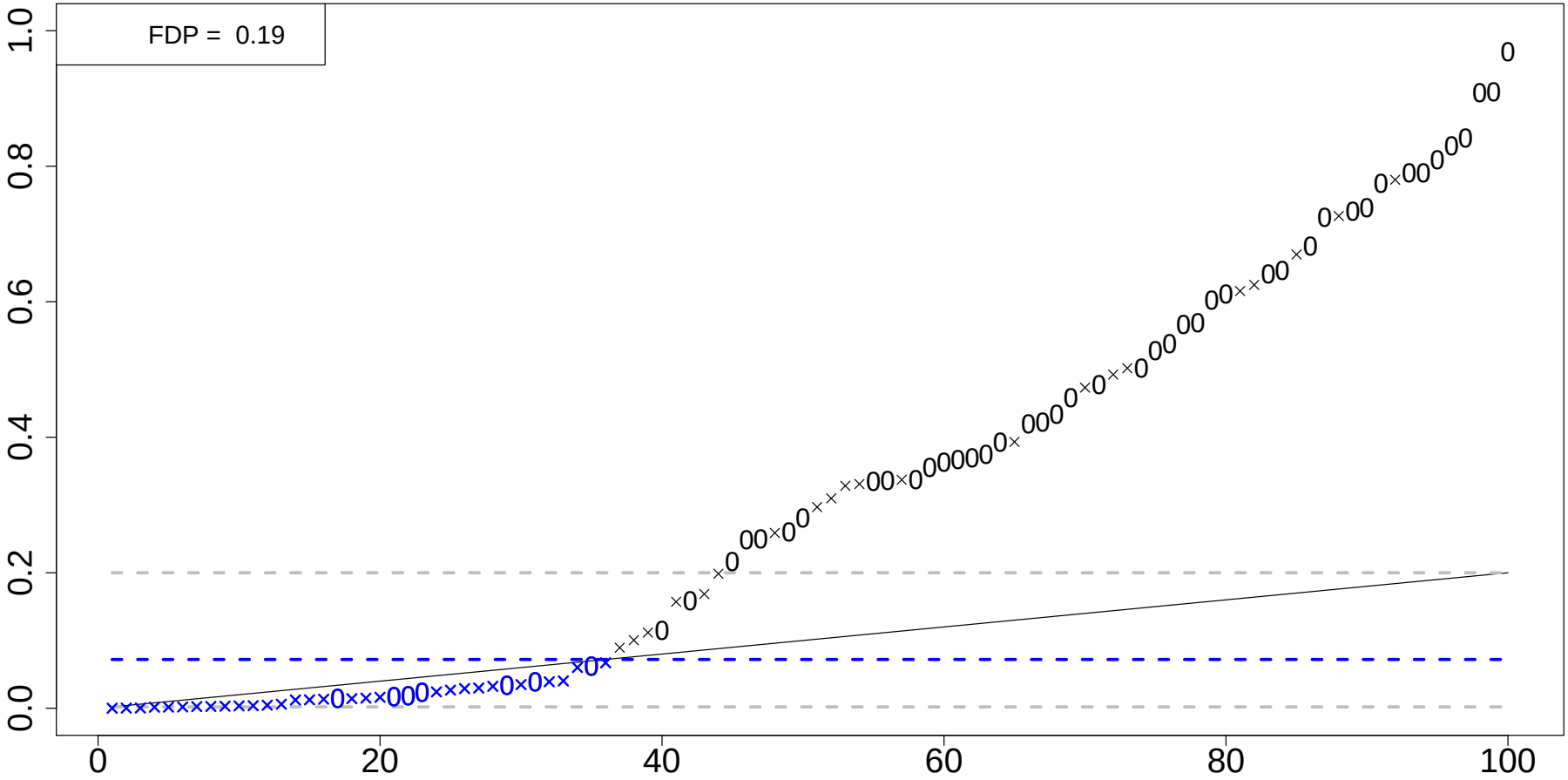
FDR(BH) illustration, $\alpha = 0.2$, $m = 100$



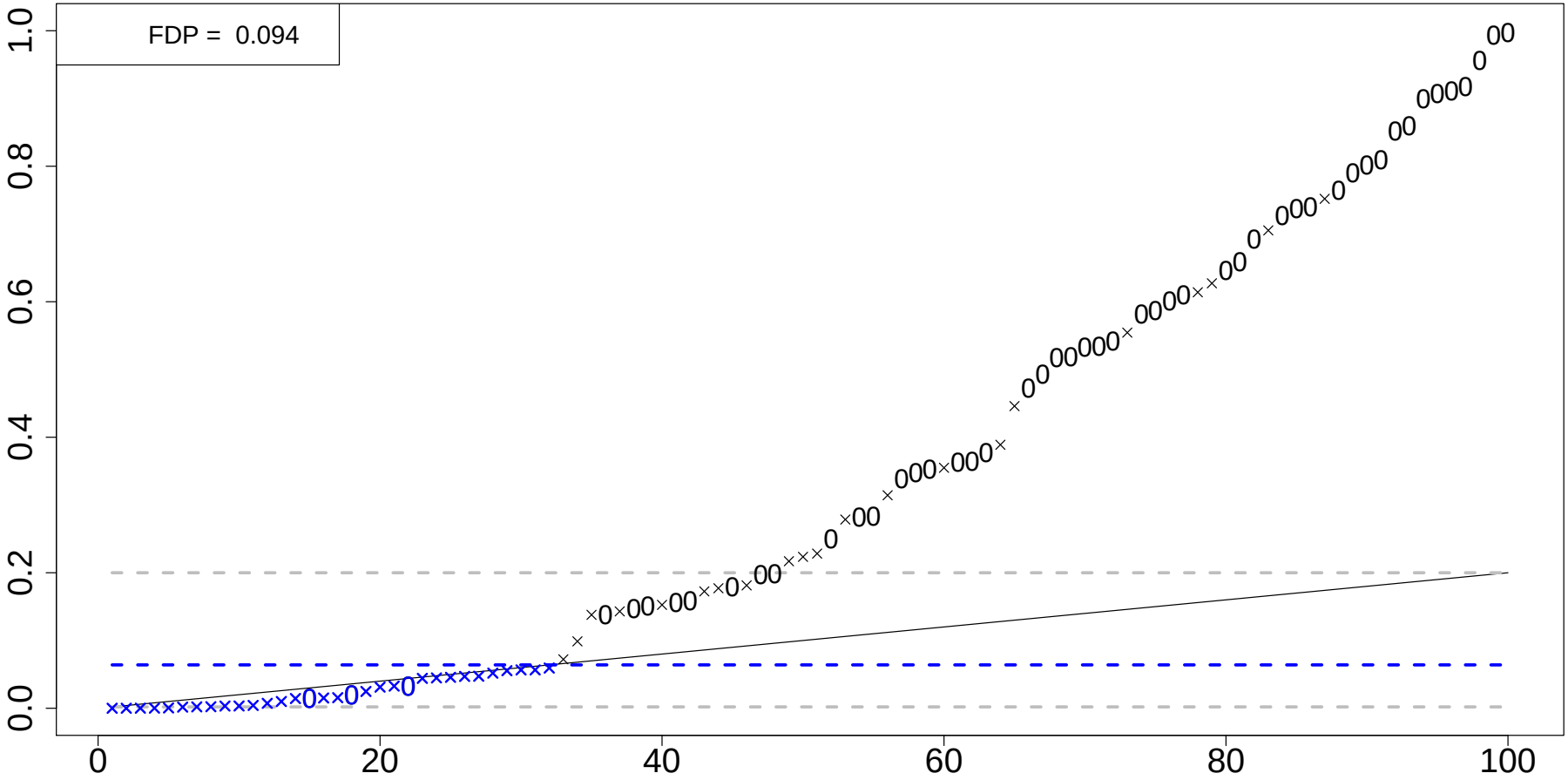
FDR(BH) illustration, $\alpha = 0.2$, $m = 100$



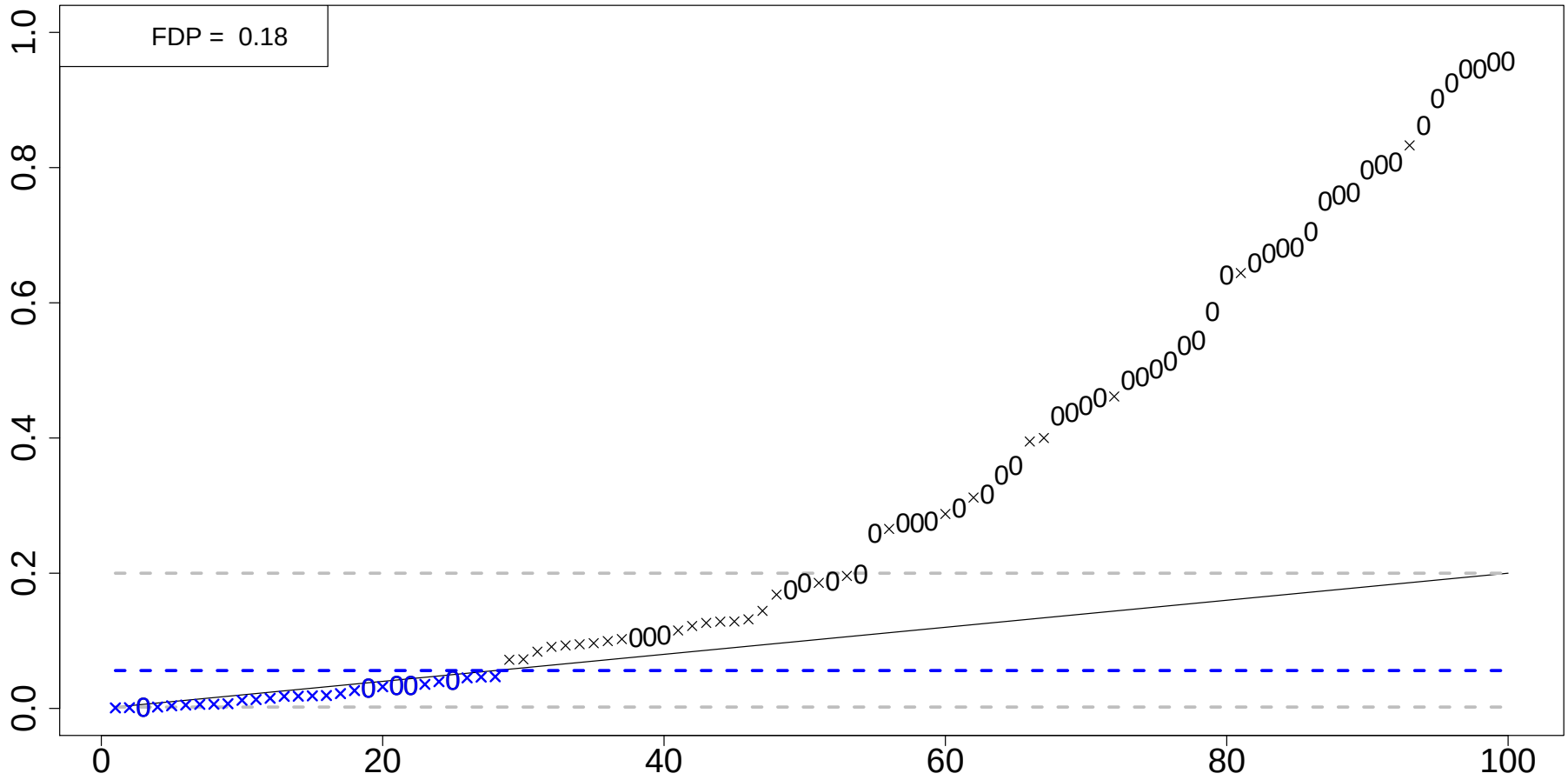
FDR(BH) illustration, $\alpha = 0.2$, $m = 100$



FDR(BH) illustration, $\alpha = 0.2$, $m = 100$



FDR(BH) illustration, $\alpha = 0.2$, $m = 100$



FDR control for BH

Theorem [Benjamini and Hochberg (1995)]

Assume

- $p_i \succeq U(0,1)$ for $\theta_i = 0$
- $(p_i, 1 \leq i \leq m)$ are mutually indep for any $P \in \mathcal{P}$.

Then for all $P \in \mathcal{P}$,

$$\mathrm{FDR}(\mathrm{BH}_\alpha, P) \leq \pi_0 \alpha \leq \alpha$$

for $\pi_0 = m_0/m$ proportion of zeros in $\theta(P)$.

Simple proof ☜ [Blanchard and R. (2008, EJS)]

Lemma :

Let $U \succeq U(0,1)$ and $g : [0,1] \to [0,\infty)$ nonincreasing measurable. Then

$$\mathbf{E}\left(\frac{\mathbf{1}\{U \leq c\, g(U)\}}{g(U)} \mathbf{1}\{g(U) > 0\}\right) \leq c, \quad \text{for any } c > 0.$$

FDR control for BH

Theorem [Benjamini and Hochberg (1995)]

Assume

- $p_i \succeq U(0,1)$ for $\theta_i = 0$
- $(p_i, 1 \leq i \leq m)$ are mutually indep for any $P \in \mathcal{P}$.

Then for all $P \in \mathcal{P}$,

$$\mathrm{FDR}(\mathrm{BH}_\alpha, P) \leq \pi_0 \alpha \leq \alpha$$

for $\pi_0 = m_0/m$ proportion of zeros in $\theta(P)$.

Simple proof ☜ [Blanchard and R. (2008, EJS)]

Lemma :

Let $U \succeq U(0,1)$ and $g : [0,1] \to [0,\infty)$ nonincreasing measurable. Then

$$\mathbf{E}\left(\frac{\mathbf{1}\{U \leq c\, g(U)\}}{g(U)} \mathbf{1}\{g(U) > 0\}\right) \leq c, \quad \text{for any } c > 0.$$

FDR control for BH

Theorem [Benjamini and Hochberg (1995)]

Assume

- $p_i \succeq U(0,1)$ for $\theta_i = 0$
- $(p_i, 1 \leq i \leq m)$ are mutually indep for any $P \in \mathcal{P}$.

Then for all $P \in \mathcal{P}$,

$$\mathrm{FDR}(\mathrm{BH}_\alpha, P) \leq \pi_0 \alpha \leq \alpha$$

for $\pi_0 = m_0/m$ proportion of zeros in $\theta(P)$.

Simple proof ☜ [Blanchard and R. (2008, EJS)]

Lemma :

Let $U \succeq U(0,1)$ and $g : [0,1] \to [0,\infty)$ nonincreasing measurable. Then

$$\mathbf{E}\left(\frac{\mathbf{1}\{U \leq c\, g(U)\}}{g(U)} \mathbf{1}\{g(U) > 0\}\right) \leq c, \quad \text{for any } c > 0.$$

Extension to positive dependence

Theorem [Benjamini and Yekutieli (2001)]

For all $P \in \mathcal{P}$, $\mathrm{FDR}(\mathrm{BH}_\alpha, P) \leq \pi_0 \alpha$ if

- $p_i \succeq U(0,1)$ for $\theta_i = 0$
- $\mathbf{p} = (p_i, 1 \leq i \leq m)$ are weak PRDS : $\forall P \in \mathcal{P}, \forall i$

$$\forall D \subset [0,1]^m \nearrow, \quad u \mapsto \mathbf{P}(\mathbf{p} \in D \,|\, p_i \leq u) \nearrow$$

Simple proof ☜ [Blanchard and R. (2008, EJS)]

Lemma :

Let $U \succeq U(0,1)$; V such that $\forall v$, $u \mapsto \mathbf{P}(V < v \,|\, U \leq u) \nearrow$ Then

$$\mathbf{E}\left(\frac{\mathbf{1}\{U \leq c\, V\}}{V}\mathbf{1}\{V > 0\}\right) \leq c, \quad \text{for any } c > 0.$$

Extension to positive dependence

Theorem [Benjamini and Yekutieli (2001)]

For all $P \in \mathcal{P}$, $\mathrm{FDR}(\mathrm{BH}_\alpha, P) \leq \pi_0 \alpha$ if

- $p_i \succeq U(0,1)$ for $\theta_i = 0$
- $\mathbf{p} = (p_i, 1 \leq i \leq m)$ are weak PRDS : $\forall P \in \mathcal{P}, \forall i$

$$\forall D \subset [0,1]^m \nearrow, \quad u \mapsto \mathbf{P}(\mathbf{p} \in D \,|\, p_i \leq u) \nearrow$$

Simple proof ☜ [Blanchard and R. (2008, EJS)]

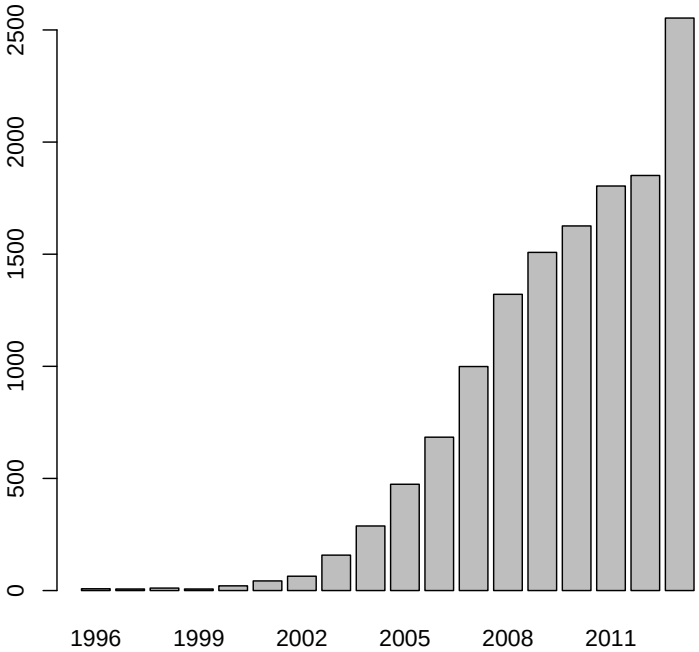
Lemma :

Let $U \succeq U(0,1)$; V such that $\forall v$, $u \mapsto \mathbf{P}(V < v \,|\, U \leq u) \nearrow$ Then

$$\mathbf{E}\left(\frac{\mathbf{1}\{U \leq c\, V\}}{V}\mathbf{1}\{V > 0\}\right) \leq c, \quad \text{for any } c > 0.$$

FDR revolution

Benjamini and Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing



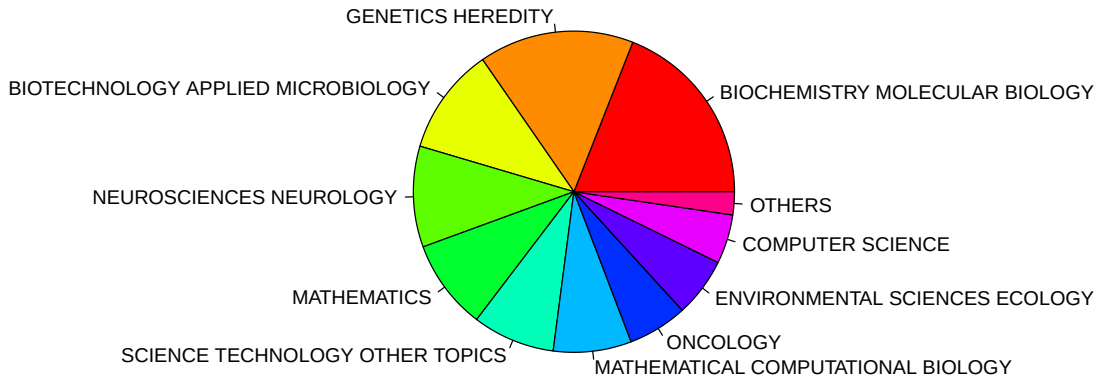


Figure: 13,427 papers citing this work (1996-2013) according to "the web of science".

$\geq 30,085$ citations on google scholar

1 Introduction

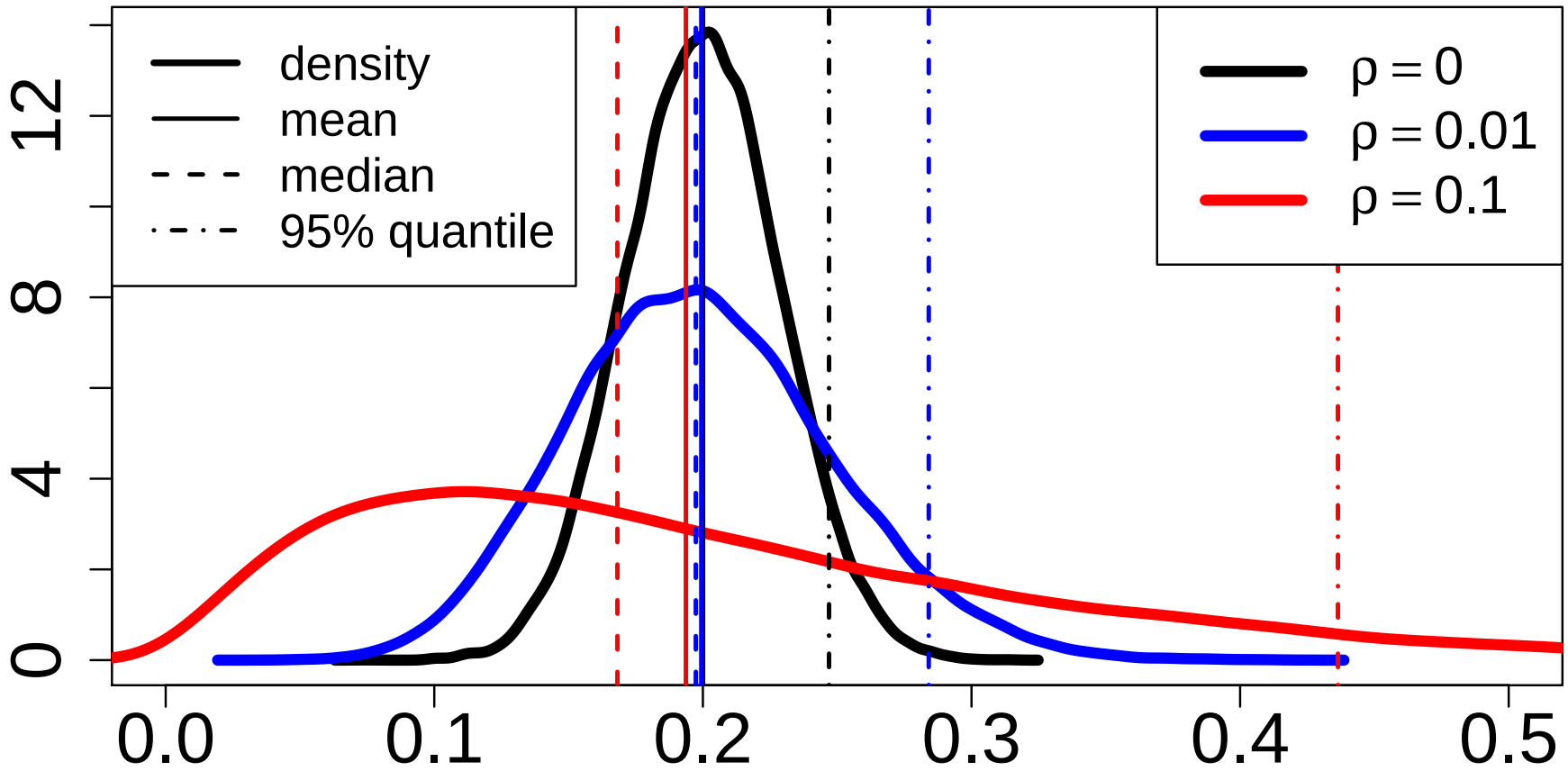
2 Classical results

3 FDP control

4 Posthoc inference

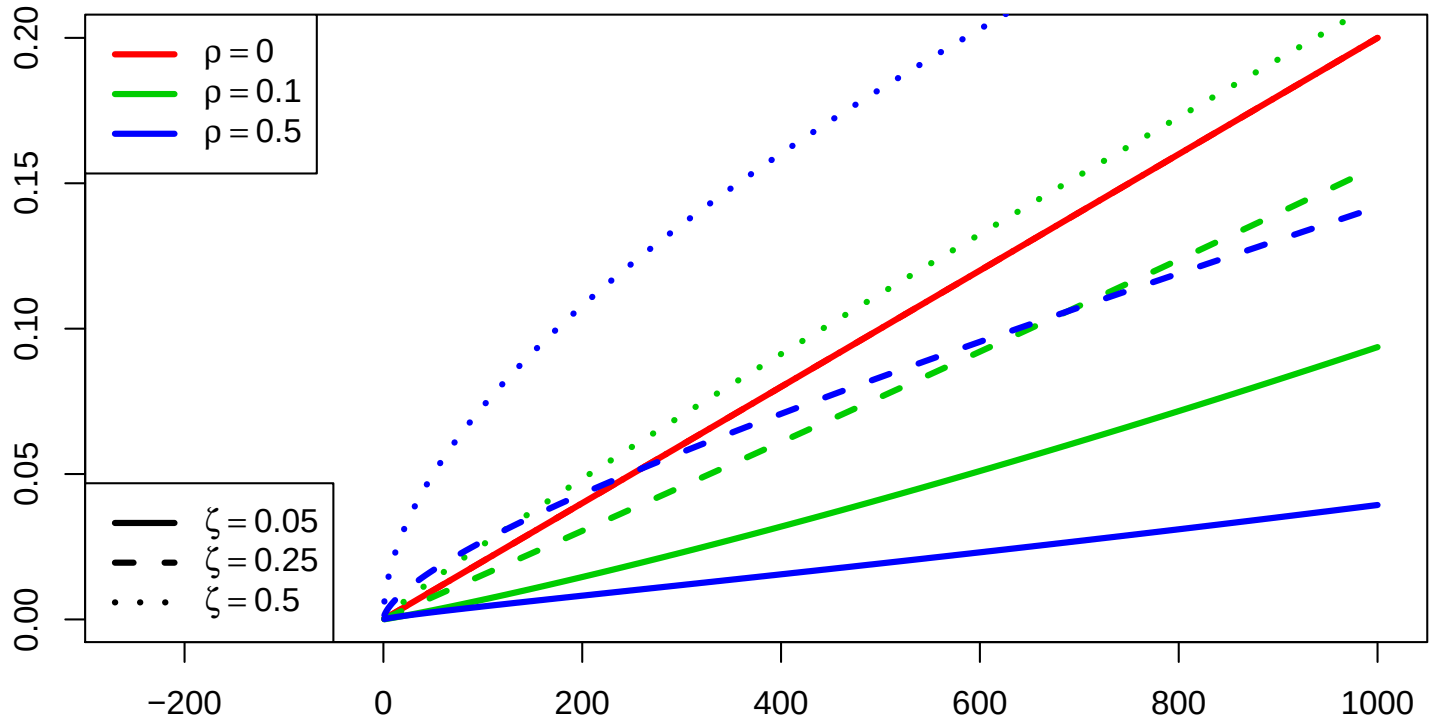
Motivation

Distribution of FDP(BH) under ρ-equicorrelation



Aim

Choose new critical values τ_ℓ, $1 \leq \ell \leq m$ (incorporating dep. structure)



such that $\widehat{t} = \tau_{\hat{\ell}}$ (same way as BH) satisfies

$$\mathbf{P}(FDP(\widehat{t}) \leq \alpha) \geq 1 - \zeta.$$

Method

[Delattre and R. *New procedures controlling the false discovery proportion via Romano-Wolf's heuristic* (2015, AoS)]

Proposition:

- $B(\ell, \ell') = \mathbf{P}(V'(\tau_\ell) \geq \lfloor \alpha \ell' \rfloor + 1)$
- $V'(t) \geq V(t)$ for all t (with $V'(t) \nearrow$ in t and $V'(0) = 0$ a.s.)

Then

$$\mathbf{P}(FDP(\widehat{t}) > \alpha) \leq \sum_{\ell=1}^{m} \big(B(\ell,\ell) - B(\ell-1,\ell)\big) \wedge \big(B(\ell,\ell-1) - B(\ell-1,\ell-1)\big)$$

- Example : $X \sim \mathcal{N}(\mu, \Gamma)$ (one-sided), Γ known

$$B(\ell, \ell') = \mathbf{P}_{Z \sim \mathcal{N}(0,\Gamma)} \left(\sum_{i=1}^{m} \mathbf{1}\{\overline{\Phi}(Z_i) \leq \tau_\ell\} \geq \lfloor \alpha \ell' \rfloor + 1 \right).$$

- Incorporate dependence, much powerful than [Guo et al. (2014, AoS)]

1 Introduction

2 Classical results

3 FDP control

4 Posthoc inference

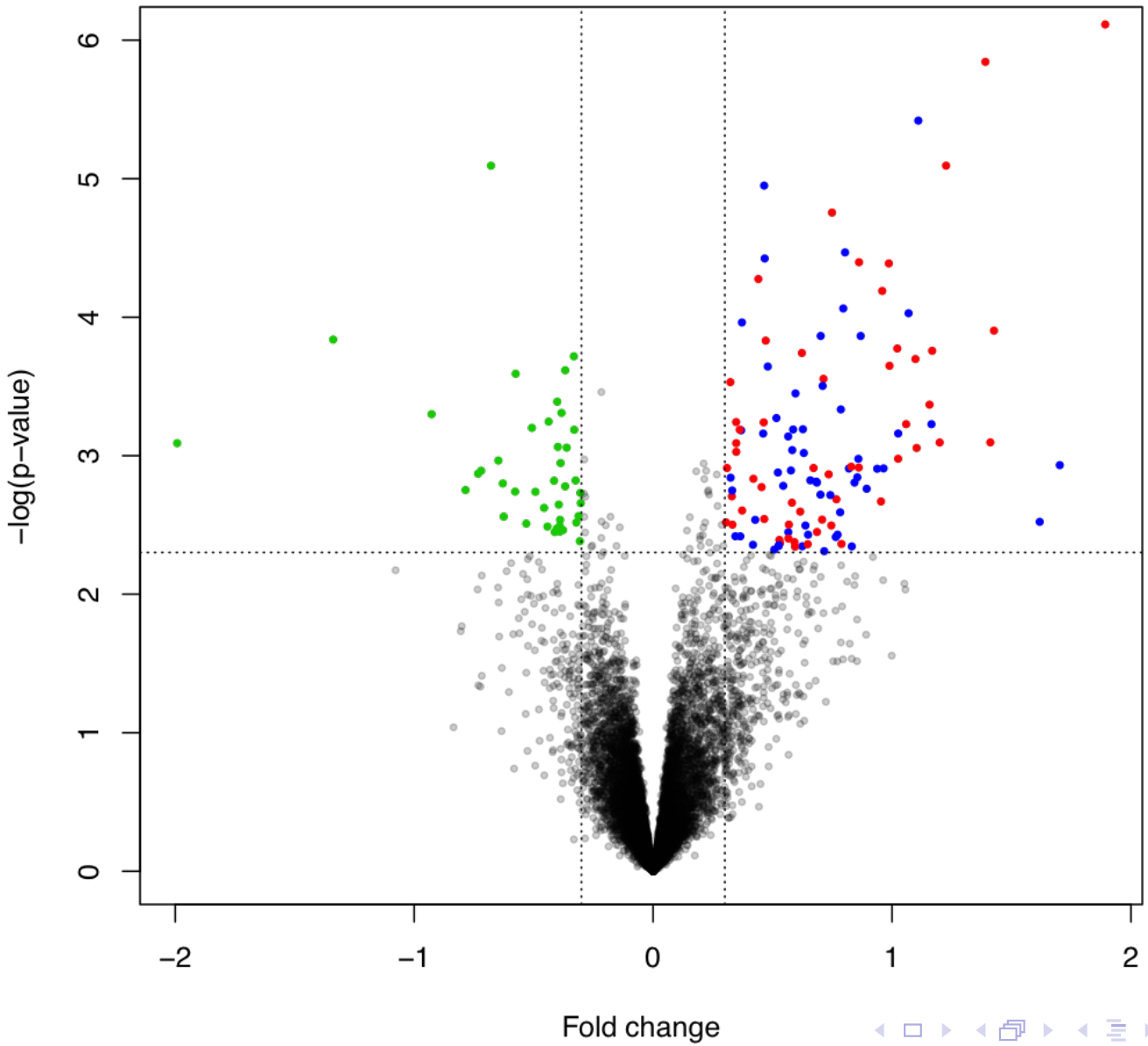
Coming back to the pizza...

Abstract of [Giovannucci et al. (1995)]

"Between 1986 and 1992, 812 new cases of prostate cancer [...] were documented. Of 46 vegetables and fruits or related products, four were significantly associated with lower prostate cancer risk; of the four–tomato sauce (P for trend = .001), tomatoes (P for trend = .03), and pizza (P for trend = .05), but not strawberries–were primary sources of lycopene. Combined intake of tomatoes, tomato sauce, tomato juice, and pizza (which accounted for 82% of lycopene intake) was inversely associated with risk of prostate cancer (multivariate RR = 0.65; 95% CI = 0.44-0.95, for consumption frequency greater than 10 versus less than 1.5 servings per week; P for trend = .01) and advanced (stages C and D) prostate cancers (multivariate RR = 0.47; 95% CI = 0.22-1.00; P for trend = .03)."

How many false positives in that abstract?

Volcano plot [Chiaretti et al. (2004)]

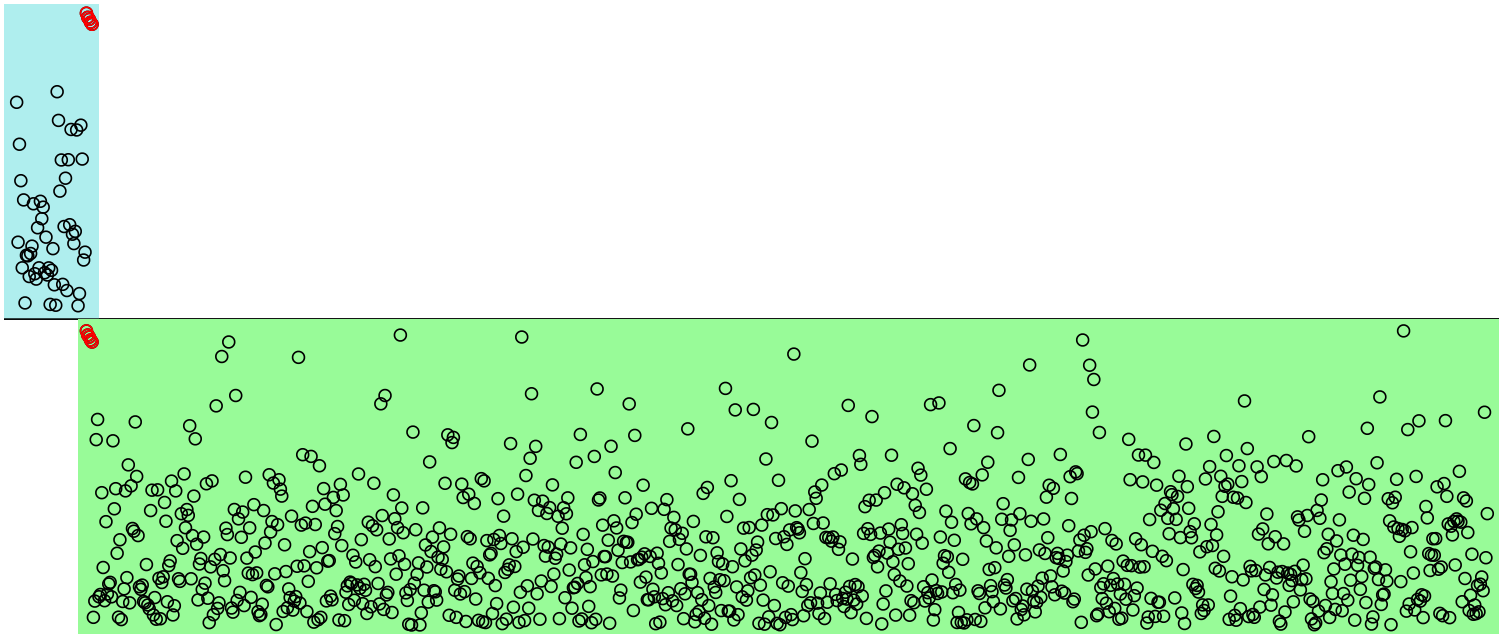


Bias selection

what means statistic values in the selected?

Simulate i.i.d. absolute values of $\mathcal{N}(0,1)$

Data 1 : 50 values



Data2 : 1000 values

Posthoc inference

- Selective inference = after some selection procedure.
  e.g., after LASSO selection [Tibshirani et al. (2015), Barber and Candes (2015,AoS)]
- 'In omnia paratus' posthoc inference = ready for any selection.

Number of false positives in $R \subset \{1, \dots, m\}$:

$$V(R) = \sum_{i=1}^{m} (1 - \theta_i) \mathbf{1}\{i \in R\}.$$

Aim: posthoc bound

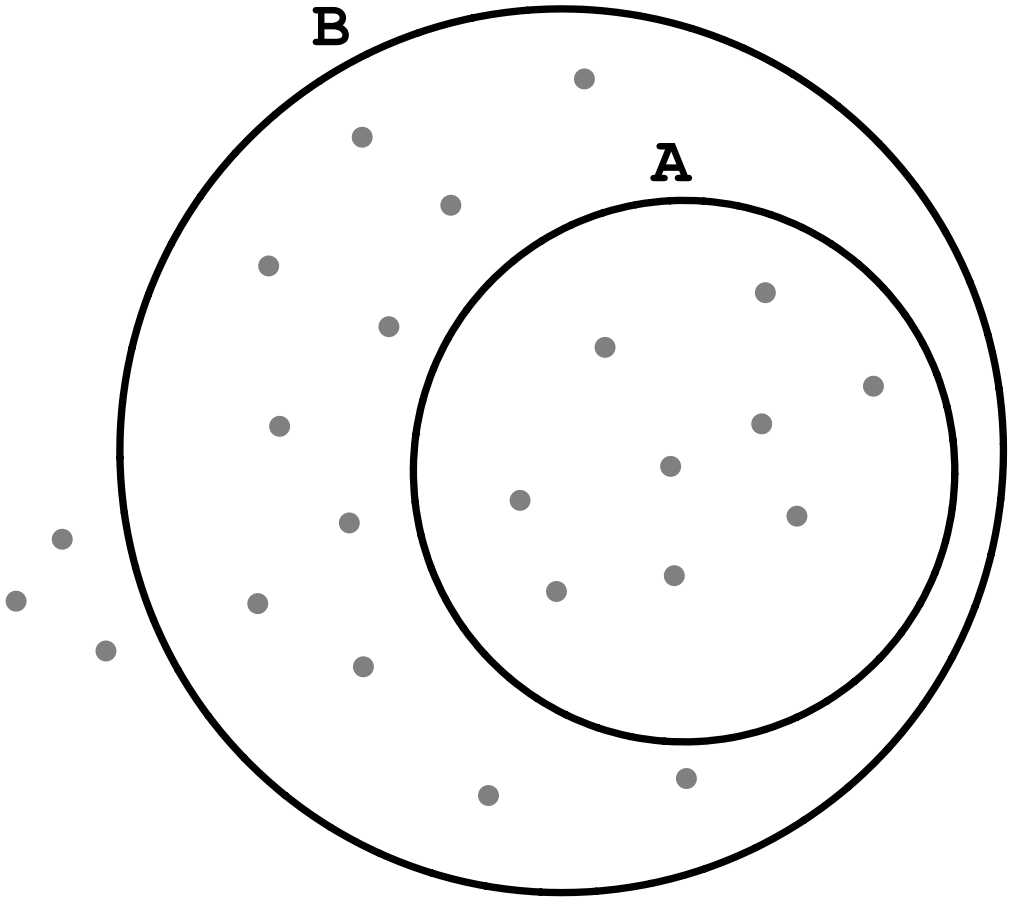
α given, find $\overline{V}_\alpha(\cdot) \in \mathbb{N}$, such that for all $P \in \mathcal{P}$,

$$\mathbf{P}\left(\forall R \subset \{1, \dots, m\} \ : \ V(R) \leq \overline{V}_\alpha(R)\right) \geq 1 - \alpha$$

Proposed in [Goeman and Solari (2011,StatScience)] (with closed testing)

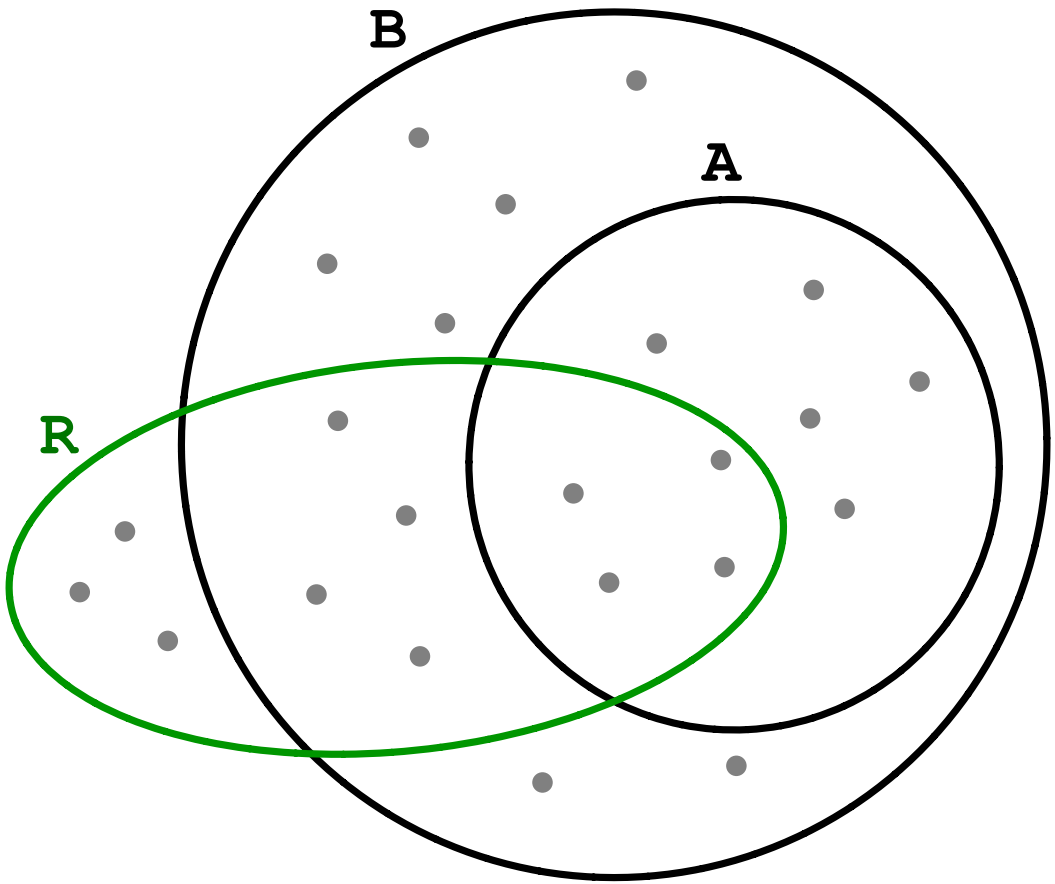
Our method: reference family

If $|A \cap \mathcal{H}_0| \leq 5$, $|B \cap \mathcal{H}_0| \leq 7$



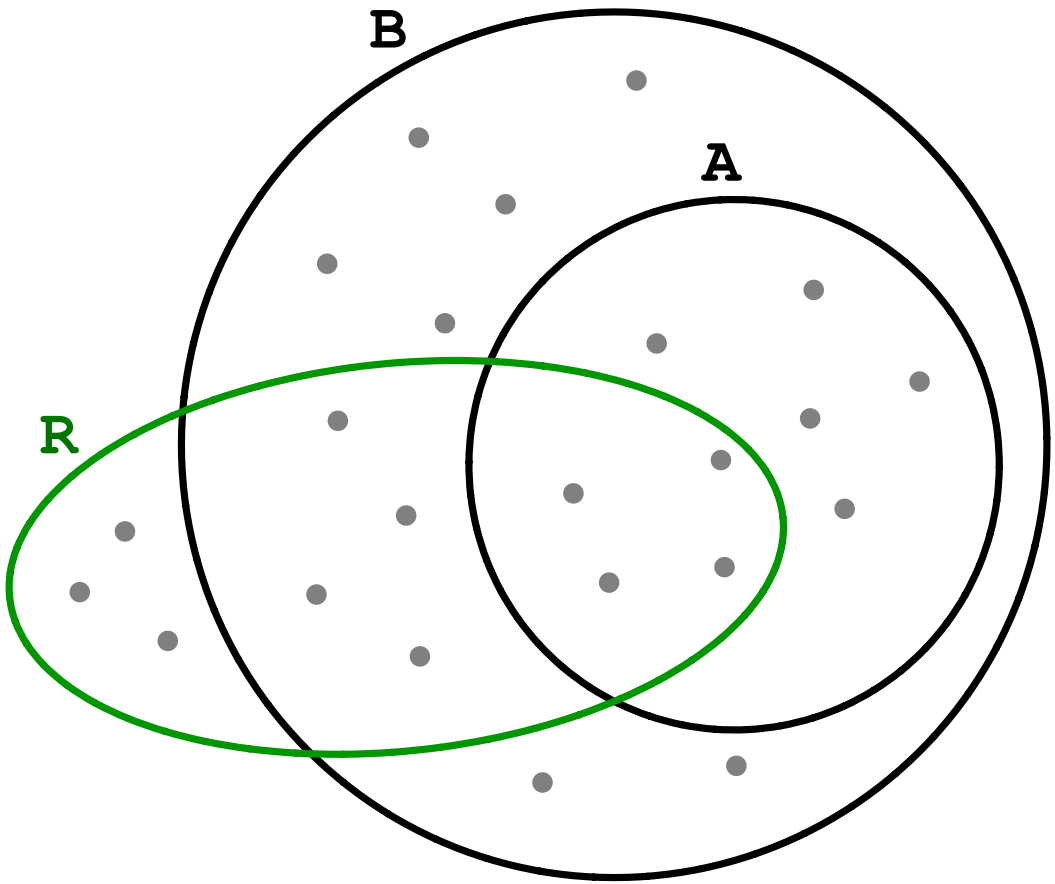
Our method: reference family

If $|A \cap \mathcal{H}_0| \leq 5$, $|B \cap \mathcal{H}_0| \leq 7$ $\quad\quad$ $|R \cap \mathcal{H}_0| \leq$?



Our method: reference family

If $|A \cap \mathcal{H}_0| \leq 5$, $|B \cap \mathcal{H}_0| \leq 7$ $\quad\quad$ $|R \cap \mathcal{H}_0| \leq 10$



JR criterion

☞ [Blanchard, Neuvial and R. *On controlled posthoc inference* (ongoing work)]

Define for $\mathcal{T} = \{t_k\}_{1 \leq k \leq m}$,

$$\begin{aligned} \mathrm{JR}(\mathcal{T}, P) &= \mathbf{P}(\exists k \in \{1, \ldots, m\} \ : \ V(t_k) \geq k) \\ &= 1 - \mathbf{P}(\forall k \in \{1, \ldots, m\} \ : \ |\{i \ : \ p_i \leq t_k\} \cap \mathcal{H}_0| \leq k - 1) \end{aligned}$$

Proposition 1

JR control at level α implies posthoc bound:

$$\overline{V}_\alpha(R) = \min_{1 \leq k \leq m} \left\{ \sum_{i \in R} \mathbf{1}\{p_i(X) > t_k\} + k - 1 \right\}, \quad R \subset \{1, \ldots, m\}.$$

JR criterion

☞ [Blanchard, Neuvial and R. *On controlled posthoc inference* (ongoing work)]

Define for $\mathcal{T} = \{t_k\}_{1 \leq k \leq m}$,

$$\begin{aligned} \mathrm{JR}(\mathcal{T}, P) &= \mathbf{P}(\exists k \in \{1, \ldots, m\} \ : \ V(t_k) \geq k) \\ &= 1 - \mathbf{P}(\forall k \in \{1, \ldots, m\} \ : \ |\{i \ : \ p_i \leq t_k\} \cap \mathcal{H}_0| \leq k - 1) \end{aligned}$$

Proposition 1

JR control at level α implies posthoc bound:

$$\overline{V}_\alpha(R) = \min_{1 \leq k \leq m} \left\{ \sum_{i \in R} \mathbf{1}\{p_i(X) > t_k\} + k - 1 \right\}, \quad R \subset \{1, \ldots, m\}.$$

JR criterion

☞ [Blanchard, Neuvial and R. *On controlled posthoc inference* (ongoing work)]

Define for $\mathcal{T} = \{t_k\}_{1 \leq k \leq m}$,

$$\begin{aligned} \mathrm{JR}(\mathcal{T}, P) &= \mathbf{P}(\exists k \in \{1, \ldots, m\} \ : \ V(t_k) \geq k) \\ &= 1 - \mathbf{P}(\forall k \in \{1, \ldots, m\} \ : \ |\{i \ : \ p_i \leq t_k\} \cap \mathcal{H}_0| \leq k - 1) \end{aligned}$$

Proposition 1

JR control at level α implies posthoc bound:

$$\overline{V}_\alpha(R) = \min_{1 \leq k \leq m} \left\{ \sum_{i \in R} \mathbf{1}\{p_i(X) > t_k\} + k - 1 \right\}, \quad R \subset \{1, \ldots, m\}.$$

Controlling JR

Proposition 2

Taking $t_k = t_k(\lambda) \nearrow$ and right-continuous in λ,

$$\mathrm{JR}(\mathcal{T}, P) = \mathbf{P}\left(\min_{1 \leq k \leq m_0} \left\{ t_k^{-1}(p_{(k:\mathcal{H}_0)}) \right\} \leq \lambda \right)$$

Example : λ-adjustement $X \sim \mathcal{N}(\mu, \Gamma)$ (one-sided), Γ known

- $t_k(\lambda) = \lambda k/m$
- λ chosen as α-quantile of

$$\mathcal{D}_{Z \sim \mathcal{N}(0,\Gamma)} \left(\min_{1 \leq k \leq m} \left\{ \frac{m}{k} \overline{\Phi}(Z)_{(k:m)} \right\} \right)$$

Outlook

Various criteria

- FWER
- FDR
- FDP
- JR

Various challenges

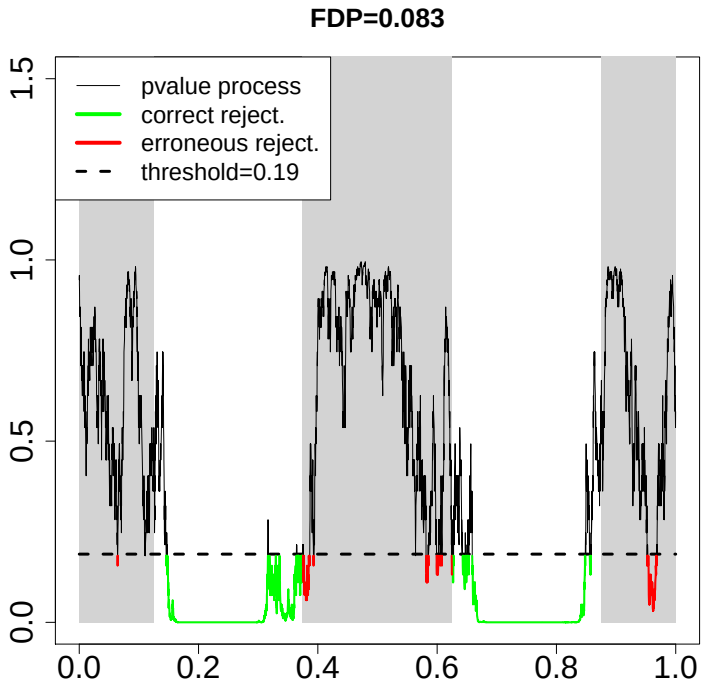
- Surviving to dependence
- Incorporating known dependence
- Adaptating to unknown dependence

Continuous testing in a nutshell

[Blanchard, Delattre and R. *Testing over a continuum of null hypotheses with False Discovery Rate control* (2014, Bernoulli)]

- $(H_{0,t}, t \in [0,1])$ continuous set of null hypotheses
- $(p_t(X), t \in [0,1])$ p-value process (jointly-measurable)
- continuous FDR

$$\mathrm{FDR}(\widehat{t}, P) = \mathbf{E}[\mathrm{FDP}(\widehat{t}, P)], \quad \mathrm{FDP}(\widehat{t}, P) = \frac{\Lambda(\{t \,:\, \theta_t = 0,\ p_t(X) \le \widehat{t}\})}{\Lambda(\{t \,:\, p_t(X) \le \widehat{t}\})}$$



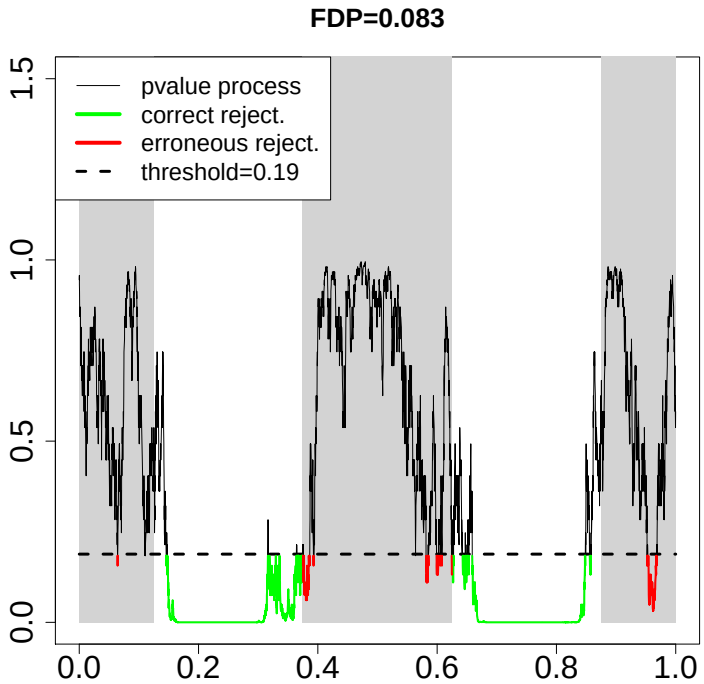
- Continuous BH procedure $\widehat{t}_\alpha$
- Finite dimensional PRDS
- Result : continuous FDR control

Continuous testing in a nutshell

[Blanchard, Delattre and R. *Testing over a continuum of null hypotheses with False Discovery Rate control* (2014, Bernoulli)]

- $(H_{0,t}, t \in [0,1])$ continuous set of null hypotheses
- $(p_t(X), t \in [0,1])$ p-value process (jointly-measurable)
- continuous FDR

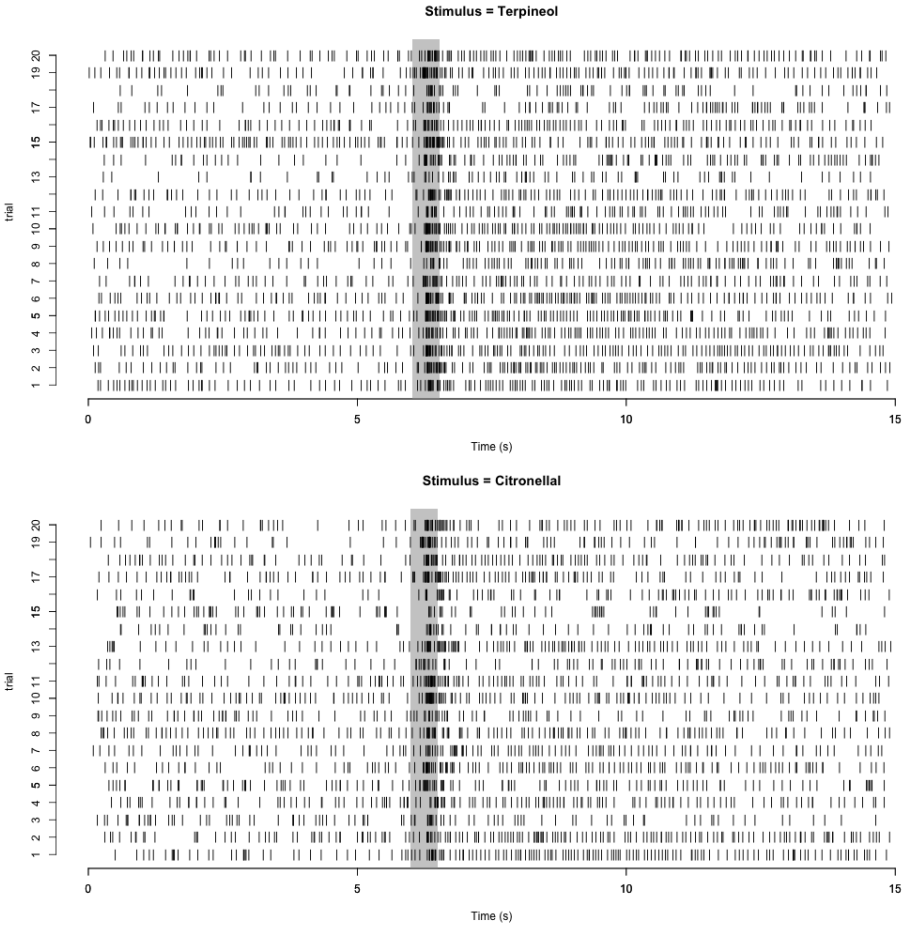
$$\mathrm{FDR}(\widehat{t}, P) = \mathbf{E}[\mathrm{FDP}(\widehat{t}, P)], \quad \mathrm{FDP}(\widehat{t}, P) = \frac{\Lambda(\{t \,:\, \theta_t = 0,\ p_t(X) \le \widehat{t}\})}{\Lambda(\{t \,:\, p_t(X) \le \widehat{t}\})}$$



- Continuous BH procedure $\widehat{t}_\alpha$
- Finite dimensional PRDS
- Result : continuous FDR control

Application

☞ [Picard et al. *Continuous testing for Poisson process intensities* (ongoing work)]



- N_A heterogeneous poisson process intensity λ_A
- N_B heterogeneous poisson process intensity λ_B
- $H_{0,t}$: "$\lambda_A = \lambda_B$ sur I_t"
- I_t, $t \in [0,1]$, sliding windows
- Basic test statistics give finite dim PRDS
- Continuous BH = data-driven weighted BH

Pouzat, package 'STAR' : Spike Train Analysis with R.