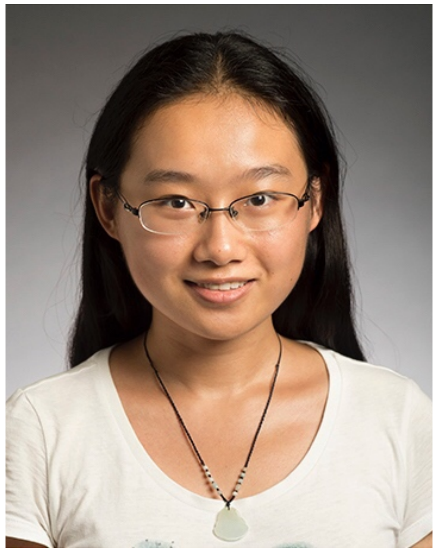
Dept. of Statistics and Data Science
Carnegie Mellon University

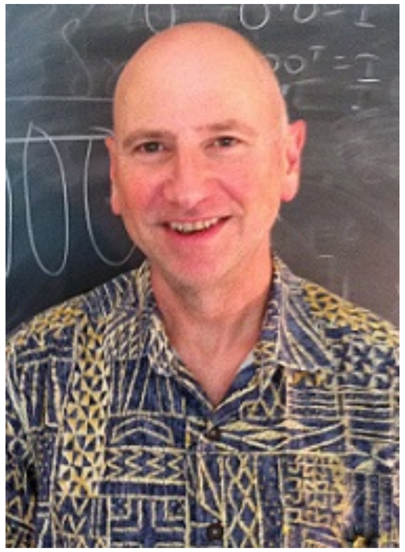
Familywise error rate control by interactive unmasking

Boyan Duan

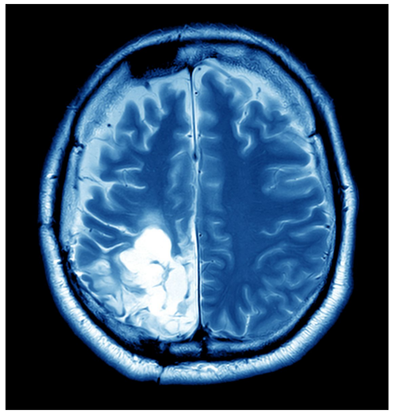
Aaditya Ramdas

Larry Wasserman

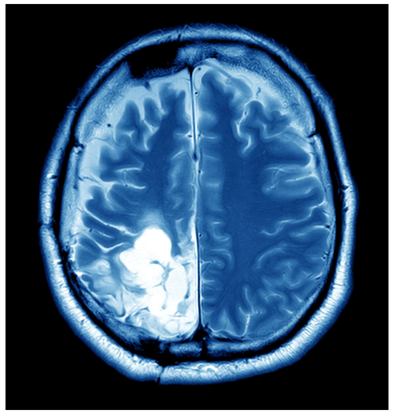




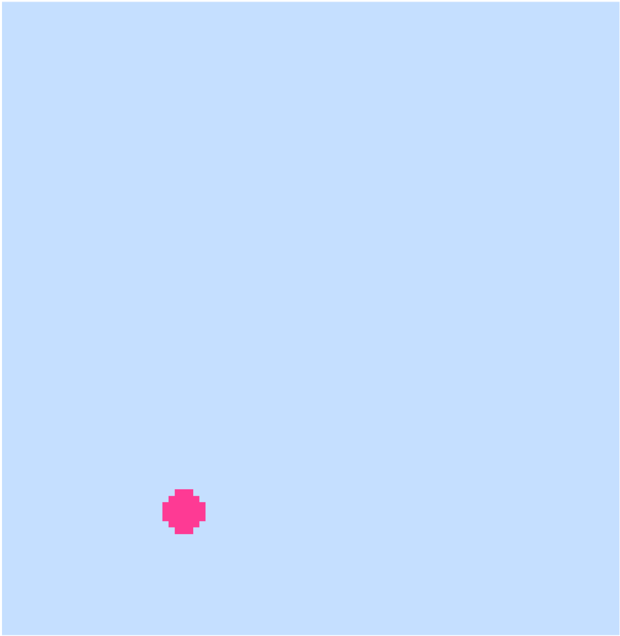
Motivating example: tumor detection in brain image



Motivating example: tumor detection in brain image

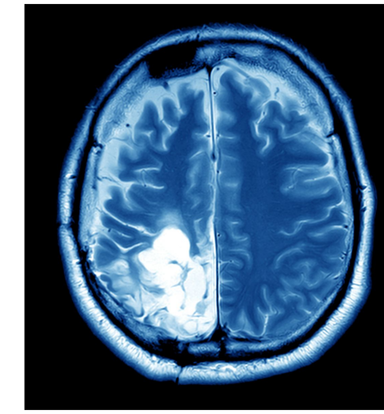


Eg. $Z_i \sim N(\mu_i, 1)$ for each pixel and test $H_i : \mu_i \leq 0$.

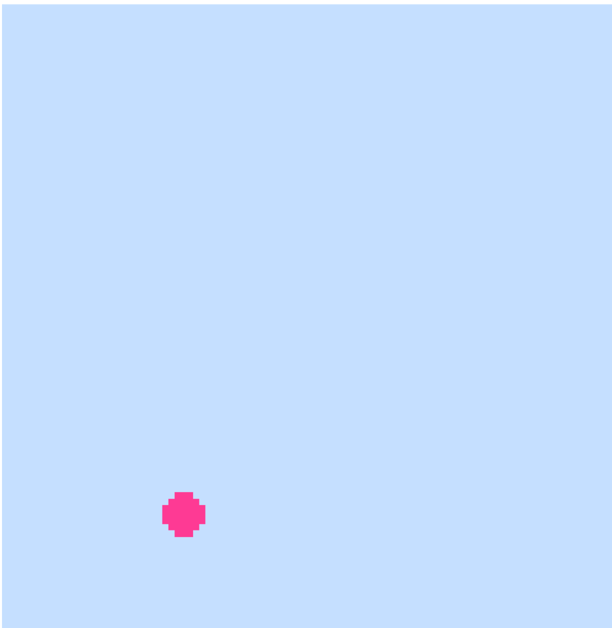


Underlying true μ_i

Motivating example: tumor detection in brain image



Eg. $Z_i \sim N(\mu_i, 1)$ for each pixel and test $H_i : \mu_i \leq 0$.

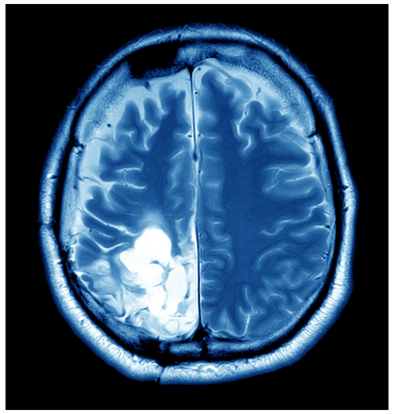


Underlying true μ_i

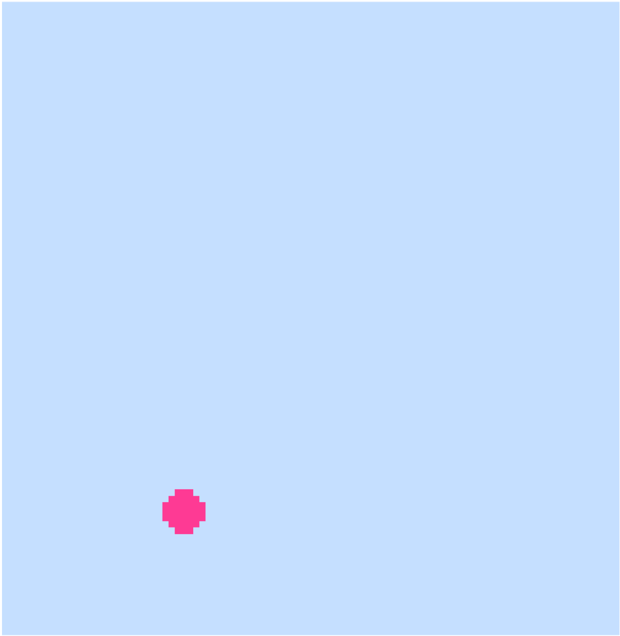
Observed p-values

$$P_i = 1 - \Phi(Z_i)$$

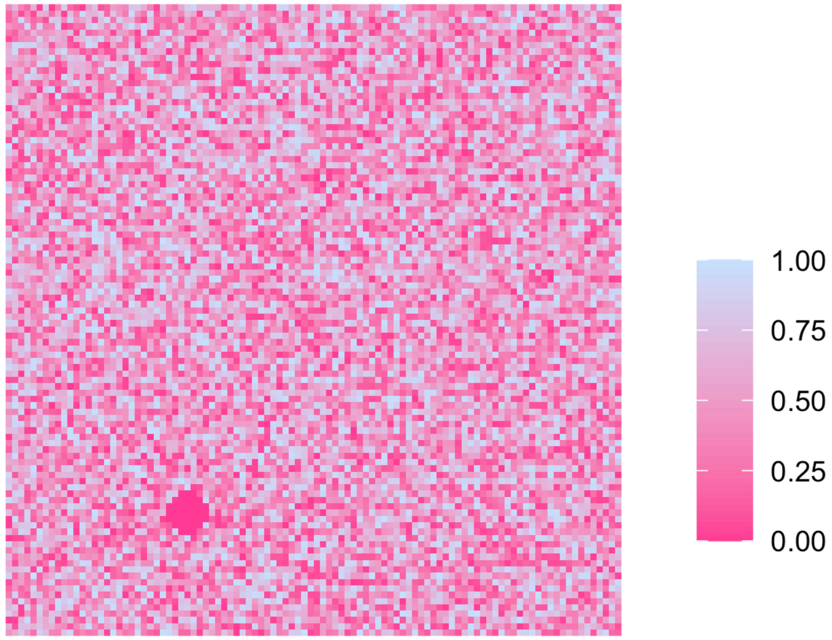
Motivating example: tumor detection in brain image



Eg. $Z_i \sim N(\mu_i, 1)$ for each pixel and test $H_i : \mu_i \leq 0$.

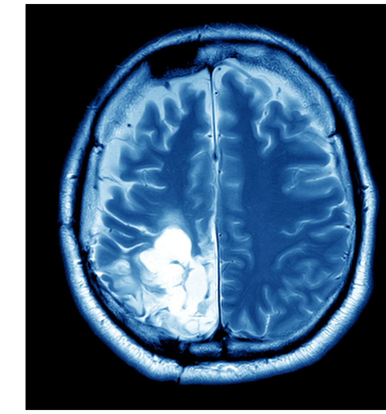


Underlying true μ_i

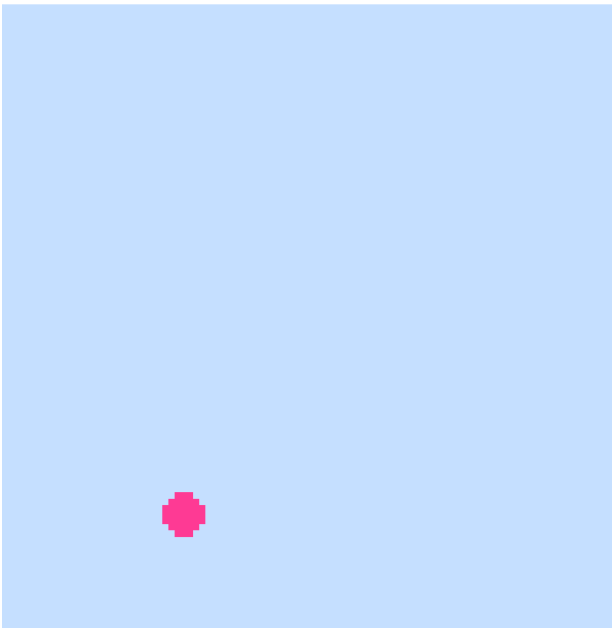


Observed p-values
$P_i = 1 - \Phi(Z_i)$

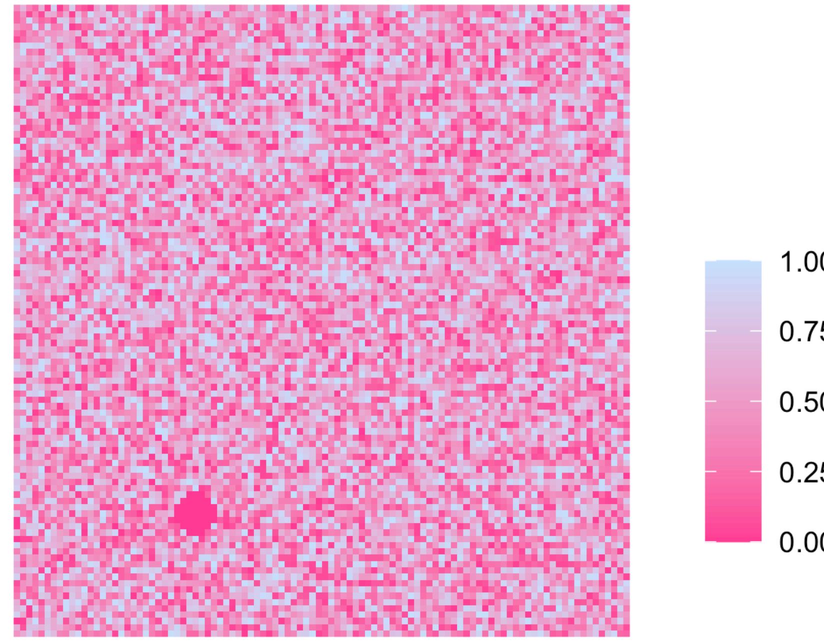
Motivating example: tumor detection in brain image



Eg. $Z_i \sim N(\mu_i, 1)$ for each pixel and test $H_i : \mu_i \leq 0$.



Underlying true μ_i

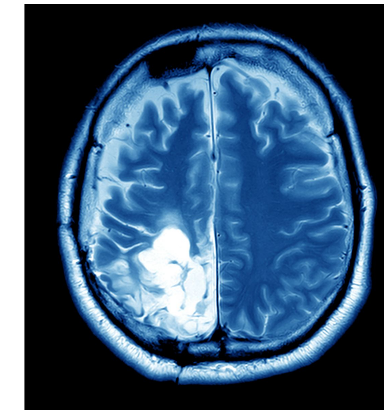


Observed p-values

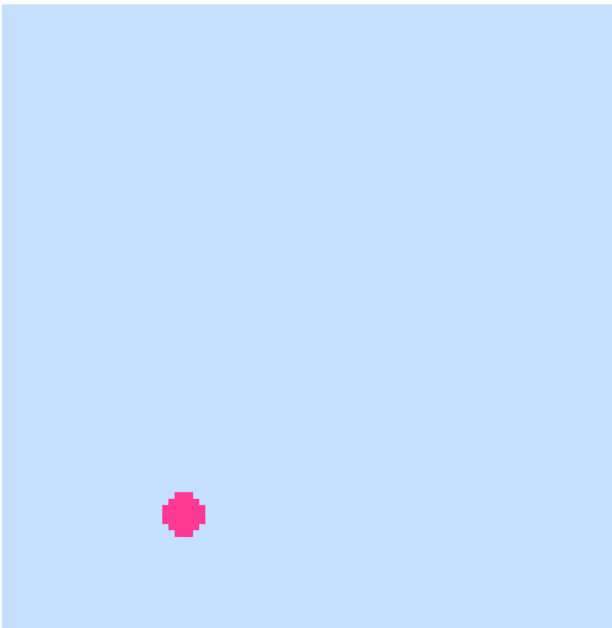
$$P_i = 1 - \Phi(Z_i)$$

Task: identify non-nulls (decide whether to reject each H_i),

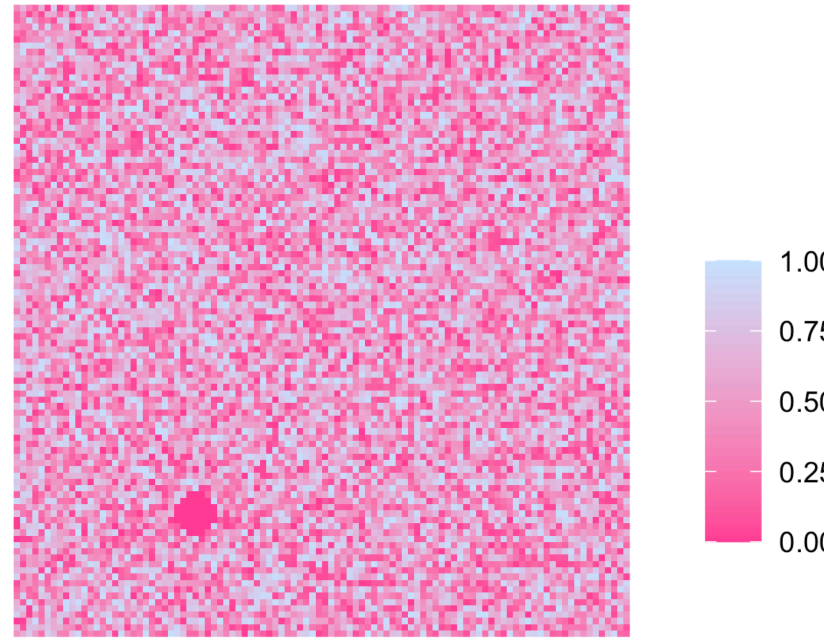
Motivating example: tumor detection in brain image



Eg. $Z_i \sim N(\mu_i, 1)$ for each pixel and test $H_i : \mu_i \leq 0$.



Underlying true μ_i



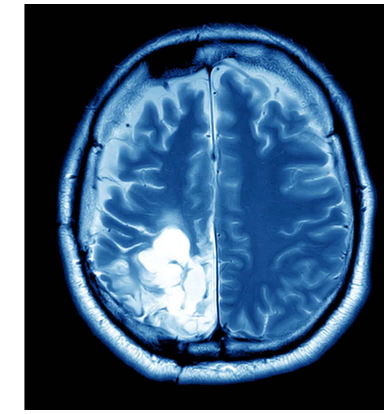
Observed p-values

$$P_i = 1 - \Phi(Z_i)$$

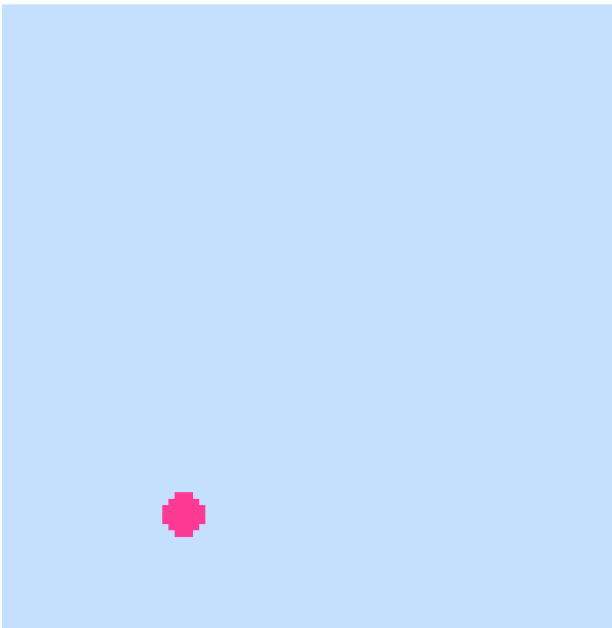
Task: identify non-nulls (decide whether to reject each H_i), with familywise error rate (FWER) control:

$$\text{FWER} := \mathbb{P}(\#\text{falsely rejected nulls} \geq 1)$$

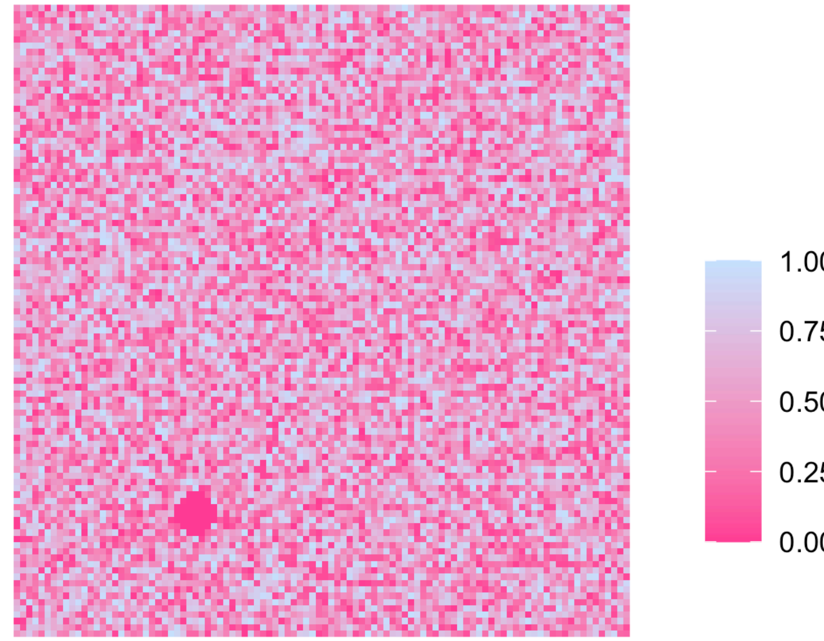
Motivating example: tumor detection in brain image



Eg. $Z_i \sim N(\mu_i, 1)$ for each pixel and test $H_i : \mu_i \leq 0$.



Underlying true μ_i



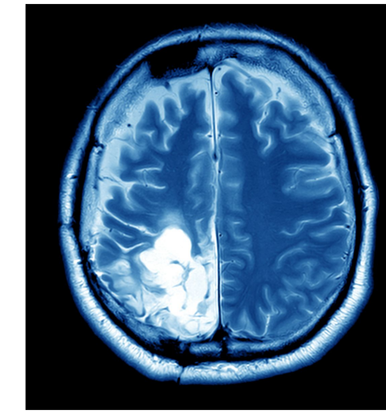
Observed p-values

$$P_i = 1 - \Phi(Z_i)$$

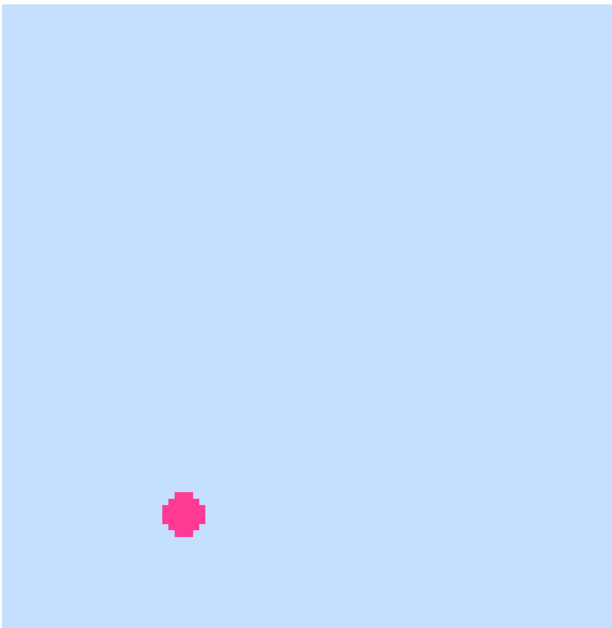
Task: identify non-nulls (decide whether to reject each H_i), with familywise error rate (FWER) control:

$$\text{FWER} := \mathbb{P}(\#\text{falsely rejected nulls} \geq 1) \leq \alpha \quad \text{Eg. } \alpha = 0.2$$

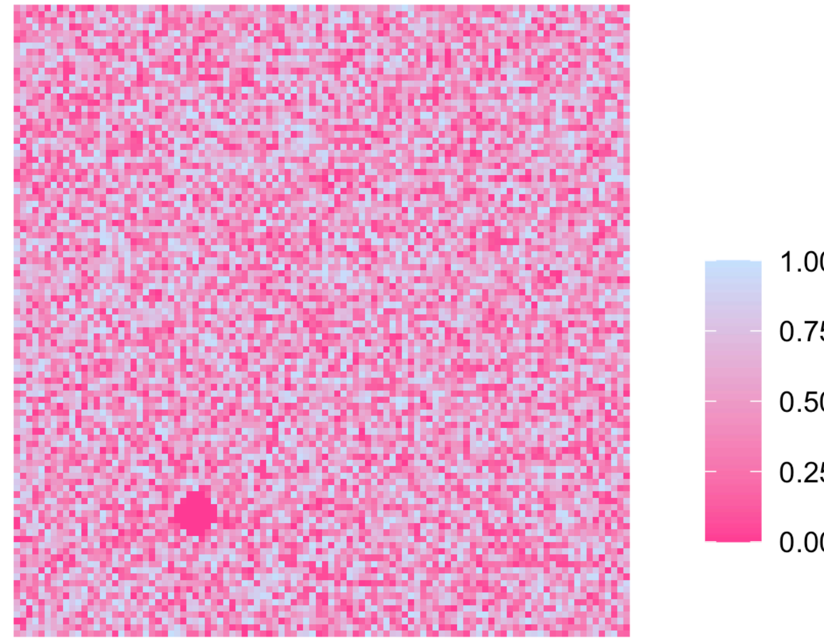
Motivating example: tumor detection in brain image



Eg. $Z_i \sim N(\mu_i, 1)$ for each pixel and test $H_i : \mu_i \leq 0$.



Underlying true μ_i



Observed p-values

$P_i = 1 - \Phi(Z_i)$

Task: identify non-nulls (decide whether to reject each H_i), with familywise error rate (FWER) control:

$$\text{FWER} := \mathbb{P}(\#\text{falsely rejected nulls} \geq 1) \leq \alpha \quad \text{Eg. } \alpha = 0.2$$

taking account of <u>side information</u>. Eg. structure, covariates…

Classical testing

Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i/w_i \leq \alpha/n$.

Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

Classical testing

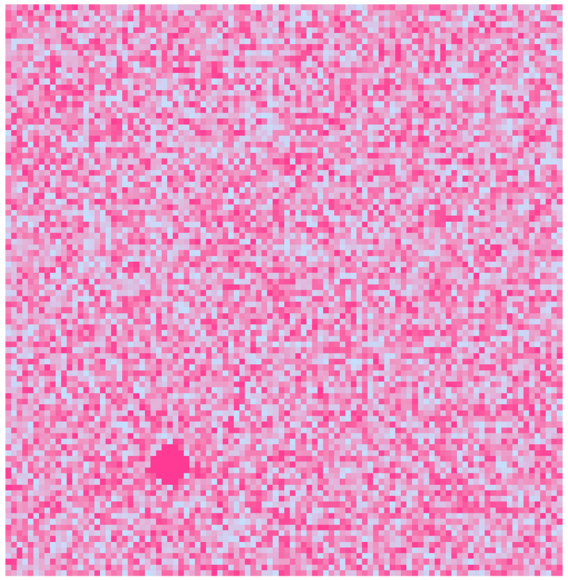
Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i / w_i \leq \alpha / n$.

Classical test
(single step)

Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

Classical testing

Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i/w_i \leq \alpha/n$.



p-values

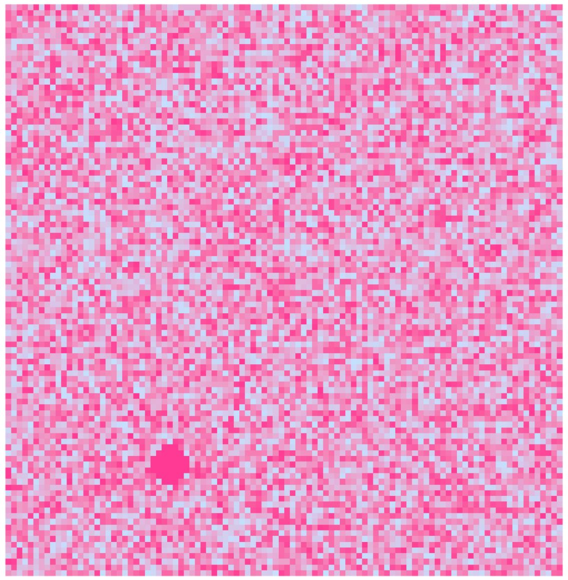
Data
(& side info)

Classical test
(single step)

Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

Classical testing

Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i / w_i \leq \alpha / n$.



p-values

Data
(& side info)

↓

Classical test
(single step)

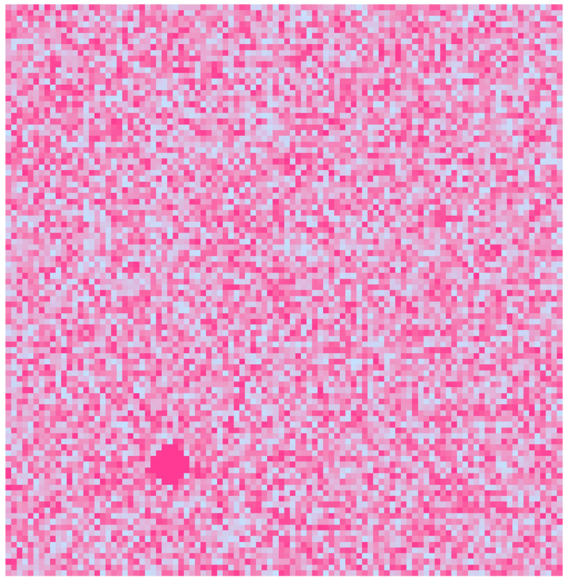
↓

Rejection
set

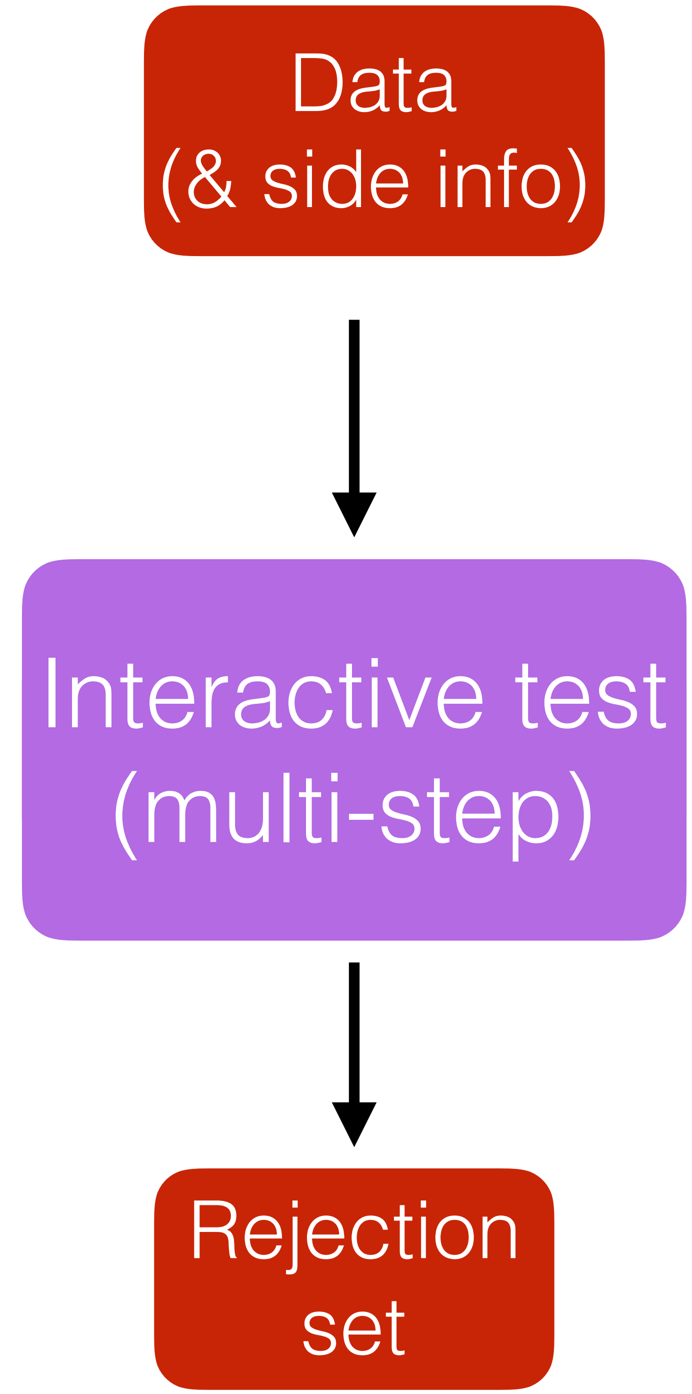
Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

Classical testing

Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i / w_i \leq \alpha / n$.



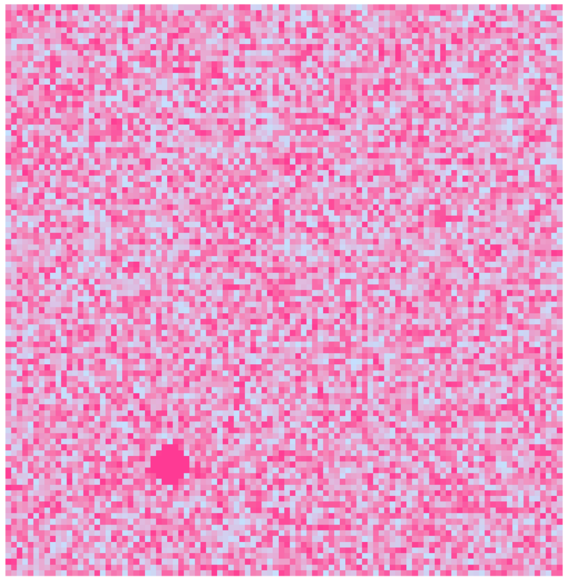
p-values



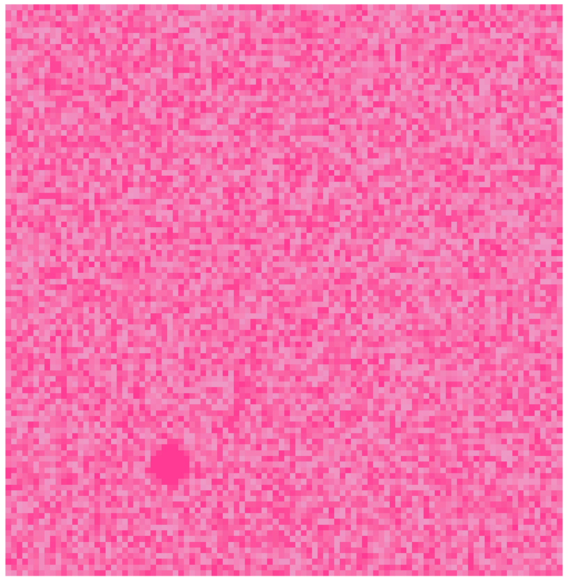
Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

Classical testing

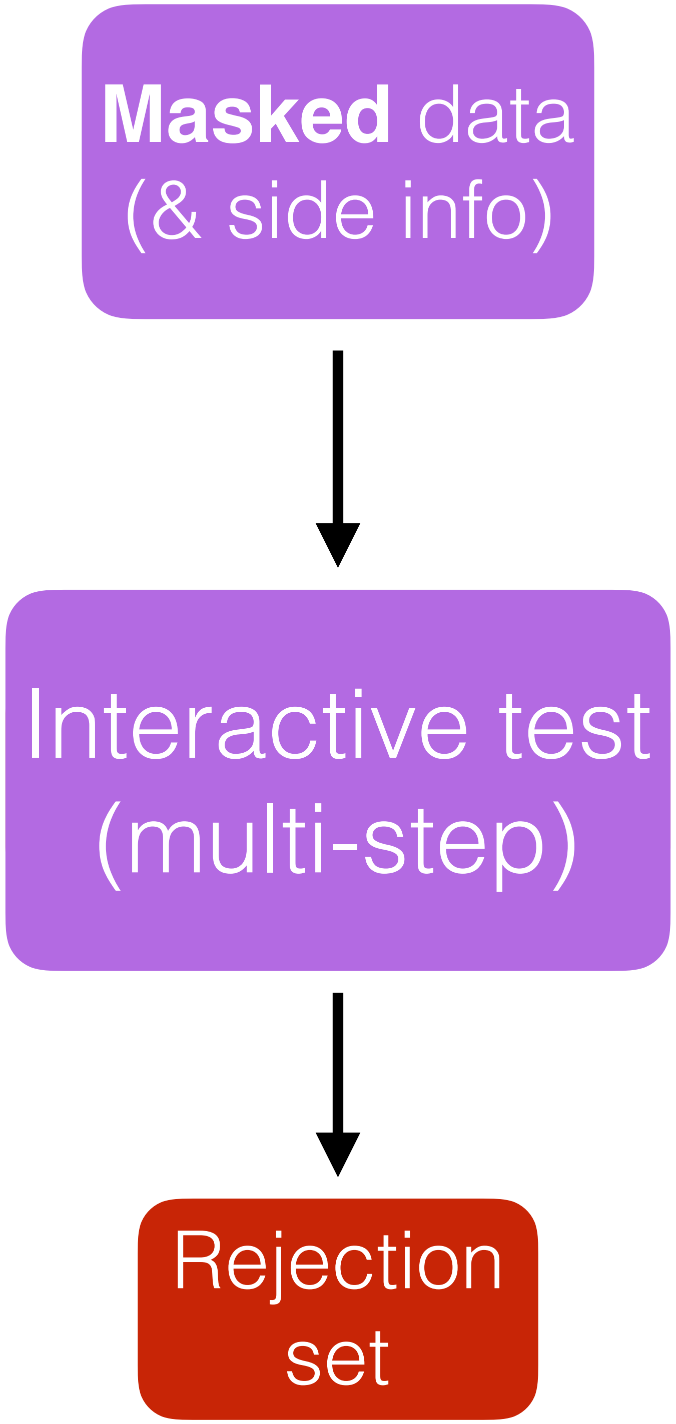
Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i/w_i \leq \alpha/n$.



p-values



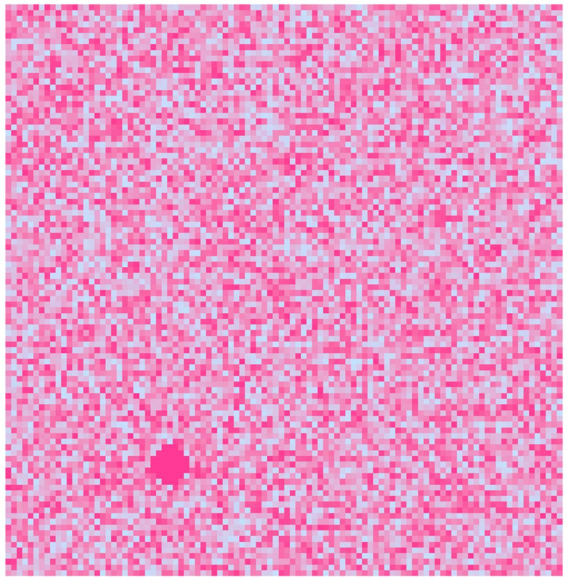
masked p-values



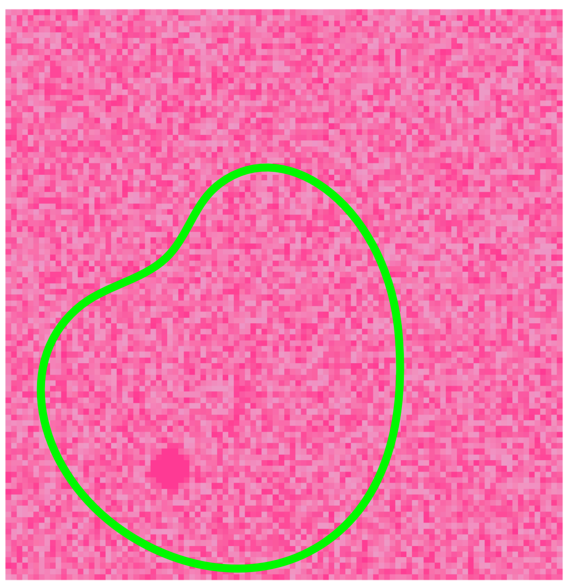
Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

Classical testing

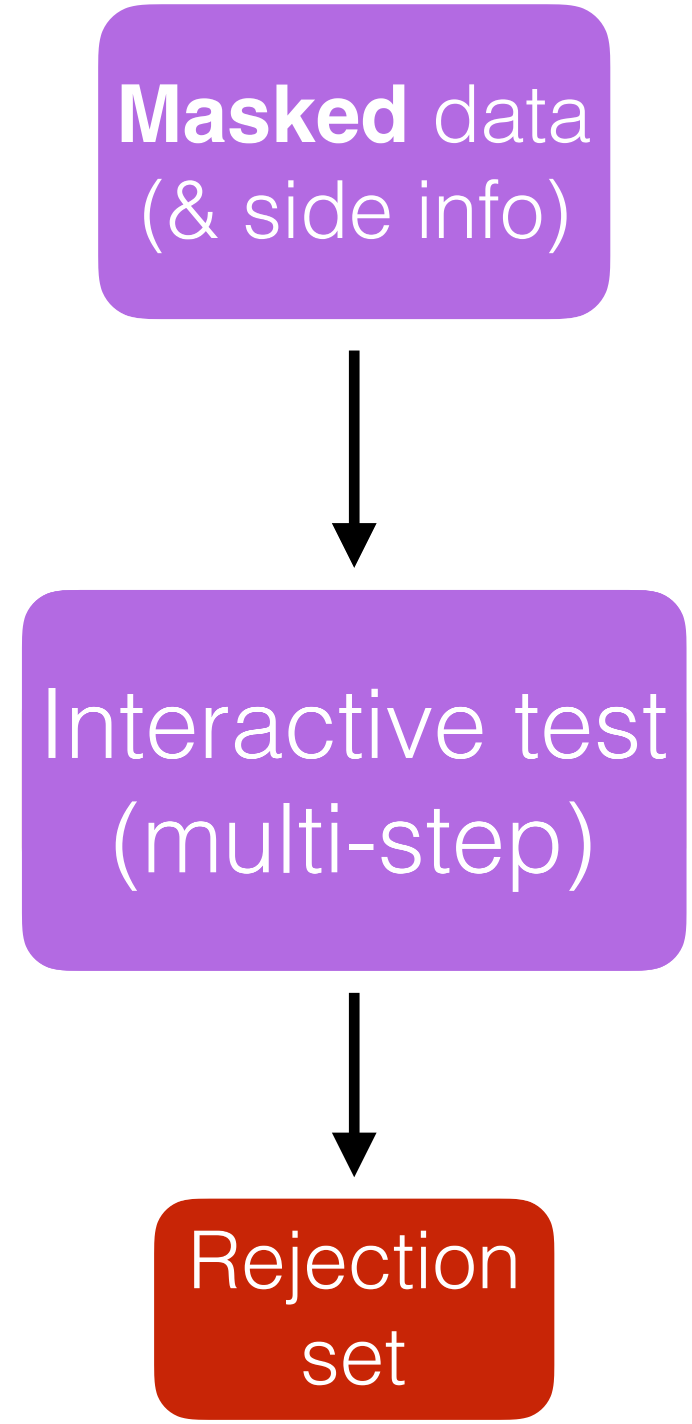
Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i/w_i \leq \alpha/n$.



p-values



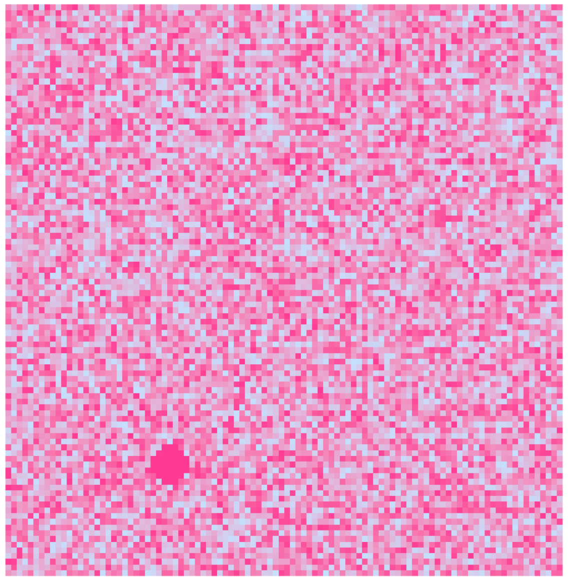
masked p-values



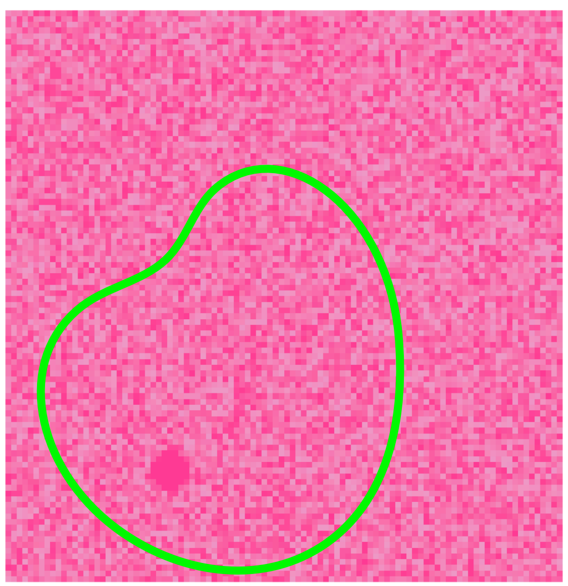
Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

Classical testing

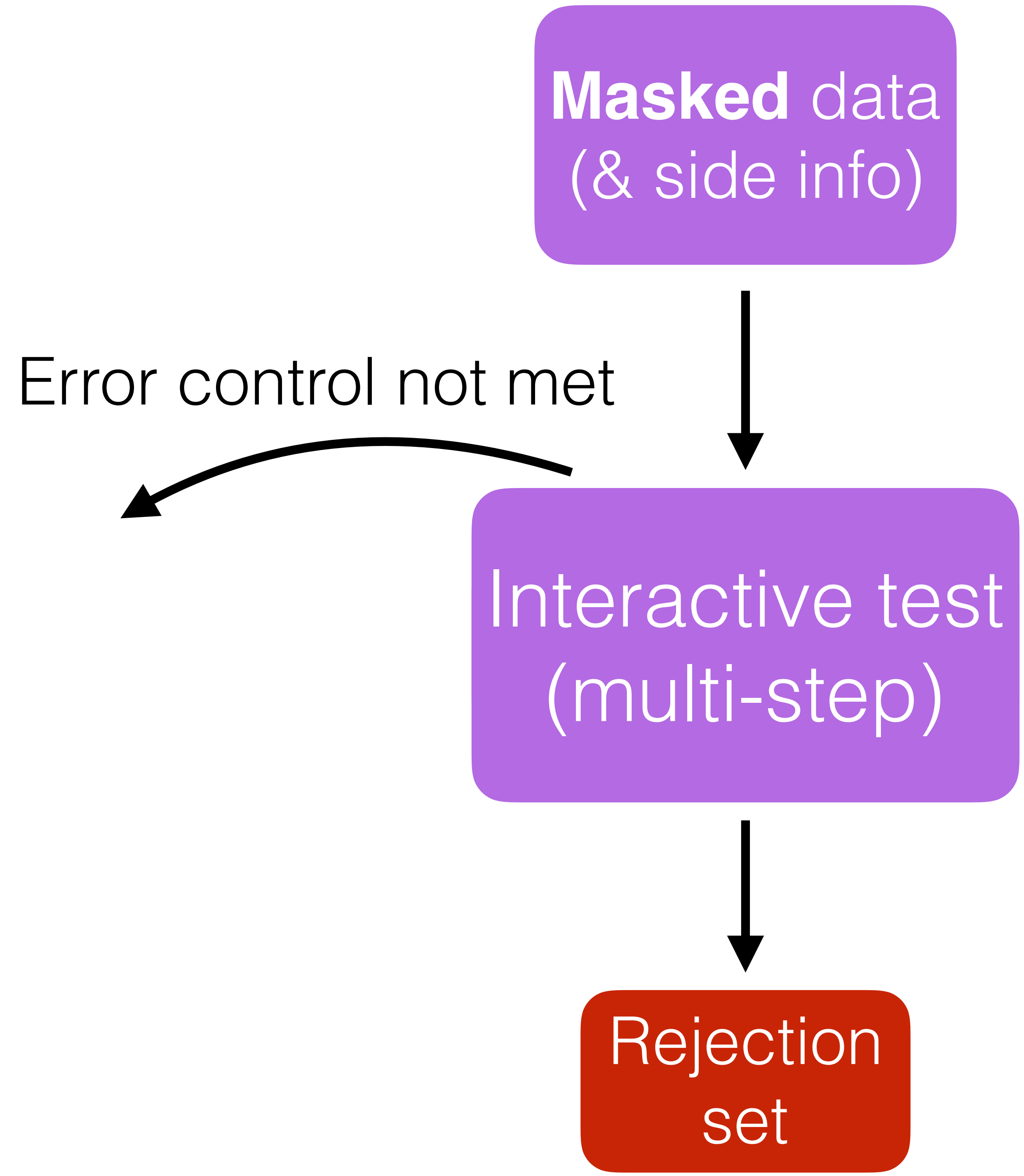
Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i / w_i \leq \alpha / n$.



p-values



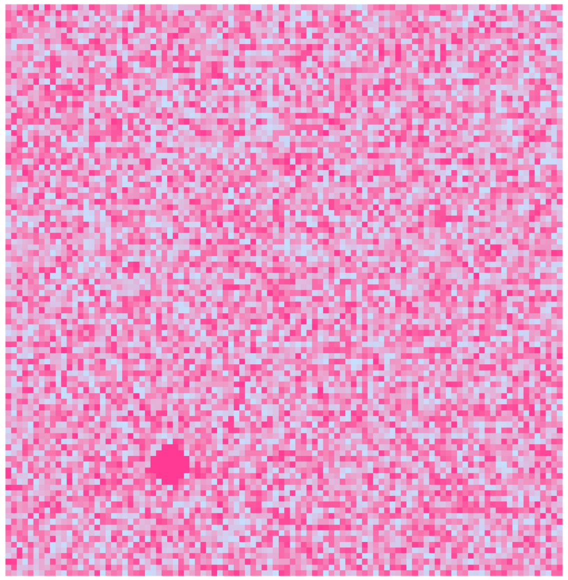
masked p-values



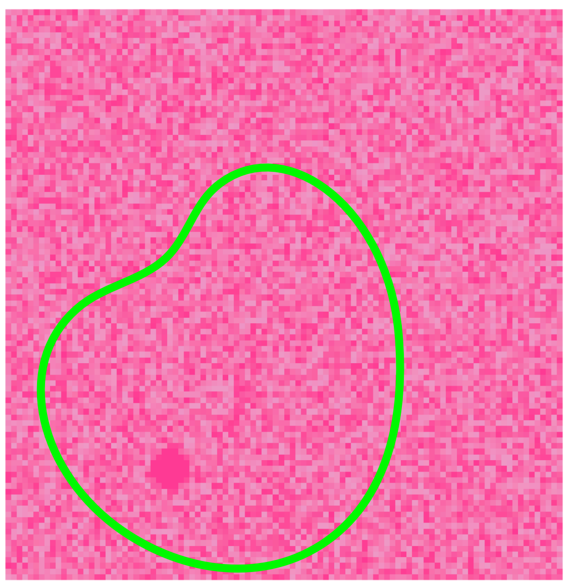
Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

Classical testing

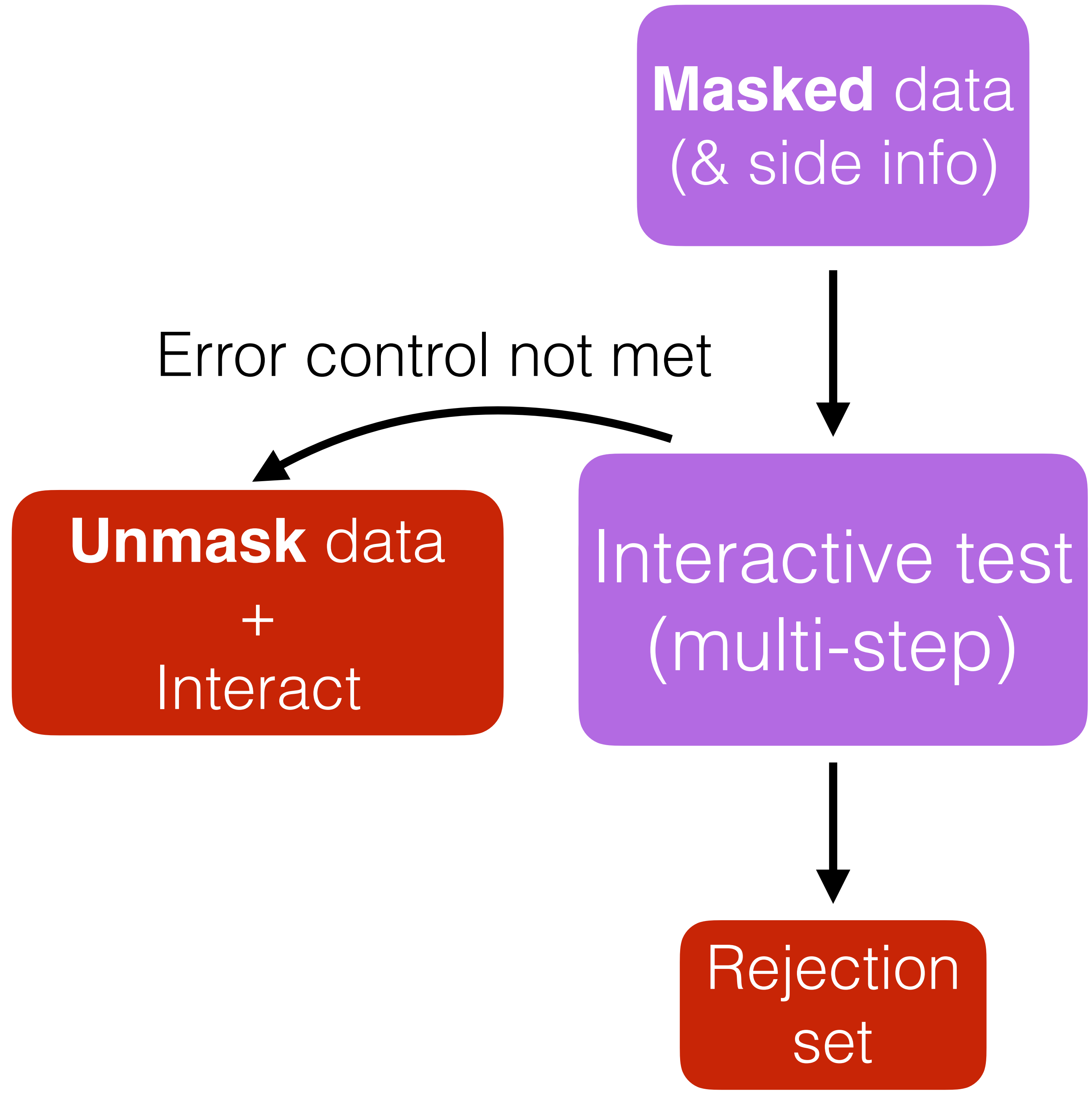
Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i/w_i \leq \alpha/n$.



p-values



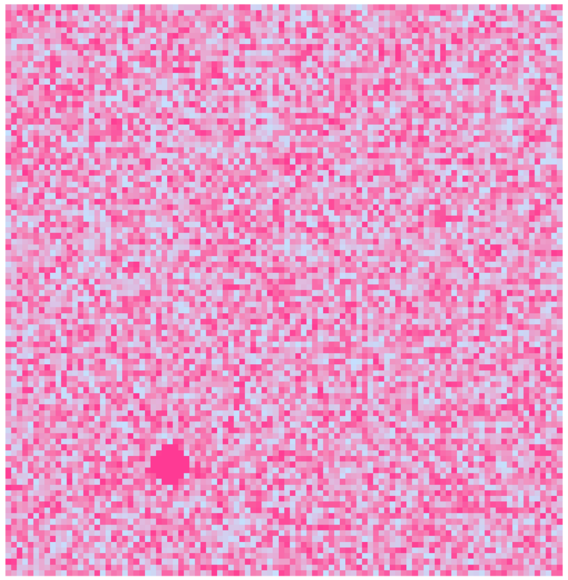
masked p-values



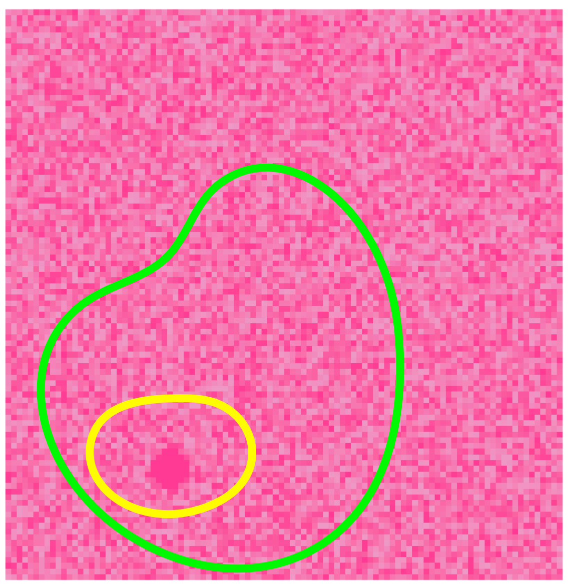
Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

Classical testing

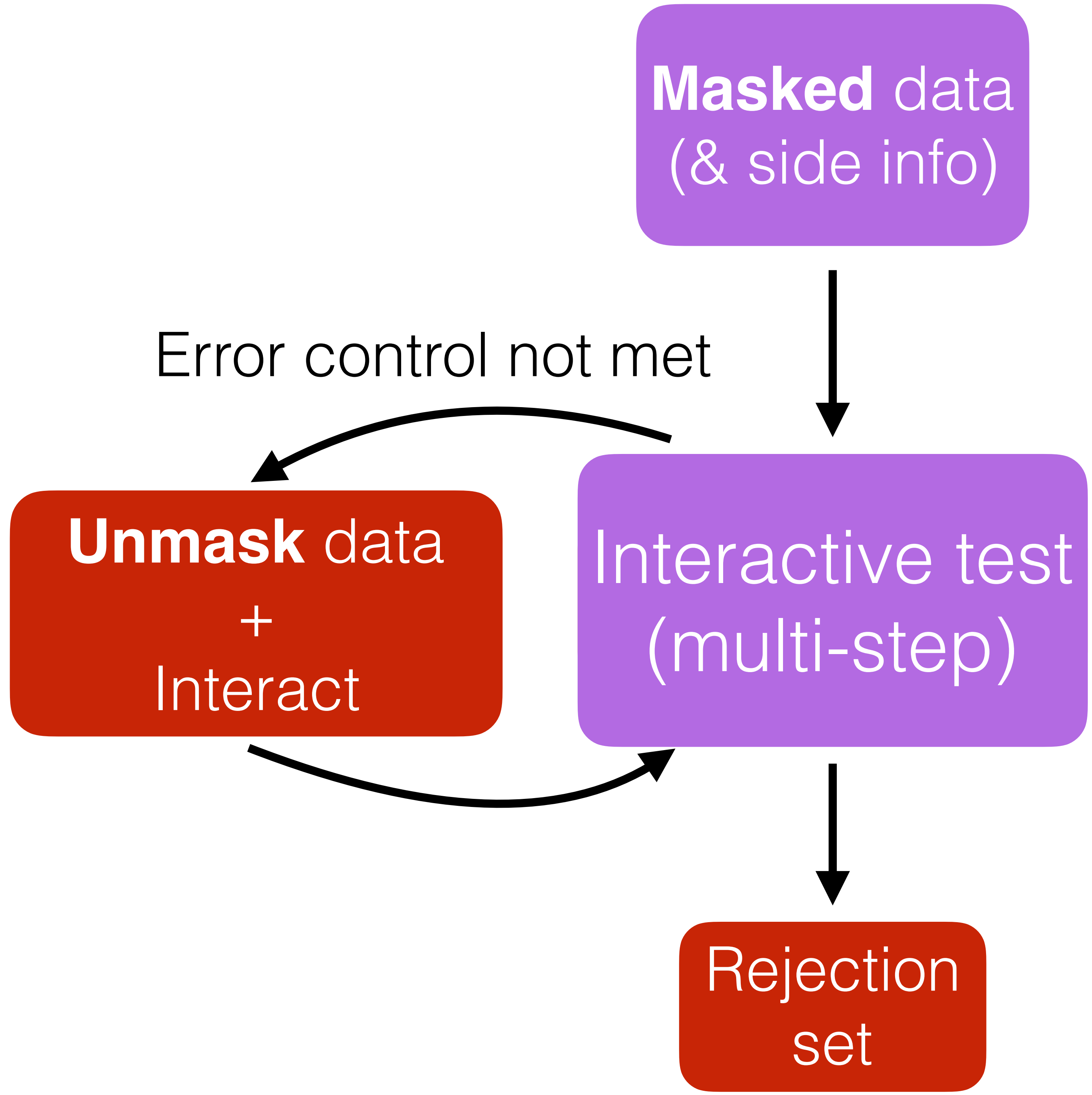
Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i/w_i \leq \alpha/n$.



p-values



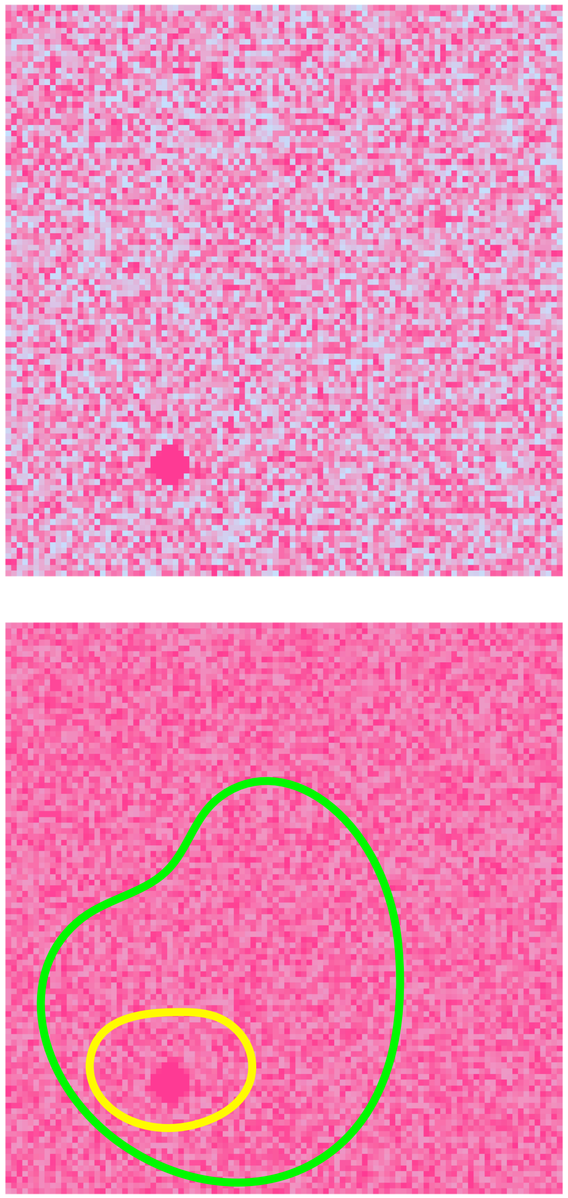
masked
p-values



Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

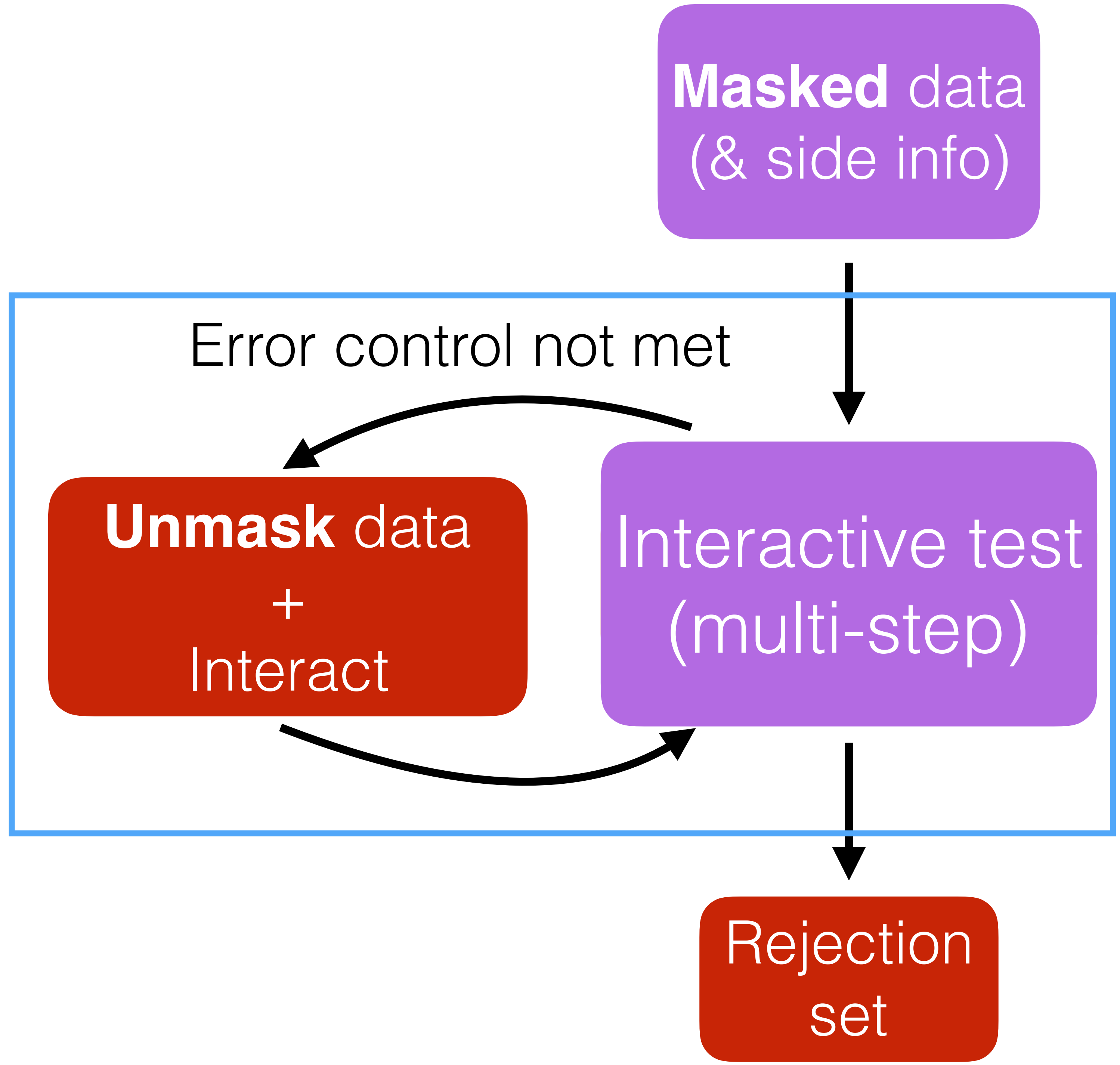
Classical testing

Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i/w_i \leq \alpha/n$.



p-values

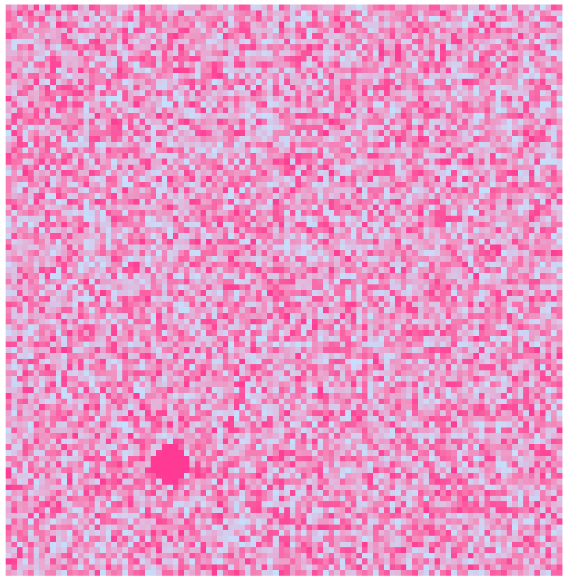
masked p-values



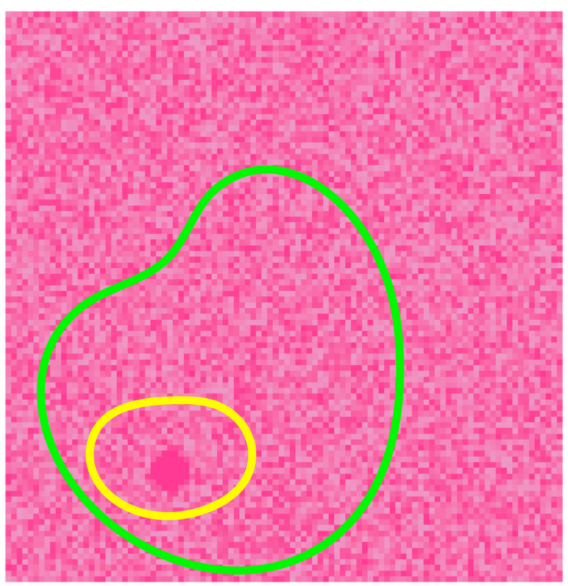
Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

Classical testing

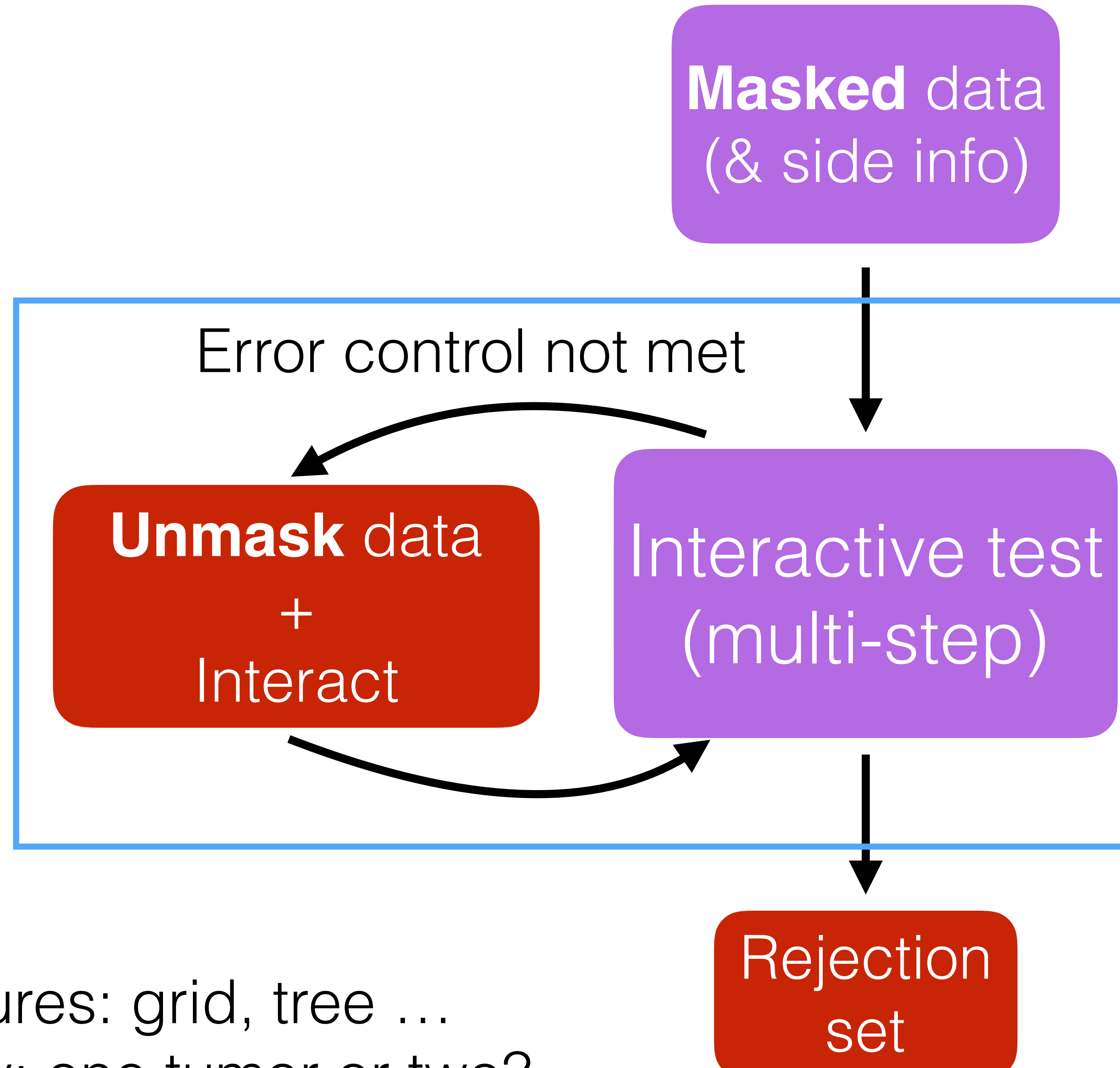
Pre-fixed procedure. Eg. weighted Bonferroni: reject H_i if $P_i/w_i \leq \alpha/n$.



p-values



masked p-values



Interactive testing

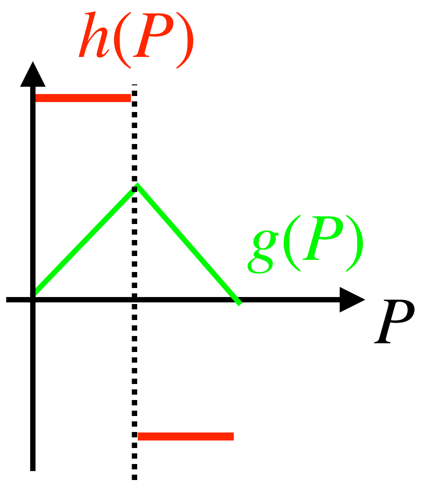
- Accommodate various structures: grid, tree …
- Be revised manually on the fly: one tumor or two?

Interactive tests with FDR control:
Lei, Fithian (2018); Lei, Ramdas, Fithian (2019)

Component 1: mask p-values

Component 1: mask p-values

$$P \longrightarrow \begin{cases} g(P) = \min\{P, 1 - P\} \\ h(P) = 2 \cdot 1\{P < 0.5\} - 1 \end{cases}$$



Component 1: mask p-values

$$P \longrightarrow \begin{cases} g(P) = \min\{P, 1 - P\} \\ h(P) = 2 \cdot 1\{P < 0.5\} - 1 \end{cases}$$

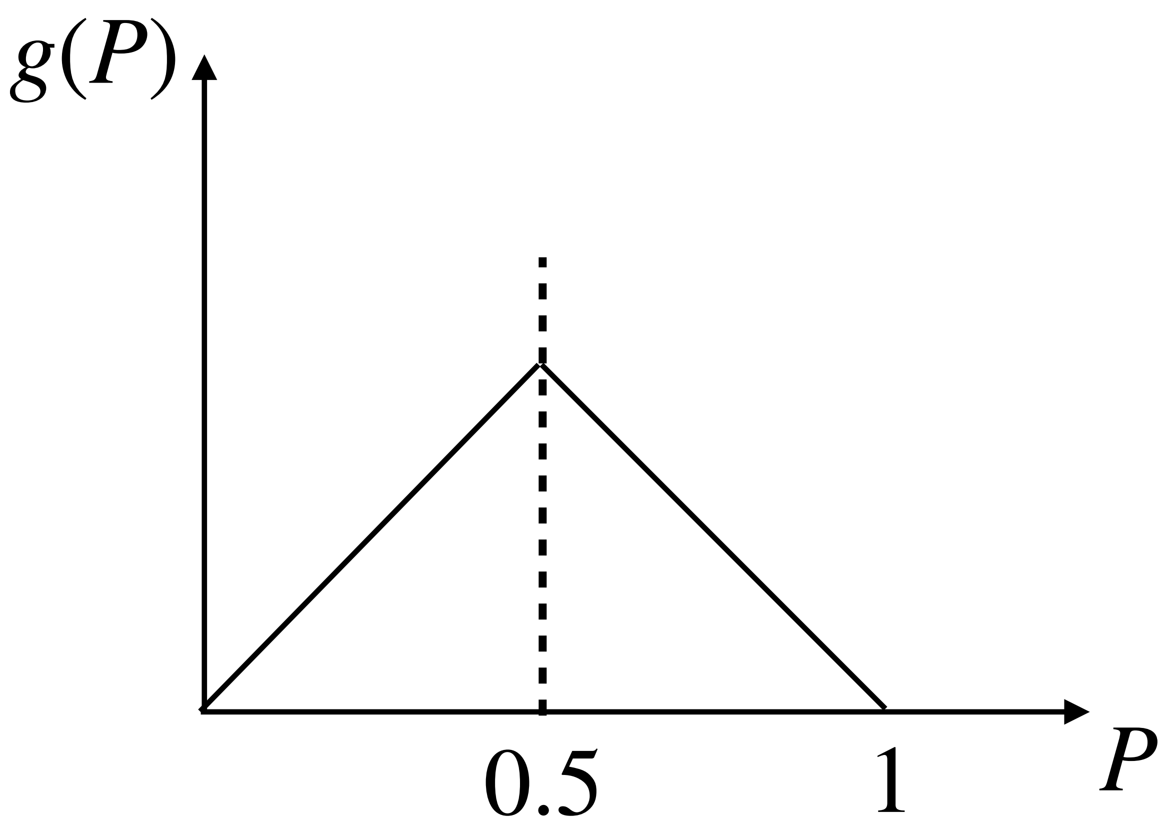
Independent for nulls

Component 1: mask p-values

$$P \longrightarrow \begin{cases} g(P) = \min\{P, 1-P\} \\ h(P) = 2 \cdot 1\{P < 0.5\} - 1 \end{cases}$$

Independent for nulls

Small $g(P)$ indicate non-nulls

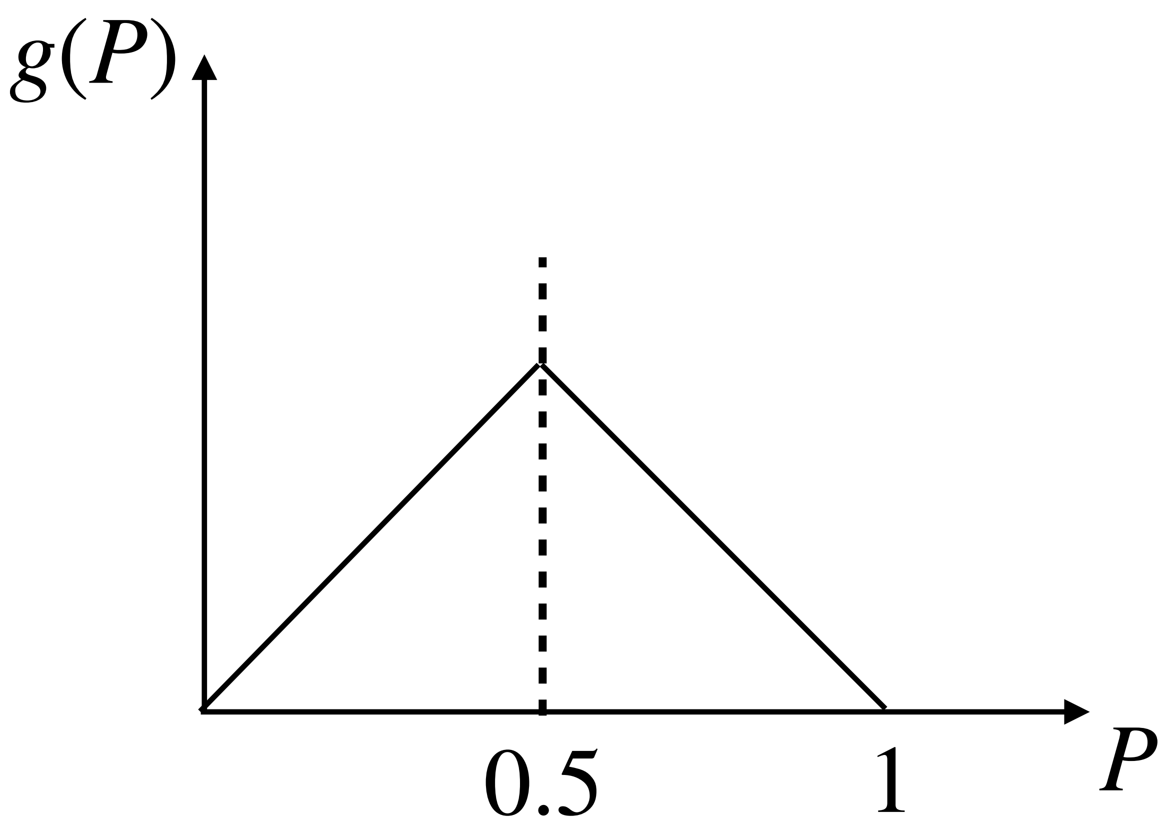


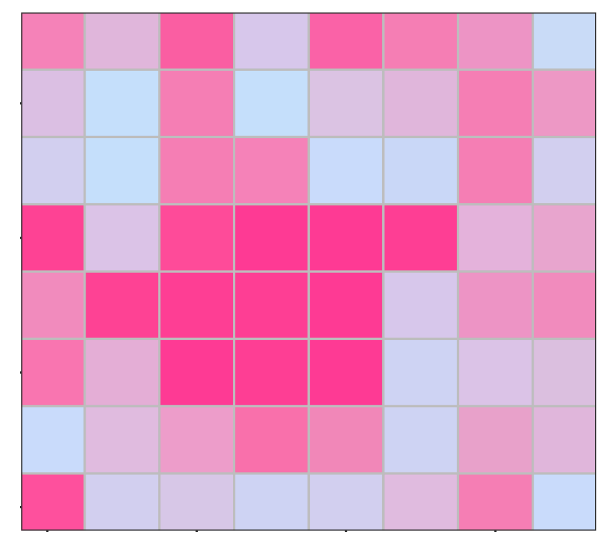
Component 1: mask p-values

$$P \longrightarrow \begin{cases} g(P) = \min\{P, 1 - P\} \\ h(P) = 2 \cdot 1\{P < 0.5\} - 1 \end{cases}$$

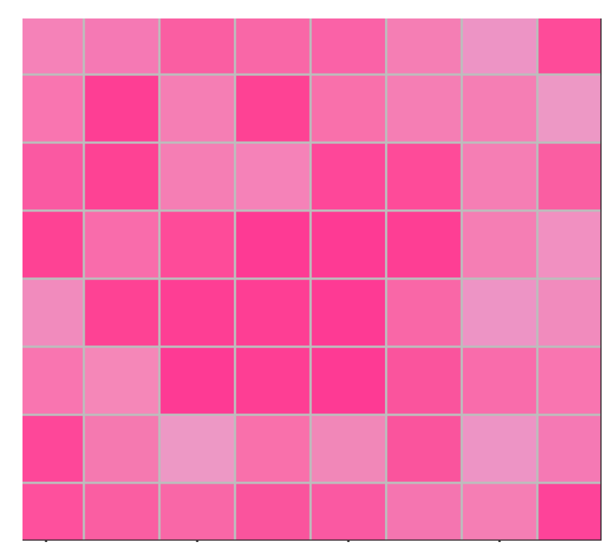
Independent for nulls

Small $g(P)$ indicate non-nulls





P



$g(P)$

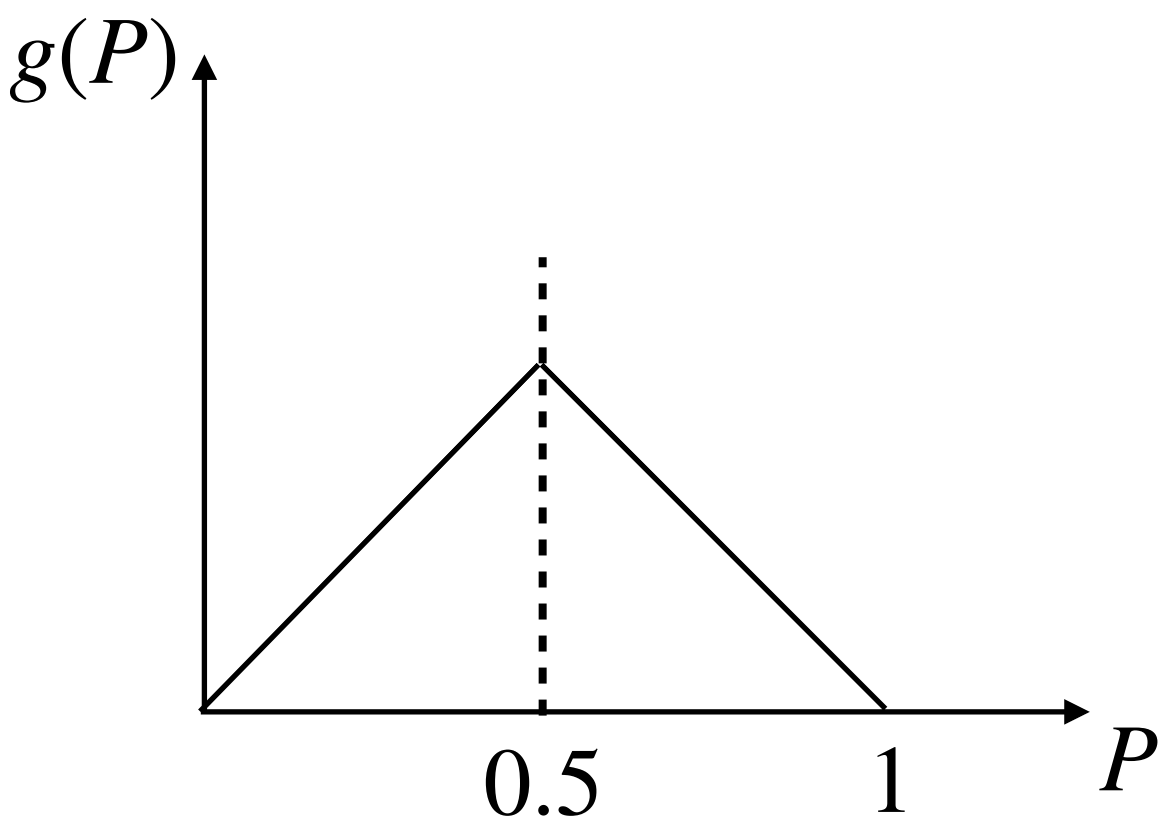
Component 1: mask p-values

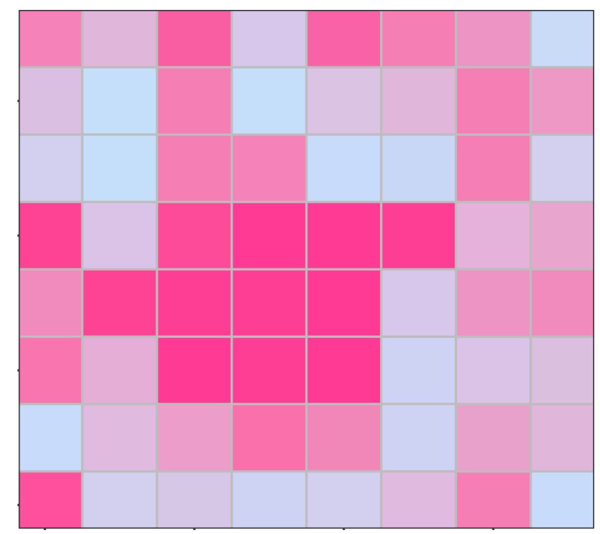
$$P \longrightarrow \begin{cases} g(P) = \min\{P, 1 - P\} \\ h(P) = 2 \cdot 1\{P < 0.5\} - 1 \end{cases}$$

Independent for nulls

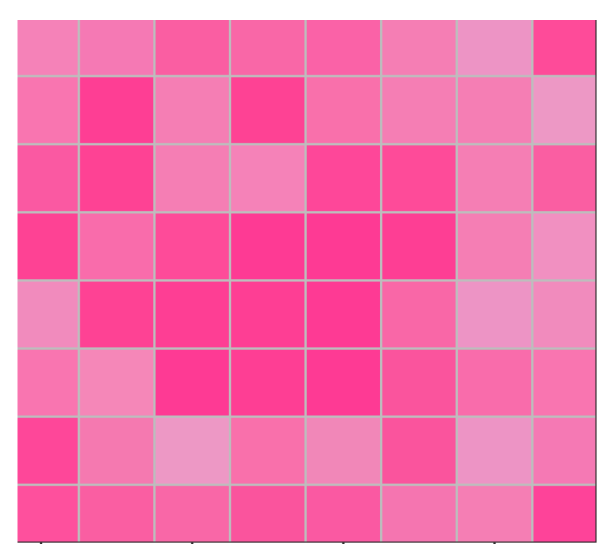
Note: above masking works for FDR control, but not *FWER* control.

Small $g(P)$ indicate non-nulls





P



$g(P)$

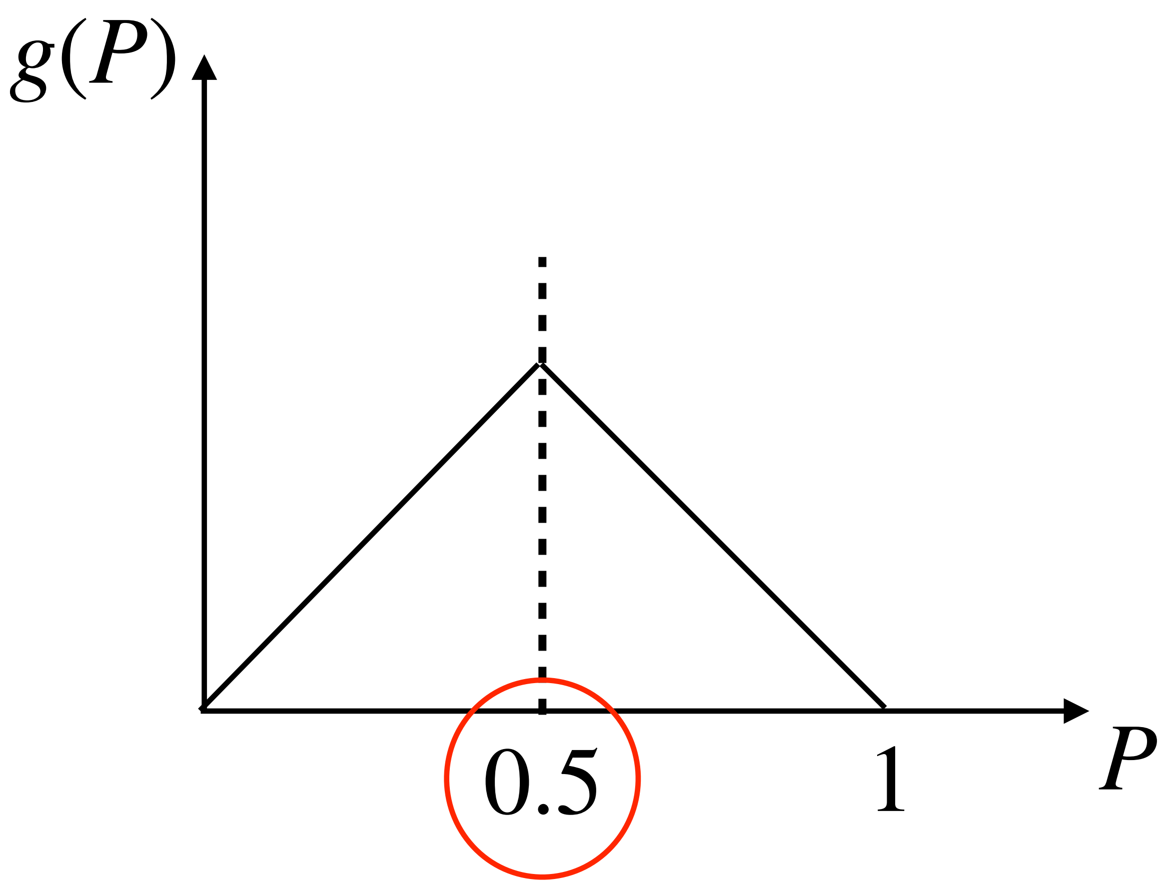
Component 1: mask p-values

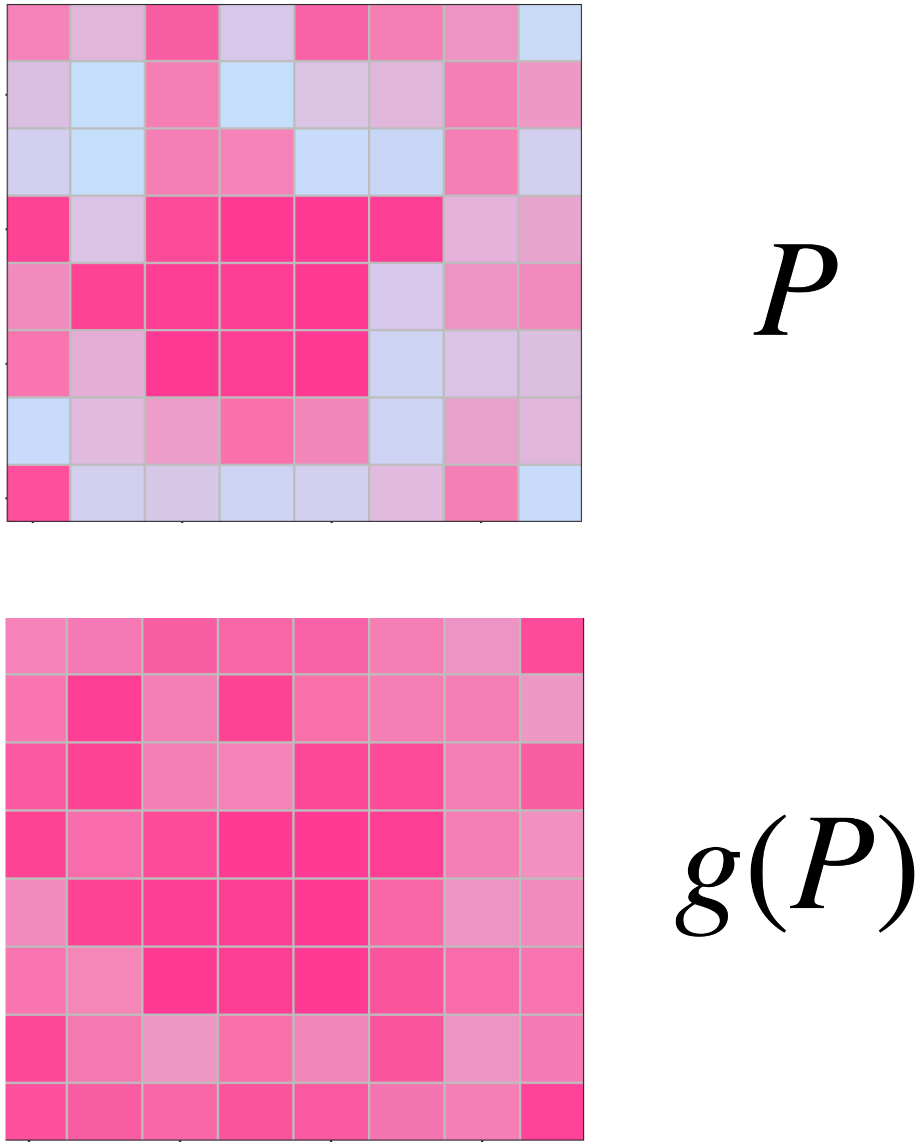
$$P \longrightarrow \begin{cases} g(P) = \min\{P, 1-P\} \\ h(P) = 2 \cdot 1\{P < 0.5\} - 1 \end{cases}$$

Independent for nulls

Note: above masking works for FDR control, but not *FWER* control.

Small $g(P)$ indicate non-nulls





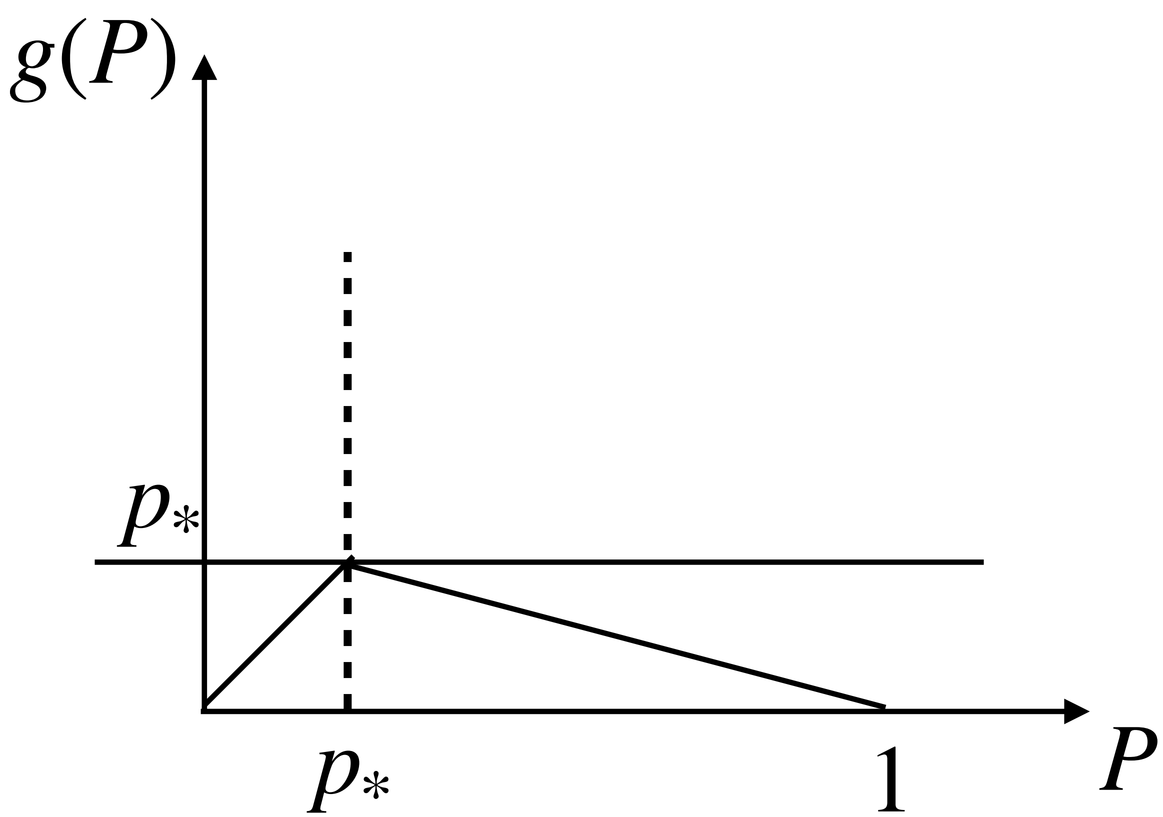
Component 1: mask p-values

$$P \longrightarrow \begin{cases} g(P) = \min\{P, 1 - P\} \\ h(P) = 2 \cdot 1\{P < 0.5\} - 1 \end{cases}$$

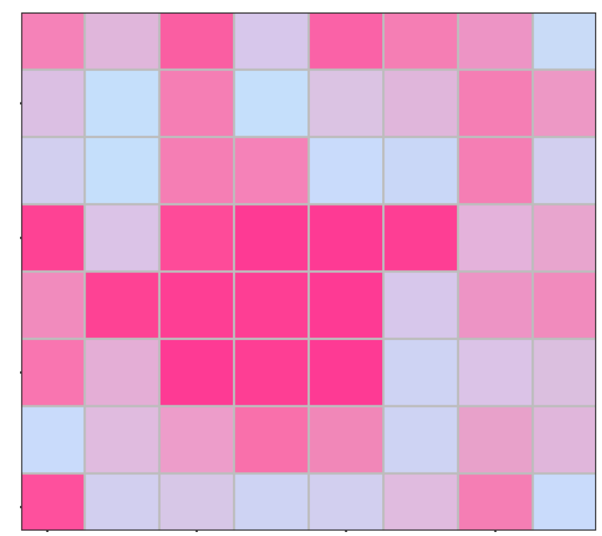
Independent for nulls

Note: above masking works for FDR control, but not *FWER* control.

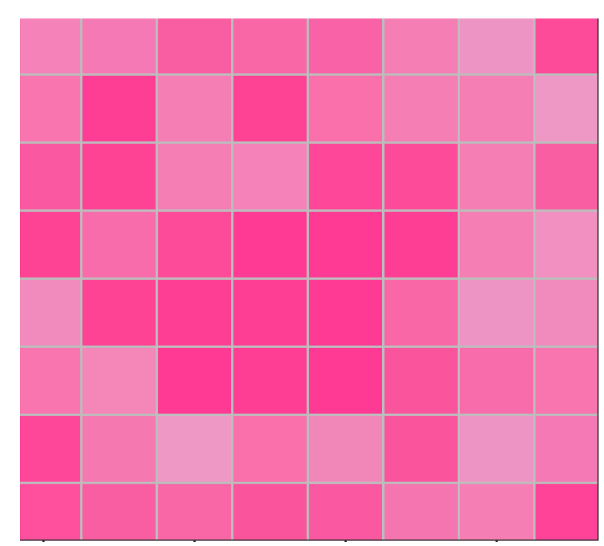
Small $g(P)$ indicate non-nulls



(Default $p_* = \alpha/2$)



P



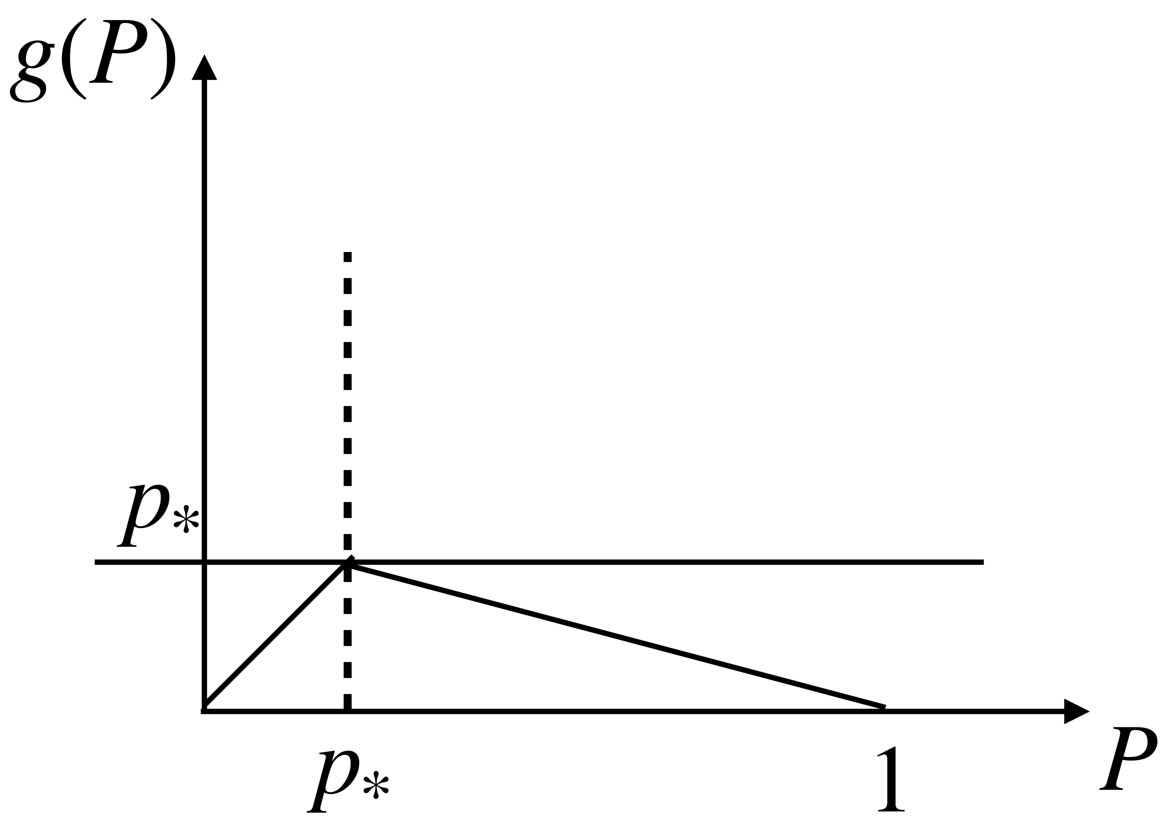
$g(P)$

Component 1: mask p-values

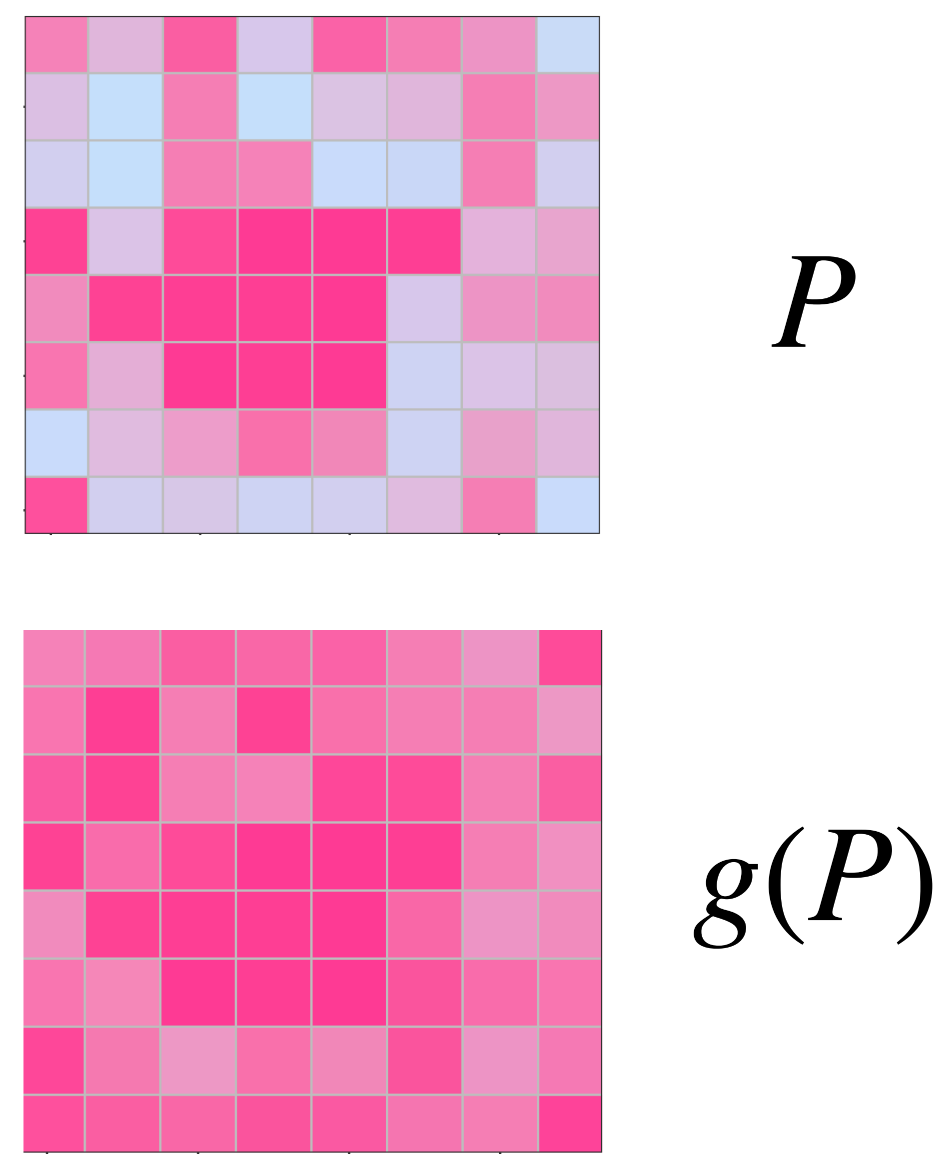
$$P \longrightarrow \begin{cases} g(P; p_*) = \min\{P, \frac{p_*}{1-p_*}(1-P)\} \\ h(P; p_*) = 2 \cdot 1\{P < p_*\} - 1 \end{cases}$$

Independent for nulls

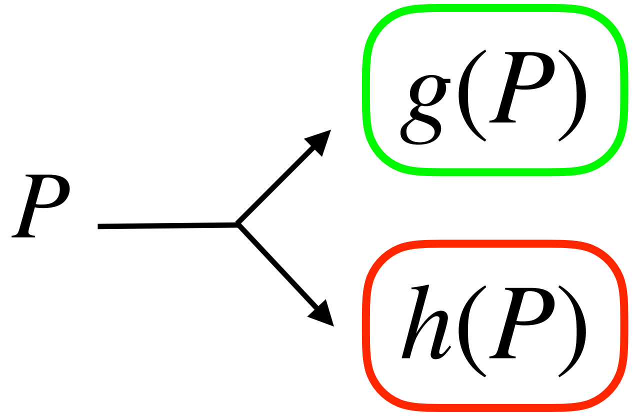
Small $g(P)$ indicate non-nulls



(Default $p_* = \alpha/2$)

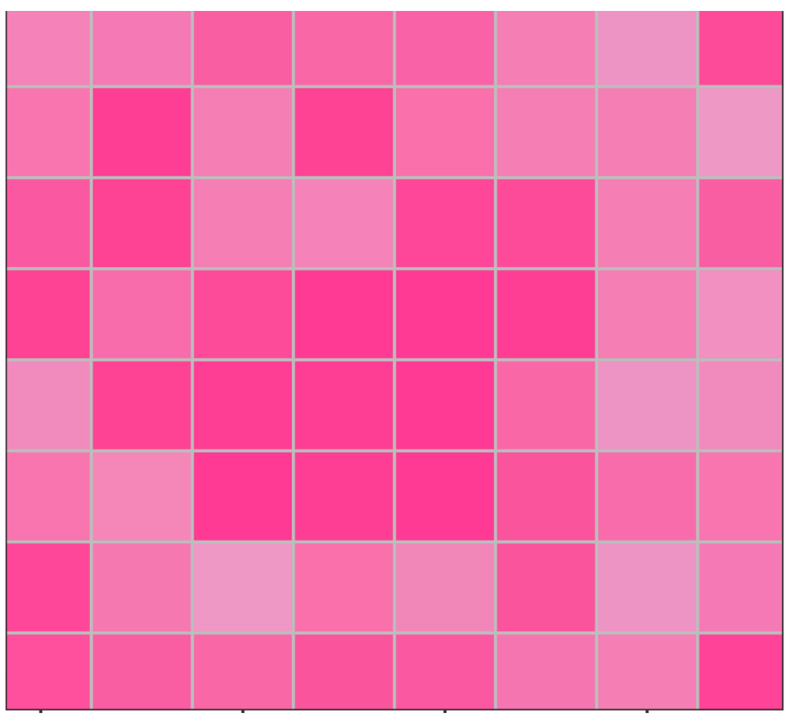


Component 2: select candidate set interactively

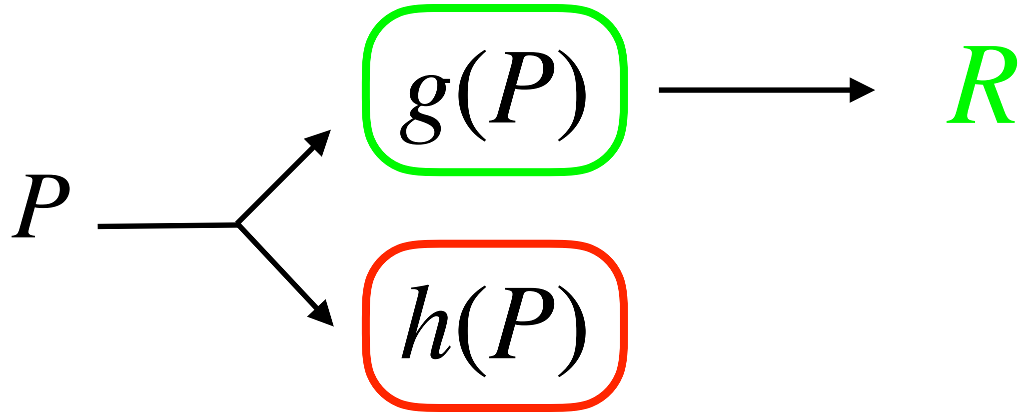


for cand. set selection

for error control

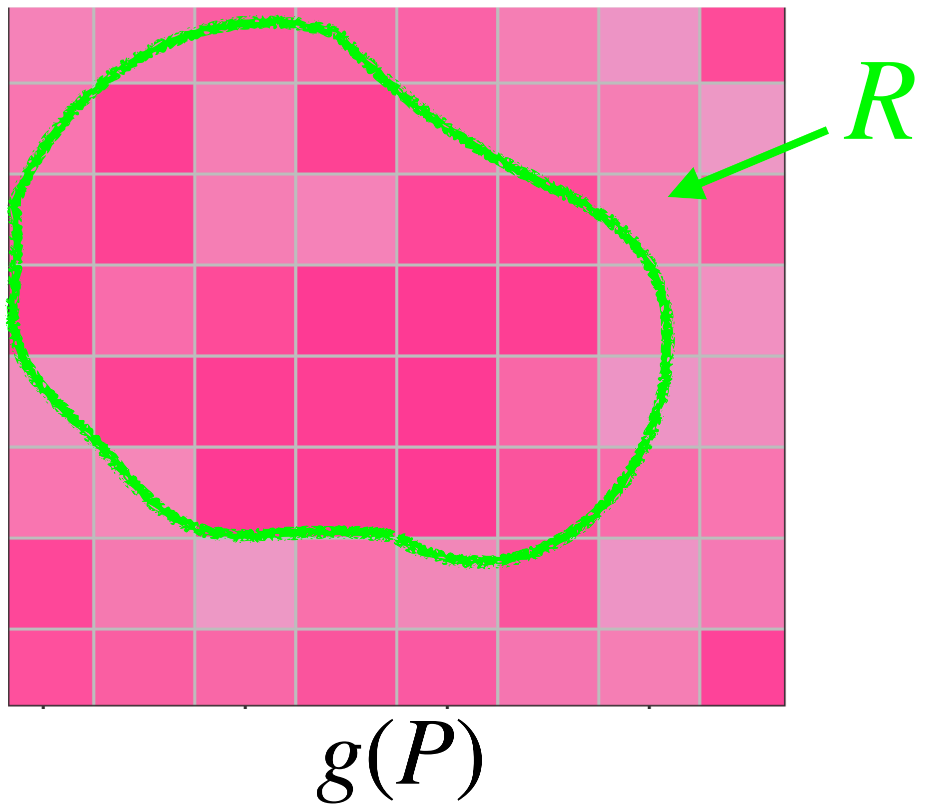


Component 2: select candidate set interactively



for cand. set selection

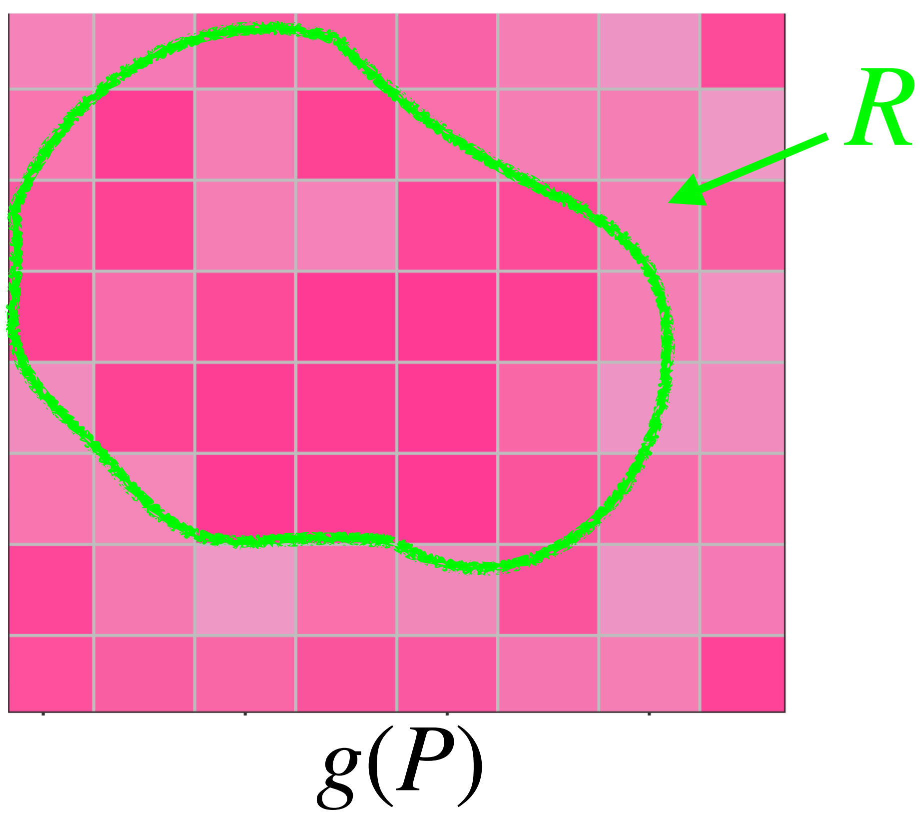
for error control



Component 2: select candidate set interactively

$P \to g(P) \to R$ for cand. set selection

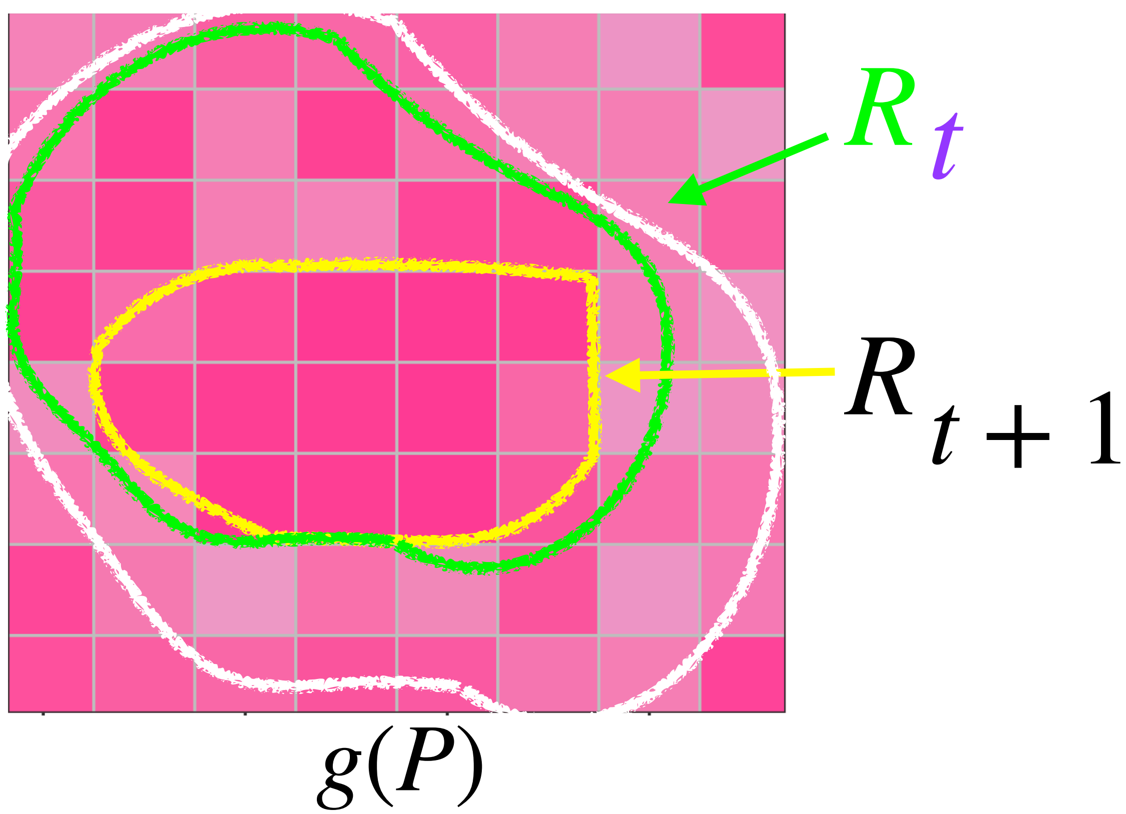
$P \to h(P) \to \widehat{\text{FWER}} \overset{?}{\leq} \alpha$ for error control



Component 2: select candidate set interactively

$P \to g(P) \longrightarrow R_t$ for cand. set selection

$P \to h(P) \longrightarrow \widehat{\text{FWER}}_t \overset{?}{\leq} \alpha$ for error control



$R_0 = \{1, \ldots, n\}$

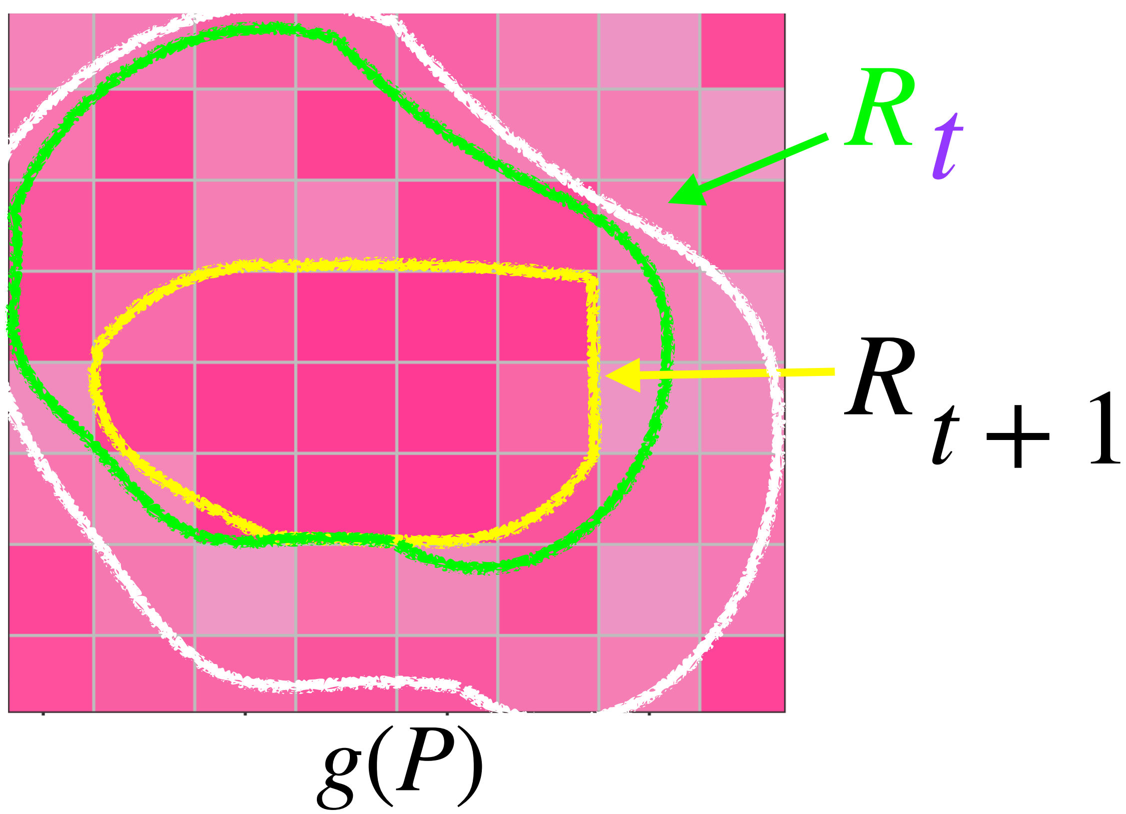
$R_0 \supseteq R_1 \supseteq \ldots$

progressively shrink R_t

Component 2: select candidate set interactively

$P \to g(P) \to R_t$ for cand. set selection

$P \to h(P) \to \widehat{\text{FWER}}_t \overset{?}{\leq} \alpha$ for error control



$R_0 = \{1, \ldots, n\}$

$R_0 \supseteq R_1 \supseteq \ldots$

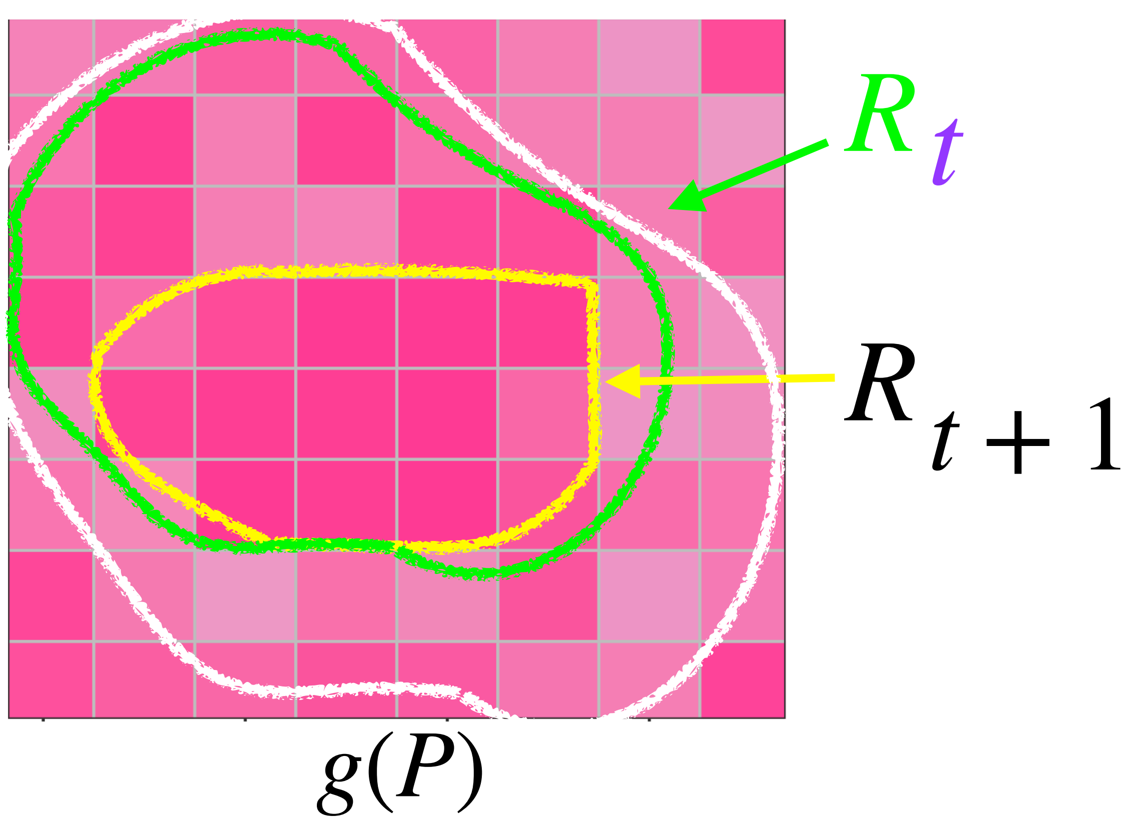
progressively shrink R_t

$\{g(P_i)\}_{i=1}^n$

Component 2: select candidate set interactively

$P \to g(P) \to R_t$ for cand. set selection

$P \to h(P) \to \widehat{\text{FWER}}_t \overset{?}{\leq} \alpha$ for error control



$R_0 = \{1, \ldots, n\}$

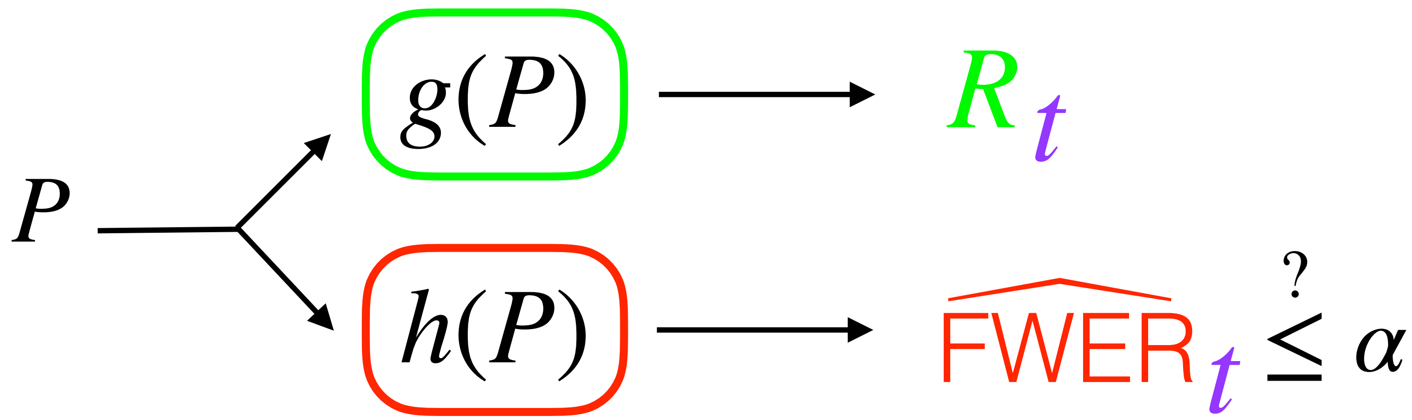
$R_0 \supseteq R_1 \supseteq \ldots$

progressively shrink R_t

$\{g(P_i)\}_{i=1}^n$

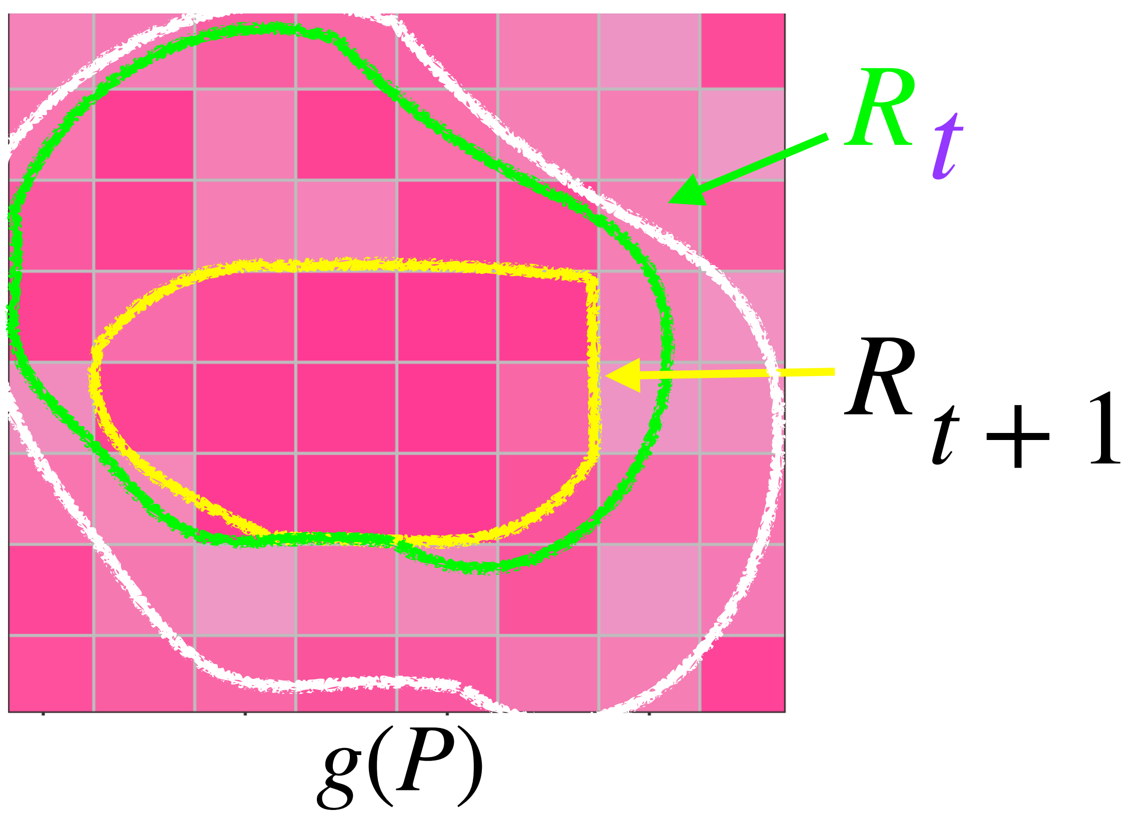
+ coordinates

Component 2: select candidate set interactively



for cand. set selection

for error control



$R_0 = \{1, \ldots, n\}$
$R_0 \supseteq R_1 \supseteq \ldots$

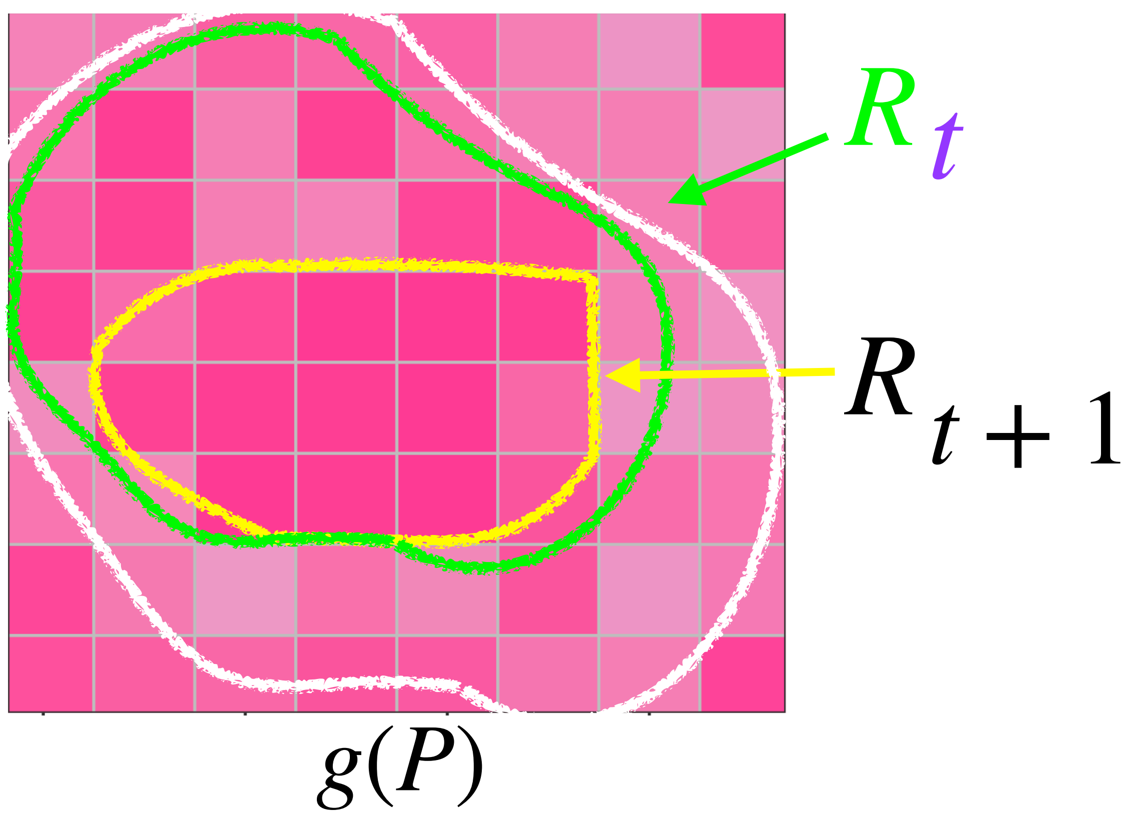
progressively shrink R_t

$\{g(P_i)\}_{i=1}^n$
\+ coordinates
(side information $\{x_i\}_{i=1}^n$)

Component 2: select candidate set interactively

$P \to g(P) \to R_t$ for cand. set selection

$P \to h(P) \to \widehat{\text{FWER}}_t \overset{?}{\leq} \alpha$ for error control

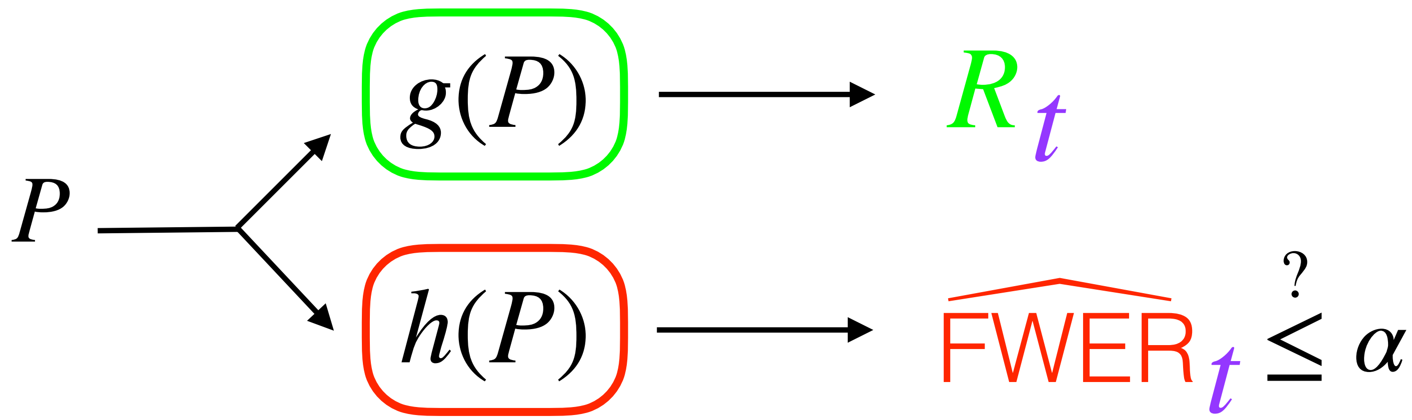


$R_0 = \{1, \ldots, n\}$
$R_0 \supseteq R_1 \supseteq \ldots$

progressively shrink R_t

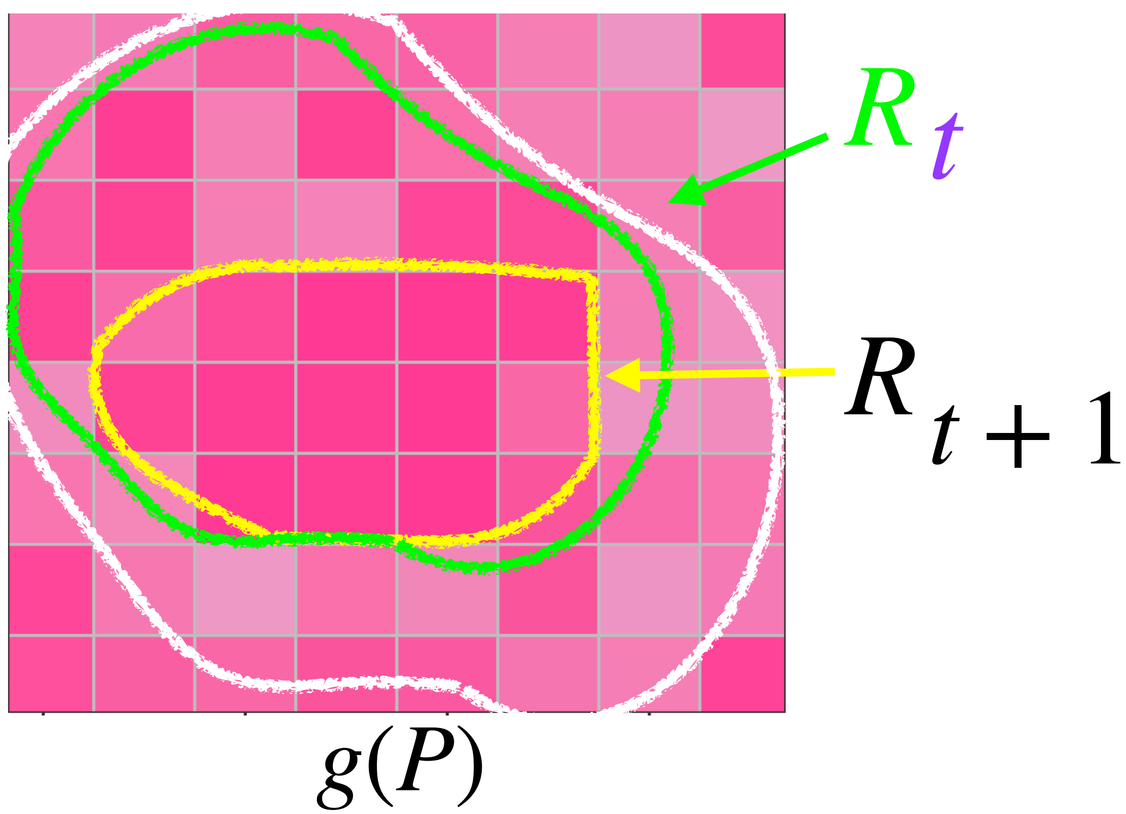
$\{g(P_i)\}_{i=1}^n$
+ coordinates
(side information $\{x_i\}_{i=1}^n$)
+ $\{h(P_i)\}_{i \notin R_t}$

Component 2: select candidate set interactively



for cand. set selection

for error control



$R_0 = \{1, \ldots, n\}$

$R_0 \supseteq R_1 \supseteq \ldots$

progressively shrink R_t using increasing information:

$\{g(P_i)\}_{i=1}^n$

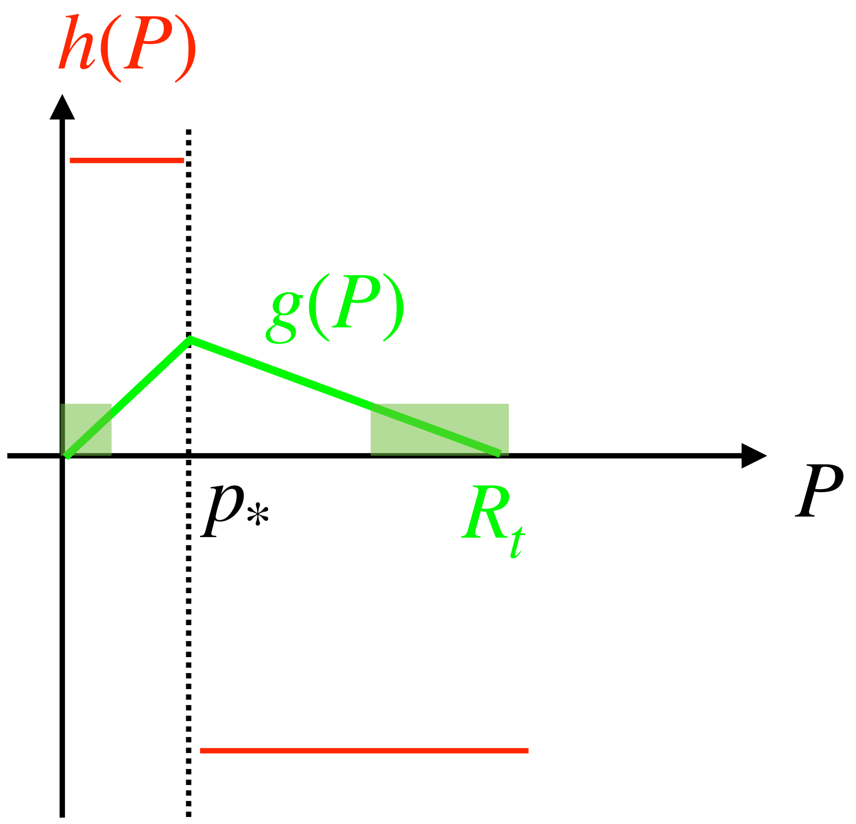
+ coordinates

(side information $\{x_i\}_{i=1}^n$)

+ $\{h(P_i)\}_{i \notin R_t}$

Component 3: error control

FWER := $\mathbb{P}$(#false rejections ≥ 1)

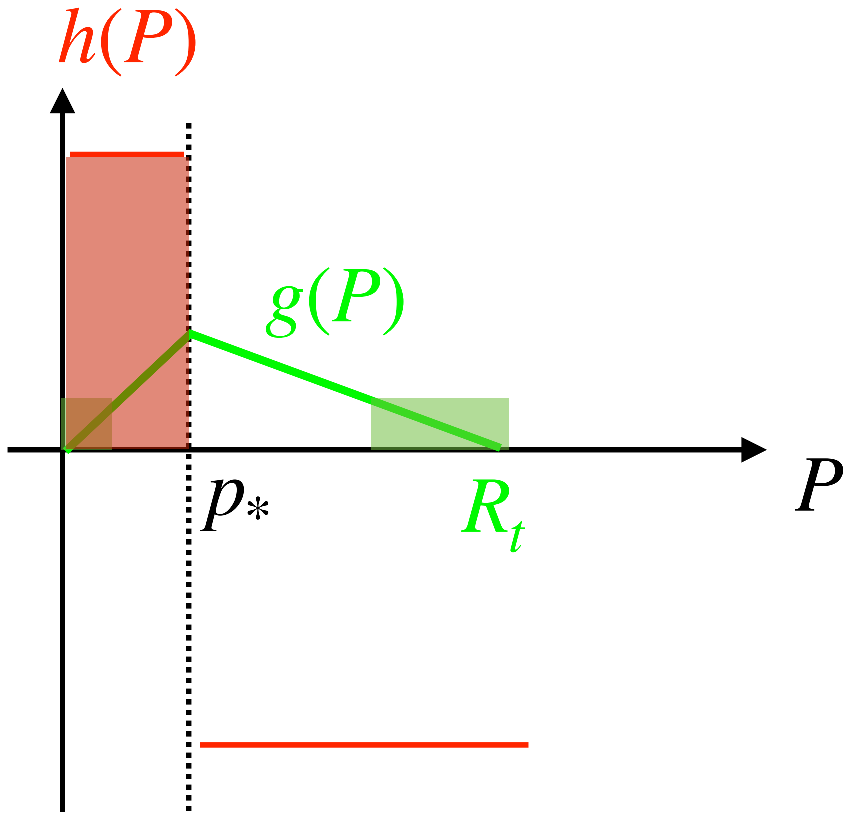


Component 3: error control

FWER := $\mathbb{P}$(#false rejections ≥ 1)

Reject the set

$$\{H_i \in R_t : h(P_i) = 1\}.$$



Component 3: error control

FWER := $\mathbb{P}$(#false rejections ≥ 1)

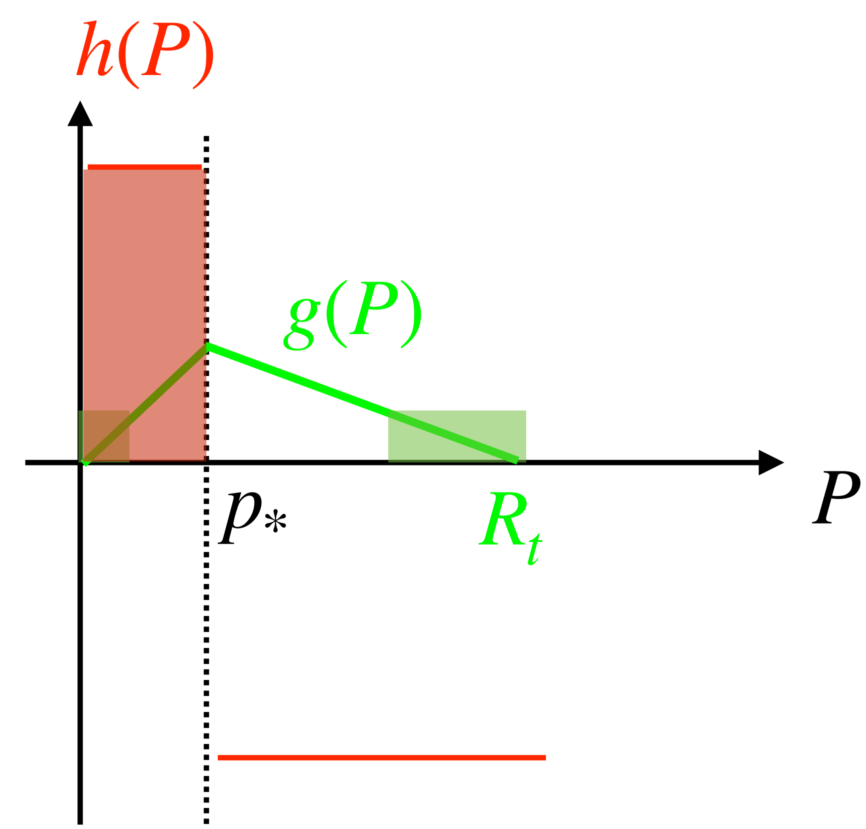
Reject the set

$$\{H_i \in R_t : h(P_i) = 1\}.$$

Count

$$R_t^- := |\{H_i \in R_t : h(P_i) = -1\}|$$

$h(P_i) \sim \mathrm{Ber}(p_*)$ if H_i is null



FWER control using binary var.: Janson & Su (2016)

Component 3: error control

FWER := $\mathbb{P}$(#false rejections ≥ 1)

Reject the set

$$\{H_i \in R_t : h(P_i) = 1\}.$$

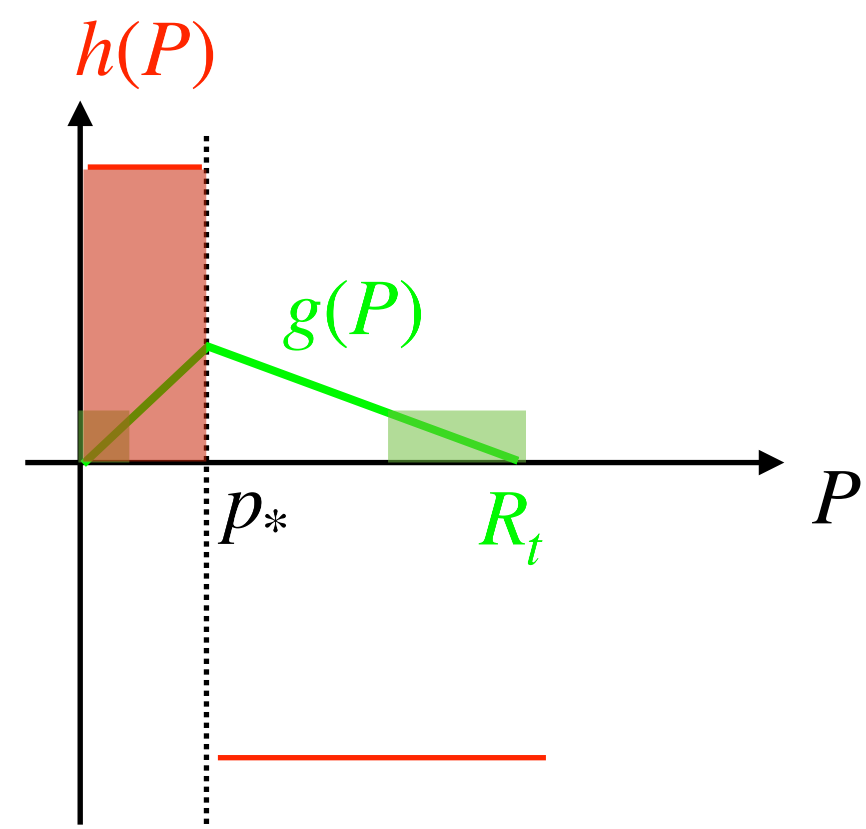
Count

$$R_t^- := |\{H_i \in R_t : h(P_i) = -1\}|$$

$h(P_i) \sim \mathrm{Ber}(p_*)$ if H_i is null

Stop shrinking when

$$\widehat{\text{FWER}}_t := 1 - (1 - p_*)^{R_t^- + 1} \leq \alpha.$$



FWER control using binary var.: Janson & Su (2016)

Component 3: error control

FWER := $\mathbb{P}$(#false rejections ≥ 1)

Reject the set

$$\{H_i \in R_t : h(P_i) = 1\}.$$

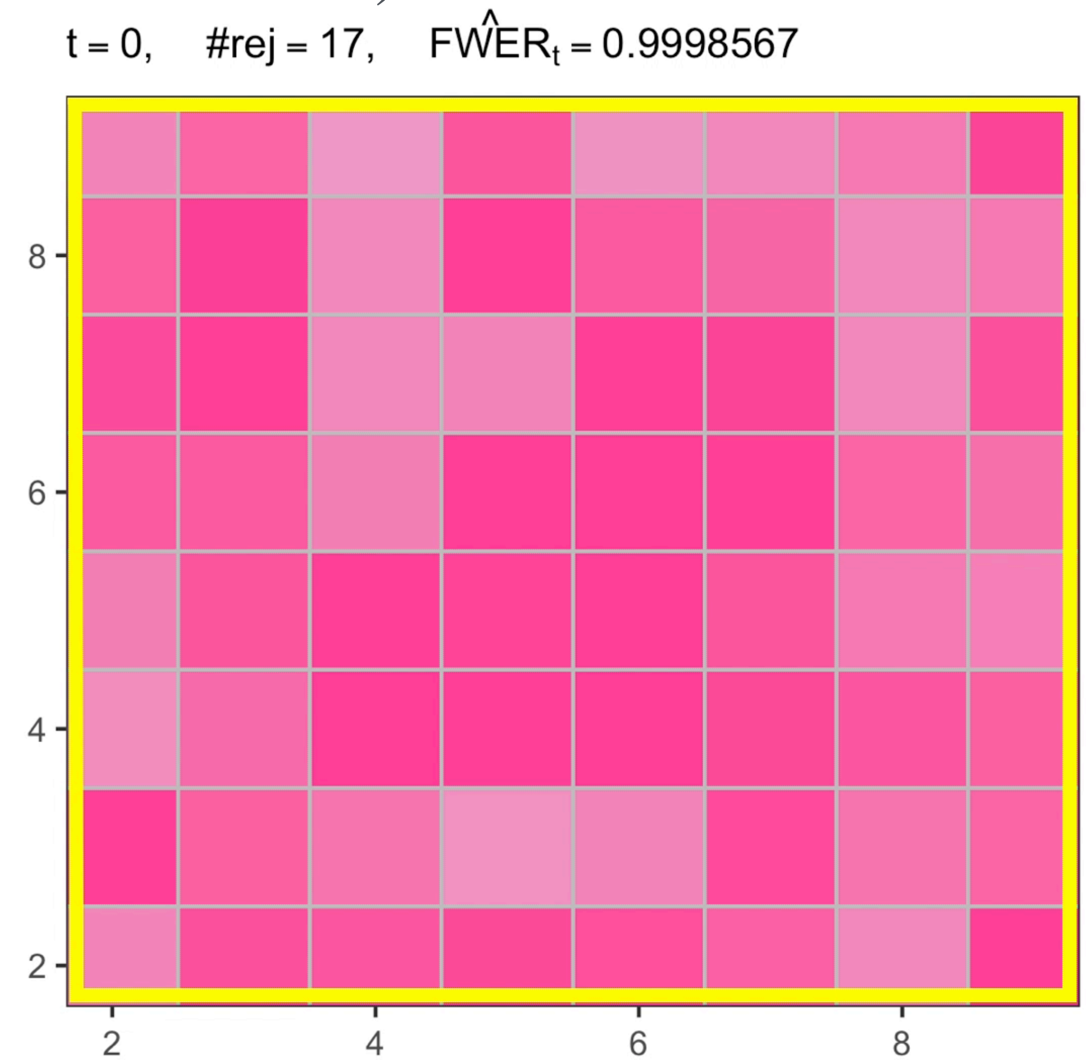
Count

$$R_t^- := |\{H_i \in R_t : h(P_i) = -1\}|$$

$h(P_i) \sim \text{Ber}(p_*)$ if H_i is null

Stop shrinking when

$$\widehat{\text{FWER}}_t := 1 - (1 - p_*)^{R_t^- + 1} \leq \alpha.$$



10×10 hypotheses with 9 non-nulls ($\alpha = 0.2$)

FWER control using binary var.: Janson & Su (2016)

Component 3: error control

FWER := $\mathbb{P}$(#false rejections ≥ 1)

Reject the set

$$\{H_i \in R_t : h(P_i) = 1\}.$$

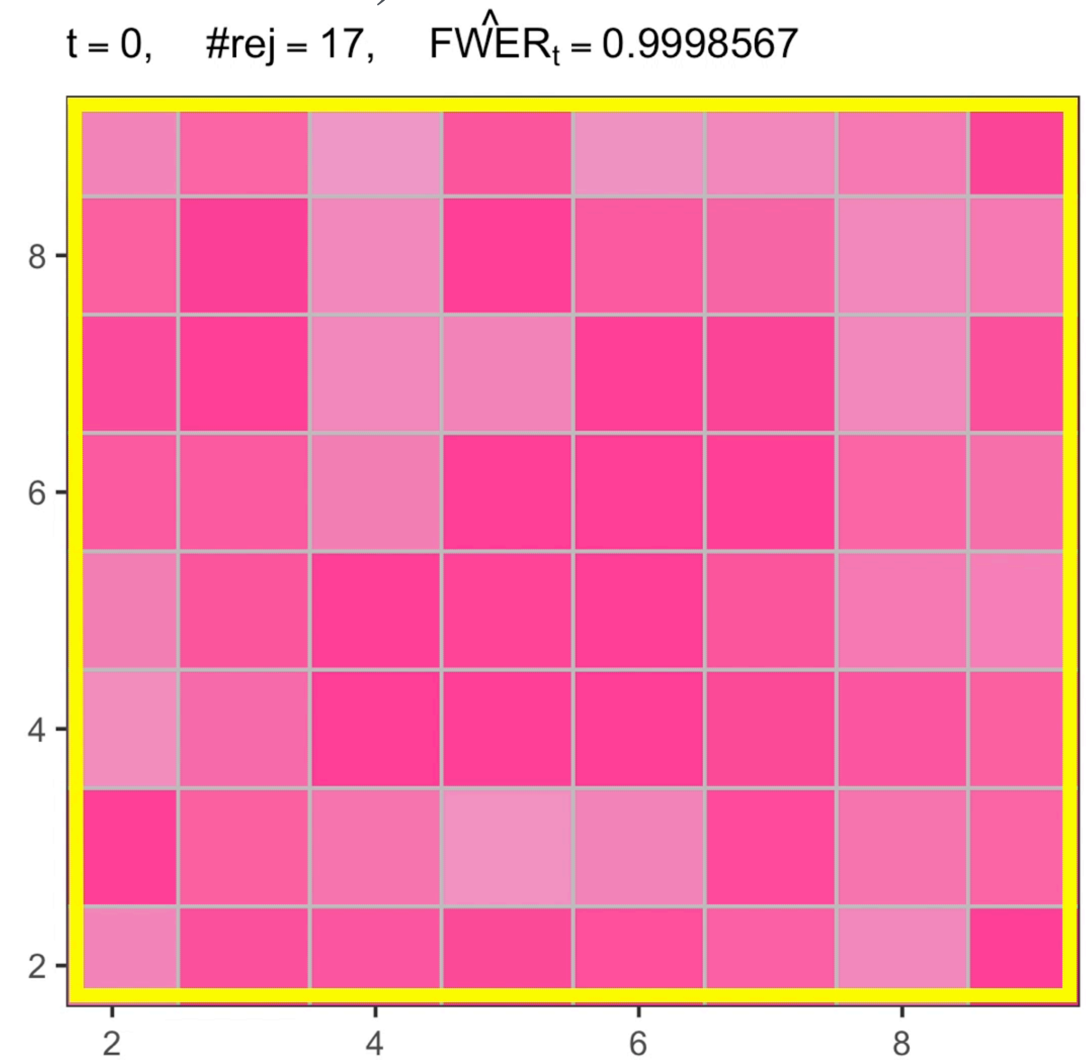
Count

$$R_t^- := |\{H_i \in R_t : h(P_i) = -1\}|$$

$h(P_i) \sim \mathrm{Ber}(p_*)$ if H_i is null

Stop shrinking when

$$\widehat{\text{FWER}}_t := 1 - (1 - p_*)^{R_t^- + 1} \leq \alpha.$$



10×10 hypotheses with 9 non-nulls ($\alpha = 0.2$)

FWER control using binary var.: Janson & Su (2016)

Component 3: error control

FWER := $\mathbb{P}$(#false rejections ≥ 1)

Reject the set

$$\{H_i \in R_t : h(P_i) = 1\}.$$

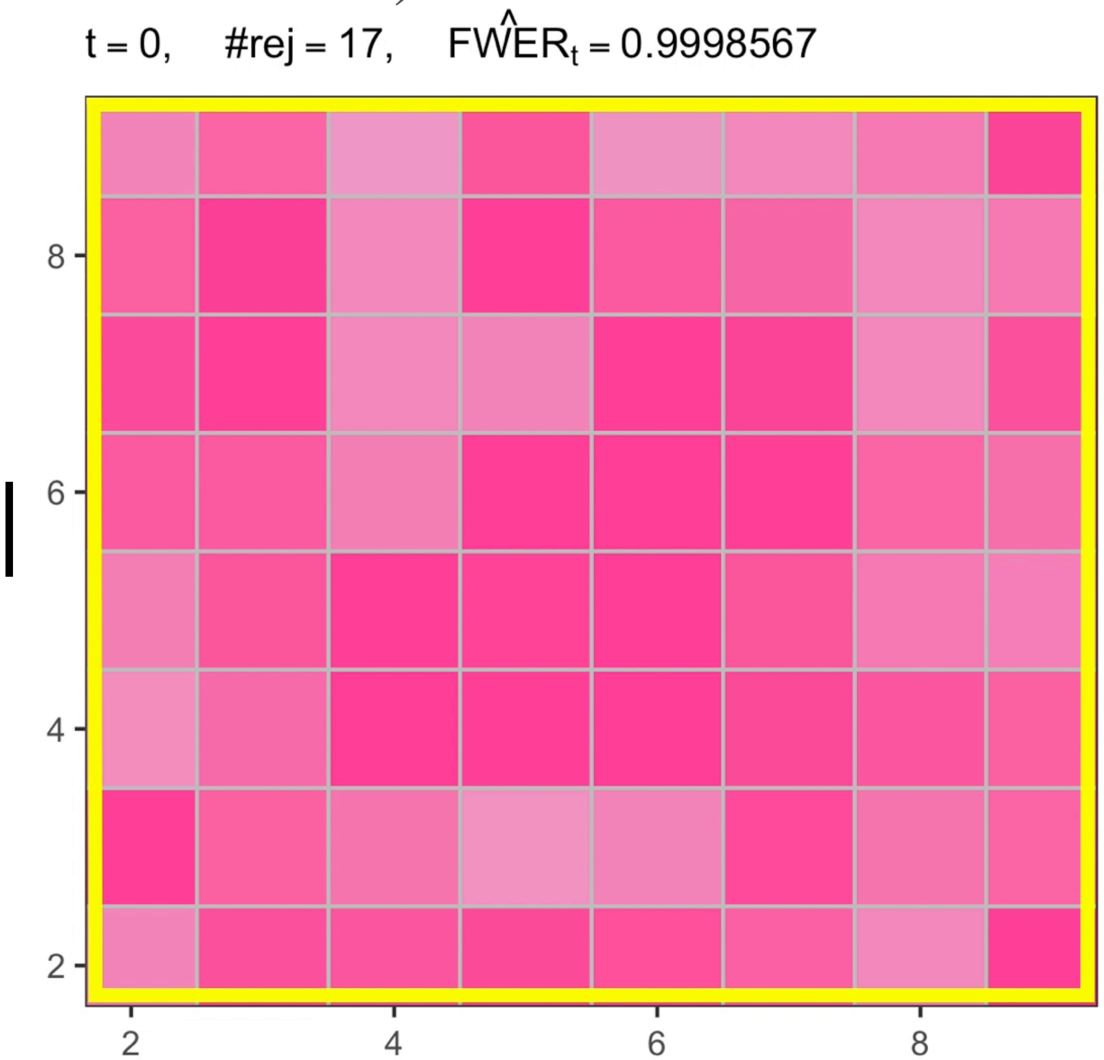
Count

$$R_t^- := |\{H_i \in R_t : h(P_i) = -1\}|$$

$h(P_i) \sim \text{Ber}(p_*)$ if H_i is null

Stop shrinking when

$$\widehat{\text{FWER}}_t := 1 - (1 - p_*)^{R_t^- + 1} \leq \alpha.$$



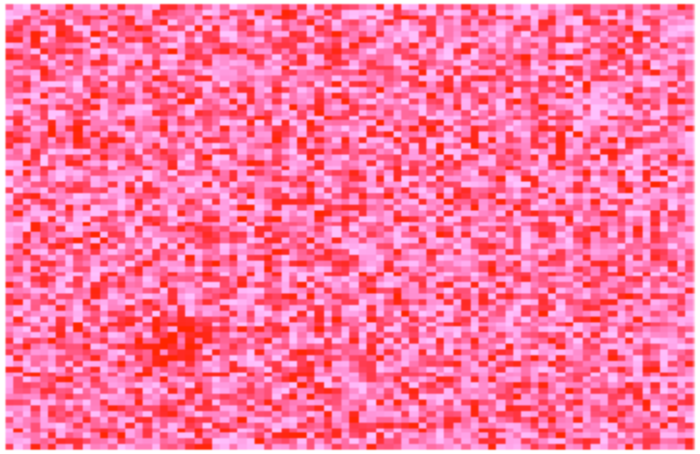
10×10 hypotheses with 9 non-nulls
($\alpha = 0.2$)

Theorem: the i-FWER method control FWER if null p-values are mutually independent, and are independently of the non-nulls.

FWER control using binary var.:
Janson & Su (2016)

Strategies to shrink R_t

- Estimate the probability of being non-null.

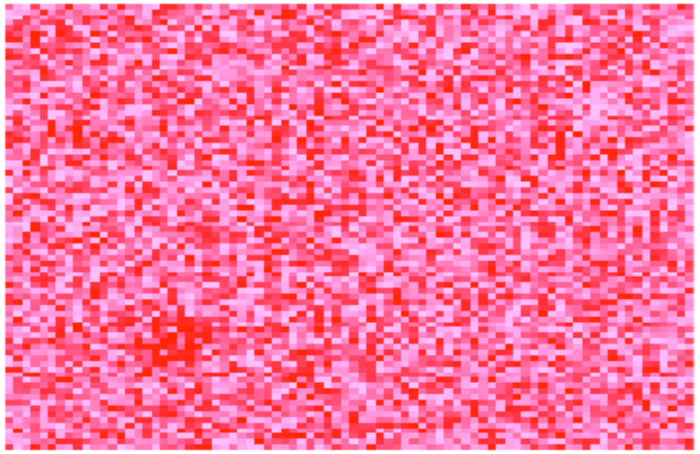


$g(P)$

smoothed probability

Strategies to shrink R_t

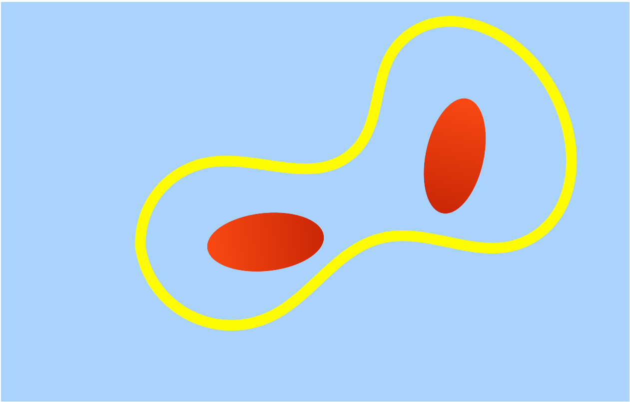
- Estimate the probability of being non-null.



$g(P)$

smoothed probability

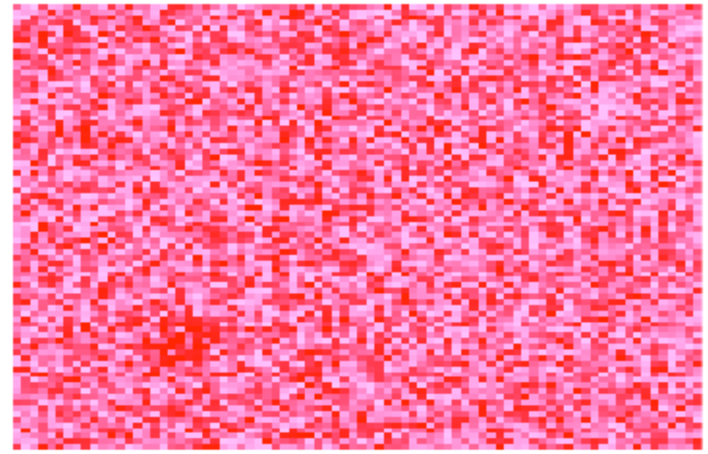
- Change the strategy at any step.



single center?

Strategies to shrink R_t

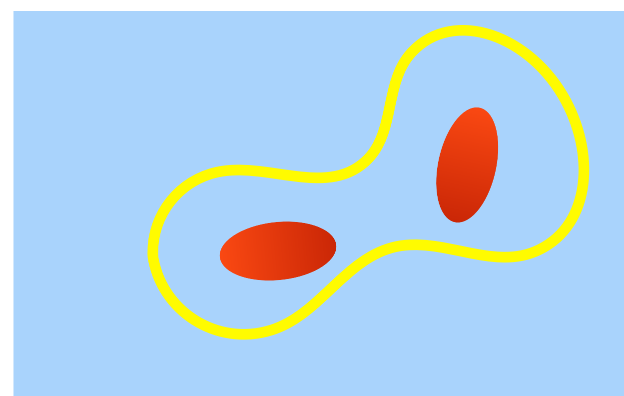
- Estimate the probability of being non-null.



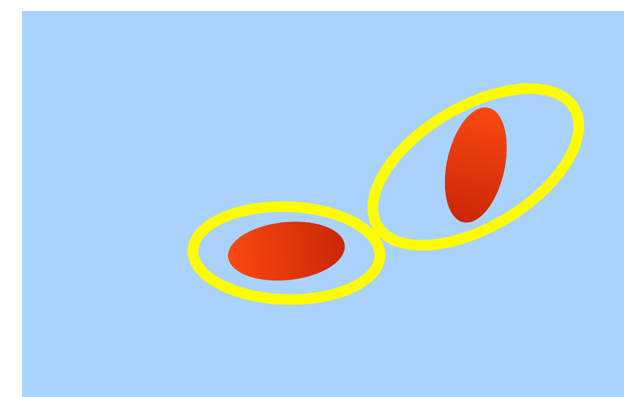
$g(P)$

smoothed probability

- Change the strategy at any step.



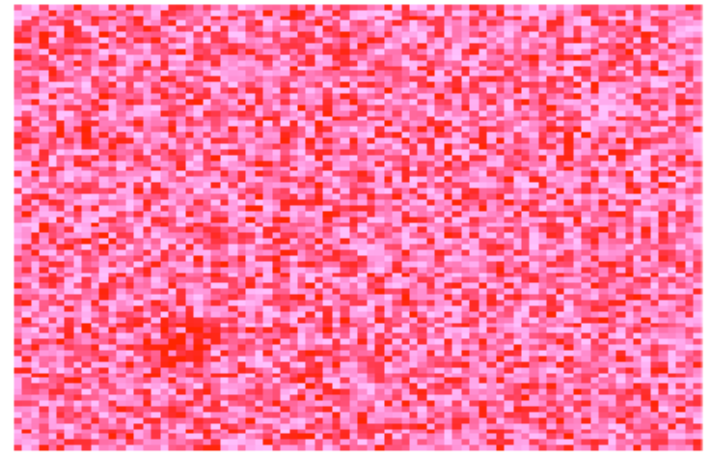
single center?



two centers

Strategies to shrink R_t

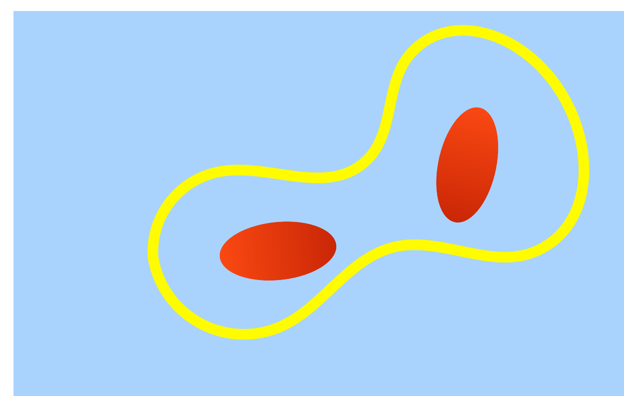
- Estimate the probability of being non-null.



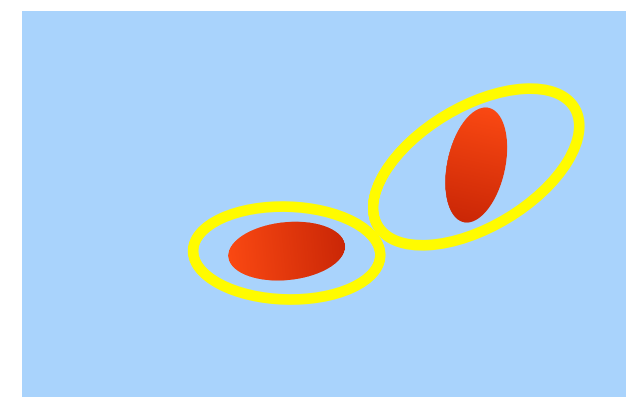
$g(P)$

smoothed probability

- Change the strategy at any step.

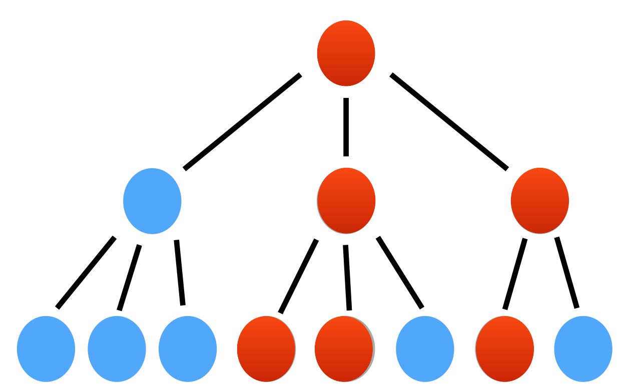


single center?



two centers

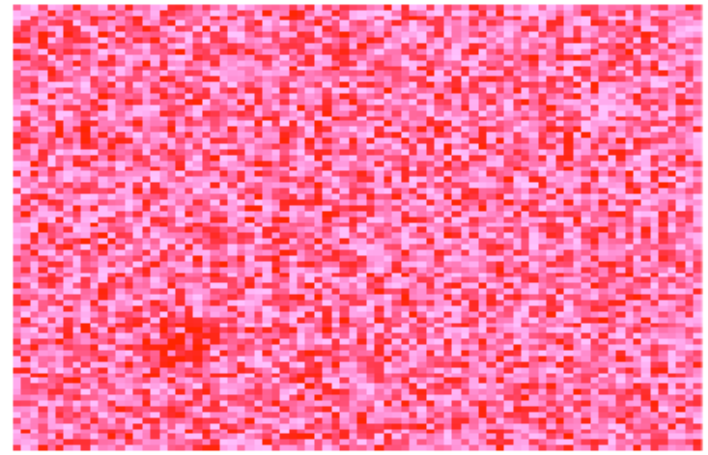
- Customize the strategy to different structures.



Hierarchical structure

Strategies to shrink R_t

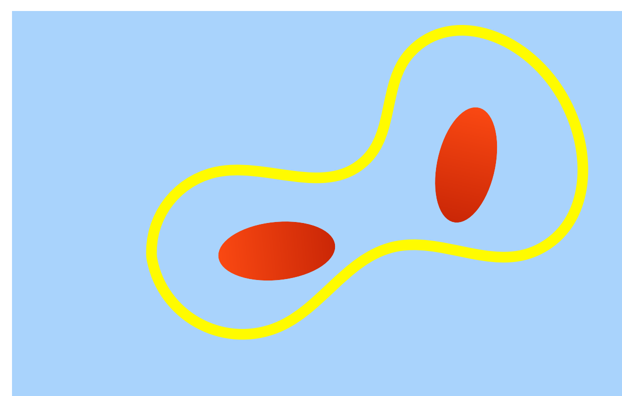
- Estimate the probability of being non-null.



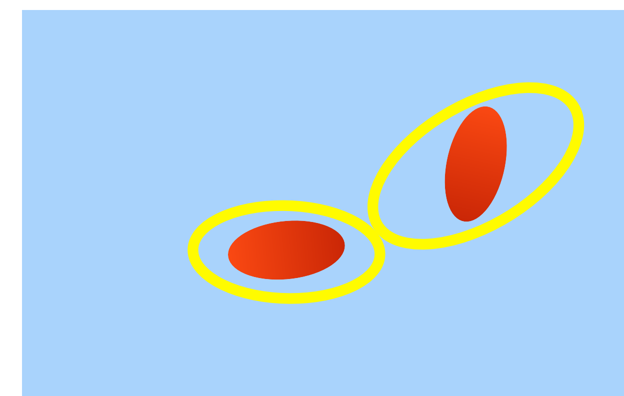
$g(P)$

smoothed probability

- Change the strategy at any step.

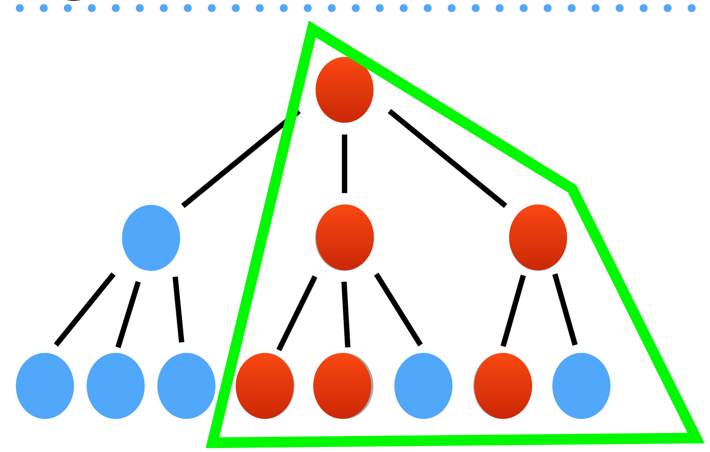


single center?



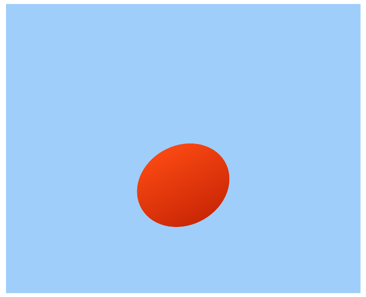
two centers

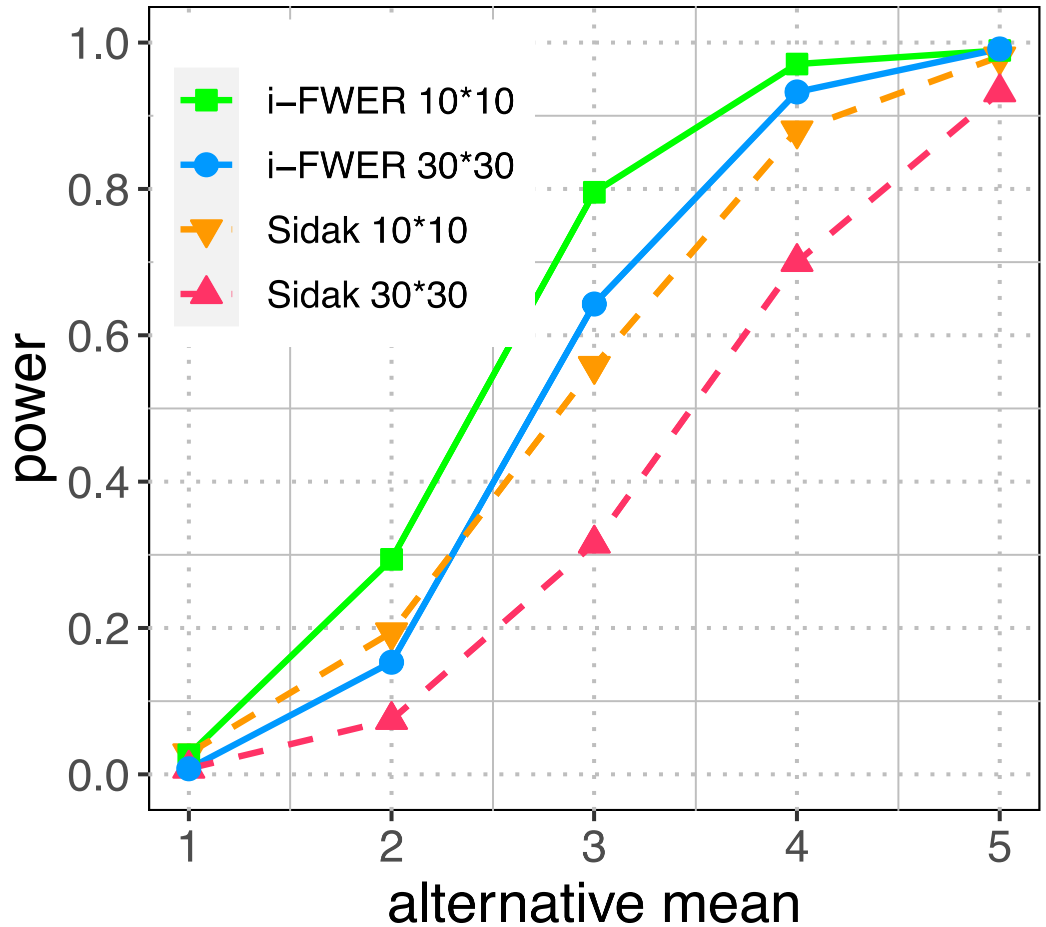
- Customize the strategy to different structures.



Hierarchical structure

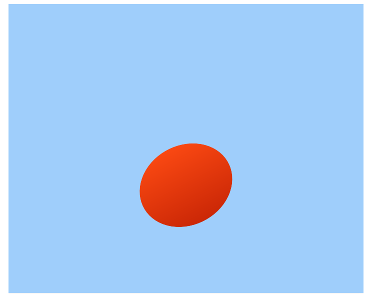
Grid

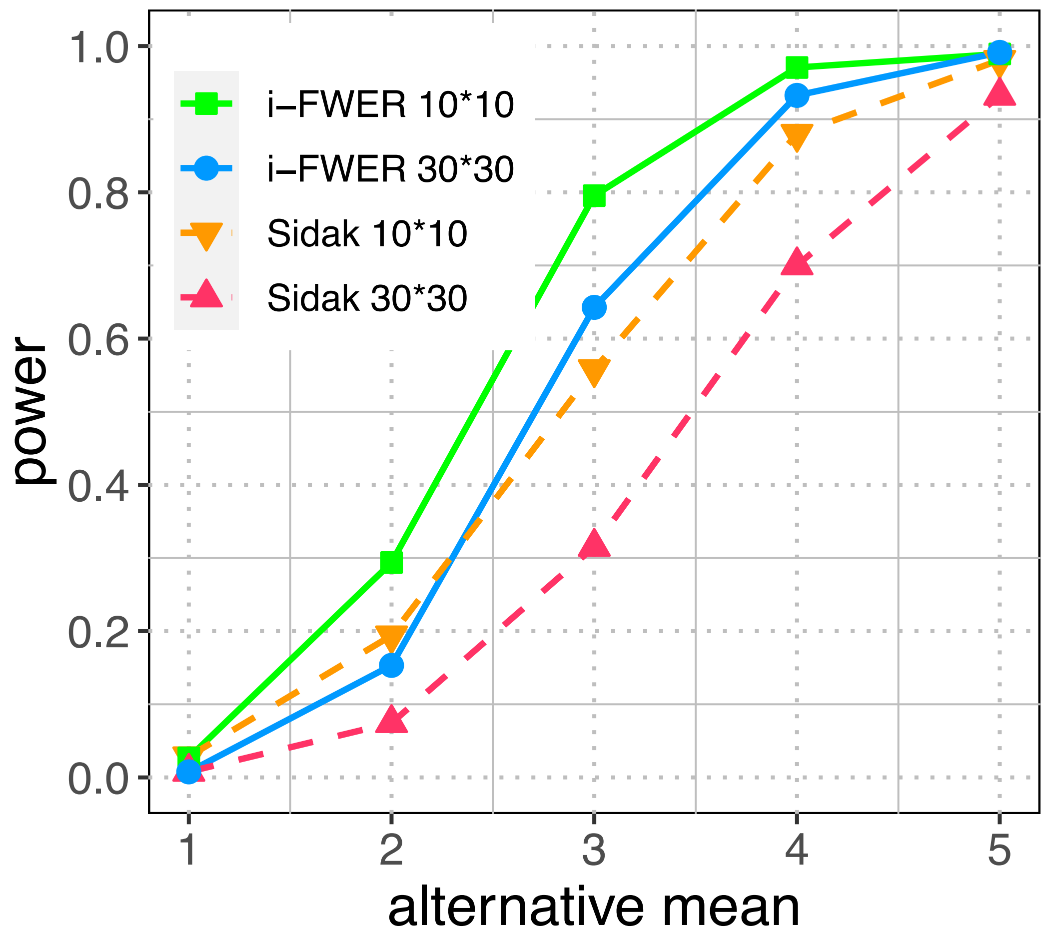




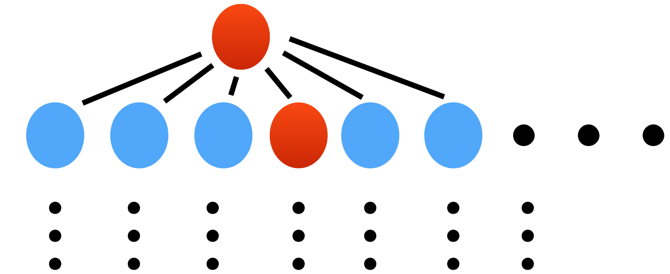
Meinshausen (2008); Goeman and Finos (2012); Ignatiadis et al. (2016); Lei, Fithian (2018)

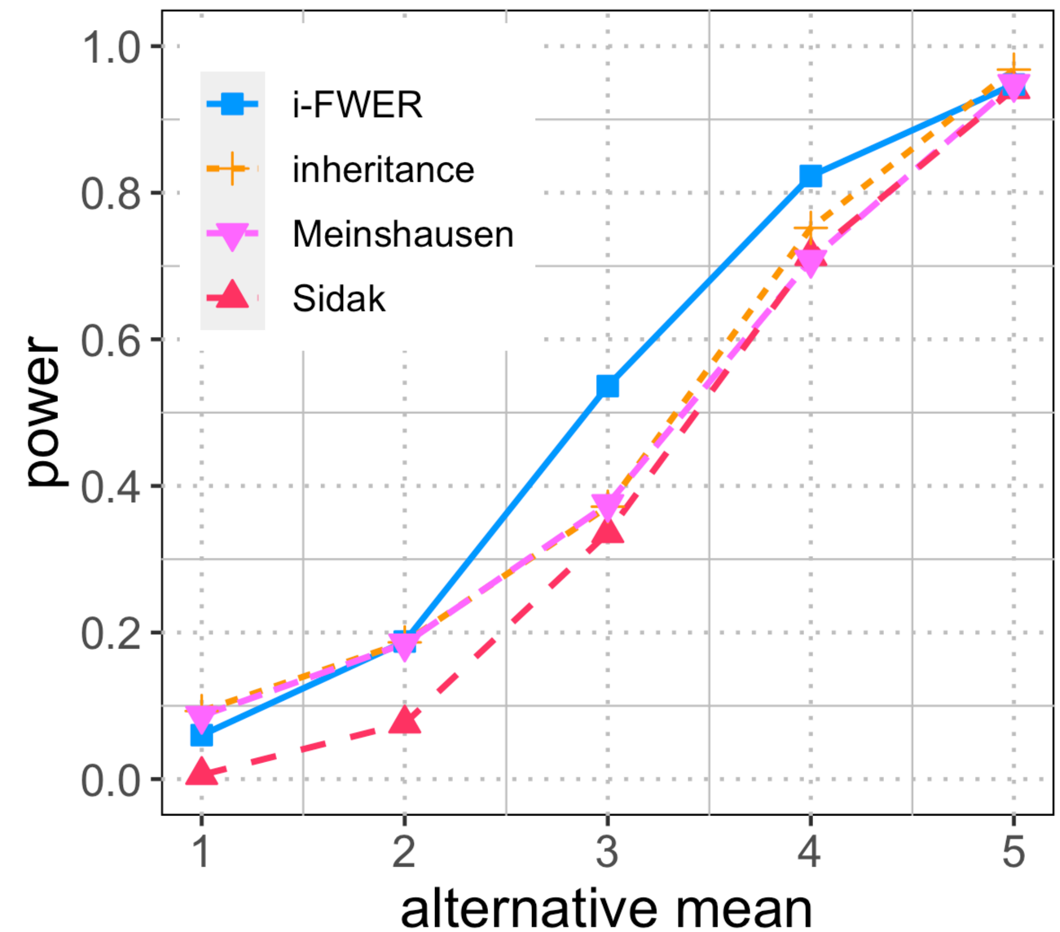
Grid





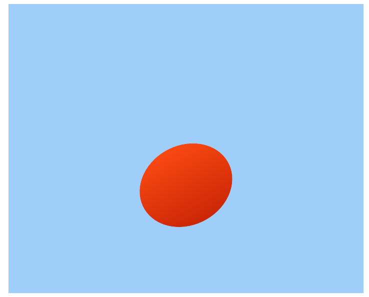
Tree

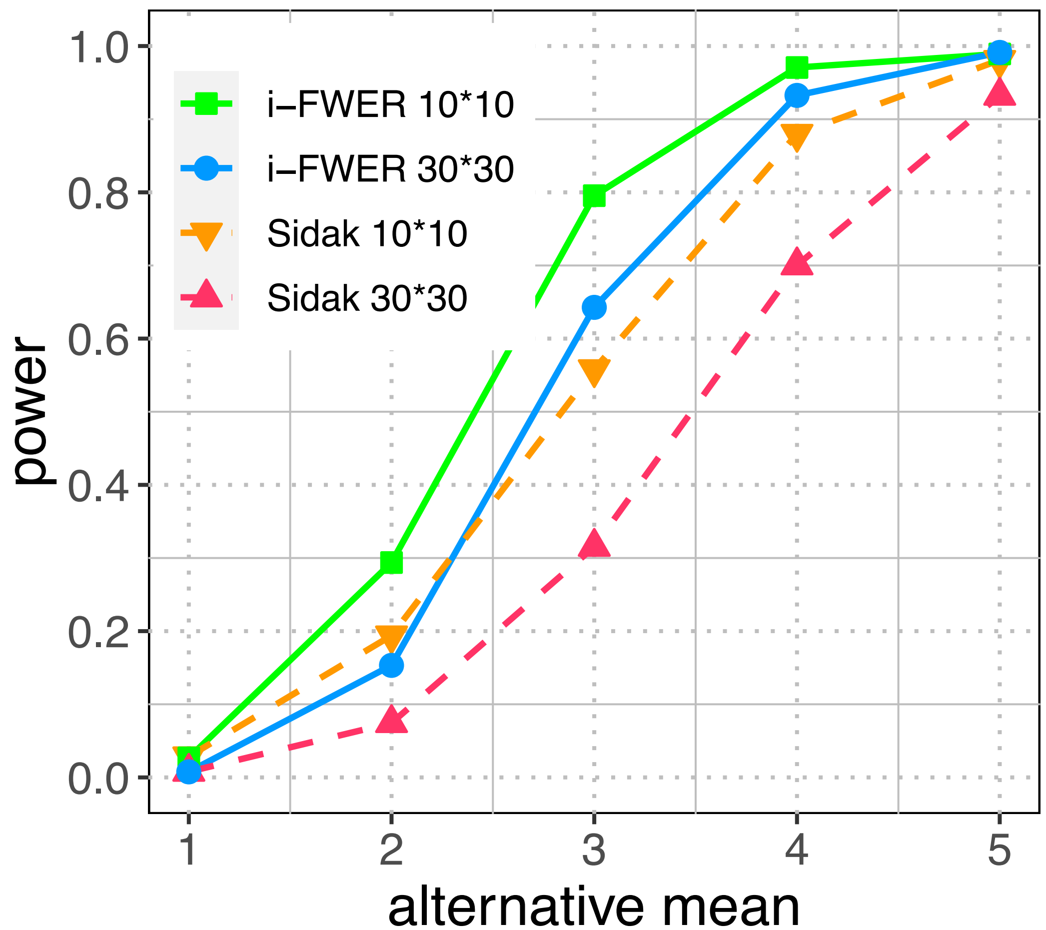




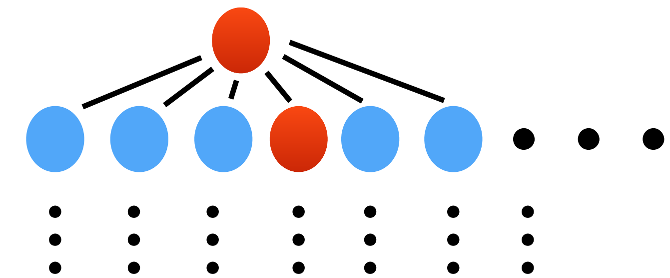
Meinshausen (2008); Goeman and Finos (2012); Ignatiadis et al. (2016); Lei, Fithian (2018)

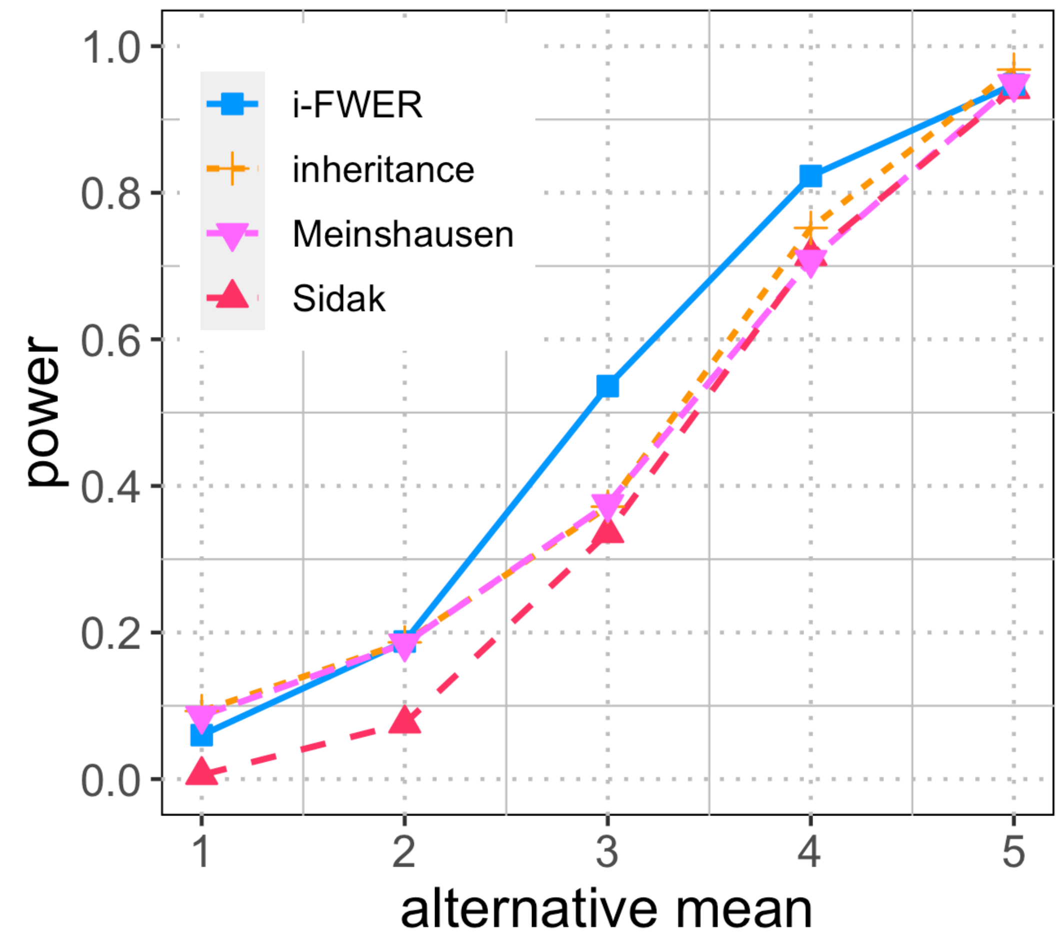
Grid





Tree





Real data: RNA sequence

FWER level	IHW★	i-FWER
0.1	1552	1613
0.2	1645	1740
0.3	1708	1844

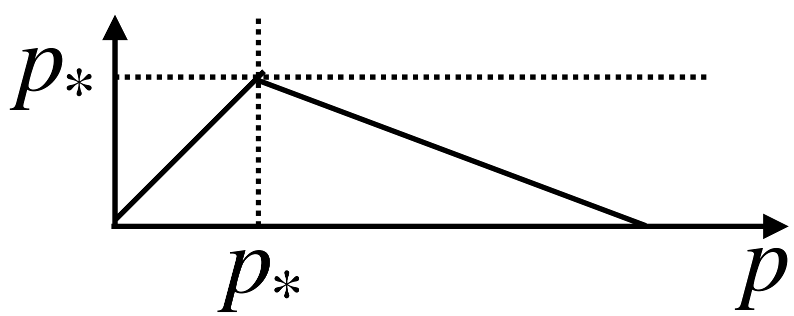
differentially expressed genes
in airway smooth muscle cell lines
in response to dexamethasone

Meinshausen (2008); Goeman and Finos (2012);
★Ignatiadis et al. (2016); Lei, Fithian (2018)

Extensions

Masking functions

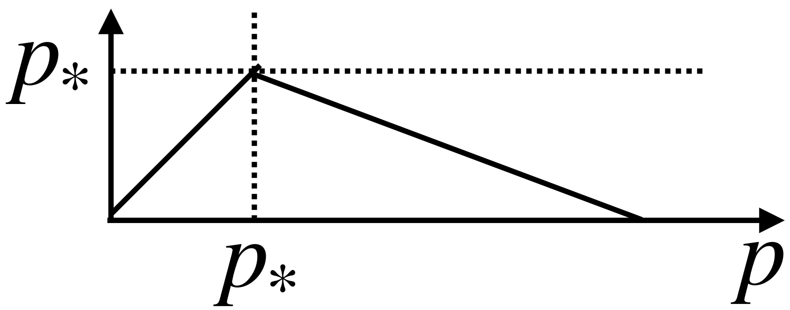
Tent function



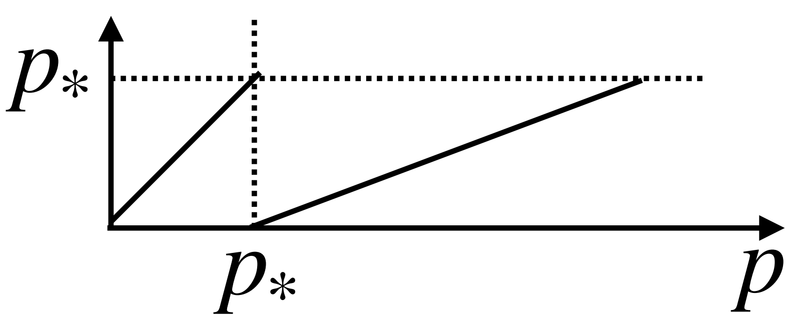
Extensions

Masking functions

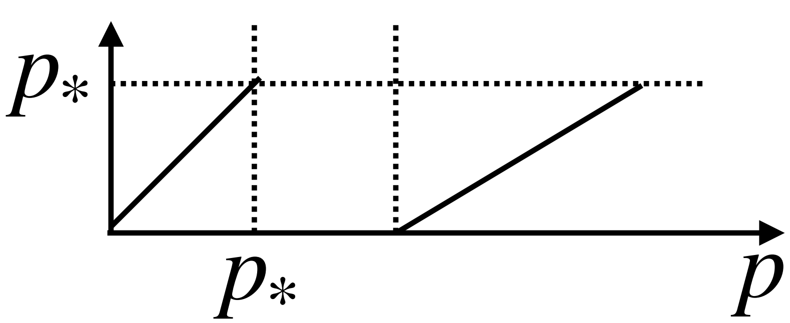
Tent function



Railway function



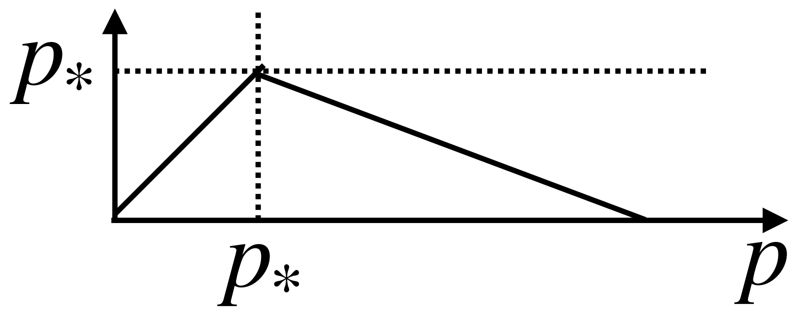
Gap function



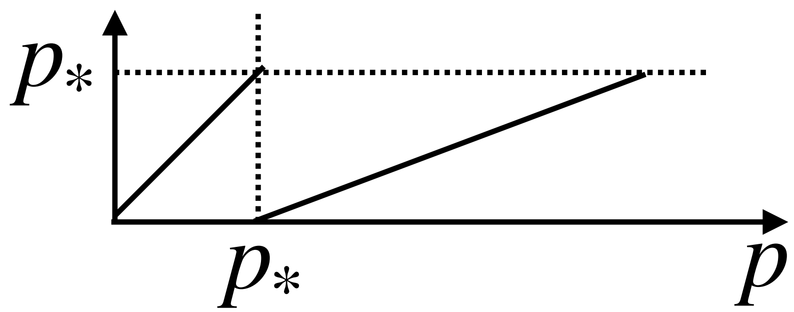
Extensions

Masking functions

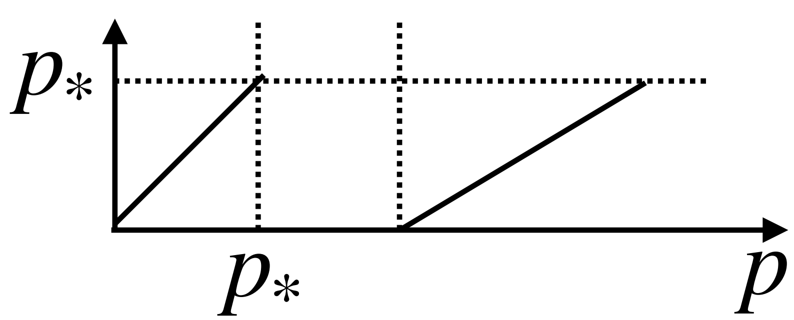
Tent function



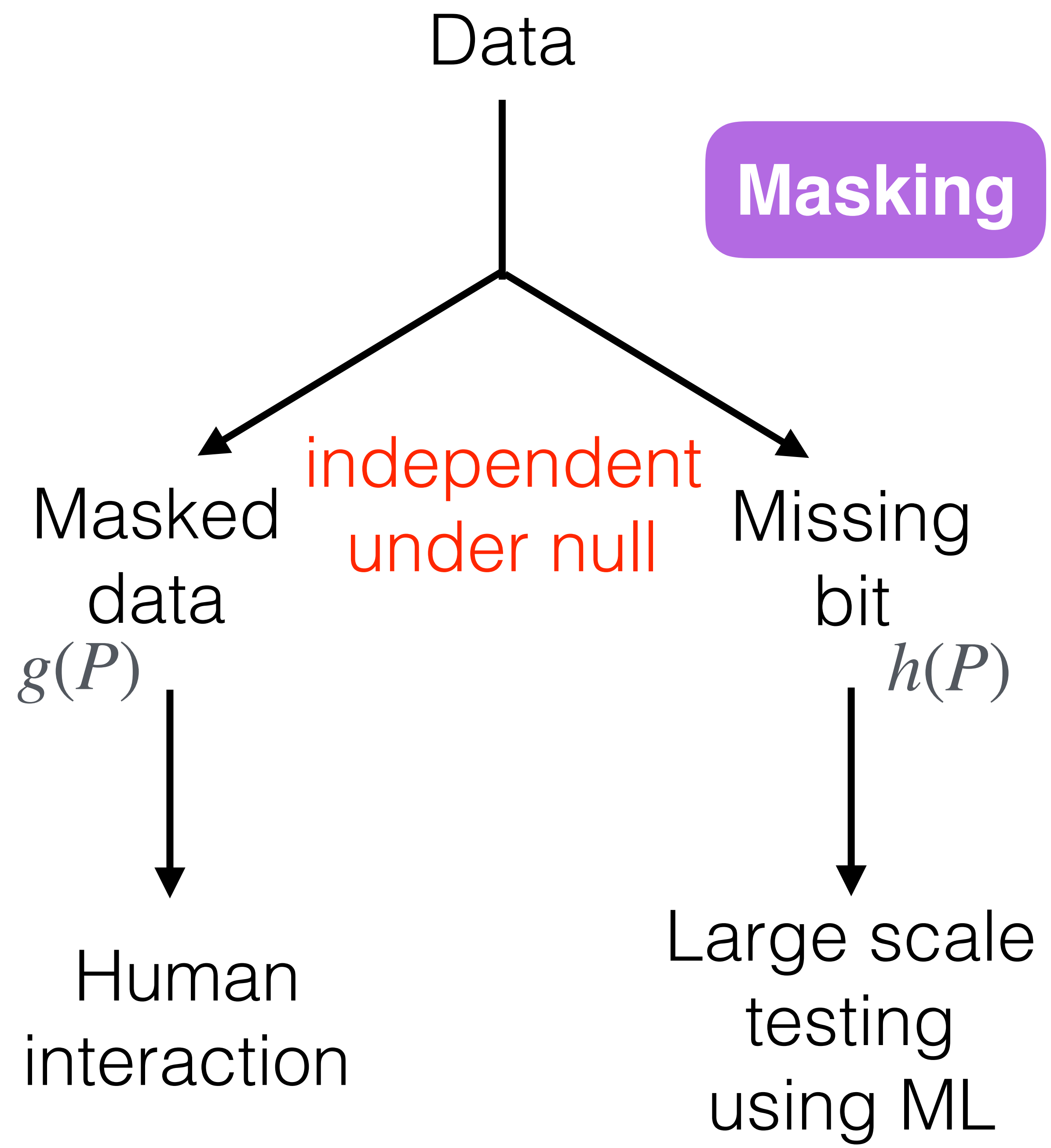
Railway function



Gap function



Interactive testing



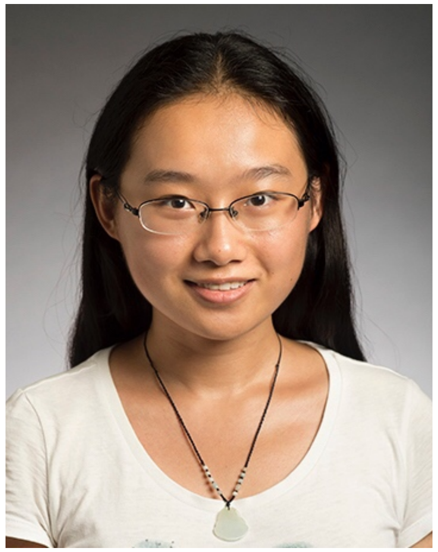
Dept. of Statistics and Data Science
Carnegie Mellon University

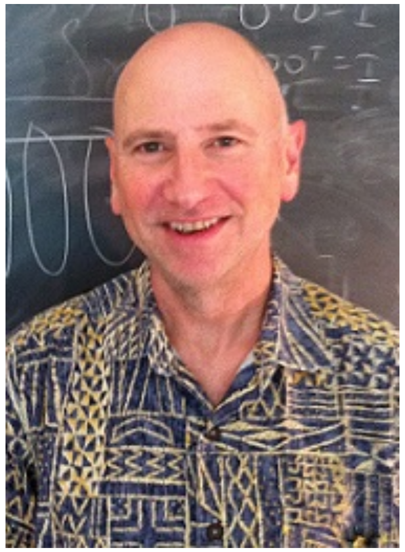
Familywise error rate control by interactive unmasking

Boyan Duan

Aaditya Ramdas

Larry Wasserman





Thank you!