I^3 : Interactive Identification of Individuals with positive treatment effect while controlling FDR

Aaditya Ramdas

Associate Professor
Dept. of Statistics and Data Science (75%)
Machine Learning Dept. (25%)
Carnegie Mellon University

Amazon Research (20%)

Data scientists *want* to interact with data, and use their intuition, priors to exploit structure in the data.

However, many classical statistical methods were not built to handle interactivity.

"double-dipping" selection bias

There is a need for principled statistical methods that work in sequential and interactive settings.

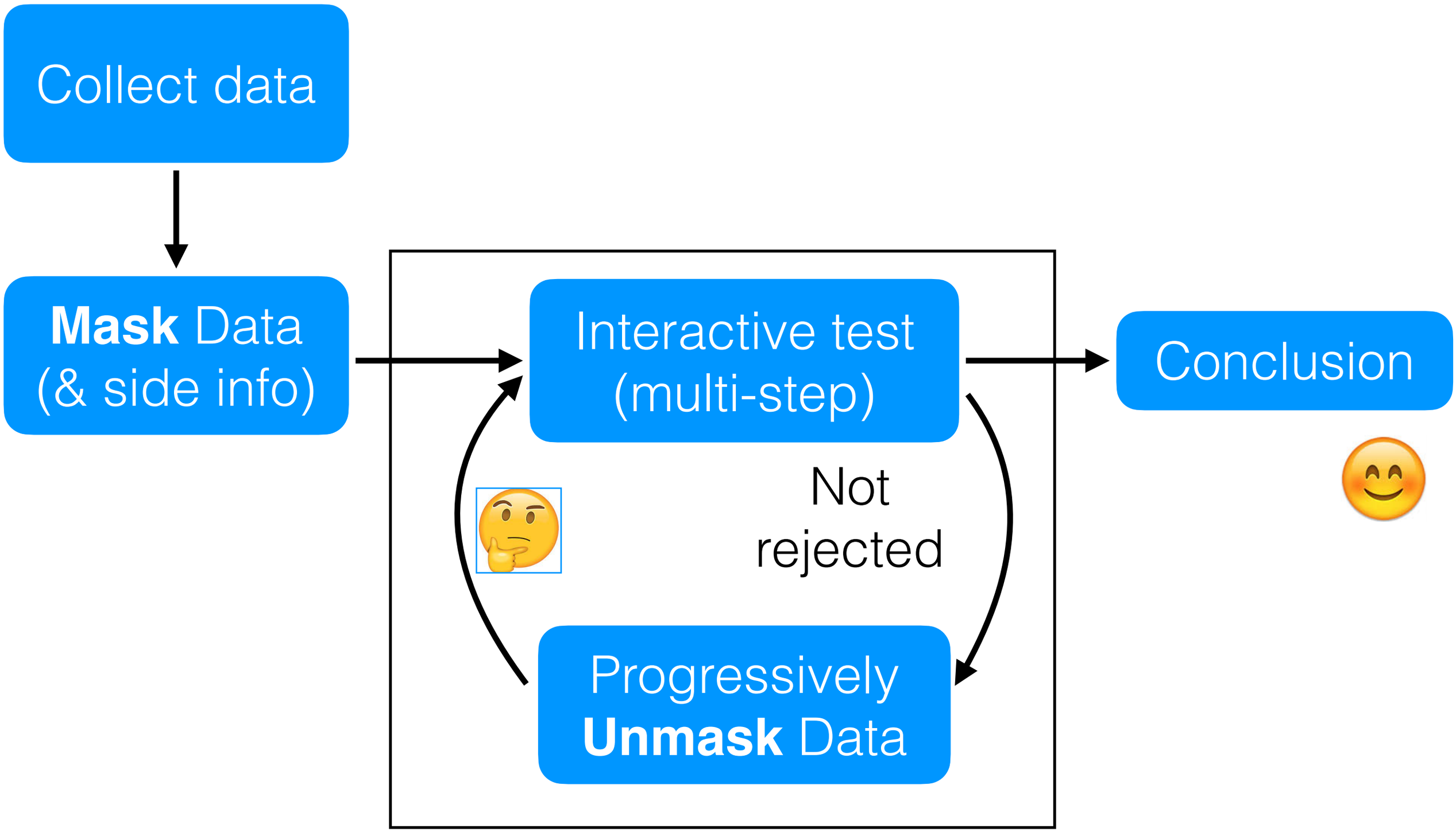
A hypothetical data scientists' wish-list:

- use prior knowledge and intuition
- incorporate structure, soft or hard constraints
- interactive exploration with human-in-the-loop
- employ flexible probabilistic modeling tools
- robust to unknown dependence

The challenge: correct "statistical inference"

This talk: some progress within multiple testing

Interactive testing allows “loops”



Test can be customized and revised after observing data.

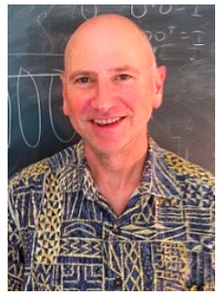
Interactive multiple testing


Time

Lei & Fithian (JRSSB'18) conceptualized masking, designed AdaPT algorithm (i-FDR)

Lei, R, Fithian (Biometrika'21) STAR algorithm (i-FDR) enforces structure, generalized masking

Duan, R, Balakrishnan, Wasserman (EJS'21) martingale tests (i-global null)

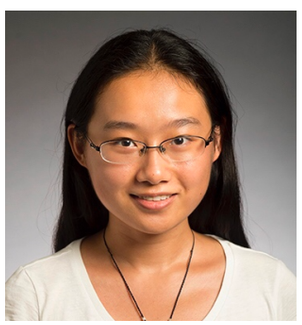


Duan, R, Wasserman (ICML'20) handles conservative nulls, better, flexible masking (i-FWER)

Duan, R, Wasserman (CLEAR'23) interactive Wilcoxon and Friedman tests (i-rank tests)

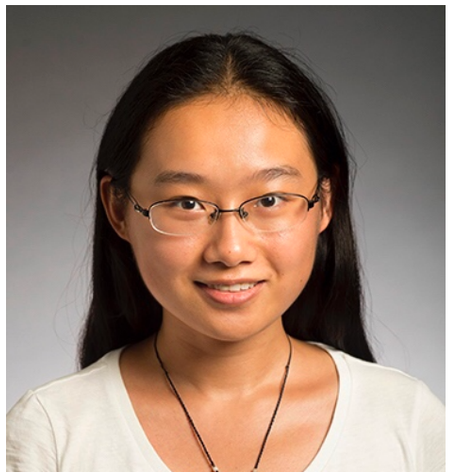
Duan, R, Wasserman (JCI'24) moves away from p-values (i-FDR-causal)

Leiner, Duan, Wasserman, R (JASA discussion paper 2024) Data fission: splitting a single data point (i-FCR)



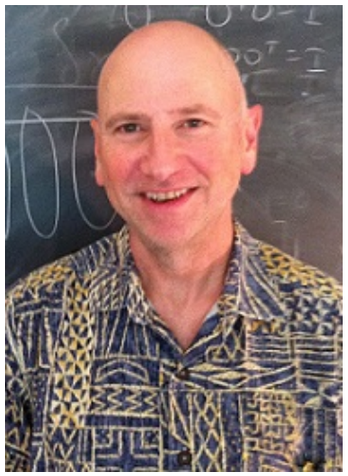
Boyan Duan (PhD 2021; currently, Google)
Thesis: Advances in interactive inference.

Interactive identification of individuals with positive treatment effect while controlling false discoveries (**FDR**)



Boyan Duan

Aaditya Ramdas



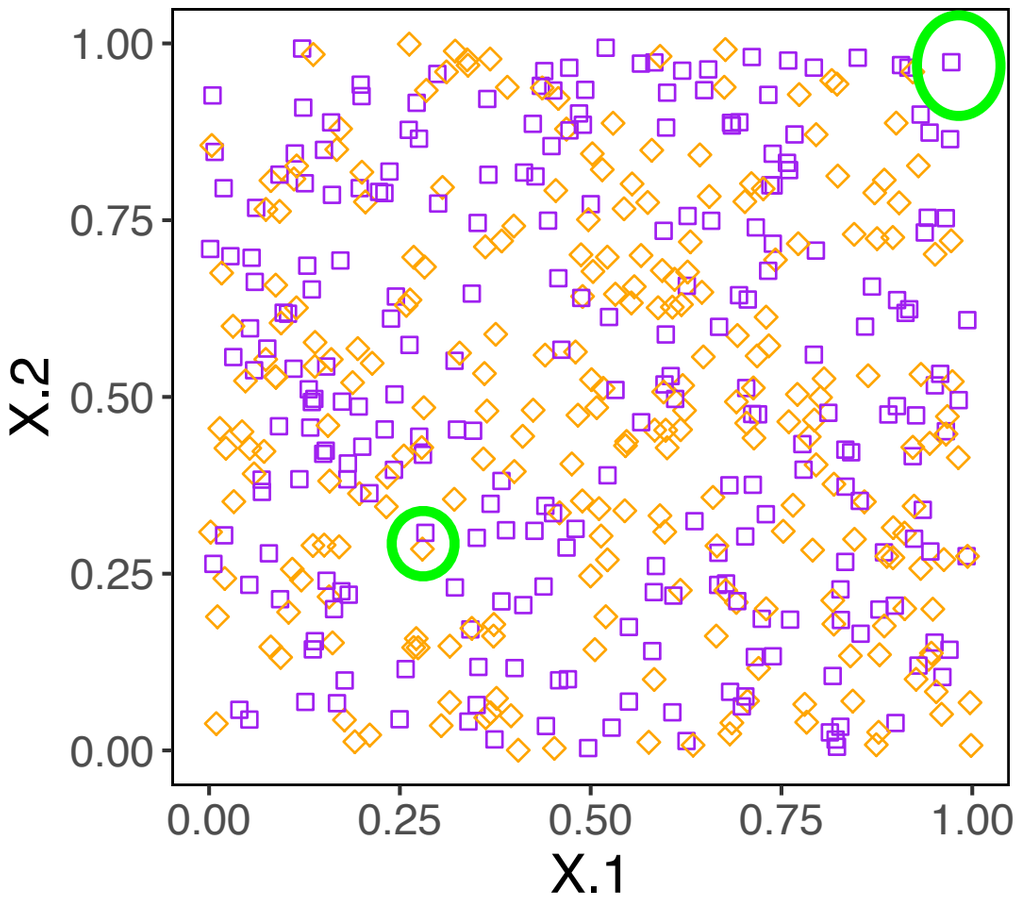
Larry Wasserman

Journal of Causal Inference (2024)

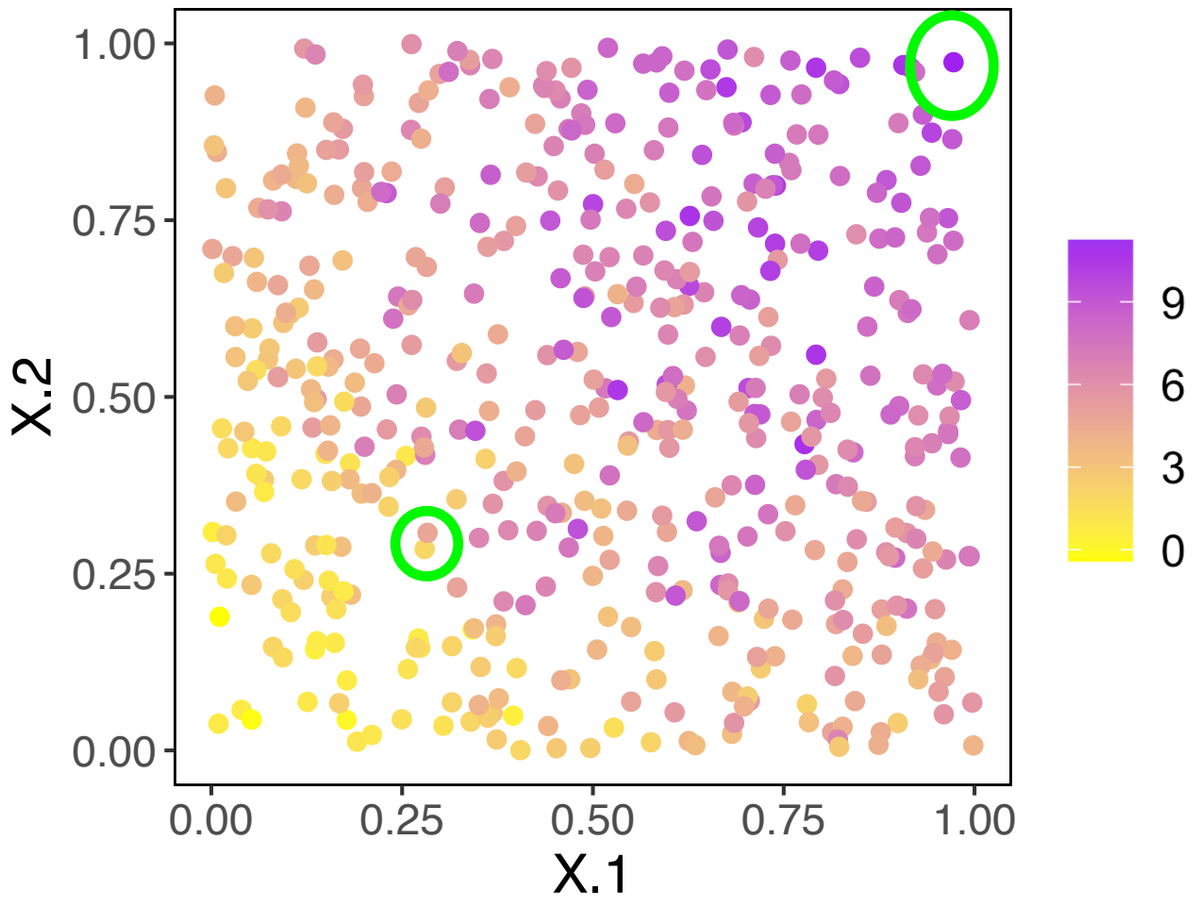
arXiv:2102.10778

Motivating example

- covariates X_i (eg.age, body weight, gender…)
- i.i.d. $\mathrm{Ber}(1/2)$ treatment assignment $A_i \in \{0,1\}$
- potential control and treated outcome (Y_i^C, Y_i^T)
- observe $Y_i = A_i Y_i^T + (1 - A_i) Y_i^C$ (consistency)



Treatment assignments A_i

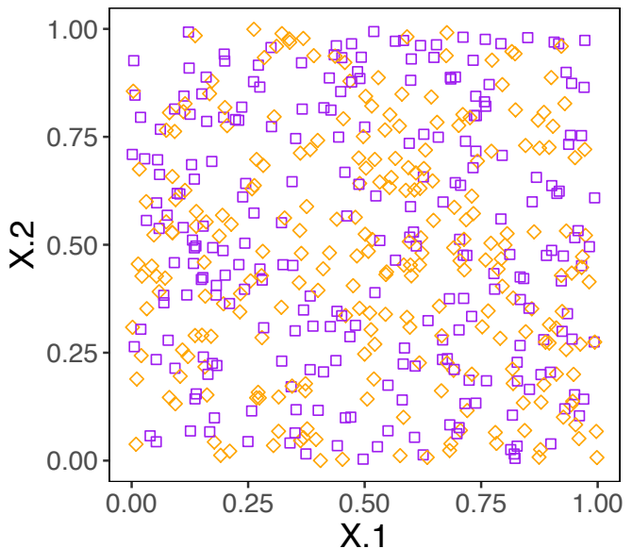


Observed outcomes Y_i

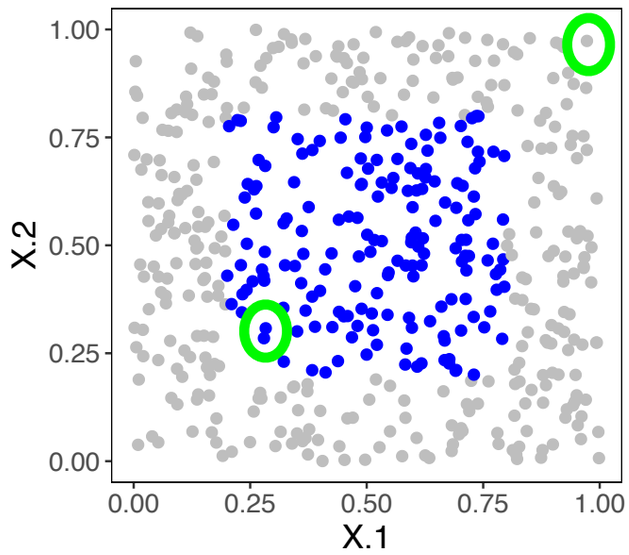
Q: which subjects have positive treatment effects $Y_i^T > Y_i^C$?
(individual level inference)

Motivating example

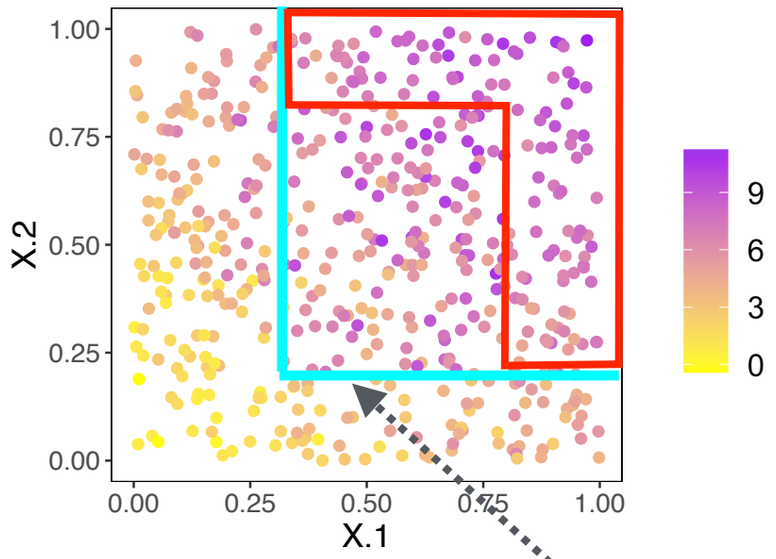
- covariates X_i (eg.age, body weight, gender…)
- i.i.d. $\mathrm{Ber}(1/2)$ treatment assignment $A_i \in \{0,1\}$
- potential control and treated outcome (Y_i^C, Y_i^T)
- observe $Y_i = A_i Y_i^T + (1 - A_i) Y_i^C$ (consistency)



Treatment assignments A_i.



True positive effects $Y_i^T > Y_i^C$.



Observed outcomes Y_i.

Goal: identify subjects with positive effect $Y_i^T > Y_i^C$, denoted as set R, with error control on the expected proportion of false identifications:

$$\mathbb{E}\left[\frac{|\{i : Y_i^T \le Y_i^C\} \cap R|}{|R| \vee 1}\right].$$

We choose to formalize the problem using the language of hypothesis testing.

Problem setup

Define a hypothesis testing problem for each subject $i \in \{1,\ldots,n\}$:

Definition 1 treating potential outcomes as random variables

$$H_{i0}^{\star}: F(Y_i^T \mid X_i) = F(Y_i^C \mid X_i) \qquad H_{i1}: F(Y_i^T \mid X_i) \succ F(Y_i^C \mid X_i)$$

zero effect
(equal in distribution)

positive effect
(stochastic dominance)

Alternative Definitions eg. treating potential outcomes and covariates as fixed values

We deal with randomized experiments without interference, assuming:

(A1) assignments are independent coin flips:

$$\mathbb{P}[(A_1, \ldots, A_n) = (a_1, \ldots, a_n) \mid X_1, \ldots, X_n] = \prod_{i=1}^{n} \mathbb{P}(A_i = a_i) = (1/2)^n;$$

(A2) conditional on the covariates, the outcome of one subject Y_{i_1} is independent of the assignment of another A_{i_2} for any $i_1 \neq i_2$.

*can be extended to error control for *nonpositive effects*:

$$\widetilde{H}_{i0}: F(Y_i^T \mid X_i) \preceq F(Y_i^C \mid X_i).$$

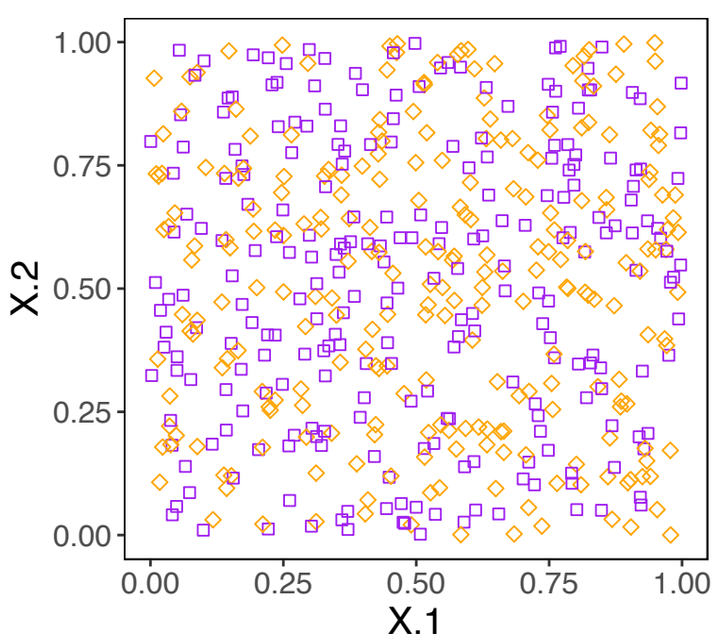
Problem setup

Let the set of nulls (subjects with zero effect) be $\mathscr{H}_0 = \{i : H_{i0} \text{ is true}\}$,
the set of rejected (identified) subjects be R.

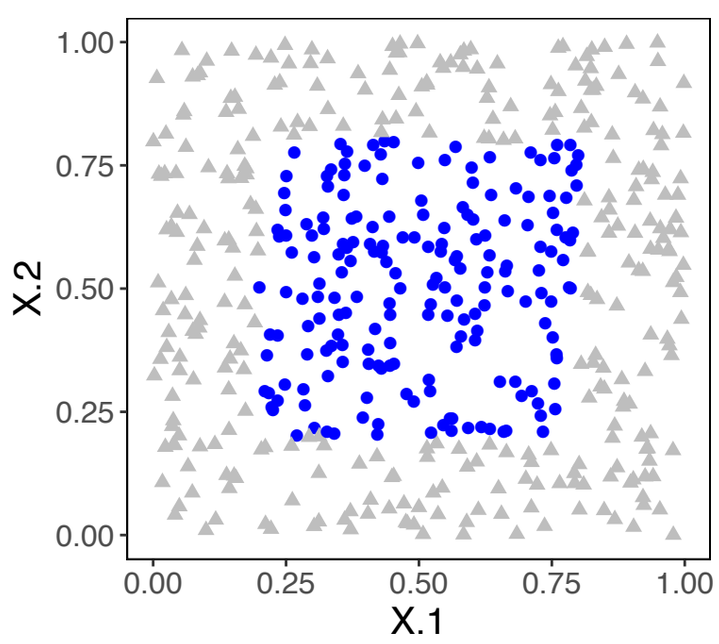
False discovery rate: expected proportion of false identifications

$$\mathrm{FDR} := \mathbb{E}\left(\frac{|\mathscr{H}_0 \cap R|}{|R| \vee 1}\right).$$

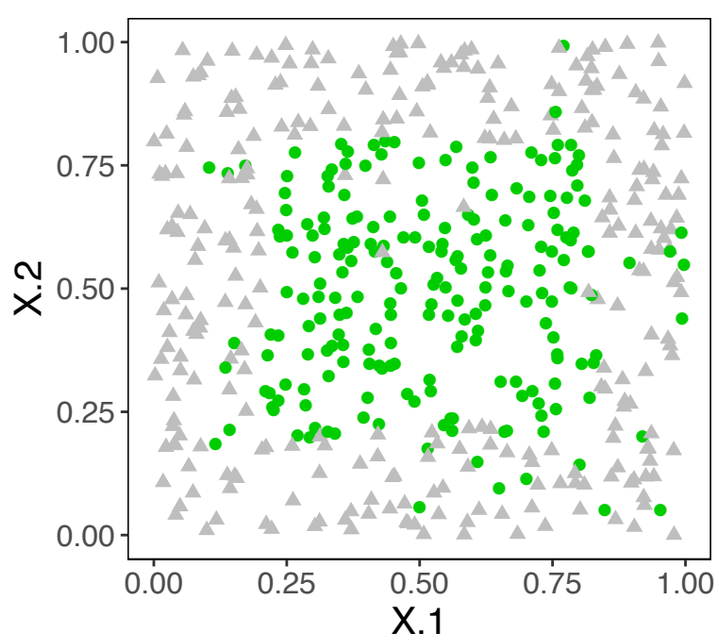
Output of proposed I^3 (for interactive identification of individual effects):
a set of identified subjects R with $\mathrm{FDR} \leq \alpha$ and reasonably high power.



Subjects (with two covariates) randomly assigned to treated and control group.



True positive effects (blue) are unknown ground truth.



Identified subjects (green) contain most true positives, regardless of treated or not.

Algorithm description of I^3

start with $R_0 = \{1,\ldots,n\}$ and $t = 0$

Prior knowledge $\{Y_i, X_i\}_{i=1}^n$ | $\{A_i\}_{i=1}^n$

Explorer	Oracle
1. Let $\widehat{Y}$ be any estimator of $\mathbb{E}(Y \mid X)$, and compute residuals $E_i = Y_i - \widehat{Y}(X_i)$	
	Inform $(E_i)_{i=1}^n$ →
	2. Compute a treatment effect estimator: $\widehat{\Delta}_i := 4(A_i - 1/2)E_i$
	3. Divide R_t into $R_t^+ := \{i \in R_t : \widehat{\Delta}_i > 0\}$ and $R_t^- := \{i \in R_t : \widehat{\Delta}_i \leq 0\}$ (to construct FDR estimator)
4.	
5.	
6.	
7.	
8.	

Repeat step 4-8 for $t = 1,\ldots,n$

Estimated effect $\widehat{\Delta}_i$

Define

$$\widehat{\Delta}_i := 4(A_i - 1/2)(Y_i - \widehat{Y}(X_i))$$

where $\widehat{Y}(\cdot)$ is an arbitrary estimator of $\mathbb{E}(Y_i \mid X_i)$ without using $\{A_i\}_{i=1}^n$.

Reasonable

- used in recent papers such as Nie and Wager (2020), Kennedy (2020); and can be traced back to Robinson (1988).
- recover true effect in simple cases. Eg. Suppose $Y_i^C = c$ and $Y_i^T = c + \delta$ for all i.
  If $\widehat{Y}$ is correctly learned: $\widehat{Y}(X_i) = c + \delta/2$, then $\widehat{\Delta}_i = \delta$.

Error control by the sign property:

$$\mathbb{P}\left(\widehat{\Delta}_i > 0 \mid \{Y_j, X_j\}_{i=1}^n\right) \leq 1/2,$$

if subject i has zero effect (null hypothesis is true), under assumption (A1) and (A2) of randomized experiments;
since $F(Y_i^T \mid X_i) \overset{d}{=} F(Y_i^C \mid X_i) \Rightarrow A_i \perp\!\!\!\perp Y_i \mid X_i \Rightarrow \mathbb{P}(A_i - 1/2 > 0 \mid Y_i, X_i) = 1/2.$

Error control **Selection**

Algorithm description of I^3

start with $R_0 = \{1,\ldots,n\}$ and $t = 0$

Prior knowledge $\{Y_i, X_i\}_{i=1}^n$ | $\{A_i\}_{i=1}^n$

Explorer | Oracle

1. Learn $\hat{Y}$ and compute $E_i := Y_i - \hat{Y}(X_i)$

2. Set $\widehat{\Delta}_i := 4(A_i - 1/2)E_i$
3. Divide R_t into R_t^+ and R_t^-

Reveal $|R_t^+|$

4. Check $\widehat{\mathrm{FDR}}(R_t) := \dfrac{|R_t^-| + 1}{|R_t^+| \vee 1} \le \alpha$

Yes → Stop at $\tau = t$, Report R_τ^+

If no ↓

5. Pick $i_t^* \in R_t$ (hopefully a null and $\widehat{\Delta}_{i_t^*} \le 0$) using explorer's current knowledge

6. Reveal $A_{i_t^*}$

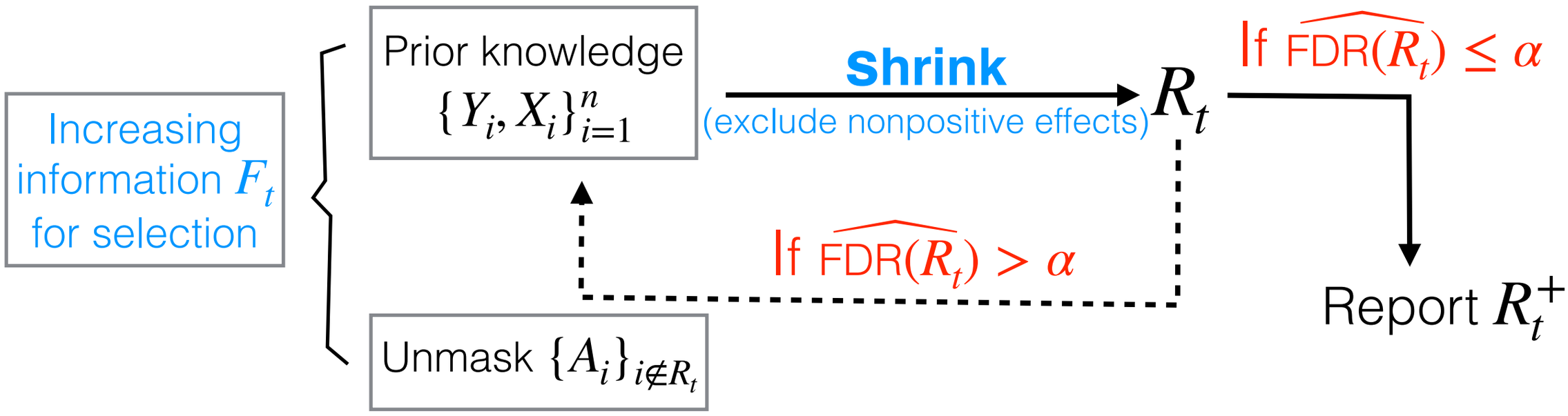
7. Calculate $\widehat{\Delta}_{i_t^*}$ (and its sign).
8. Update $R_{t+1} = R_t \backslash \{i_t^*\}$ also $|R_{t+1}^+|, |R_{t+1}^-|$

Repeat step 4-8

FDR control using binary variables:
Barber and Candes (2015)
Lei, Fithian (2018)
Lei, Ramdas, Fithian (2021)

Summary of I^3 (in the perspective of the explorer)

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



Theorem:

I^3controls FDR under the standard causal assumptions (A1) and (A2) for randomized experiments.

Issue in the selection step

Goal: exclude subjects that are likely to have negative $\widehat{\Delta}_i$
given revealed information — but no A_i at the first iteration!

Illustrative example

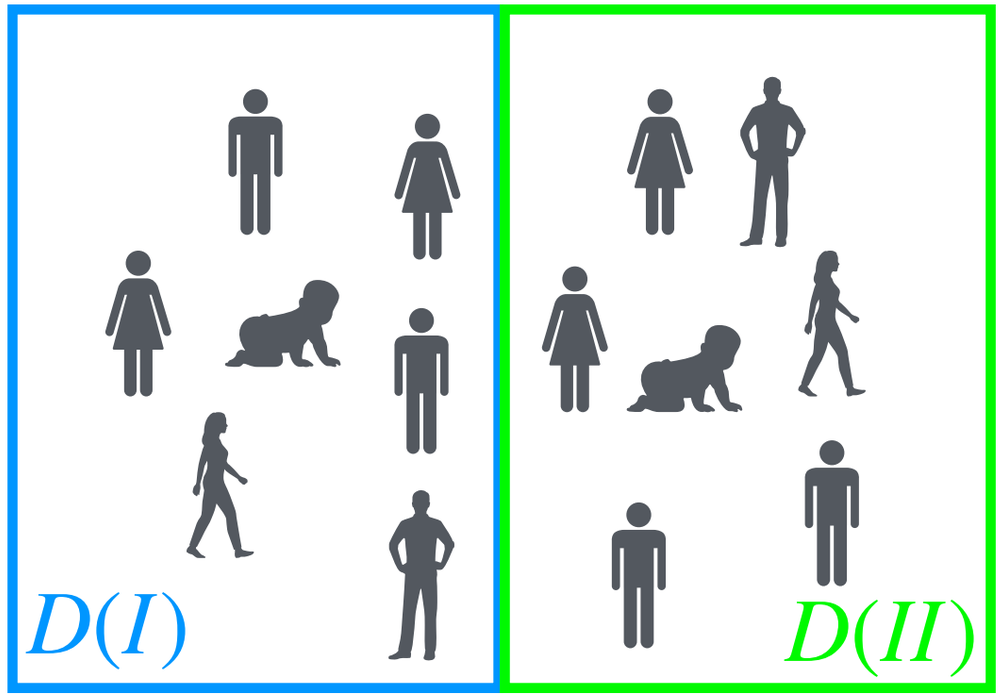
X_i	Y_i	A_i	**sign**
F	5	1	+
F	0	0	+
F	10	1	+
F	0	0	+
M	0	1	-
M	5	0	-
M	0	1	-
M	10	0	-

When all assignments are masked, we cannot tell whether we have guessed the correct signs or the opposite signs for all subjects, since both are equally plausible.
→ exclude true positive effect
→ low identification power

Fix: wrap around $\mathbf{I}^3$ by cross-fitting framework

Schematic of Crossfit-I^3

2. Apply I^3 on $D(I)$ where all information in $D(II)$ is revealed to the analyst, get a rejection set $R(I) \subseteq D(I)$ at level $\alpha/2$.
3. Apply I^3 on $D(II)$ where all information in $D(I)$ is revealed to the analyst, get a rejection set $R(II) \subseteq D(II)$ at level $\alpha/2$.
4. Final rejection set is union $R(I) \cup R(II)$.



1. Random split

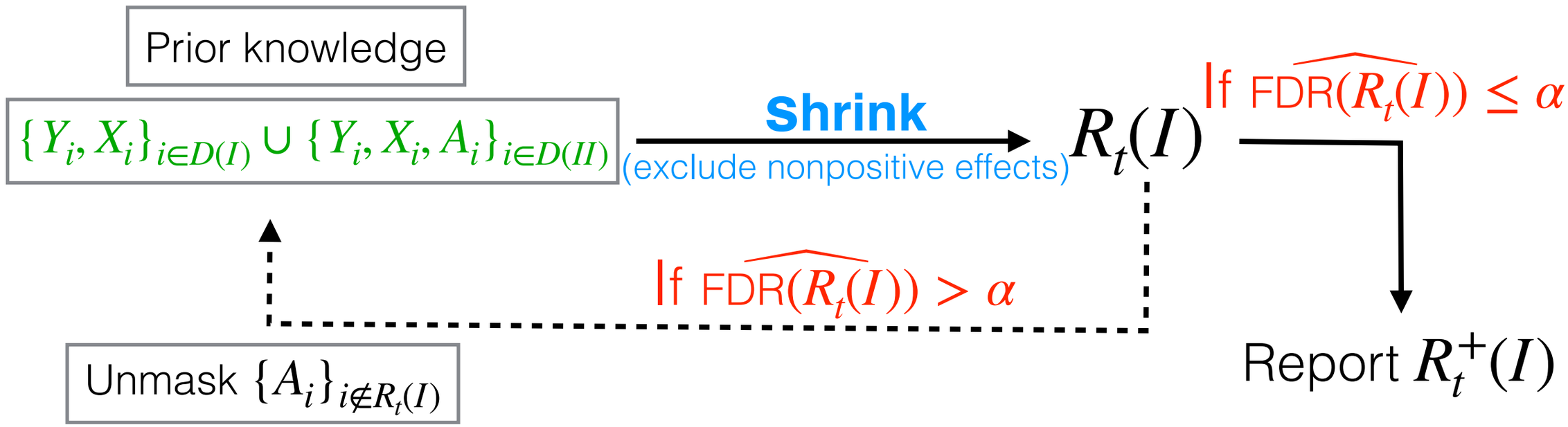
Cont. eg.

X_i	Y_i	A_i	sign	A_i'	sign
F	5	1	+	1	+
F	0	0	+	0	+
F	10	1	+	0	-
F	0	0	+	1	-
M	0	1	-	1	-
M	5	0	-	0	-
M	0	1	-	0	+
M	10	0	-	1	+

Subjects with same covariates should be more likely to have effect in the same direction.

Based on the effect signs of revealed samples, we would prefer the correct signs over the opposite signs for all subjects.

$R_0(I) = D(I)$ and $R_0 \supseteq R_1 \supseteq \ldots$



An example of automated algorithm to shrink $R_t(I)$

1. Using non-candidate subjects $j \notin R_t(I)$ with complete data, train a random forest classifier where the label is $\mathbf{1}\left(\widehat{\Delta}_j > 0\right) \equiv \mathbf{1}\left((A_j - 1/2)(Y_j - \hat{Y}(X_j)) > 0\right)$ and the predictors are Y_j, X_j and $Y_j - \hat{Y}(X_j)$.
2. For candidate subjects $i \in R_t(I)$ without A_i hence without label, predict the probability of $\widehat{\Delta}_i$ being positive, denoted as $\hat{p}(i,t)$.
3. Find $i_t^* = \text{argmin}\{\hat{p}(i,t) : i \in R_t(I)\}$ and $R_{t+1}(I) = R_t(I) \backslash \{i_t^*\}$.

Note: the analyst can choose to **update** or **change** to some parametric modeling (could lead to higher power if correct) at any step.

Numerical experiments

Generating model:

$X_i = (X_i(1), X_i(2), X_i(3)) \in \{0,1\}^2 \times \mathbb{R}$ and $\epsilon_i \sim N(0,1)$

$Y_i^C = 5(X_i(1) + X_i(2) + X_i(3)) + \epsilon_i$

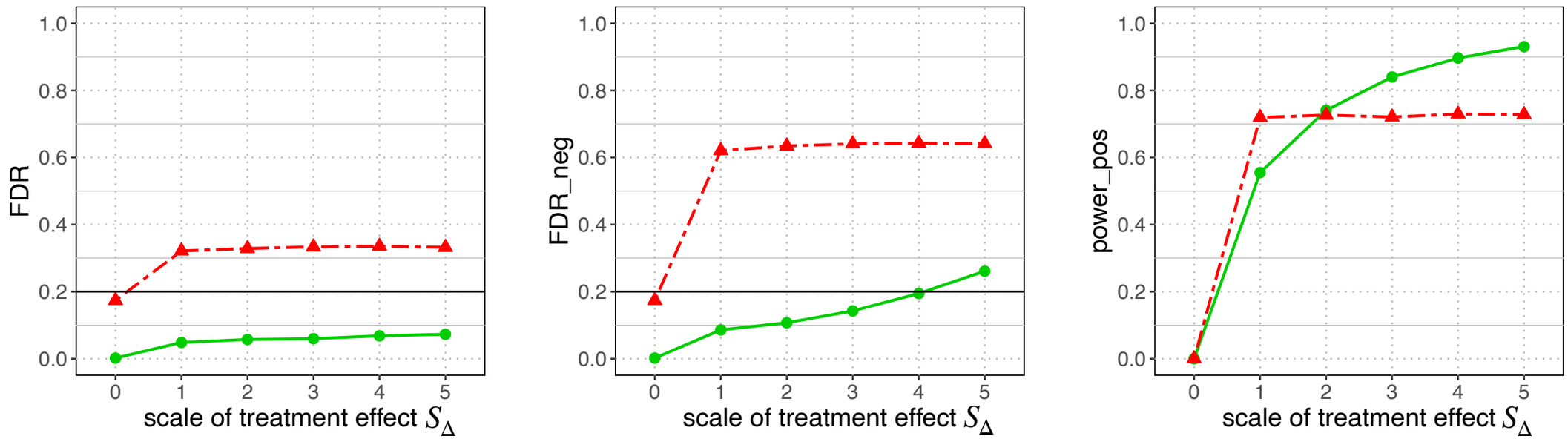
$Y_i^T = \Delta(X_i) + 5(X_i(1) + X_i(2) + X_i(3)) + \epsilon_i$

Positive-biased effect:

$\Delta(X_i) = S_\Delta \cdot [5X_i^3(3)\mathbf{1}\{X_i(3) > 1\} - X_i(1)/2]$

- 15 % positive effects with size $20S_\Delta$ and
- 45 % negative effects with size $0.5S_\Delta$.

Crossfit-I^3 Linear-BH (asymptotic FDR control under linear assumption)



$$\mathbb{E}\left[\frac{|\{i : \Delta(X_i) = 0\} \cap R_t^+|}{|R_t^+| \vee 1}\right] \qquad \mathbb{E}\left[\frac{|\{i : \Delta(X_i) \leq 0\} \cap R_t^+|}{|R_t^+| \vee 1}\right] \qquad \mathbb{E}\left[\frac{|\{i : \Delta(X_i) > 0\} \cap R_t^+|}{|\{i : \Delta(X_i) > 0\}|}\right]$$

- Crossfit-$\mathbf{I}^3$ guarantees FDR control; the parametric method (Linear-BH) does not.
- Less than $25\,\%$ identified by Crossfit-$\mathbf{I}^3$ are nonpositive; more than $\mathbf{50}\,\%$ by Linear-BH.
- Crossfit-$\mathbf{I}^3$ has good power to identify true positive effect.

Doubly-robust FDR for observational studies

$$0 < \pi_{\min} \leq \pi_i \equiv \mathbb{P}(A_i = 1 \mid \{X_i\}_{i=1}^n) \leq \pi_{\max} < 1 \text{ for all } i \in \{1,\ldots,n\}$$

Recall original FDR estimator: $\widehat{\mathrm{FDR}}(R_t) \equiv \dfrac{|R_t^-| + 1}{|R_t^+| \vee 1}.$

When bounds are known, $\mathbf{I}^3$ with FDR estimator:

$$\widehat{\mathrm{FDR}}(R_t) := \left(\frac{1}{\min\{\pi_{\min}, 1 - \pi_{\max}\}} - 1 \right) \frac{|R_t^-| + 1}{|R_t^+| \vee 1}.$$

When the interactive algorithm stops, $\mathrm{FDR}(R_\tau^+) \leq \alpha$

When bounds are unknown, use

$$\widehat{\mathrm{FDR}}(R_t) := \left(\frac{1}{\min\{\hat{\pi}_{\min}, 1 - \hat{\pi}_{\max}\}} - 1 \right) \frac{|R_t^-| + 1}{|R_t^+| \vee 1},$$

where $\hat{\pi}_{\min}$ and $\hat{\pi}_{\max}$ are estimated from the data revealed to the explorer.

Asymptotic FDR control when either $\pi_{\min}, \pi_{\max}$ are well estimated or $\hat{Y}(X_i)$ approximates $\mathbb{E}(Y_i \mid X_i)$ — double robustness.

Recall $\hat{Y}(X_i)$ is used in $\widehat{\Delta}_i := 4(A_i - 1/2)(Y_i - \hat{Y}(X_i))$

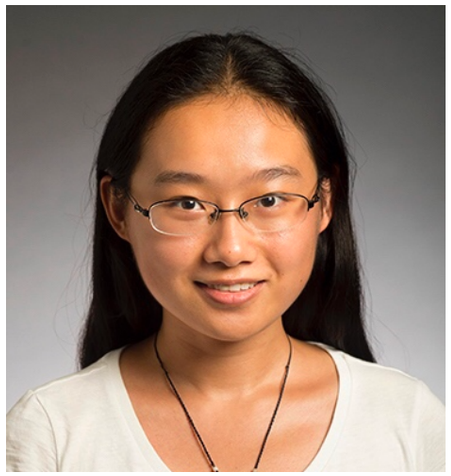
Summary

- Crossfit-$\mathbf{I}^3$ outputs a set of subjects whose effects are positive, with a guaranteed FDR control in randomized experiments, and a doubly-robust FDR control in observational settings.
- Crossfit-$\mathbf{I}^3$ allows an analyst to incorporating various types of covariates and prior knowledge in a flexible manner.
- It can identify subjects that even if they are not treated.

Extensions in the paper

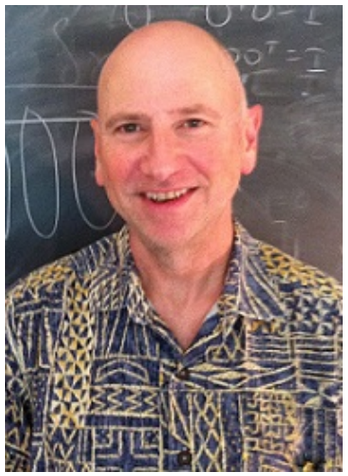
- Relax the null to control false identification of **nonpositive effects**.
- Extend the experiment setup to **paired samples** and **subgroup identification**.

Interactive identification of individuals with positive treatment effect while controlling false discoveries (**FDR**)



Boyan Duan

Aaditya Ramdas



Larry Wasserman

Journal of Causal Inference (2024)

arXiv:2102.10778