Assumption-lean weak limits and tests for two-stage adaptive experiments

International Seminar on Selective Inference
June 3, 2025

Ziang Niu
University of Pennsylvania
https://ziangniu6.github.io

Co-author

Zhimei Ren

RCT versus adaptive experiment

RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]

1: Ambrosius et al. (2014)

RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]

Control : 

Treatment: 

1: Ambrosius et al. (2014)

RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]

Control : 

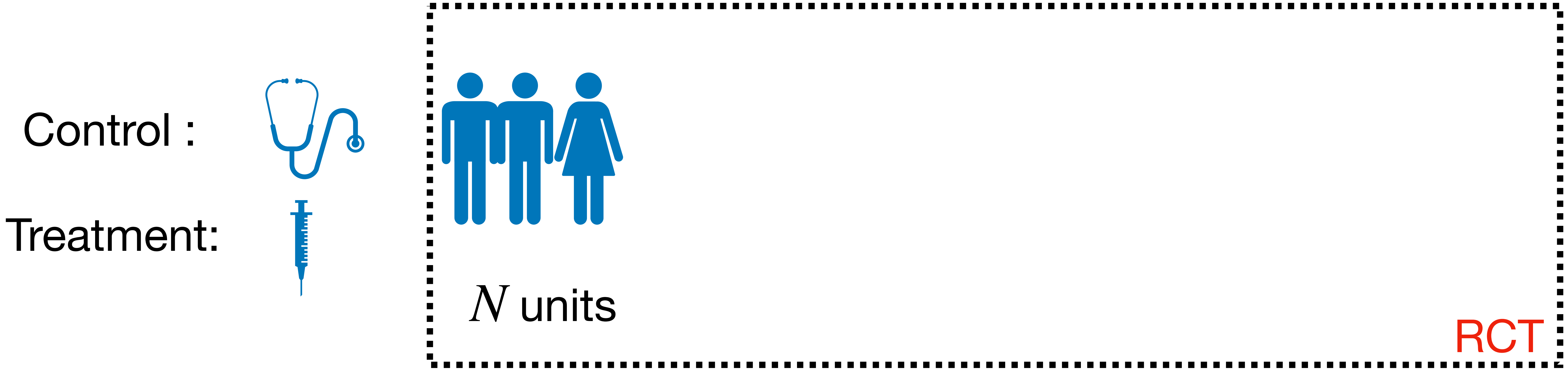
Treatment: 

RCT

1: Ambrosius et al. (2014)

RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]



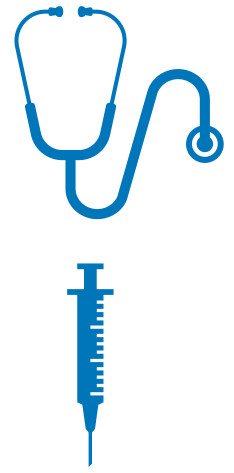
1: Ambrosius et al. (2014)

RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]

Control :

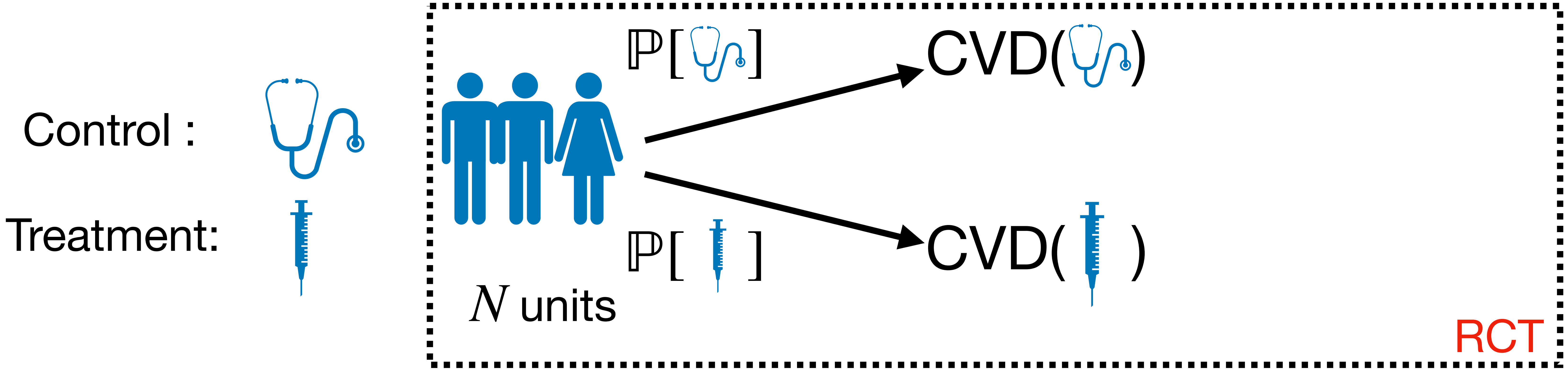
Treatment:



1: Ambrosius et al. (2014)

RCT versus adaptive experiment

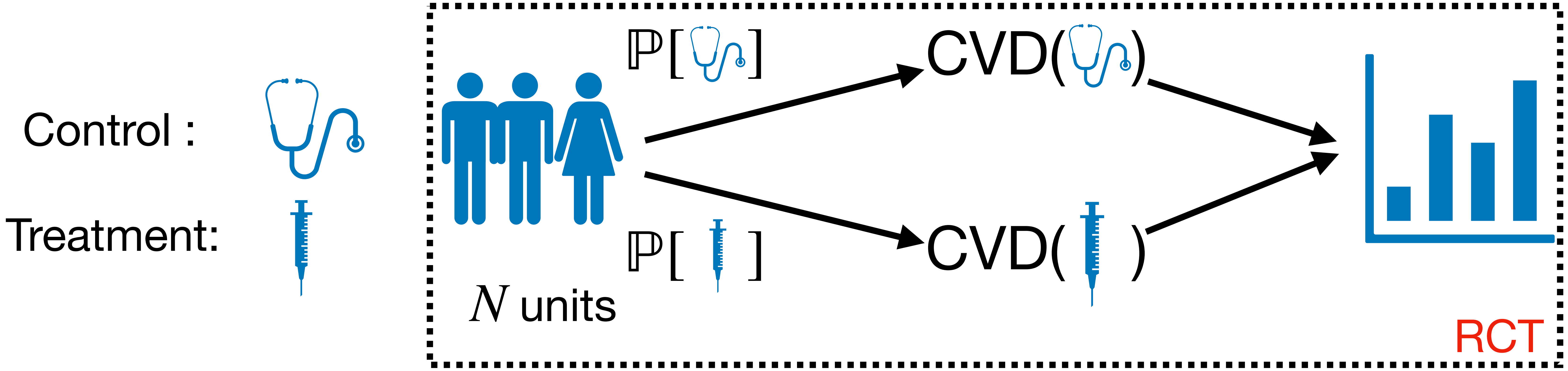
Motivating example: new treatment development for lowering CVD risk.[1]



1: Ambrosius et al. (2014)

RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]



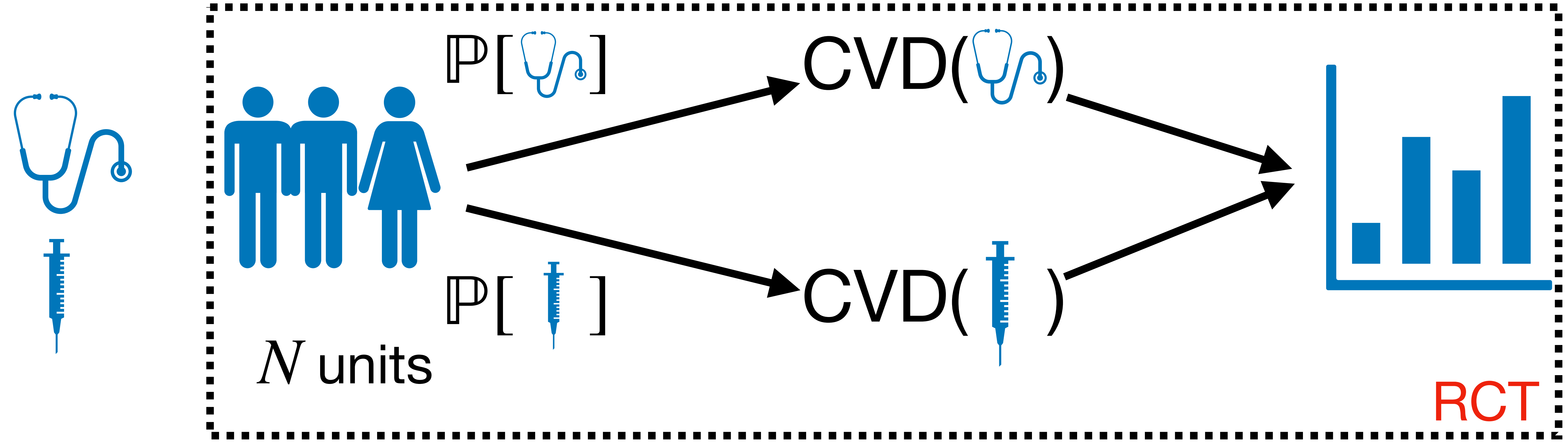
1: Ambrosius et al. (2014)

RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]

Control :

Treatment:



Ethical consideration: reduce the risk of subjects exposing to inferior treatment.

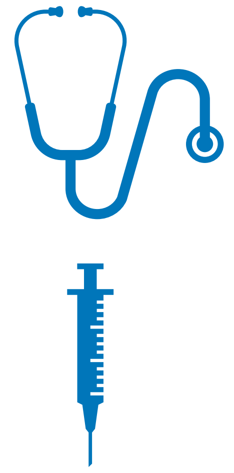
1: Ambrosius et al. (2014)

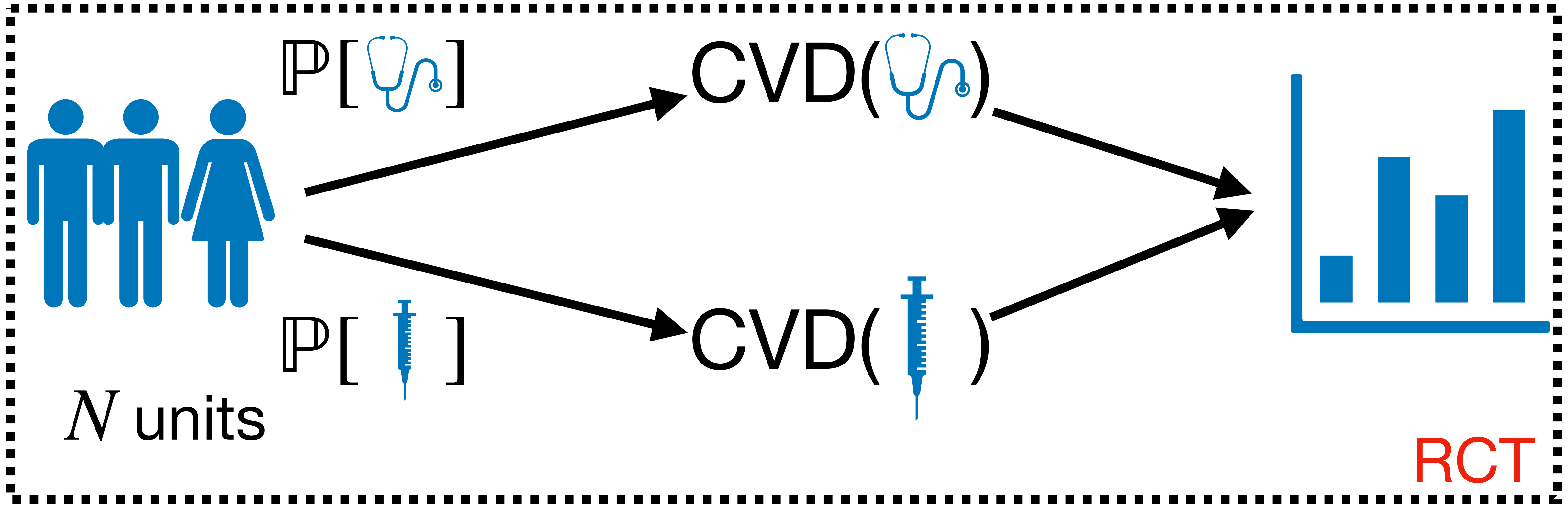
RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]

Control :

Treatment:





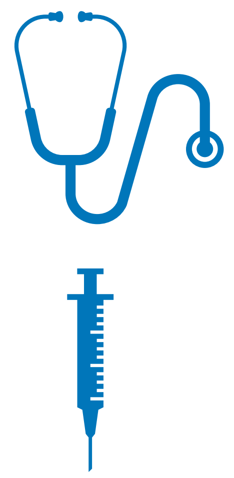
Ethical consideration: reduce the risk of subjects exposing to inferior treatment.

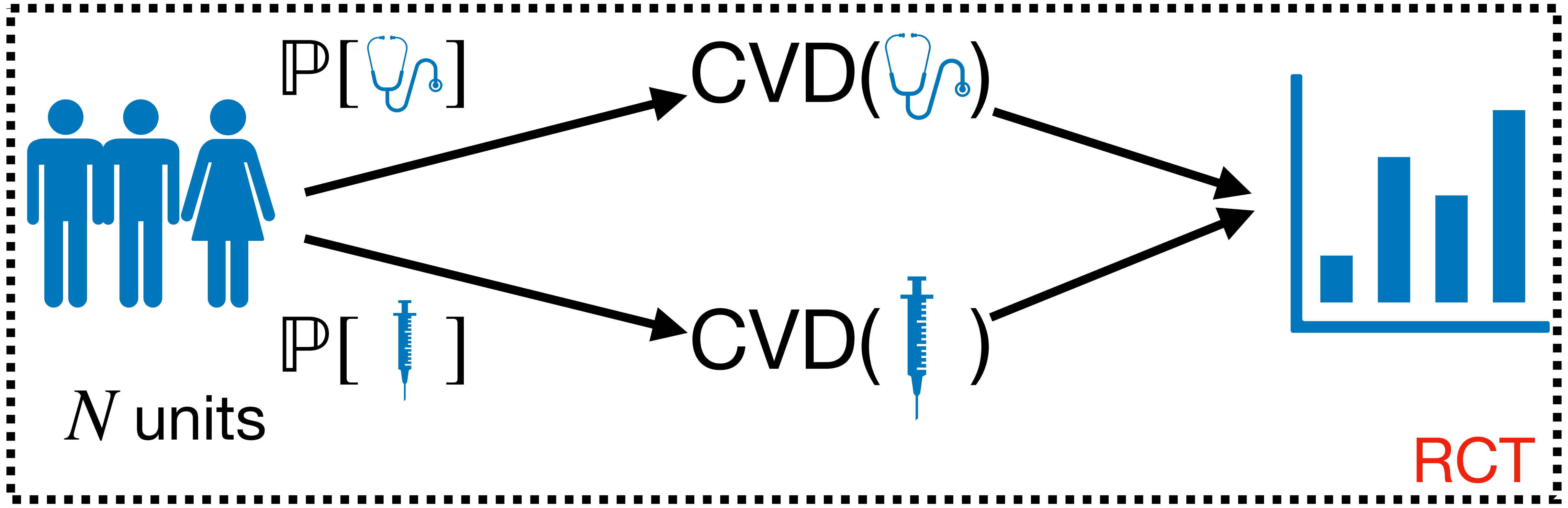
1: Ambrosius et al. (2014)

RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]

Control :
Treatment:



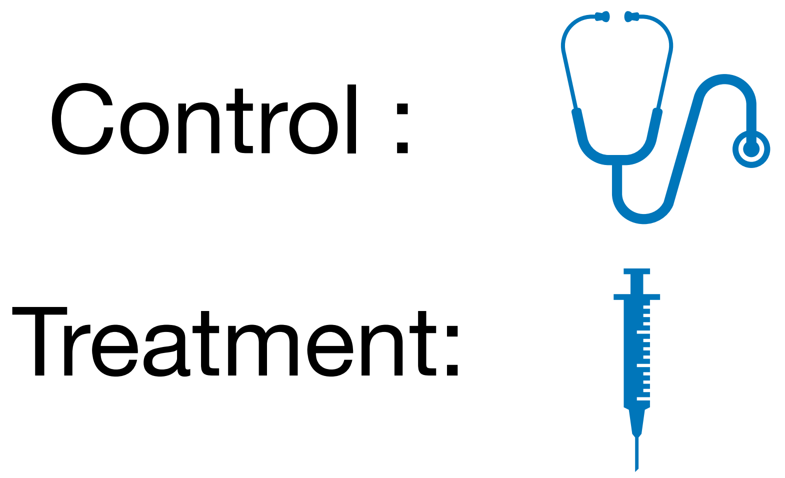


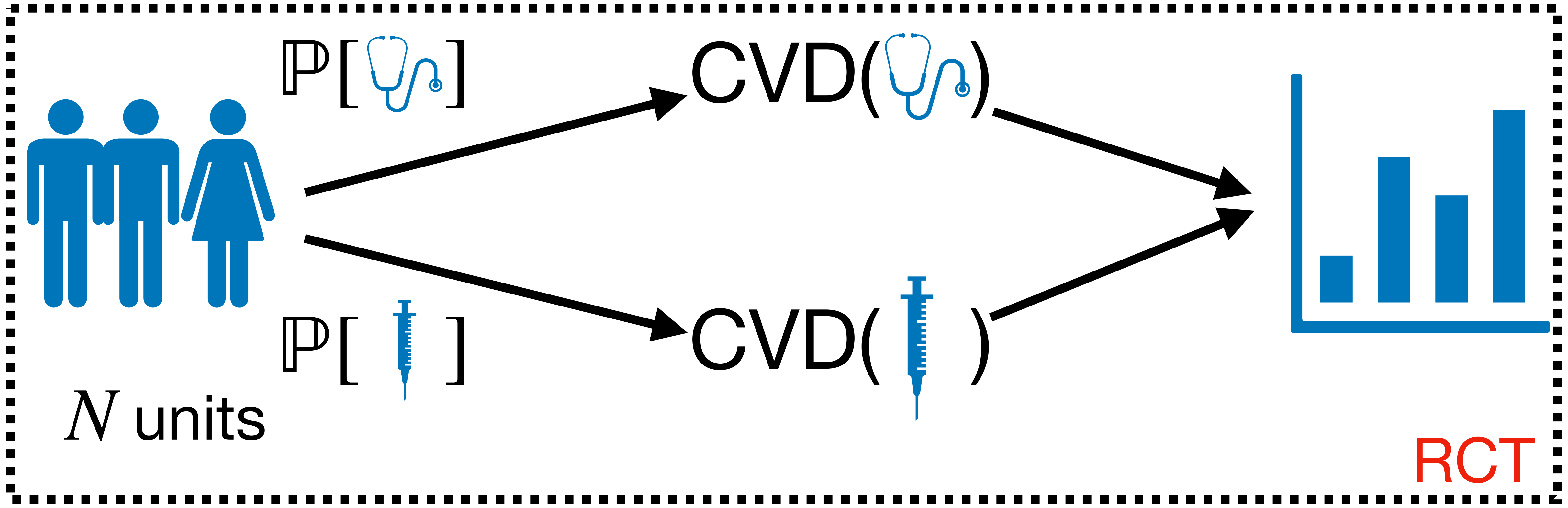
Ethical consideration: reduce the risk of subjects exposing to inferior treatment.

1: Ambrosius et al. (2014)

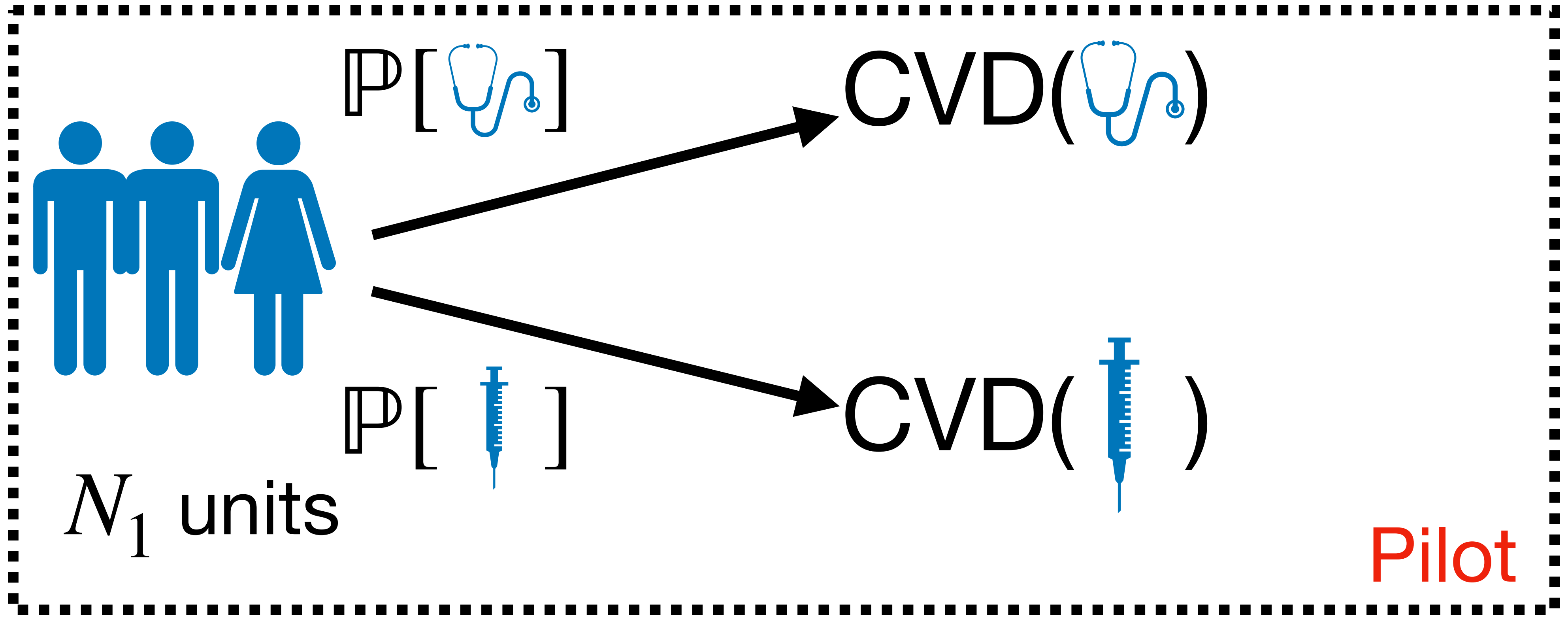
RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]





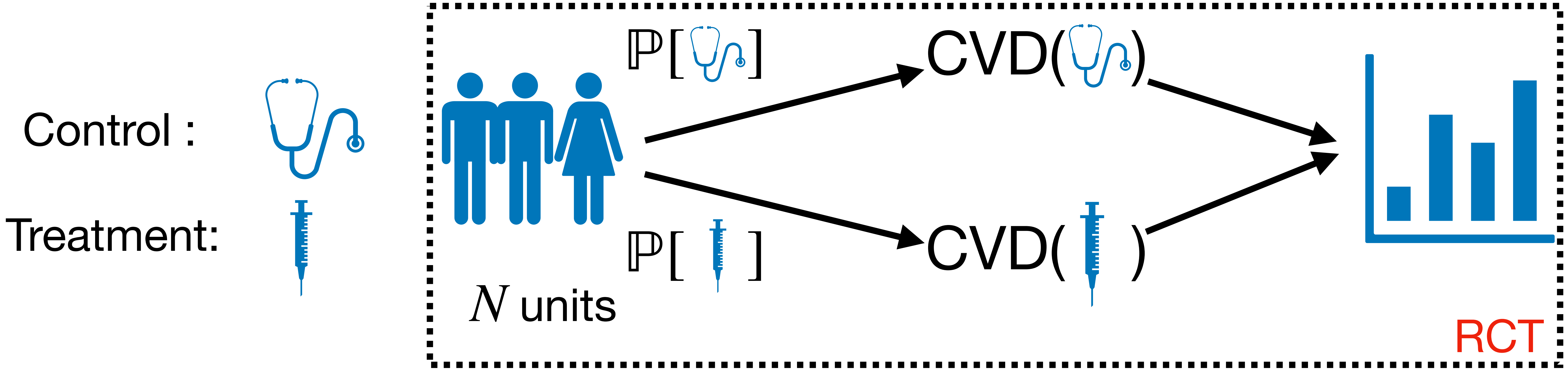
Ethical consideration: reduce the risk of subjects exposing to inferior treatment.



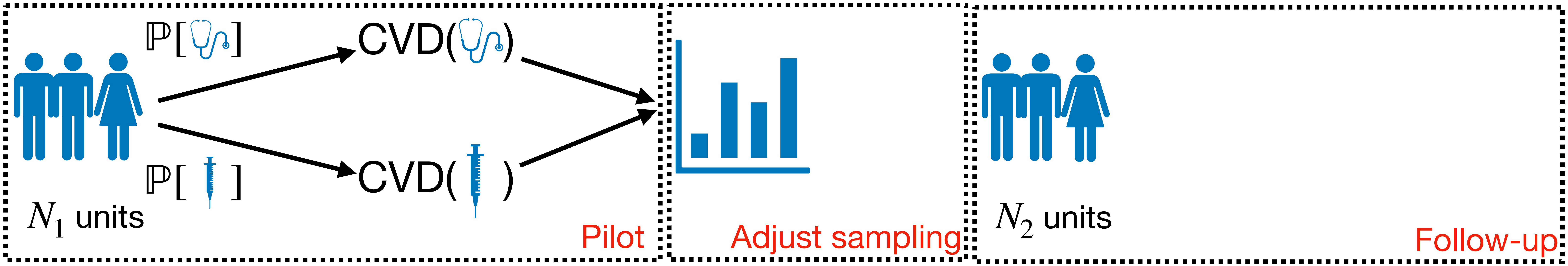
1: Ambrosius et al. (2014)

RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]



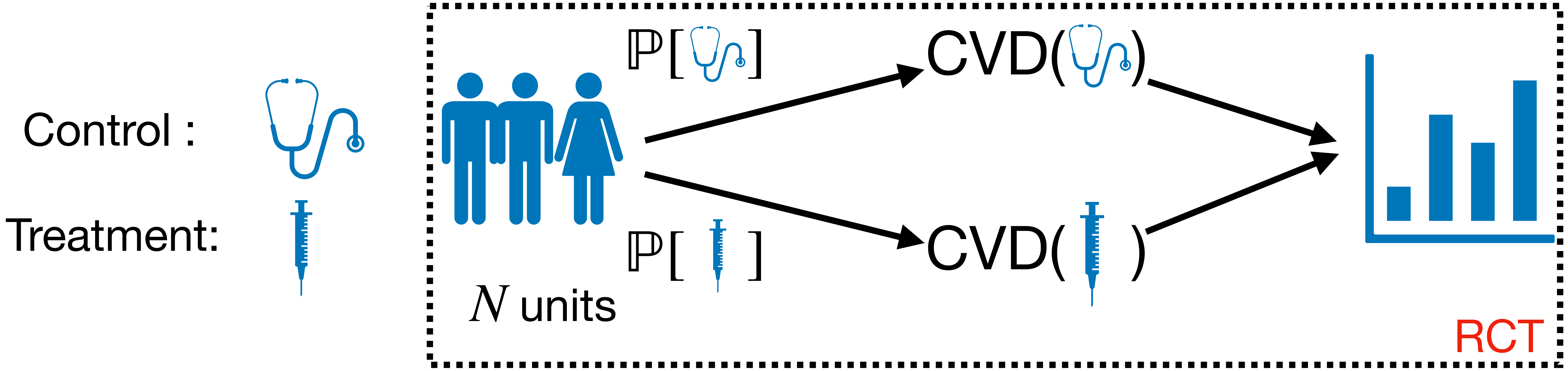
Ethical consideration: reduce the risk of subjects exposing to inferior treatment.



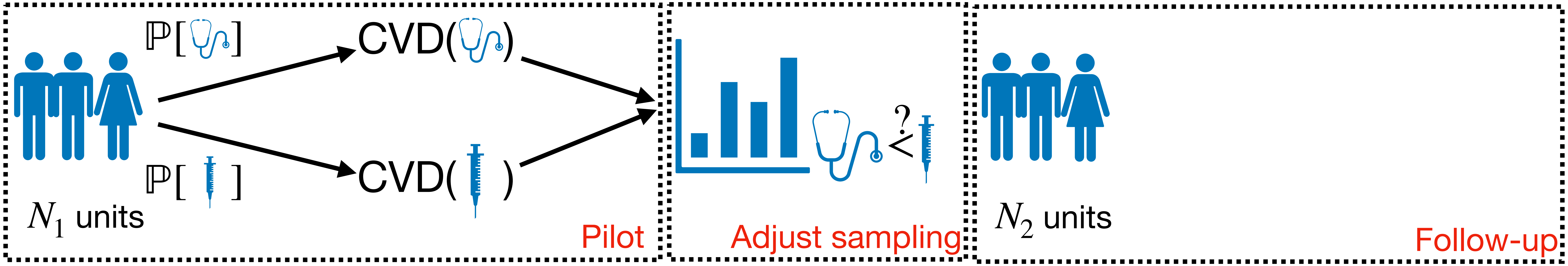
1: Ambrosius et al. (2014)

RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]



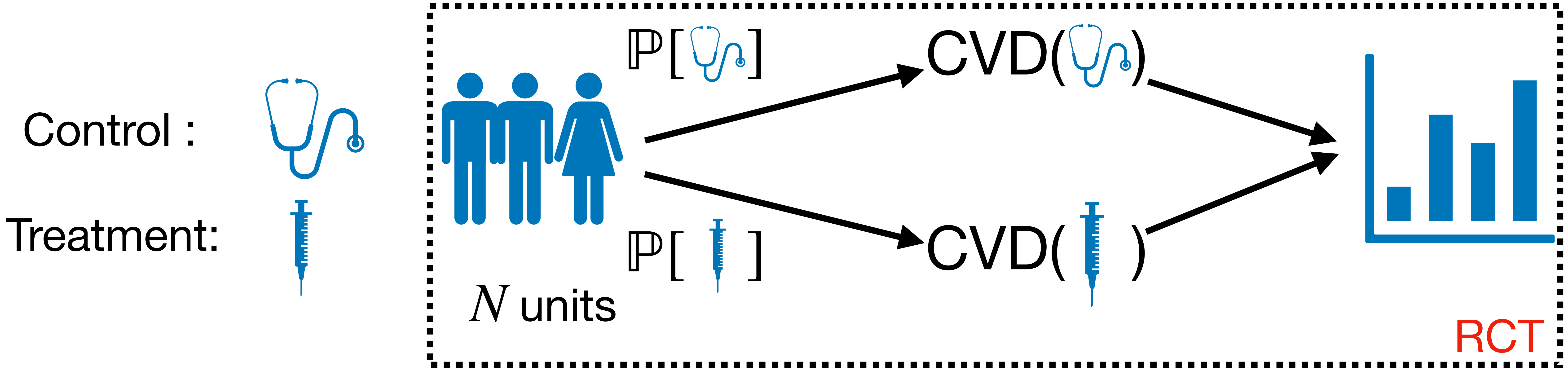
Ethical consideration: reduce the risk of subjects exposing to inferior treatment.



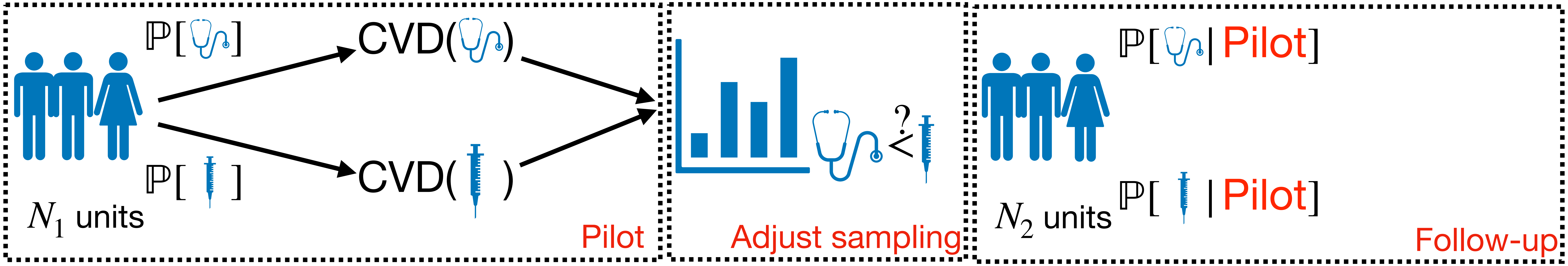
1: Ambrosius et al. (2014)

RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]



Ethical consideration: reduce the risk of subjects exposing to inferior treatment.



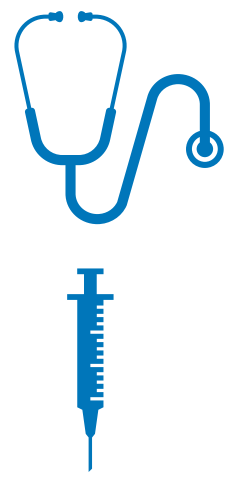
1: Ambrosius et al. (2014)

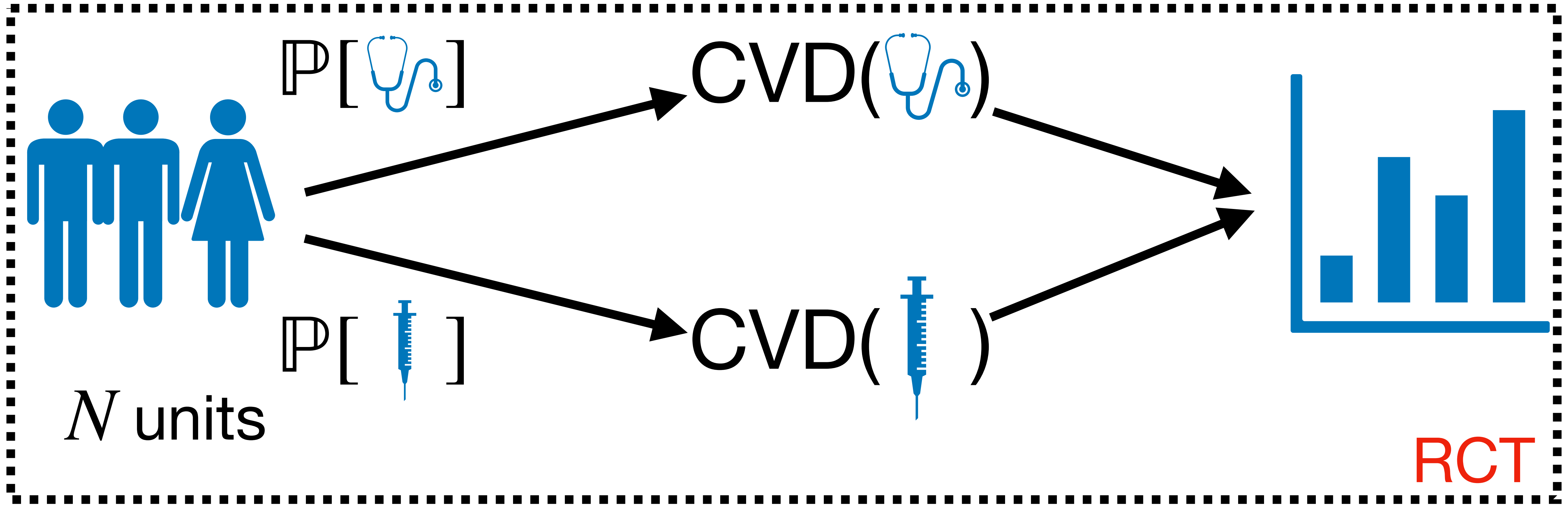
RCT versus adaptive experiment

Motivating example: new treatment development for lowering CVD risk.[1]

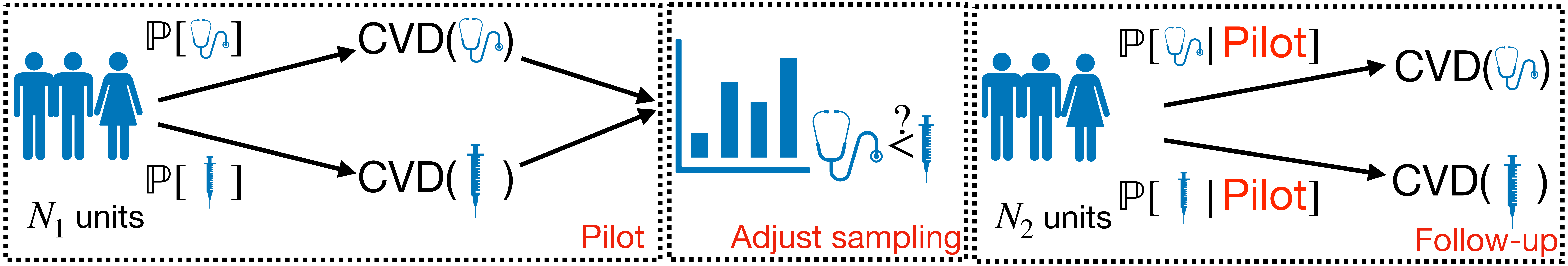
Control :

Treatment:





Ethical consideration: reduce the risk of subjects exposing to inferior treatment.



1: Ambrosius et al. (2014)

Motivating example (cont’d)

Motivating example (cont'd)

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Motivating example (cont'd)

Control : 

Treatment: 

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Adaptive sampling: adjust sampling by DIM $\equiv \widehat{\mathbb{E}}[\text{CVD}($  $)] - \widehat{\mathbb{E}}[\text{CVD}($  $)]$.

Control : 

Treatment: 

Motivating example (cont'd)

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Adaptive sampling: adjust sampling by DIM $\equiv \widehat{\mathbb{E}}[\text{CVD}($  $)] - \widehat{\mathbb{E}}[\text{CVD}($  $)]$.

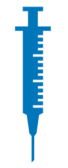
Adaptive algorithms: $\mathbb{P}[$  $|\,\text{Pilot}] = f(\text{DIM})$.

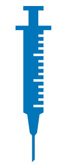
Motivating example (cont'd)

Control : 

Treatment: 

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Adaptive sampling: adjust sampling by DIM $\equiv \widehat{\mathbb{E}}$ [CVD()] $-$ $\widehat{\mathbb{E}}$ [CVD()] .

Adaptive algorithms: $\mathbb{P}[$  $|$ Pilot$] = f(\mathrm{DIM})$.

- Greedy algorithm:

Motivating example (cont'd)

Control : 

Treatment: 

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Adaptive sampling: adjust sampling by DIM $\equiv \widehat{\mathbb{E}}$[CVD( )] $-$ $\widehat{\mathbb{E}}$[CVD( )] .

Adaptive algorithms: $\mathbb{P}[$ | Pilot$] = f(\mathrm{DIM})$.

- Greedy algorithm:

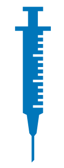
$$\mathbb{P}[\quad | \text{Pilot}] = \begin{cases} 0.9 \text{ if DIM} > 0 \\ 0.1 \text{ if DIM} \leq 0 \end{cases}.$$

Motivating example (cont'd)

Control : 

Treatment: 

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Adaptive sampling: adjust sampling by DIM $\equiv \widehat{\mathbb{E}}$[CVD()] $-$ $\widehat{\mathbb{E}}$[CVD()] .

Adaptive algorithms: $\mathbb{P}[$  $\mid$ Pilot$] = f(\mathrm{DIM})$.

- Greedy algorithm:

$\mathbb{P}[$  $\mid$ Pilot$] = \begin{cases} 0.9 \text{ if DIM} > 0 \\ 0.1 \text{ if DIM} \leq 0 \end{cases}.$

- (Modified) Thompson sampling:

Control : 🩺

Treatment: 💉

Motivating example (cont'd)

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Adaptive sampling: adjust sampling by $\text{DIM} \equiv \widehat{\mathbb{E}}[\text{CVD}(\text{💉})] - \widehat{\mathbb{E}}[\text{CVD}(\text{🩺})]$.

Adaptive algorithms: $\mathbb{P}[\text{💉} \mid \text{Pilot}] = f(\text{DIM})$.

- Greedy algorithm:

$$\mathbb{P}[\text{💉} \mid \text{Pilot}] = \begin{cases} 0.9 \text{ if DIM} > 0 \\ 0.1 \text{ if DIM} \leq 0 \end{cases}.$$

- (Modified) Thompson sampling:

$$\mathbb{P}[\text{💉} \mid \text{Pilot}] = \Phi(\text{DIM}).$$

Control : 

Treatment: 

Motivating example (cont'd)

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Adaptive sampling: adjust sampling by DIM $\equiv \widehat{\mathbb{E}}$[CVD()] – $\widehat{\mathbb{E}}$[CVD()] .

Adaptive algorithms: $\mathbb{P}$[ | Pilot] = f(DIM) .

- Greedy algorithm:

$\mathbb{P}$[ | Pilot] $= \begin{cases} 0.9 \text{ if DIM} > 0 \\ 0.1 \text{ if DIM} \leq 0 \end{cases}.$

- (Modified) Thompson sampling:

$\mathbb{P}$[ | Pilot] = Φ(DIM) .

- Interim p-value: define $T = \text{DIM}/\hat{\sigma}$ with $p_L = \Phi(T)$ and $p_R = 1 - p_L$.

Motivating example (cont'd)

Control : 

Treatment: 

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Adaptive sampling: adjust sampling by $\text{DIM} \equiv \widehat{\mathbb{E}}[\text{CVD}(\;)] - \widehat{\mathbb{E}}[\text{CVD}(\;)]$.

Adaptive algorithms: $\mathbb{P}[\;\mid \text{Pilot}] = f(\text{DIM})$.

- Greedy algorithm:

$$\mathbb{P}[\;\mid \text{Pilot}] = \begin{cases} 0.9 \text{ if DIM} > 0 \\ 0.1 \text{ if DIM} \leq 0 \end{cases}.$$

- (Modified) Thompson sampling:

$$\mathbb{P}[\;\mid \text{Pilot}] = \Phi(\text{DIM}).$$

- Interim p-value: define $T = \text{DIM}/\hat{\sigma}$ with $p_L = \Phi(T)$ and $p_R = 1 - p_L$.

$$\mathbb{P}[\;\mid \text{Pilot}] = 0.9 \text{ if } p_L \leq \alpha;$$

Motivating example (cont'd)

Control : 

Treatment: 

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Adaptive sampling: adjust sampling by DIM $\equiv \widehat{\mathbb{E}}$ [CVD()] $-$ $\widehat{\mathbb{E}}$ [CVD()] .

Adaptive algorithms: $\mathbb{P}[$  $\mid$ Pilot$] = f(\mathrm{DIM})$.

- Greedy algorithm:

$\mathbb{P}[$  $\mid$ Pilot$] = \begin{cases} 0.9 \text{ if DIM} > 0 \\ 0.1 \text{ if DIM} \leq 0 \end{cases}.$

- (Modified) Thompson sampling:

$\mathbb{P}[$  $\mid$ Pilot$] = \Phi(\mathrm{DIM})$.

- Interim p-value: define $T = \mathrm{DIM}/\hat{\sigma}$ with $p_L = \Phi(T)$ and $p_R = 1 - p_L$.

$\mathbb{P}[$  $\mid$ Pilot$] = 0.9$ if $p_L \leq \alpha; \quad 0.1$ if $p_L \geq 1 - \alpha;$

Motivating example (cont'd)

Control : 

Treatment: 

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Adaptive sampling: adjust sampling by DIM $\equiv \widehat{\mathbb{E}}$[CVD()] – $\widehat{\mathbb{E}}$[CVD()] .

Adaptive algorithms: $\mathbb{P}$[ | Pilot] $= f(\text{DIM})$.

- Greedy algorithm:

$\mathbb{P}$[ | Pilot] $= \begin{cases} 0.9 \text{ if DIM} > 0 \\ 0.1 \text{ if DIM} \leq 0 \end{cases}.$

- (Modified) Thompson sampling:

$\mathbb{P}$[ | Pilot] $= \Phi(\text{DIM})$.

- Interim p-value: define $T = \text{DIM}/\hat{\sigma}$ with $p_L = \Phi(T)$ and $p_R = 1 - p_L$.

$\mathbb{P}$[ | Pilot] $= 0.9$ if $p_L \leq \alpha$; $\quad 0.1$ if $p_L \geq 1 - \alpha$; $\quad 0.5$ if $p_L \in (\alpha, 1 - \alpha)$.

Motivating example (cont'd)

Control : 

Treatment: 

Response adaptive experiments: drop-the-loser experiment; batched bandit experiment.

Adaptive sampling: adjust sampling by DIM $\equiv \widehat{\mathbb{E}}$ [CVD( )] $-$ $\widehat{\mathbb{E}}$ [CVD( )] .

Adaptive algorithms: $\mathbb{P}[\ \ |$ Pilot$] = f(\mathrm{DIM})$.

- Greedy algorithm:

$$\mathbb{P}[\ \ |\ \text{Pilot}] = \begin{cases} 0.9 \text{ if DIM} > 0 \\ 0.1 \text{ if DIM} \leq 0 \end{cases}.$$

- (Modified) Thompson sampling:

$$\mathbb{P}[\ \ |\ \text{Pilot}] = \Phi(\text{DIM}) .$$

- Interim p-value: define $T = \text{DIM}/\hat{\sigma}$ with $p_L = \Phi(T)$ and $p_R = 1 - p_L$.

$$\mathbb{P}[\ \ |\ \text{Pilot}] = 0.9 \text{ if } p_L \leq \alpha; \quad 0.1 \text{ if } p_L \geq 1 - \alpha; \quad 0.5 \text{ if } p_L \in (\alpha, 1 - \alpha) .$$

Post-experiment evaluation

Post-experiment evaluation

Control : 

Treatment: 

Decision-related questions:

Post-experiment evaluation

Control : 

Treatment: 

Decision-related questions:

- Is the new treatment more effective? ($\mathbb{E}[\text{Diff}] \equiv \mathbb{E}[\text{CVD}($  $) - \text{CVD}($$)] < 0$?)

Control : 

Treatment: 

Post-experiment evaluation

Decision-related questions:

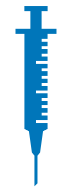
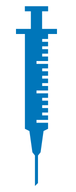
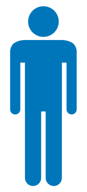
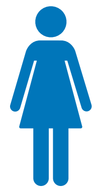
- Is the new treatment more effective? ($\mathbb{E}[\text{Diff}] \equiv \mathbb{E}[\text{CVD}($  $) - \text{CVD}($$)] < 0$?)
- Is the treatment effect significant? ($\mathbb{E}[\text{CVD}($  $)] < 0.03$?)

Control : 

Treatment: 

Post-experiment evaluation

Decision-related questions:

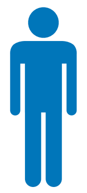
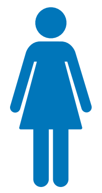
- Is the new treatment more effective? ($\mathbb{E}[\text{Diff}] \equiv \mathbb{E}[\text{CVD}($  $) - \text{CVD}($  $)] < 0$?)
- Is the treatment effect significant? ($\mathbb{E}[\text{CVD}($  $)] < 0.03$?)
- Is the treatment more effective for particular subgroup? ($\mathbb{E}[\text{Diff} \mid$  $] > \mathbb{E}[\text{Diff} \mid$  $]$?)

Post-experiment evaluation

Control : 

Treatment: 

Decision-related questions:

- Is the new treatment more effective? ($\mathbb{E}[\text{Diff}] \equiv$ 𝔼[CVD() − CVD()] < 0?)
- Is the treatment effect significant? (𝔼[CVD()] < 0.03?)
- Is the treatment more effective for particular subgroup? (𝔼[Diff | ] > 𝔼[Diff | ]?)

Data collection: treatment A, outcome Y.

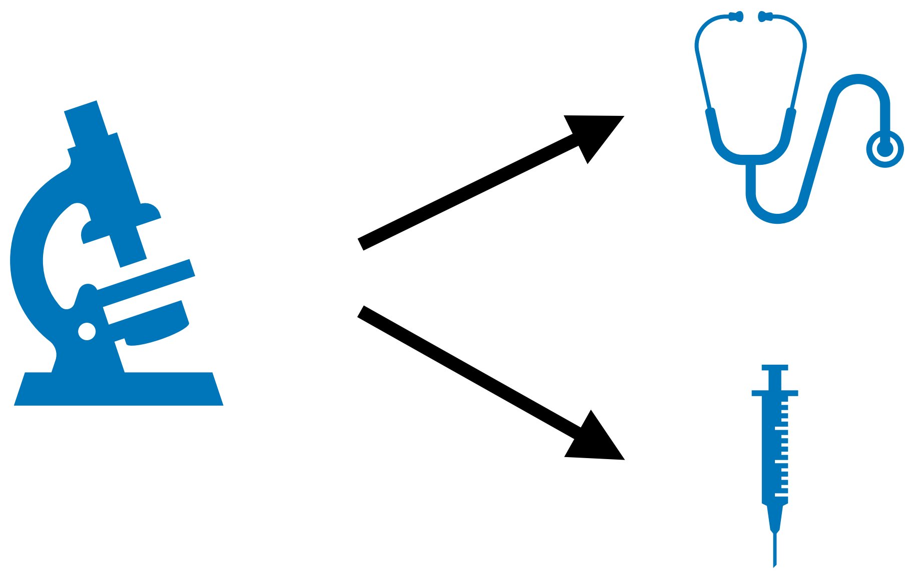
Post-experiment evaluation

Control :

Treatment:

Decision-related questions:

- Is the new treatment more effective? ($\mathbb{E}[\text{Diff}] \equiv \mathbb{E}[\text{CVD}(\ ) - \text{CVD}(\ )] < 0$?)
- Is the treatment effect significant? ($\mathbb{E}[\text{CVD}(\ )] < 0.03$?)
- Is the treatment more effective for particular subgroup? ($\mathbb{E}[\text{Diff} \mid \ ] > \mathbb{E}[\text{Diff} \mid \ ]$?)



Data collection: treatment A, outcome Y.

Stage 1: N_1 sample, $(A_1^{(1)}, Y_1^{(1)}), \ldots, (A_{N_1}^{(1)}, Y_{N_1}^{(1)})$.

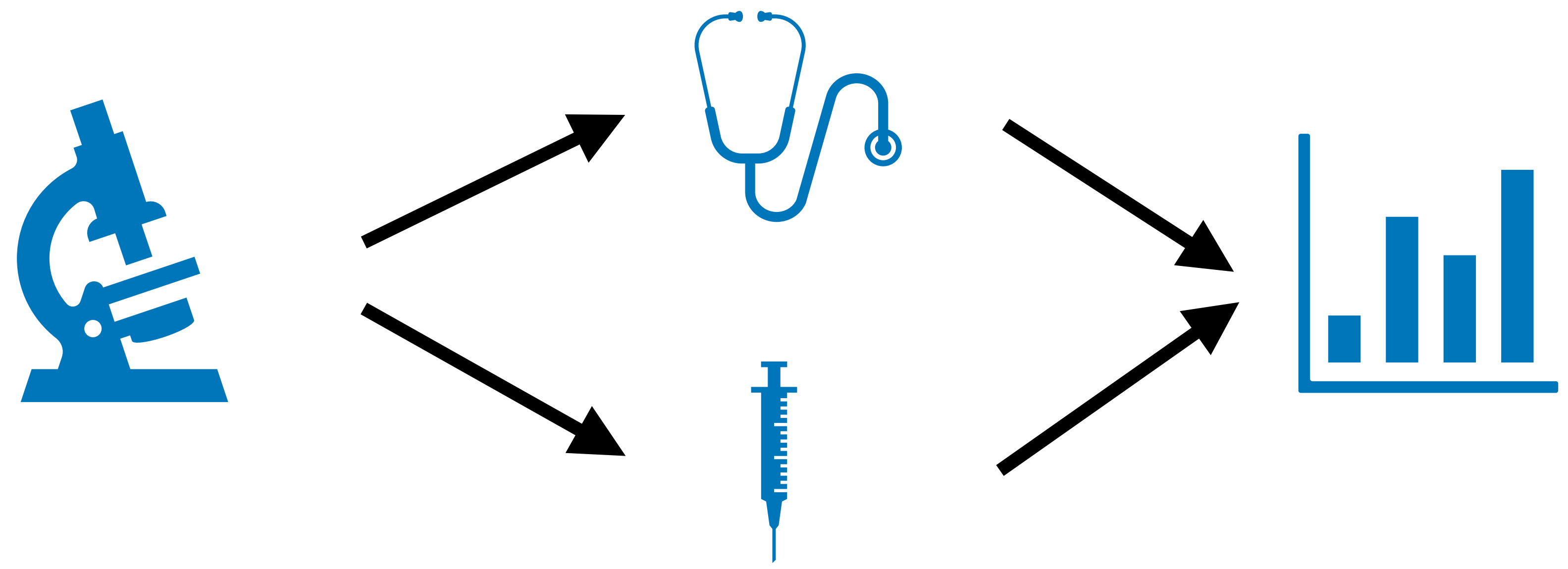
Control :

Treatment:

Post-experiment evaluation

Decision-related questions:

- Is the new treatment more effective? ($\mathbb{E}[\text{Diff}] \equiv \mathbb{E}[\text{CVD}(\ ) - \text{CVD}(\ )] < 0$?)
- Is the treatment effect significant? ($\mathbb{E}[\text{CVD}(\ )] < 0.03$?)
- Is the treatment more effective for particular subgroup? ($\mathbb{E}[\text{Diff} \mid \ ] > \mathbb{E}[\text{Diff} \mid \ ]$?)



Data collection: treatment A, outcome Y.

Stage 1: N_1 sample, $(A_1^{(1)}, Y_1^{(1)}), \ldots, (A_{N_1}^{(1)}, Y_{N_1}^{(1)})$.

Adaptive sampling: greedy algorithm, Thompson sampling, interim p-value…

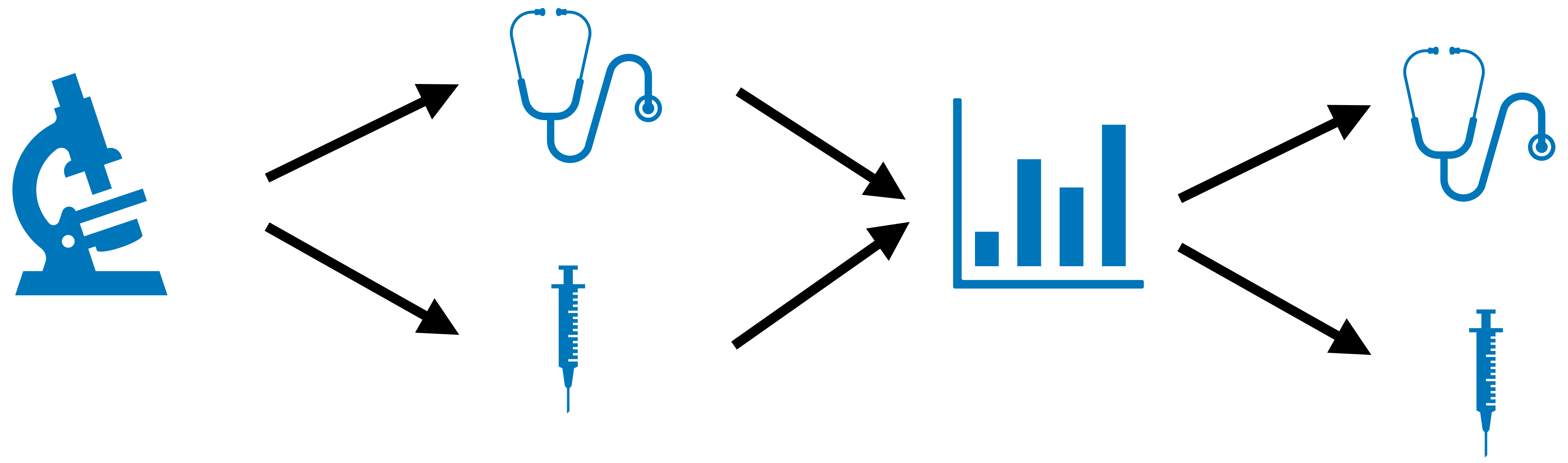
Control :

Treatment:

Post-experiment evaluation

Decision-related questions:

- Is the new treatment more effective? ($\mathbb{E}$[Diff] ≡ $\mathbb{E}$[CVD( ) − CVD( )] < 0?)
- Is the treatment effect significant? ($\mathbb{E}$[CVD( )] < 0.03?)
- Is the treatment more effective for particular subgroup? ($\mathbb{E}$[Diff |] $>$ $\mathbb{E}$[Diff |]?)



Data collection: treatment A, outcome Y.

Stage 1: N_1 sample, $(A_1^{(1)}, Y_1^{(1)}), \ldots, (A_{N_1}^{(1)}, Y_{N_1}^{(1)})$.

Adaptive sampling: greedy algorithm, Thompson sampling, interim p-value…

Stage 2: N_2 sample, $(A_1^{(2)}, Y_1^{(2)}), \ldots, (A_{N_2}^{(2)}, Y_{N_2}^{(2)})$.

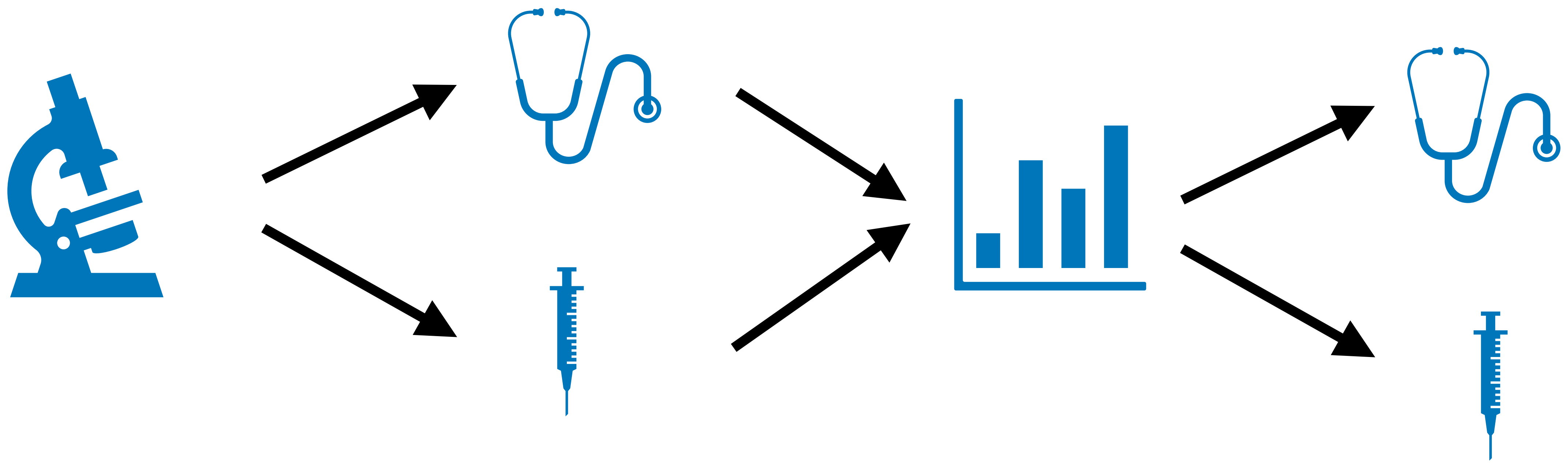
Control :

Treatment:

Post-experiment evaluation

Decision-related questions:

- Is the new treatment more effective? ($\mathbb{E}$[Diff] $\equiv$ $\mathbb{E}$[CVD( ) − CVD( )] < 0?)
- Is the treatment effect significant? ($\mathbb{E}$[CVD( )] < 0.03?)
- Is the treatment more effective for particular subgroup? ($\mathbb{E}$[Diff |] $>$ $\mathbb{E}$[Diff |]?)



Data collection: treatment A, outcome Y.

Stage 1: N_1 sample, $(A_1^{(1)}, Y_1^{(1)}), \ldots, (A_{N_1}^{(1)}, Y_{N_1}^{(1)})$.

Adaptive sampling: greedy algorithm, Thompson sampling, interim p-value…

Stage 2: N_2 sample, $(A_1^{(2)}, Y_1^{(2)}), \ldots, (A_{N_2}^{(2)}, Y_{N_2}^{(2)})$.

Challenge: pooled data $(A_u^{(t)}, Y_u^{(t)})_{u \in [N_t], t \in [2]}$ are highly dependent!

Literature review

Literature review

Conditional inference: conditional on output of adaptive algorithm.

Literature review

Conditional inference: conditional on output of adaptive algorithm.

- Data-splitting/data-carving: Chen and Andrews (2023).

Literature review

Conditional inference: conditional on output of adaptive algorithm.

- Data-splitting/data-carving: Chen and Andrews (2023).
- Randomization inference: Freidling et al. (2024).

Literature review

Conditional inference: conditional on output of adaptive algorithm.

- Data-splitting/data-carving: Chen and Andrews (2023).
- Randomization inference: Freidling et al. (2024).

Marginal inference: our interest.

Literature review

Conditional inference: conditional on output of adaptive algorithm.

- Data-splitting/data-carving: Chen and Andrews (2023).
- Randomization inference: Freidling et al. (2024).

Marginal inference: our interest.

- Randomization inference: Nair and Janson (2023).

Literature review

Conditional inference: conditional on output of adaptive algorithm.

- Data-splitting/data-carving: Chen and Andrews (2023).
- Randomization inference: Freidling et al. (2024).

Marginal inference: our interest.

- Randomization inference: Nair and Janson (2023).
- Anytime-valid inference: Howard et al. (2021); Maharaj et al. (2023).

Literature review

Conditional inference: conditional on output of adaptive algorithm.

- Data-splitting/data-carving: Chen and Andrews (2023).
- Randomization inference: Freidling et al. (2024).

Marginal inference: our interest.

- Randomization inference: Nair and Janson (2023).
- Anytime-valid inference: Howard et al. (2021); Maharaj et al. (2023).
- Asymptotic-based inference: Zhang et al. (2020); Hadad et al. (2021); Hirano and Porter (2023); Adusumilli (2023).

Literature review

Conditional inference: conditional on output of adaptive algorithm.

- Data-splitting/data-carving: Chen and Andrews (2023).
- Randomization inference: Freidling et al. (2024).

Marginal inference: our interest.

- Randomization inference: Nair and Janson (2023).
- Anytime-valid inference: Howard et al. (2021); Maharaj et al. (2023).
- Asymptotic-based inference: Zhang et al. (2020); Hadad et al. (2021); Hirano and Porter (2023); Adusumilli (2023).

Asymptotic inference frameworks

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Efforts to recover normality

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Efforts to recover normality

Zhang et al. (2020):

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Efforts to recover normality

Zhang et al. (2020):

- Stage-wise normalized estimator.

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Efforts to recover normality

Zhang et al. (2020):

- Stage-wise normalized estimator.
- Power loss because of normalization. (Hirano and Porter, 2023)

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Efforts to recover normality

Zhang et al. (2020):

- Stage-wise normalized estimator.
- Power loss because of normalization. (Hirano and Porter, 2023)

Hadad et al. (2021):

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Efforts to recover normality

Zhang et al. (2020):

- Stage-wise normalized estimator.
- Power loss because of normalization. (Hirano and Porter, 2023)

Hadad et al. (2021):

- Variance-stabilizing (A)IPW estimator.

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Efforts to recover normality

Zhang et al. (2020):

- Stage-wise normalized estimator.
- Power loss because of normalization. (Hirano and Porter, 2023)

Hadad et al. (2021):

- Variance-stabilizing (A)IPW estimator.
- Unknown asymptotic distribution under null.

Asymptotic inference frameworks

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Efforts to recover normality

Zhang et al. (2020):

- Stage-wise normalized estimator.
- Power loss because of normalization. (Hirano and Porter, 2023)

Hadad et al. (2021):

- Variance-stabilizing (A)IPW estimator.
- Unknown asymptotic distribution under null.
- Type-I error inflation when misusing the normal for testing.

Take one step back?

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Efforts to recover normality

Zhang et al. (2020):

- Stage-wise normalized estimator.
- Power loss because of normalization. (Hirano and Porter, 2023)

Hadad et al. (2021):

- Variance-stabilizing (A)IPW estimator.
- Unknown asymptotic distribution under null.
- Type-I error inflation when misusing the normal for testing.

Take one step back?

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Efforts to recover normality

Zhang et al. (2020):

- Stage-wise normalized estimator.
- Power loss because of normalization. (Hirano and Porter, 2023)

Hadad et al. (2021):

- Variance-stabilizing (A)IPW estimator.
- Unknown asymptotic distribution under null.
- Type-I error inflation when misusing the normal for testing.

Take one step back?

$$H_0 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] = 0 \quad \text{versus} \quad H_1 : \mathbb{E}[Y_u(0)] - \mathbb{E}[Y_u(1)] \neq 0$$

Non-normal sampling behavior

- Sample mean/OLS: Zhang et al. (2020, 2022).
- IPW estimator: Hadad et al. (2021).

Can we explicitly characterize these non-normal weak limits?

Efforts to recover normality

Zhang et al. (2020):

- Stage-wise normalized estimator.
- Power loss because of normalization. (Hirano and Porter, 2023)

Hadad et al. (2021):

- Variance-stabilizing (A)IPW estimator.
- Unknown asymptotic distribution under null.
- Type-I error inflation when misusing the normal for testing.

Our contributions

Our contributions

1. **Theoretically:** New assumption-lean weak limits of commonly used test statistics under minimal assumptions.

Our contributions

1. **Theoretically:** New assumption-lean weak limits of commonly used test statistics under minimal assumptions.
2. **Methodologically:** A fast bootstrap procedure for sampling from the limiting distribution and apply it for hypothesis testing.

A general class: weighted (A)IPW estimator

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathscr{H}_{t-1}]$, $\mathscr{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathcal{H}_{t-1}], \mathcal{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{e_N^{(t)}(s)} + \widehat{\mathbb{E}}(Y_u^{(t)})$$

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathcal{H}_{t-1}]$, $\mathcal{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{e_N^{(t)}(s)} + \widehat{\mathbb{E}}(Y_u^{(t)}) \qquad \text{AIPW}(s) \equiv \sum_{t=1}^{2} \frac{1}{2} \text{AIPW}^{(t)}(s)$$

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathcal{H}_{t-1}], \mathcal{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{e_N^{(t)}(s)} + \widehat{\mathbb{E}}(Y_u^{(t)}) \quad \text{WAIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{AIPW}^{(t)}(s)$$

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathscr{H}_{t-1}], \mathscr{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{\boxed{e_N^{(t)}(s)}} + \widehat{\mathbb{E}}(Y_u^{(t)}) \quad \text{WAIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{AIPW}^{(t)}(s)$$

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathcal{H}_{t-1}], \mathcal{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{\boxed{e_N^{(t)}(s)}} + \widehat{\mathbb{E}}(Y_u^{(t)}) \quad \text{WAIPW}(s) \equiv \sum_{t=1}^{2} \frac{\boxed{(e_N^{(t)}(s))^m}}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{AIPW}^{(t)}(s)$$

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathscr{H}_{t-1}]$, $\mathscr{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{\boxed{e_N^{(t)}(s)}} + \widehat{\mathbb{E}}(Y_u^{(t)}) \quad \text{WAIPW}(s) \equiv \sum_{t=1}^{2} \frac{\boxed{(e_N^{(t)}(s))^m}}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{AIPW}^{(t)}(s)$$

Compensate high variance in IPW for positive m

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathcal{H}_{t-1}], \mathcal{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{\boxed{e_N^{(t)}(s)}} + \widehat{\mathbb{E}}(Y_u^{(t)}) \quad \text{WAIPW}(s) \equiv \sum_{t=1}^{2} \frac{\boxed{(e_N^{(t)}(s))^m}}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{AIPW}^{(t)}(s)$$

Compensate high variance in IPW for positive m

Special cases:

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathscr{H}_{t-1}]$, $\mathscr{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{\boxed{e_N^{(t)}(s)}} + \widehat{\mathbb{E}}(Y_u^{(t)}) \quad \text{WAIPW}(s) \equiv \sum_{t=1}^{2} \frac{\boxed{(e_N^{(t)}(s))^m}}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{AIPW}^{(t)}(s)$$

Compensate high variance
in IPW for positive m

Special cases:

- $m = 0$: familiar AIPW estimator.

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathcal{H}_{t-1}]$, $\mathcal{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{\boxed{e_N^{(t)}(s)}} + \widehat{\mathbb{E}}(Y_u^{(t)}) \qquad \text{WAIPW}(s) \equiv \sum_{t=1}^{2} \frac{\boxed{(e_N^{(t)}(s))^m}}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{AIPW}^{(t)}(s)$$

Compensate high variance in IPW for positive m

Special cases:

- $m = 0$: familiar AIPW estimator.
- $m = 1/2$: variance-stabilizing estimator.[1]

1: Luedtke and Van Der Laan (2016); Hadad et al., (2021)

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathscr{H}_{t-1}]$, $\mathscr{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{e_N^{(t)}(s)} + \widehat{\mathbb{E}}(Y_u^{(t)}) \quad \text{WAIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \text{AIPW}^{(t)}(s)$$

Compensate high variance in IPW for positive m

Special cases:

- $m = 0$: familiar AIPW estimator.
- $m = 1/2$: variance-stabilizing estimator.[1]
- $m = 1$: (asymptotically) equivalent to sample mean.

1: Luedtke and Van Der Laan (2016); Hadad et al., (2021)

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathcal{H}_{t-1}], \mathcal{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{e_N^{(t)}(s)} + \widehat{\mathbb{E}}(Y_u^{(t)}) \quad \text{WAIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{AIPW}^{(t)}(s)$$

Special cases:

- $m = 0$: familiar AIPW estimator.
- $m = 1/2$: variance-stabilizing estimator.[1]
- $m = 1$: (asymptotically) equivalent to sample mean.

1: Luedtke and Van Der Laan (2016); Hadad et al., (2021)

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathscr{H}_{t-1}], \mathscr{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{e_N^{(t)}(s)} + \widehat{\mathbb{E}}(Y_u^{(t)}) \quad \text{WAIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{AIPW}^{(t)}(s)$$

Special cases:

- $m = 0$: familiar AIPW estimator.
- $m = 1/2$: variance-stabilizing estimator.[1]
- $m = 1$: (asymptotically) equivalent to sample mean.

Our focus:

1: Luedtke and Van Der Laan (2016); Hadad et al., (2021)

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \,|\, \mathscr{H}_{t-1}], \mathscr{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{AIPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s)(Y_u^{(t)} - \widehat{\mathbb{E}}(Y_u^{(t)}))}{e_N^{(t)}(s)} + \widehat{\mathbb{E}}(Y_u^{(t)}) \quad \text{WAIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{AIPW}^{(t)}(s)$$

Special cases:

Our focus:

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathcal{H}_{t-1}], \mathcal{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{IPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s) Y_u^{(t)}}{e_N^{(t)}(s)}$$

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Special cases:

Our focus:

Weighted IPW (WIPW) estimator, $\widehat{\mathbb{E}}(Y_u^{(t)}) = 0$

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathscr{H}_{t-1}], \mathscr{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{IPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s) Y_u^{(t)}}{e_N^{(t)}(s)}$$

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Special cases:

Our focus:

Weighted IPW (WIPW) estimator, $\widehat{\mathbb{E}}(Y_u^{(t)}) = 0$

$$m \in \{0, 1/2\}$$

A general class: weighted (A)IPW estimator

Test statistics: suppose equal batch size $N_1 = N_2 = N/2$.

- Denote the sampling probabilities $e_N^{(t)}(s) \equiv \mathbb{P}[A_u^{(t)} = s \mid \mathscr{H}_{t-1}], \mathscr{H}_t \equiv \sigma((A_1^{(t)}, Y_1^{(t)}), \ldots, (A_{N_t}^{(t)}, Y_{N_t}^{(t)}))$.

$$\text{IPW}^{(t)}(s) \equiv \frac{1}{N_t} \frac{\sum_{u=1}^{N_t} \mathbf{1}(A_u^{(t)} = s) Y_u^{(t)}}{e_N^{(t)}(s)} \qquad \text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Special cases:

- $m = 0$: familiar IPW estimator.
- $m = 1/2$: variance-stabilizing IPW estimator.

Our focus:

Weighted IPW (WIPW) estimator, $\widehat{\mathbb{E}}(Y_u^{(t)}) = 0$

$m \in \{0, 1/2\}$

A peek at the sampling distribution

A peek at the sampling distribution

Simulation setup:

A peek at the sampling distribution

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N}, 9)$ where $c_N \in \{0, -5, -10, -15\}$.

A peek at the sampling distribution

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N},9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.

A peek at the sampling distribution

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N},9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.
- Second stage: sampling based on DIM = IPW$^{(1)}(0)$ − IPW$^{(1)}(1)$ and greedy algorithm.

A peek at the sampling distribution

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N}, 9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.
- Second stage: sampling based on DIM = $\text{IPW}^{(1)}(0) - \text{IPW}^{(1)}(1)$ and greedy algorithm.

$$m = 1/2$$
$$N = 1000,\ N_1 = N_2 = 500$$

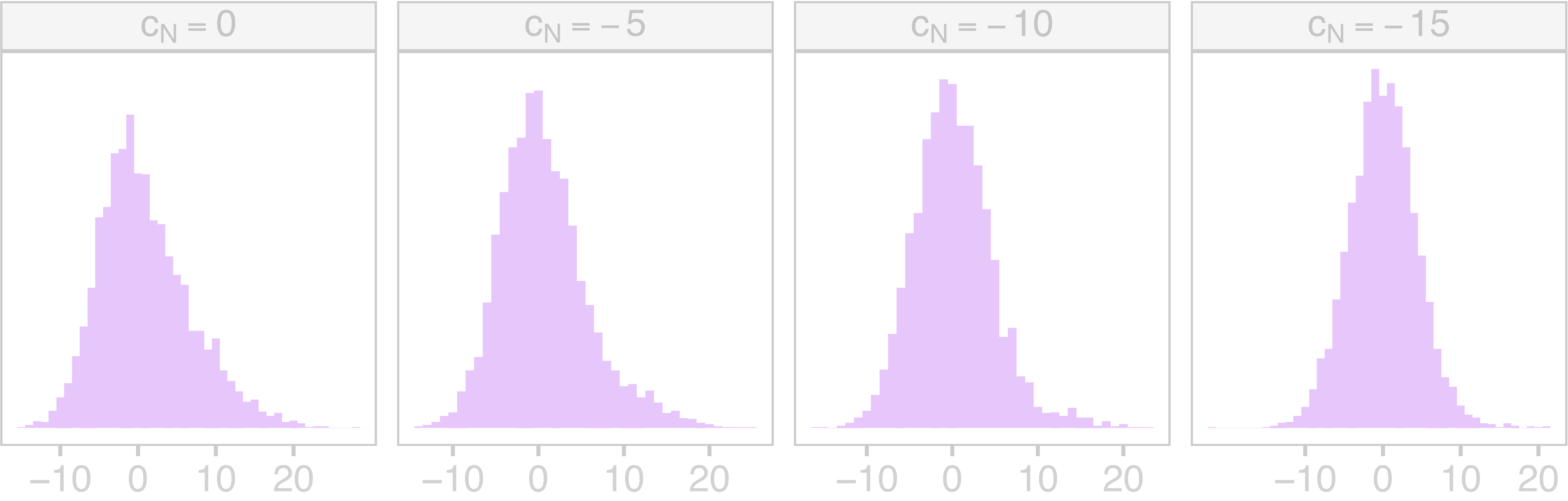
A peek at the sampling distribution

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N}, 9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.
- Second stage: sampling based on DIM = $\text{IPW}^{(1)}(0) - \text{IPW}^{(1)}(1)$ and greedy algorithm.

$$m = 1/2$$
$$N = 1000,\ N_1 = N_2 = 500$$



A peek at the sampling distribution

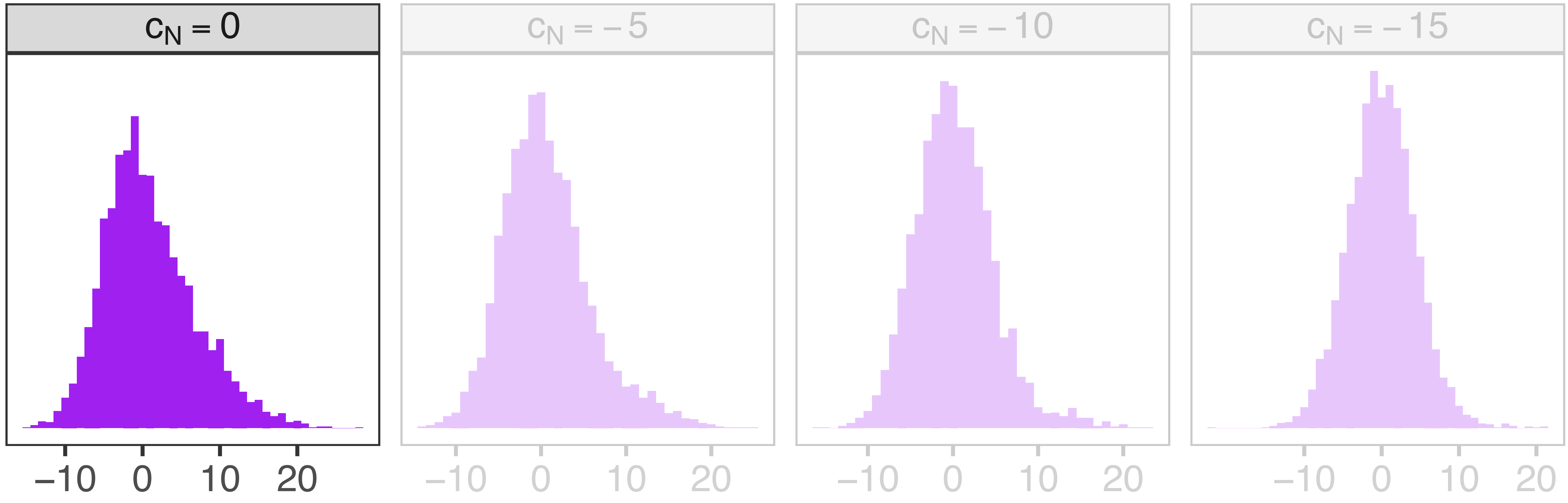
$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N}, 9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.
- Second stage: sampling based on DIM = $\text{IPW}^{(1)}(0) - \text{IPW}^{(1)}(1)$ and greedy algorithm.

$m = 1/2$

$N = 1000,\ N_1 = N_2 = 500$



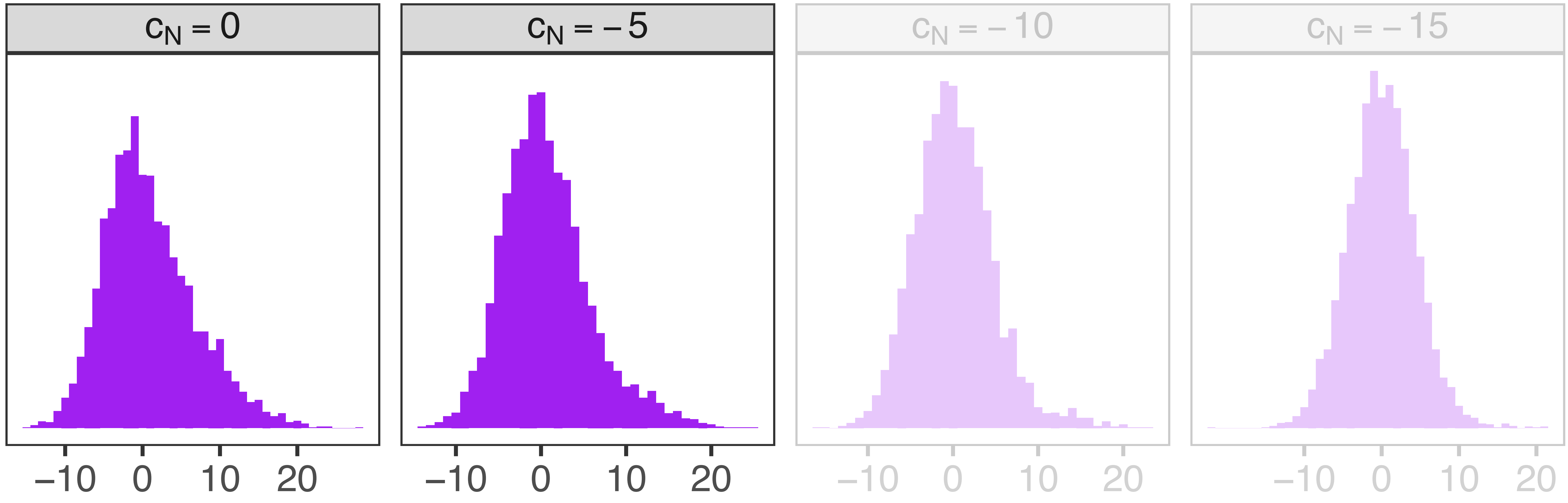
A peek at the sampling distribution

$$\mathrm{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \mathrm{IPW}^{(t)}(s)$$

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N}, 9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.
- Second stage: sampling based on DIM = IPW$^{(1)}(0)$ − IPW$^{(1)}(1)$ and greedy algorithm.

$$m = 1/2$$
$$N = 1000,\ N_1 = N_2 = 500$$



A peek at the sampling distribution

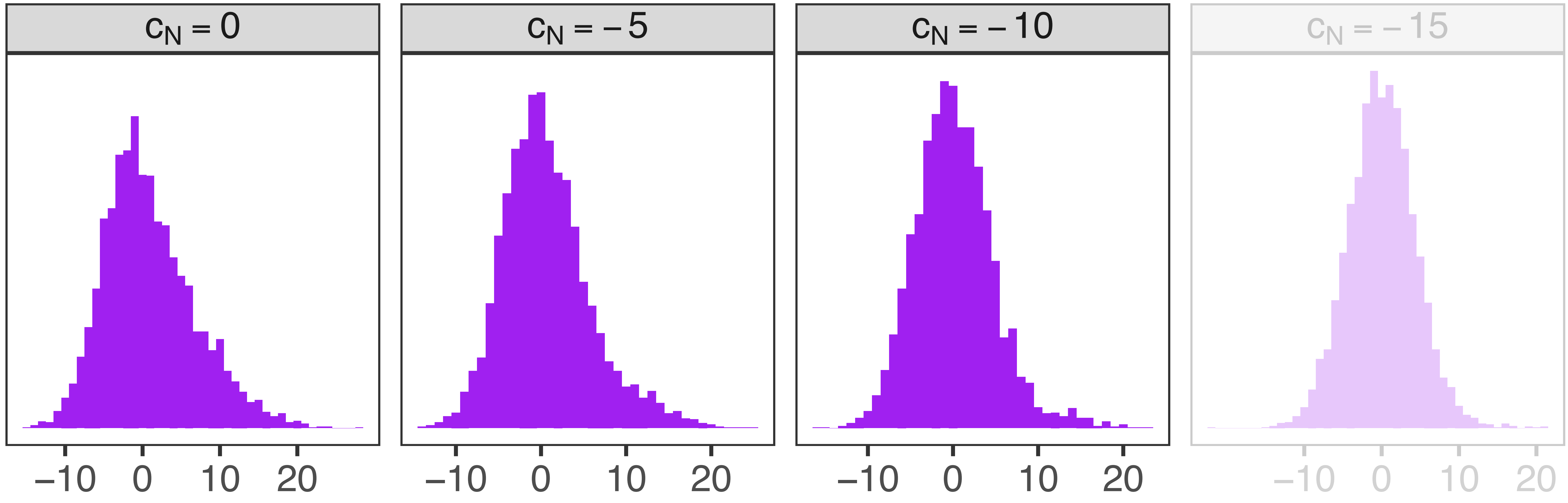
$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N}, 9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.
- Second stage: sampling based on DIM = IPW$^{(1)}(0)$ − IPW$^{(1)}(1)$ and greedy algorithm.

$m = 1/2$

$N = 1000,\ N_1 = N_2 = 500$



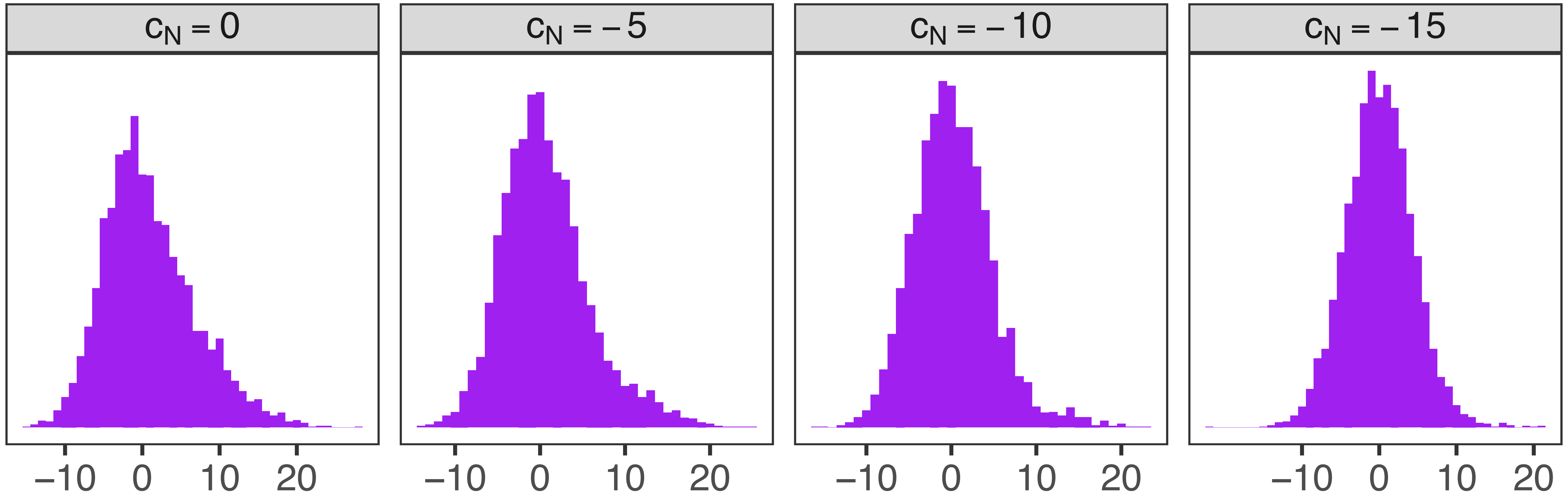
A peek at the sampling distribution

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N}, 9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.
- Second stage: sampling based on DIM = $\text{IPW}^{(1)}(0) - \text{IPW}^{(1)}(1)$ and greedy algorithm.

$$m = 1/2$$
$$N = 1000,\ N_1 = N_2 = 500$$



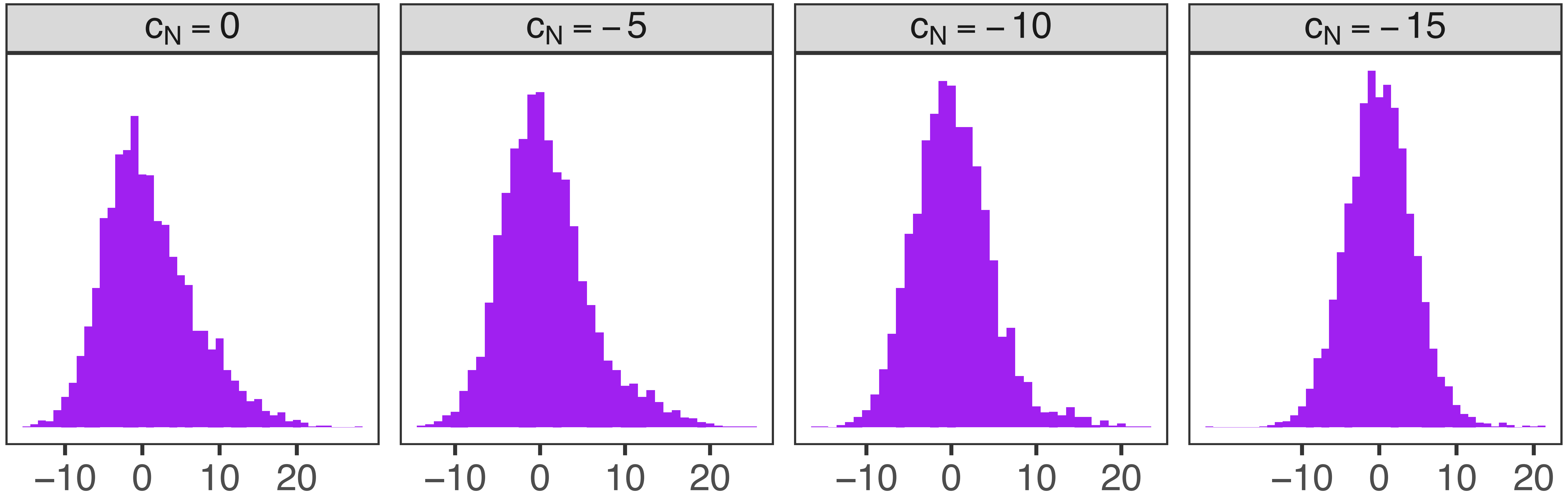
A peek at the sampling distribution

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N}, 9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.
- Second stage: sampling based on DIM = $\text{IPW}^{(1)}(0) - \text{IPW}^{(1)}(1)$ and greedy algorithm.

$$m = 1/2$$
$$N = 1000,\ N_1 = N_2 = 500$$



Weak signal

A peek at the sampling distribution

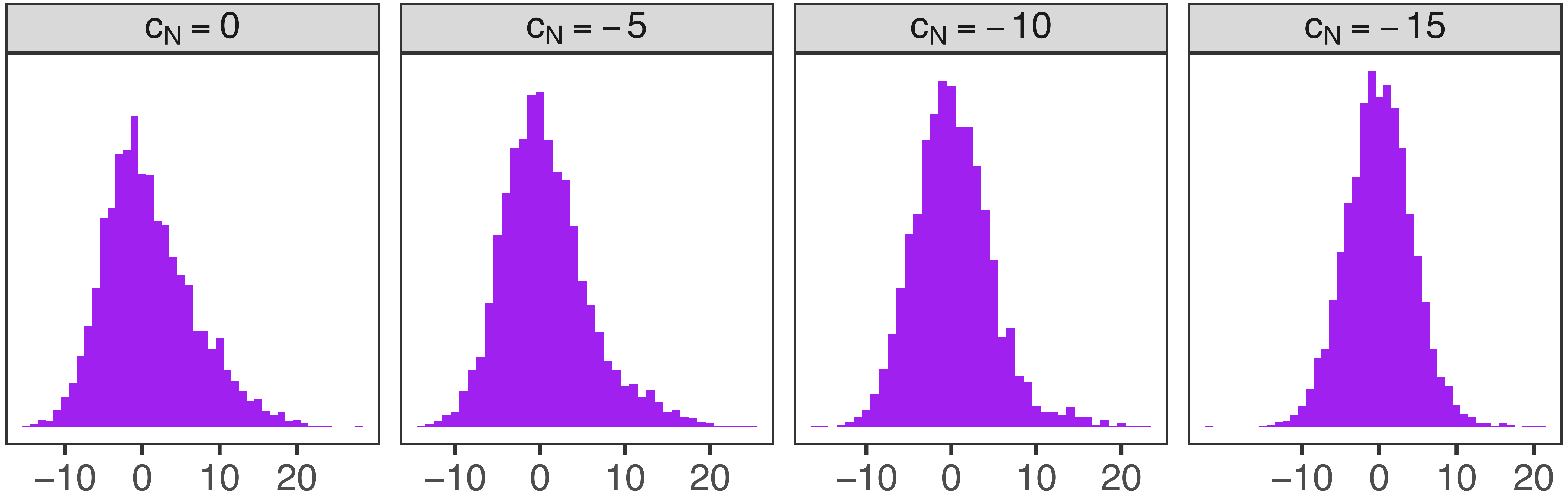
$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N}, 9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.
- Second stage: sampling based on DIM = $\text{IPW}^{(1)}(0) - \text{IPW}^{(1)}(1)$ and greedy algorithm.

$m = 1/2$

$N = 1000,\ N_1 = N_2 = 500$



Weak signal

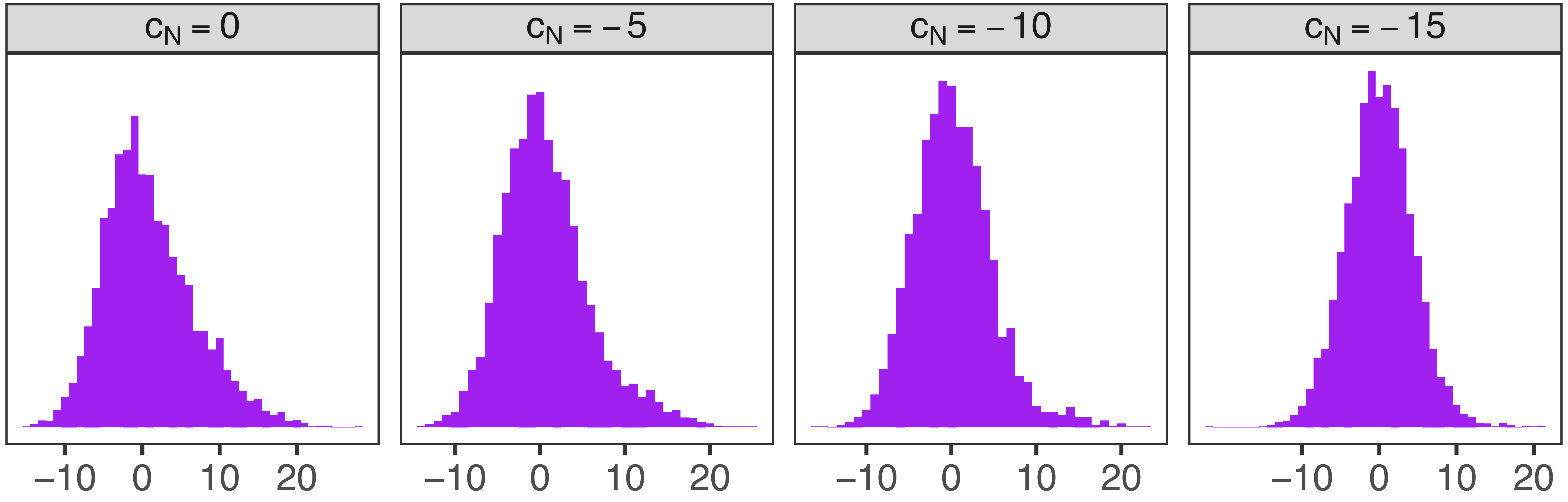
A peek at the sampling distribution

$$\mathrm{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \mathrm{IPW}^{(t)}(s)$$

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N}, 9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.
- Second stage: sampling based on DIM = $\mathrm{IPW}^{(1)}(0) - \mathrm{IPW}^{(1)}(1)$ and greedy algorithm.

$m = 1/2$
$N = 1000,\ N_1 = N_2 = 500$



Weak signal → Strong signal

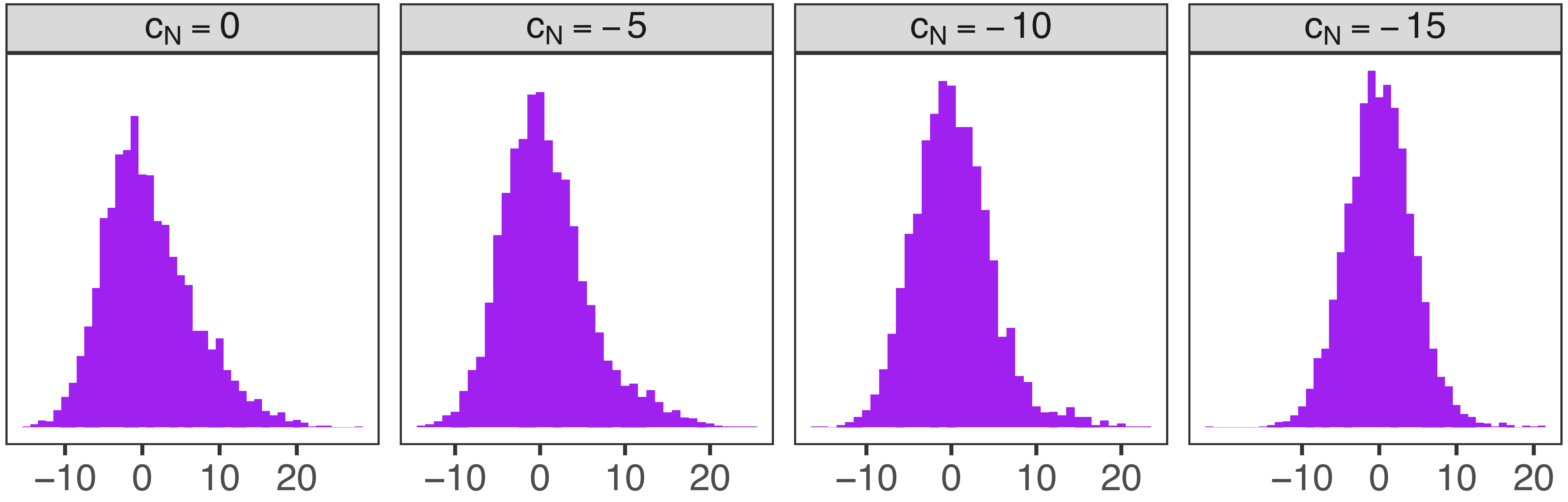
A peek at the sampling distribution

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Simulation setup:

- $Y_u(0) \sim N(0,1),\ Y_u(1) \sim N(-c_N/\sqrt{N}, 9)$ where $c_N \in \{0, -5, -10, -15\}$.
- First stage: a complete randomization, with $e_N^{(1)}(0) = e_N^{(1)}(1) = 0.5$.
- Second stage: sampling based on DIM = $\text{IPW}^{(1)}(0) - \text{IPW}^{(1)}(1)$ and greedy algorithm.

$$m = 1/2$$
$$N = 1000,\ N_1 = N_2 = 500$$



Weak signal → **Strong signal**

What role does signal strength play here?

New weak limits for WIPW

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0,1/2\}$$

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0, 1/2\}$$

as long as c_N satisfies $\lim_{N \to \infty} c_N = c$ (c can be either finite or infinite).

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0,1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

Note:

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0, 1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

Note:

- Different limit behaviors of $\mathbb{W}(c)$ on signal c.

New weak limits for WIPW

$$\mathrm{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \mathrm{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\mathrm{WIPW}(0) - \mathrm{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0,1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

Note:

- Different limit behaviors of $\mathbb{W}(c)$ on signal c.
- Highly nontrivial proof (requires extend some classical normal approximation results).[1]

1: Chatterjee and Meckes (2008)

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0,1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0, 1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

Comparison to Hirano and Porter (2023):

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0,1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

Comparison to Hirano and Porter (2023):

- No requirement on potential outcome distribution.

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0,1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

Comparison to Hirano and Porter (2023):

- No requirement on potential outcome distribution.
- Transparency in assumption.

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0,1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

Comparison to Hirano and Porter (2023):

- No requirement on potential outcome distribution.
- Transparency in assumption.
- Tailored towards to a class of estimators.

New weak limits for WIPW

$$\mathrm{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \mathrm{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\mathrm{WIPW}(0) - \mathrm{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0,1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

Comparison to Hirano and Porter (2023):

- No requirement on potential outcome distribution.
- Transparency in assumption.
- Tailored towards to a class of estimators.

Comparison to Hadad et al. (2021):

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2} (e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0,1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

Comparison to Hirano and Porter (2023):

- No requirement on potential outcome distribution.
- Transparency in assumption.
- Tailored towards to a class of estimators.

Comparison to Hadad et al. (2021):

- Same test statistic considered in both papers ($m = 1/2$).

New weak limits for WIPW

$$\mathrm{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \mathrm{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\mathrm{WIPW}(0) - \mathrm{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0,1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

Comparison to Hirano and Porter (2023):

- No requirement on potential outcome distribution.
- Transparency in assumption.
- Tailored towards to a class of estimators.

Comparison to Hadad et al. (2021):

- Same test statistic considered in both papers ($m = 1/2$).
- Full-spectrum signal strength.

New weak limits for WIPW

$$\text{WIPW}(s) \equiv \sum_{t=1}^{2} \frac{(e_N^{(t)}(s))^m}{\sum_{t=1}^{2}(e_N^{(t)}(s))^m} \text{IPW}^{(t)}(s)$$

Triangular array setup: $Y_u(s) = Y_{uN}(s)$ to include local alternative.

Theorem 1 (Niu and Ren (2025): Define $c_N \equiv \sqrt{N}(\mathbb{E}[Y_{uN}(0)] - \mathbb{E}[Y_{uN}(1)])$. Under moment and boundedness assumptions,

$$\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) - c_N \xrightarrow{d} \mathbb{W}(c) \quad \text{for} \quad m \in \{0, 1/2\}$$

as long as c_N satisfies $\lim_{N\to\infty} c_N = c$ (c can be either finite or infinite).

Comparison to Hirano and Porter (2023):

- No requirement on potential outcome distribution.
- Transparency in assumption.
- Tailored towards to a class of estimators.

Comparison to Hadad et al. (2021):

- Same test statistic considered in both papers ($m = 1/2$).
- Full-spectrum signal strength.
- Tailored towards to two-stage experiments.

Phase transition of limiting distributions

Phase transition of limiting distributions

Phase transition of limiting distributions

Theorem 2 (Niu and Ren (2025), informal result): Under the assumptions of Theorem 1, then we have

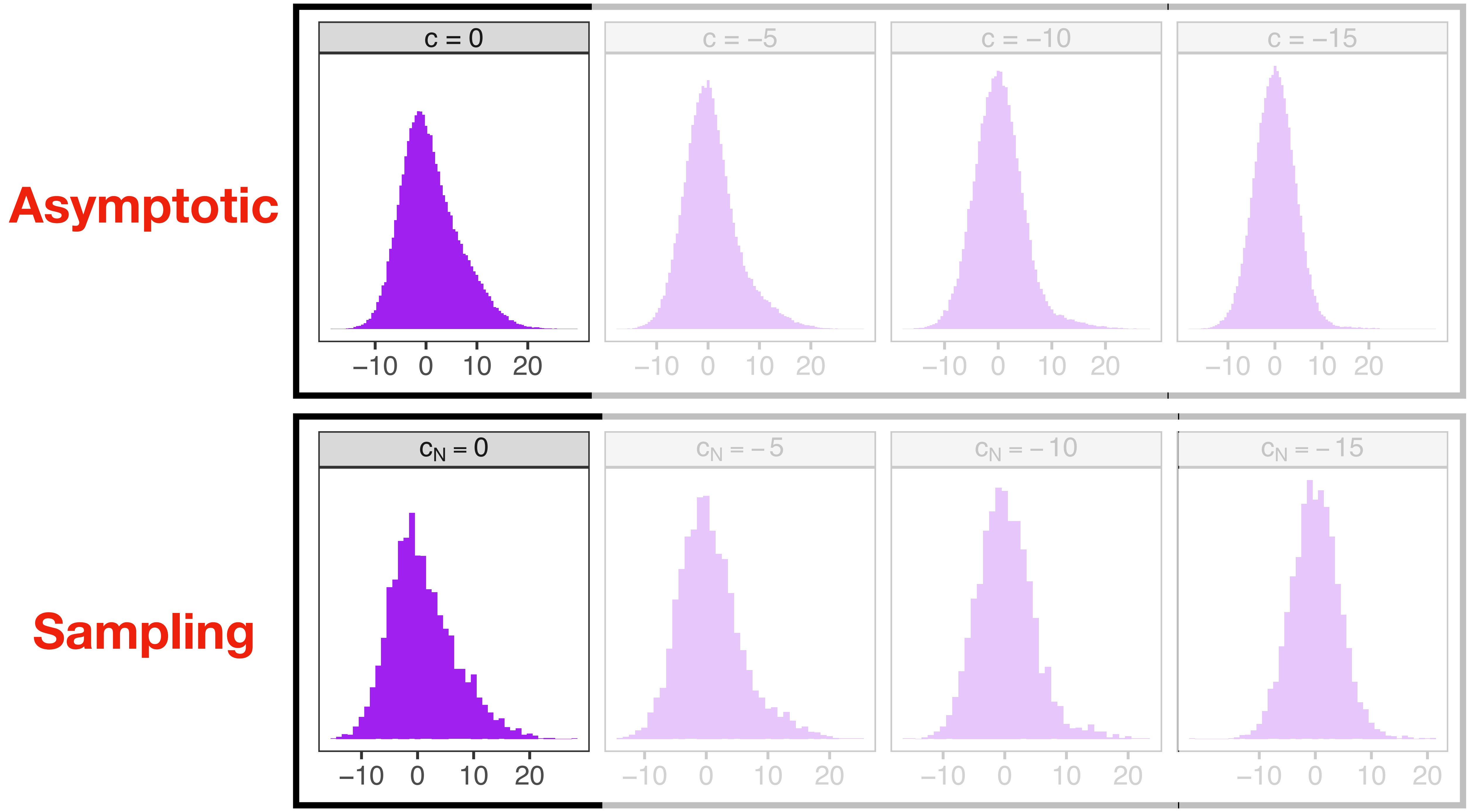
Asymptotic

Sampling

Phase transition of limiting distributions

Theorem 2 (Niu and Ren (2025), informal result): Under the assumptions of Theorem 1, then we have

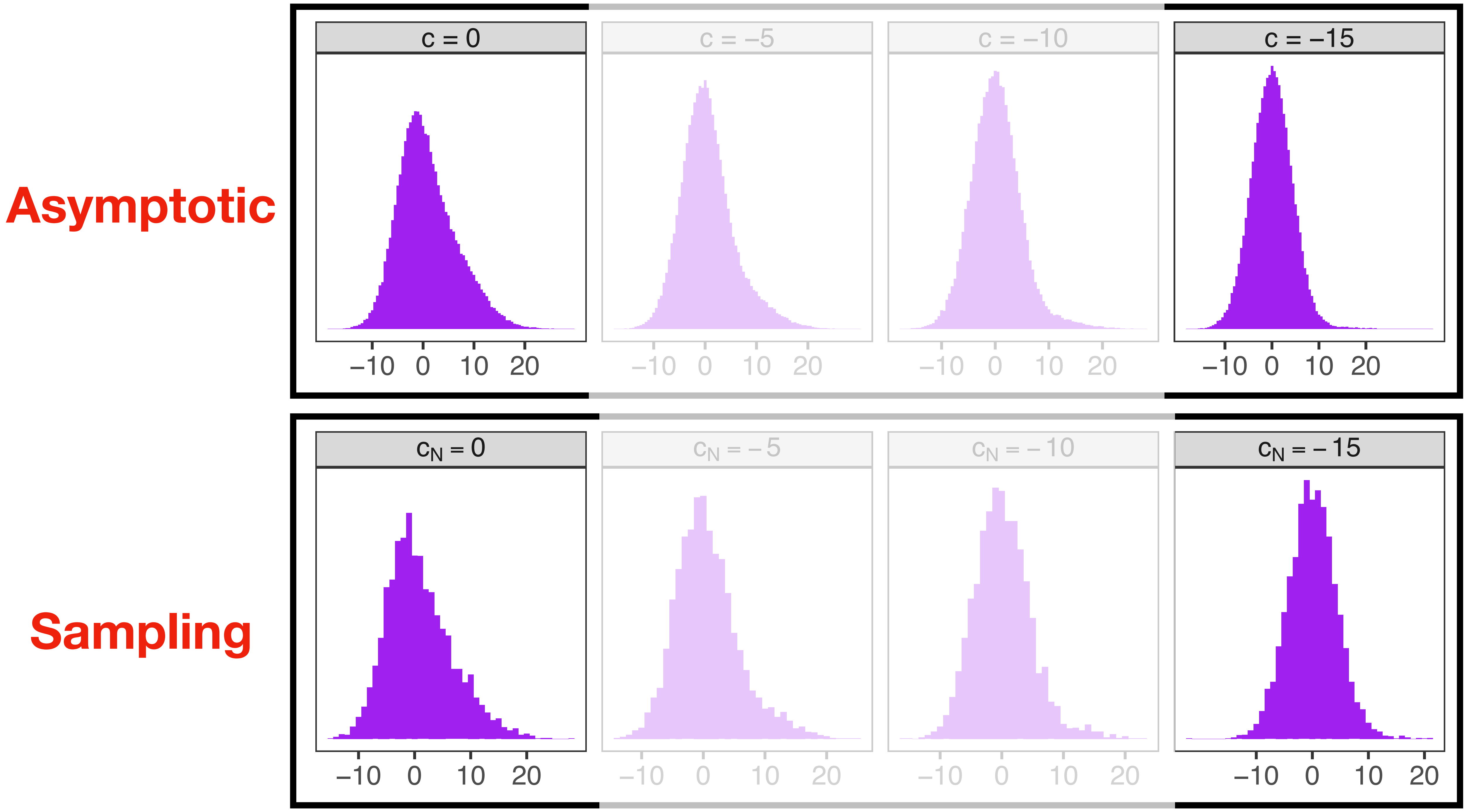
- When $c \in (-\infty, \infty)$, $\mathbb{W}(c)$ is non-normal.



Phase transition of limiting distributions

Theorem 2 (Niu and Ren (2025), informal result): Under the assumptions of Theorem 1, then we have

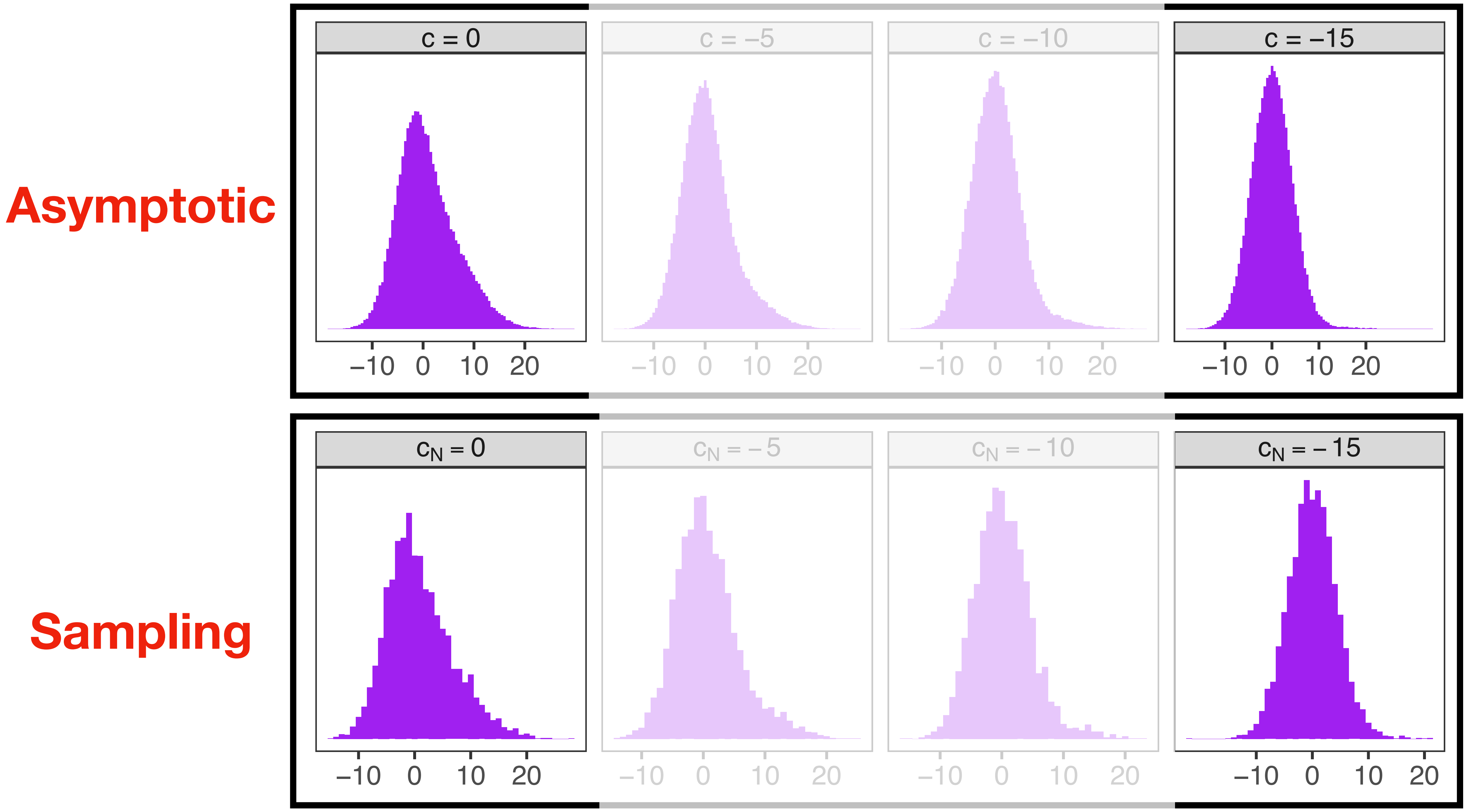
- When $c \in (-\infty, \infty)$, $\mathbb{W}(c)$ is non-normal.
- When $c \in \{-\infty, \infty\}$, $\mathbb{W}(c)$ is normal.



Phase transition of limiting distributions

Theorem 2 (Niu and Ren (2025), informal result): Under the assumptions of Theorem 1, then we have

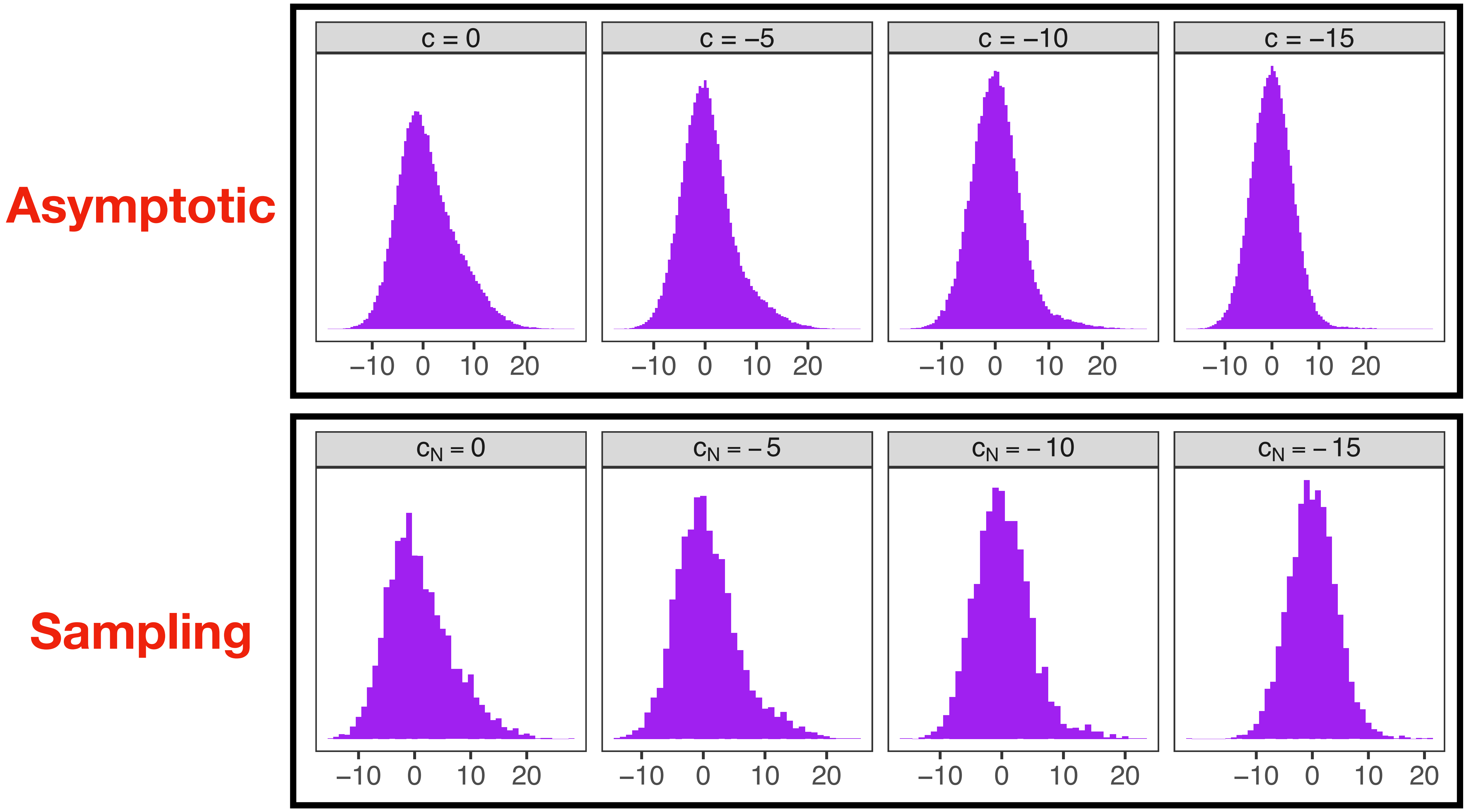
- When $c \in (-\infty, \infty)$, $\mathbb{W}(c)$ is non-normal.
- When $c \in \{-\infty, \infty\}$, $\mathbb{W}(c)$ is normal.
- Smooth transition in two regimes: $W_1(\mathbb{W}(c), \mathbb{W}(-\infty)) \to 0$ as $c \to -\infty$ (W_1: 1-Wasserstein distance).



Phase transition of limiting distributions

Theorem 2 (Niu and Ren (2025), informal result): Under the assumptions of Theorem 1, then we have

- When $c \in (-\infty, \infty)$, $\mathbb{W}(c)$ is non-normal.
- When $c \in \{-\infty, \infty\}$, $\mathbb{W}(c)$ is normal.
- Smooth transition in two regimes: $W_1(\mathbb{W}(c), \mathbb{W}(-\infty)) \to 0$ as $c \to -\infty$ (W_1: 1-Wasserstein distance).



Implication on hypothesis testing

Implication on hypothesis testing

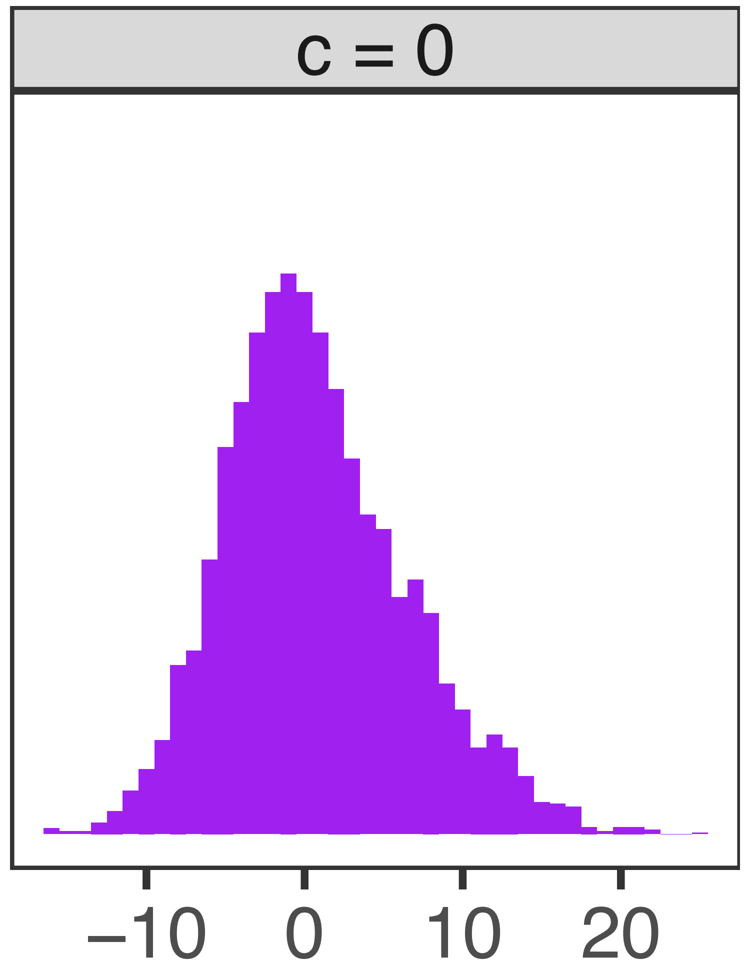
An asymptotically valid, yet impractical test:

Implication on hypothesis testing

An asymptotically valid, yet impractical test:

$$\phi \equiv \mathbf{1}(\sqrt{N}(\mathrm{WIPW}(0) - \mathrm{WIPW}(1)) > \mathbb{Q}_{1-\alpha}(\mathbb{W}(0))).$$

Implication on hypothesis testing

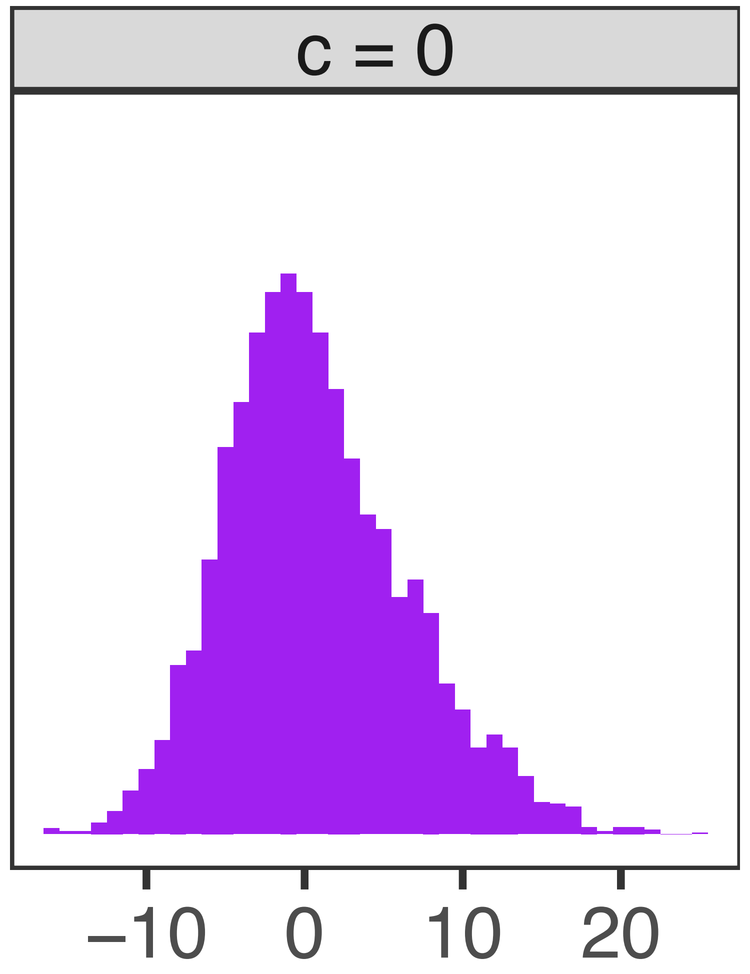


An asymptotically valid, yet impractical test:

$$\phi \equiv \mathbf{1}(\sqrt{N}(\mathrm{WIPW}(0) - \mathrm{WIPW}(1)) > \mathbb{Q}_{1-\alpha}(\mathbb{W}(0))).$$

Statistical challenge: intractable quantile $\mathbb{Q}_{1-\alpha}(\mathbb{W}(0))$.

Implication on hypothesis testing



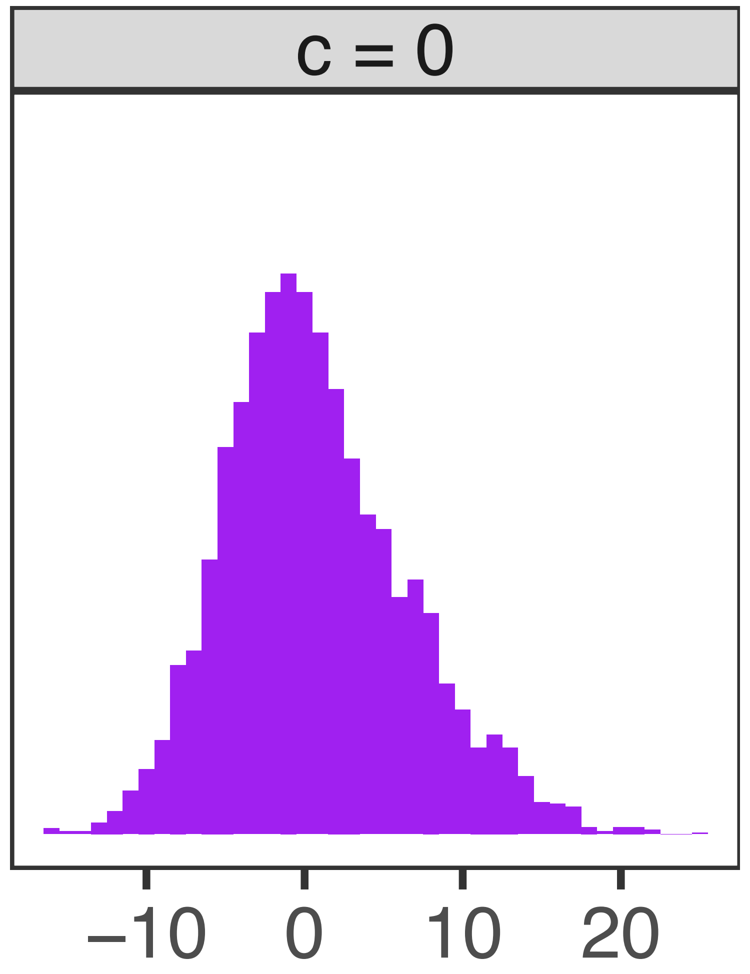
An asymptotically valid, yet impractical test:

$$\phi \equiv \mathbf{1}(\sqrt{N}(\mathrm{WIPW}(0) - \mathrm{WIPW}(1)) > \mathbb{Q}_{1-\alpha}(\mathbb{W}(0))).$$

Statistical challenge: intractable quantile $\mathbb{Q}_{1-\alpha}(\mathbb{W}(0))$.

A key observation:

Implication on hypothesis testing



An asymptotically valid, yet impractical test:

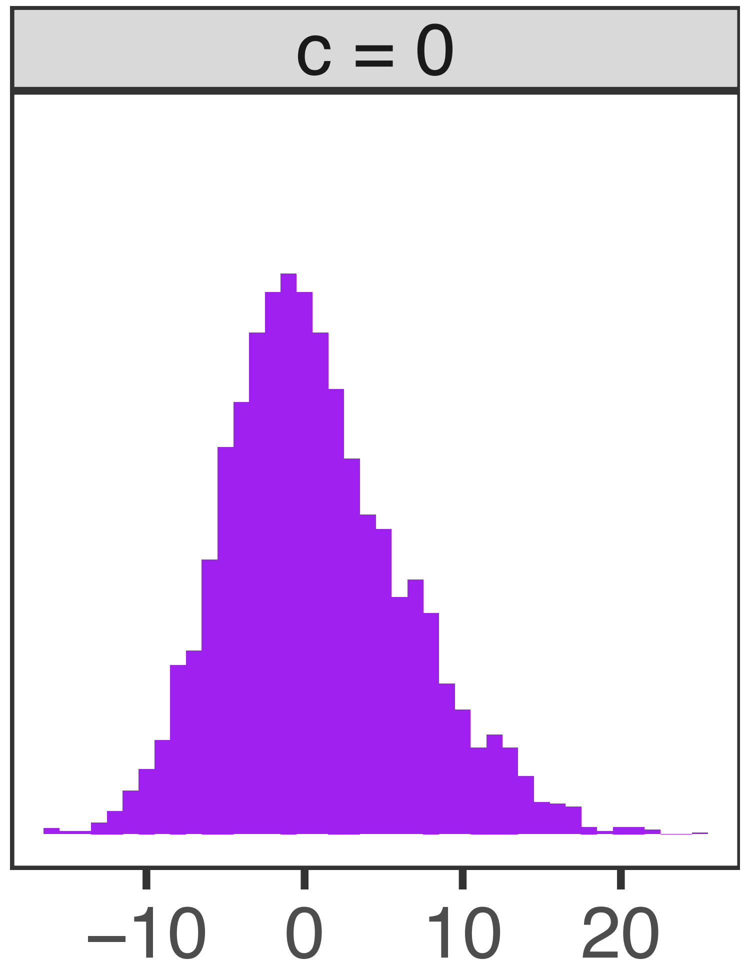
$$\phi \equiv \mathbf{1}(\sqrt{N}(\mathrm{WIPW}(0) - \mathrm{WIPW}(1)) > \mathbb{Q}_{1-\alpha}(\mathbb{W}(0))).$$

Statistical challenge: intractable quantile $\mathbb{Q}_{1-\alpha}(\mathbb{W}(0))$.

A key observation:

$$\mathbb{W}(0) \stackrel{d}{=} \sum_{t=1}^{2} G_t(A^{(t)}(0), A^{(t)}(1)).$$

Implication on hypothesis testing



An asymptotically valid, yet impractical test:

$$\phi \equiv \mathbf{1}(\sqrt{N}(\mathrm{WIPW}(0) - \mathrm{WIPW}(1)) > \mathbb{Q}_{1-\alpha}(\mathbb{W}(0))).$$

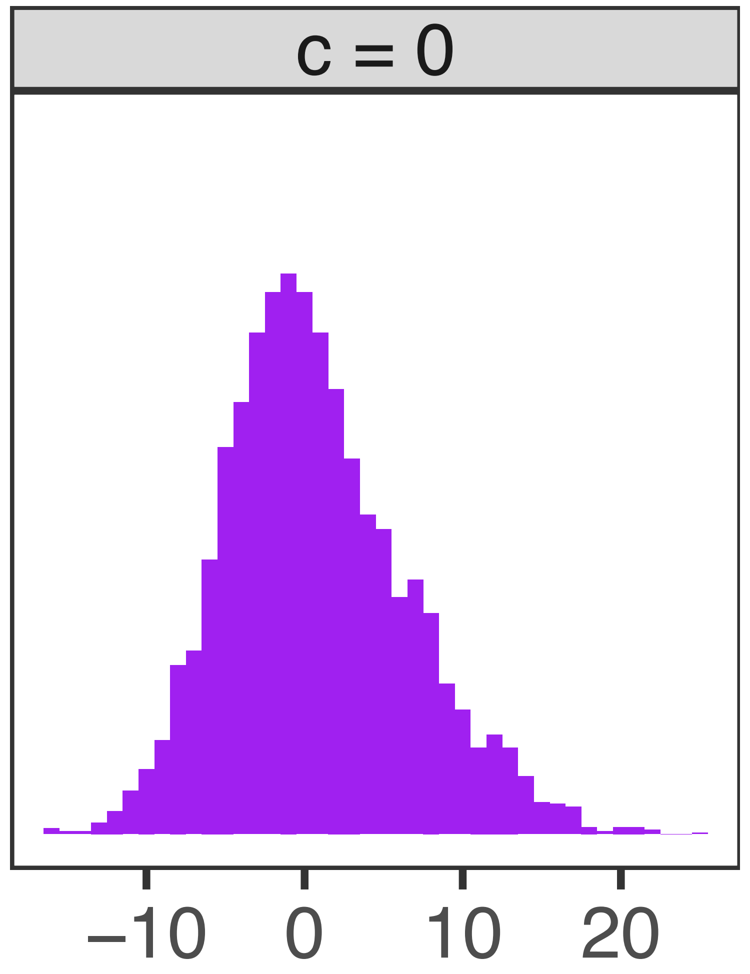
Statistical challenge: intractable quantile $\mathbb{Q}_{1-\alpha}(\mathbb{W}(0))$.

A key observation:

$$\mathbb{W}(0) \overset{d}{=} \sum_{t=1}^{2} G_t(A^{(t)}(0), A^{(t)}(1)).$$

- $A^{(1)} = (A^{(1)}(0), A^{(1)}(1)) \sim \mathrm{N}(\mathbf{0}, \mathbf{\Sigma}^{(1)})$.

Implication on hypothesis testing



An asymptotically valid, yet impractical test:

$$\phi \equiv \mathbf{1}(\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) > \mathbb{Q}_{1-\alpha}(\mathbb{W}(0))).$$

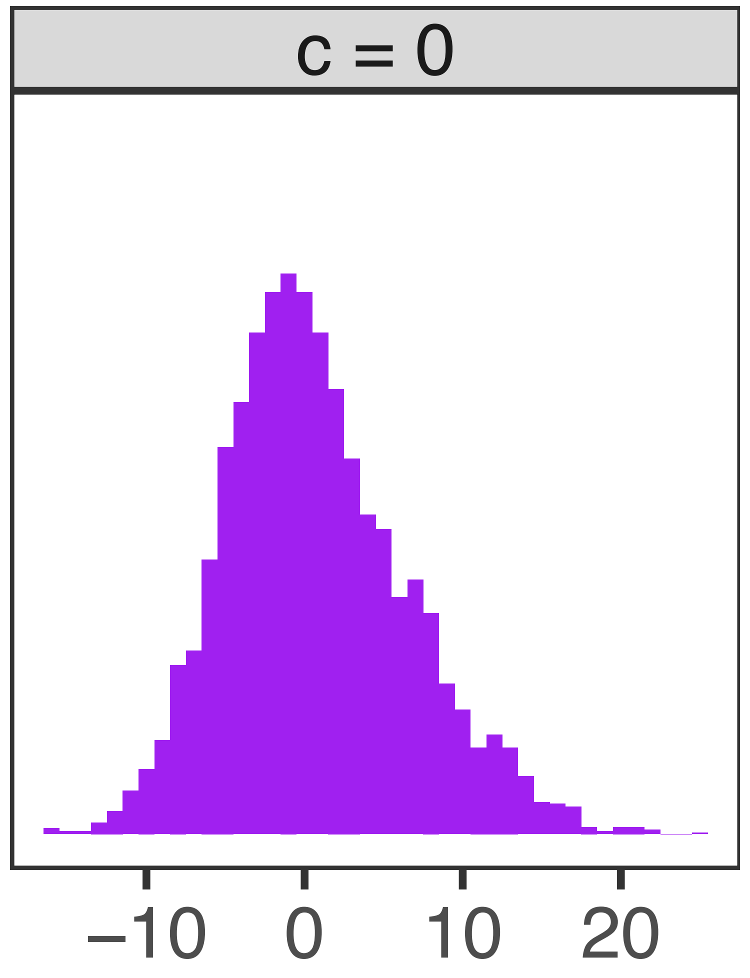
Statistical challenge: intractable quantile $\mathbb{Q}_{1-\alpha}(\mathbb{W}(0))$.

A key observation:

$$\mathbb{W}(0) \stackrel{d}{=} \sum_{t=1}^{2} G_t(A^{(t)}(0), A^{(t)}(1)).$$

- $A^{(1)} = (A^{(1)}(0), A^{(1)}(1)) \sim \mathrm{N}(\mathbf{0}, \mathbf{\Sigma}^{(1)})$.
- $A^{(2)} = (A^{(2)}(0), A^{(2)}(1)) \mid A^{(1)} \sim \mathrm{N}(\mathbf{0}, \mathbf{\Sigma}^{(2)}(A^{(1)}))$.

Implication on hypothesis testing



An asymptotically valid, yet impractical test:

$$\phi \equiv \mathbf{1}(\sqrt{N}(\mathrm{WIPW}(0) - \mathrm{WIPW}(1)) > \mathbb{Q}_{1-\alpha}(\mathbb{W}(0))).$$

Statistical challenge: intractable quantile $\mathbb{Q}_{1-\alpha}(\mathbb{W}(0))$.

A key observation:

$$\mathbb{W}(0) \stackrel{d}{=} \sum_{t=1}^{2} G_t(A^{(t)}(0), A^{(t)}(1)).$$

- $A^{(1)} = (A^{(1)}(0), A^{(1)}(1)) \sim \mathsf{N}(\mathbf{0}, \mathbf{\Sigma}^{(1)})$.
- $A^{(2)} = (A^{(2)}(0), A^{(2)}(1)) \mid A^{(1)} \sim \mathsf{N}(\mathbf{0}, \mathbf{\Sigma}^{(2)}(A^{(1)}))$.
- G_t: involving nuisance parameter but can be estimated.

A scalable bootstrap

A scalable bootstrap

A two-stage plug-in bootstrap

A scalable bootstrap

A two-stage plug-in bootstrap

- Stage 1: sample $\tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}^{(1)})$.

A scalable bootstrap

A two-stage plug-in bootstrap

- Stage 1: sample $\tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\mathbf{\Sigma}}^{(1)})$.
- Stage 2: sample $\tilde{A}^{(2)} \,|\, \tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\mathbf{\Sigma}}^{(2)}(\tilde{A}^{(1)}))$.

A scalable bootstrap

A two-stage plug-in bootstrap

- Stage 1: sample $\tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\mathbf{\Sigma}}^{(1)})$.
- Stage 2: sample $\tilde{A}^{(2)} \mid \tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\mathbf{\Sigma}}^{(2)}(\tilde{A}^{(1)}))$.
- Obtain bootstrap sample:

$$\mathbb{W}(0) \overset{d}{=} \sum_{t=1}^{2} G_t(A^{(t)}(0), A^{(t)}(1))$$

A scalable bootstrap

A two-stage plug-in bootstrap

- Stage 1: sample $\tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}^{(1)})$.
- Stage 2: sample $\tilde{A}^{(2)} \,|\, \tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}^{(2)}(\tilde{A}^{(1)}))$.
- Obtain bootstrap sample:

$$\mathscr{D} = \sum_{t=1}^{2} \widehat{G}_t(\tilde{A}^{(1)}(0), \tilde{A}^{(1)}(1)).$$

A scalable bootstrap

$$\mathbb{W}(0) \overset{d}{=} \sum_{t=1}^{2} G_t(A^{(t)}(0), A^{(t)}(1))$$

A two-stage plug-in bootstrap

- Stage 1: sample $\tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\mathbf{\Sigma}}^{(1)})$.
- Stage 2: sample $\tilde{A}^{(2)} \,|\, \tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\mathbf{\Sigma}}^{(2)}(\tilde{A}^{(1)}))$.
- Obtain bootstrap sample:

$$\mathscr{D} = \sum_{t=1}^{2} \widehat{G}_t(\tilde{A}^{(1)}(0), \tilde{A}^{(1)}(1)).$$

Note:

A scalable bootstrap

$$\mathbb{W}(0) \overset{d}{=} \sum_{t=1}^{2} G_t(A^{(t)}(0), A^{(t)}(1))$$

A two-stage plug-in bootstrap

- Stage 1: sample $\tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}^{(1)})$.
- Stage 2: sample $\tilde{A}^{(2)} \,|\, \tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}^{(2)}(\tilde{A}^{(1)}))$.
- Obtain bootstrap sample:

$$\mathscr{D} = \sum_{t=1}^{2} \widehat{G}_t(\tilde{A}^{(1)}(0), \tilde{A}^{(1)}(1)).$$

Note:

- Highly scalable, time complexity $O(1)$.

A scalable bootstrap

$$\mathbb{W}(0) \stackrel{d}{=} \sum_{t=1}^{2} G_t(A^{(t)}(0), A^{(t)}(1))$$

A two-stage plug-in bootstrap

- Stage 1: sample $\tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}^{(1)})$.
- Stage 2: sample $\tilde{A}^{(2)} \,|\, \tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}^{(2)}(\tilde{A}^{(1)}))$.
- Obtain bootstrap sample:

$$\mathscr{D} = \sum_{t=1}^{2} \widehat{G}_t(\tilde{A}^{(1)}(0), \tilde{A}^{(1)}(1)).$$

Note:

- Highly scalable, time complexity $O(1)$.
- A nonparametric sampling algorithm, agnostic to data generating procedure.

A scalable bootstrap

$$\mathbb{W}(0) \overset{d}{=} \sum_{t=1}^{2} G_t(A^{(t)}(0), A^{(t)}(1))$$

A two-stage plug-in bootstrap

- Stage 1: sample $\tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}^{(1)})$.
- Stage 2: sample $\tilde{A}^{(2)} \mid \tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}^{(2)}(\tilde{A}^{(1)}))$.
- Obtain bootstrap sample:

$$\mathscr{D} = \sum_{t=1}^{2} \widehat{G}_t(\tilde{A}^{(1)}(0), \tilde{A}^{(1)}(1)).$$

Note:

- Highly scalable, time complexity $O(1)$.
- A nonparametric sampling algorithm, agnostic to data generating procedure.

Theorem 3 (N. and Ren (2025)): Suppose $\mathscr{G}_N \equiv \sigma$(two-stage data). Then the bootstrap procedure is valid:

A scalable bootstrap

$$\mathbb{W}(0) \overset{d}{=} \sum_{t=1}^{2} G_t(A^{(t)}(0), A^{(t)}(1))$$

A two-stage plug-in bootstrap

- Stage 1: sample $\tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}^{(1)})$.
- Stage 2: sample $\tilde{A}^{(2)} \,|\, \tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}^{(2)}(\tilde{A}^{(1)}))$.
- Obtain bootstrap sample:

$$\mathscr{D} = \sum_{t=1}^{2} \widehat{G}_t(\tilde{A}^{(1)}(0), \tilde{A}^{(1)}(1)).$$

Note:

- Highly scalable, time complexity $O(1)$.
- A nonparametric sampling algorithm, agnostic to data generating procedure.

Theorem 3 (N. and Ren (2025)): Suppose $\mathscr{G}_N \equiv \sigma$(two-stage data). Then the bootstrap procedure is valid:

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}[\mathscr{D} \leq t \,|\, \mathscr{G}_n] - \mathbb{P}[\mathbb{W}(0) \leq t] \right| = o_p(1).$$

A scalable bootstrap

$$\mathbb{W}(0) \overset{d}{=} \sum_{t=1}^{2} G_t(A^{(t)}(0), A^{(t)}(1))$$

A two-stage plug-in bootstrap

- Stage 1: sample $\tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\mathbf{\Sigma}}^{(1)})$.
- Stage 2: sample $\tilde{A}^{(2)} \mid \tilde{A}^{(1)} \sim \mathrm{N}(\mathbf{0}, \widehat{\mathbf{\Sigma}}^{(2)}(\tilde{A}^{(1)}))$.
- Obtain bootstrap sample:

$$\mathscr{D} = \sum_{t=1}^{2} \widehat{G}_t(\tilde{A}^{(1)}(0), \tilde{A}^{(1)}(1)).$$

Note:

- Highly scalable, time complexity $O(1)$.
- A nonparametric sampling algorithm, agnostic to data generating procedure.

Theorem 3 (N. and Ren (2025)): Suppose $\mathscr{G}_N \equiv \sigma$(two-stage data). Then the bootstrap procedure is valid:

$$\sup_{t \in \mathbb{R}} |\mathbb{P}[\mathscr{D} \leq t \mid \mathscr{G}_n] - \mathbb{P}[\mathbb{W}(0) \leq t]| = o_p(1).$$

Moreover, the test $\phi_b \equiv \mathbf{1}(\sqrt{N}(\text{WIPW}(0) - \text{WIPW}(1)) > \mathbb{Q}_{1-\alpha}(\mathscr{D} \mid \mathscr{G}_N))$ is asymptotically valid.

Numerical simulation

Numerical simulation

Data generation details:

Numerical simulation

Data generation details:

- Potential outcome distributions:

Numerical simulation

Data generation details:

- Potential outcome distributions:

$$Y(0) \sim N(\theta, 1), \; Y(1) \sim N(0, 0.25);$$

Numerical simulation

Data generation details:

- Potential outcome distributions:

$$Y(0) \sim N(\theta, 1),\ Y(1) \sim N(0, 0.25);$$

$$Y(0) \sim \mathrm{Ber}(\theta + 0.5),\ Y(1) \sim \mathrm{Ber}(0.5).$$

Numerical simulation

Data generation details:

- Potential outcome distributions:

$$Y(0) \sim N(\theta,1),\ Y(1) \sim N(0,0.25);$$

$$Y(0) \sim \mathrm{Ber}(\theta + 0.5),\ Y(1) \sim \mathrm{Ber}(0.5).$$

- Balanced first stage sampling:

Numerical simulation

Data generation details:

- Potential outcome distributions:

$$Y(0) \sim N(\theta, 1),\ Y(1) \sim N(0, 0.25);$$

$$Y(0) \sim \mathrm{Ber}(\theta + 0.5),\ Y(1) \sim \mathrm{Ber}(0.5).$$

- Balanced first stage sampling:

Equal probability sampling.

Numerical simulation

Data generation details:

- Potential outcome distributions:

$$Y(0) \sim N(\theta, 1),\ Y(1) \sim N(0, 0.25);$$

$$Y(0) \sim \mathrm{Ber}(\theta + 0.5),\ Y(1) \sim \mathrm{Ber}(0.5).$$

- Balanced first stage sampling:

Equal probability sampling.

- Sampling algorithm:

Numerical simulation

Data generation details:

- Potential outcome distributions:

$$Y(0) \sim N(\theta,1),\ Y(1) \sim N(0,0.25);$$

$$Y(0) \sim \text{Ber}(\theta + 0.5),\ Y(1) \sim \text{Ber}(0.5).$$

- Balanced first stage sampling:

Equal probability sampling.

- Sampling algorithm:

Modified Thompson sampling with clipping ε.

Numerical simulation

Data generation details:

- Potential outcome distributions:

$$Y(0) \sim N(\theta, 1),\ Y(1) \sim N(0, 0.25);$$

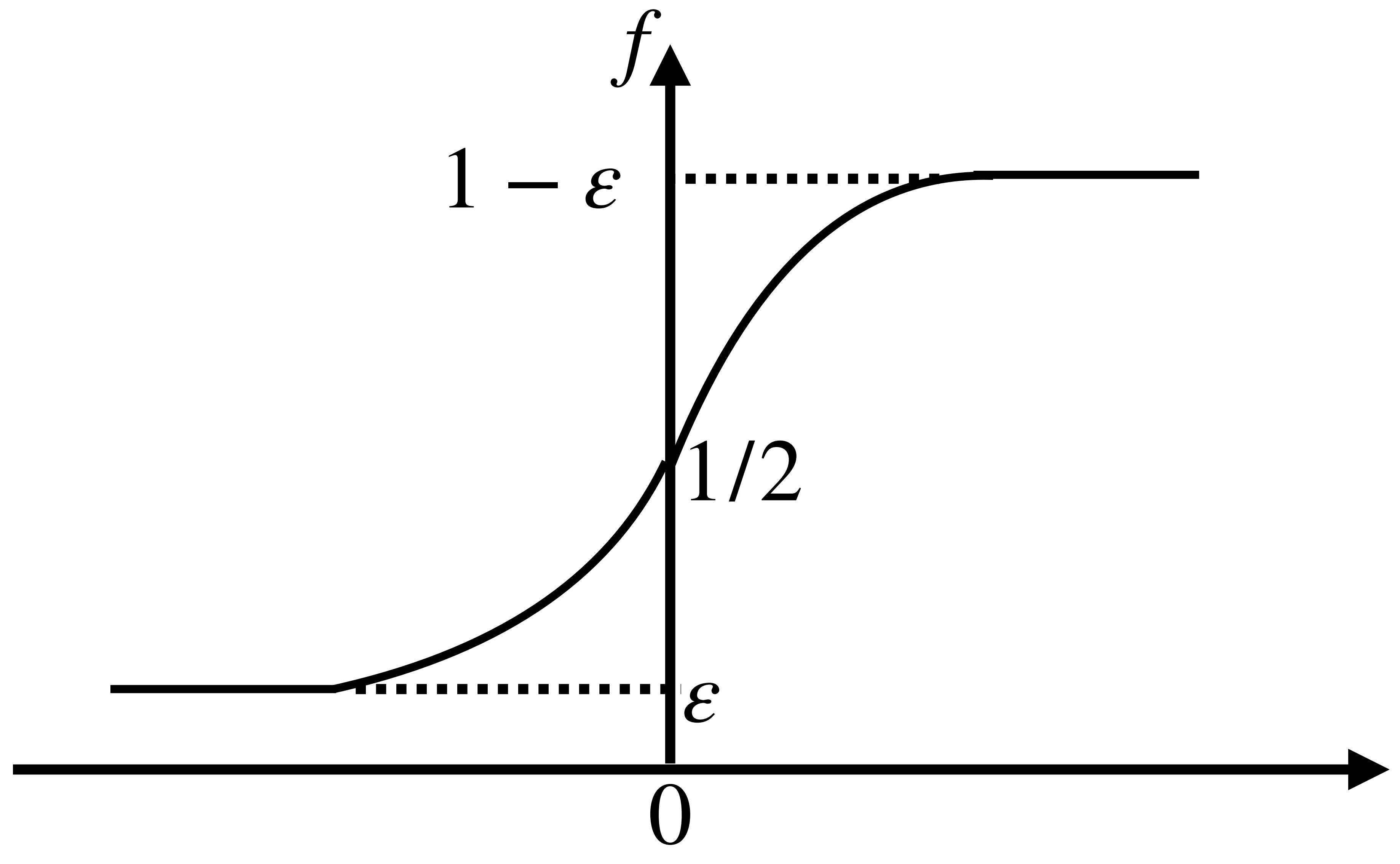
$$Y(0) \sim \mathrm{Ber}(\theta + 0.5),\ Y(1) \sim \mathrm{Ber}(0.5).$$

- Balanced first stage sampling:

Equal probability sampling.

- Sampling algorithm:

Modified Thompson sampling with clipping ε.



Numerical simulation

Data generation details:

- Potential outcome distributions:

$$Y(0) \sim N(\theta,1),\ Y(1) \sim N(0,0.25);$$

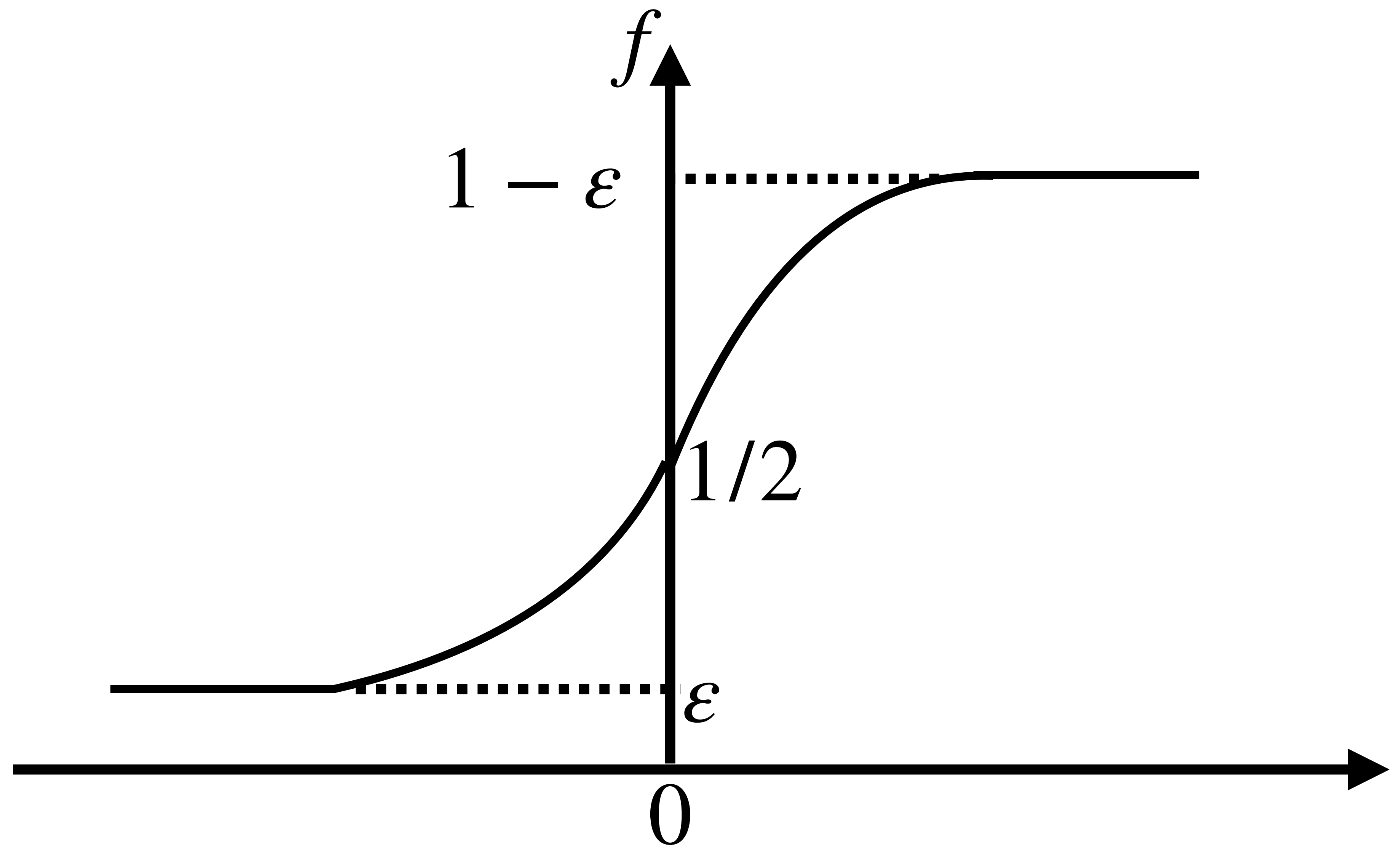
$$Y(0) \sim \mathrm{Ber}(\theta + 0.5),\ Y(1) \sim \mathrm{Ber}(0.5).$$

- Balanced first stage sampling:

Equal probability sampling.

- Sampling algorithm:

Modified Thompson sampling with clipping ε.



Test statistic choices:

Numerical simulation

Data generation details:

- Potential outcome distributions:

$$Y(0) \sim N(\theta, 1),\ Y(1) \sim N(0, 0.25);$$

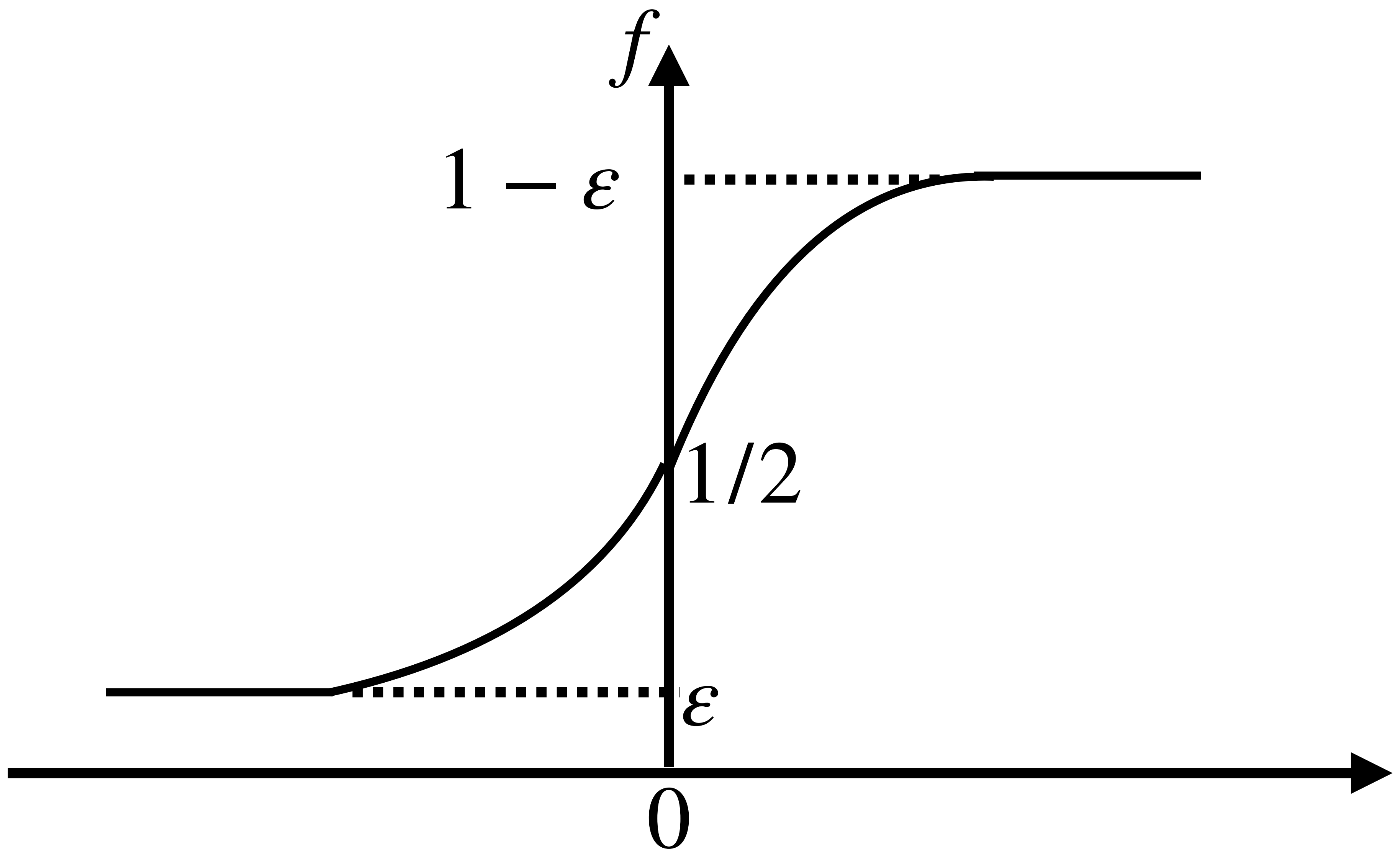
$$Y(0) \sim \mathrm{Ber}(\theta + 0.5),\ Y(1) \sim \mathrm{Ber}(0.5).$$

- Balanced first stage sampling:

Equal probability sampling.

- Sampling algorithm:

Modified Thompson sampling with clipping ε.



Test statistic choices:

1. $\{(\mathrm{WIPW}(0) - \mathrm{WIPW}(1))\} \times \{m = 0,\ m = 1/2\}$.

Numerical simulation

Data generation details:

- Potential outcome distributions:

$$Y(0) \sim N(\theta, 1),\ Y(1) \sim N(0, 0.25);$$

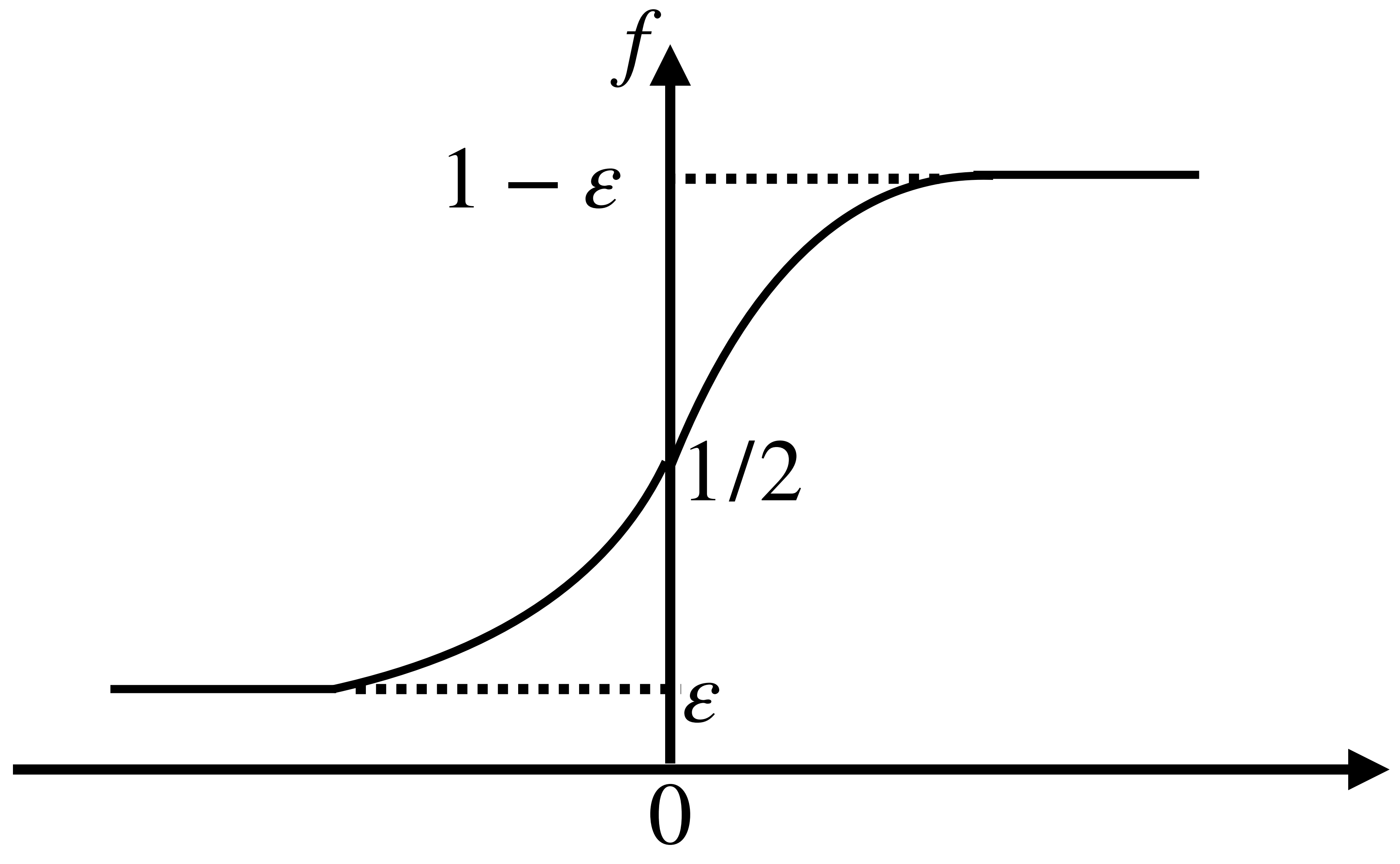
$$Y(0) \sim \mathrm{Ber}(\theta + 0.5),\ Y(1) \sim \mathrm{Ber}(0.5).$$

- Balanced first stage sampling:

  Equal probability sampling.

- Sampling algorithm:

  Modified Thompson sampling with clipping ε.



Test statistic choices:

1. $\{(\mathrm{WIPW}(0) - \mathrm{WIPW}(1))\} \times \{m = 0,\ m = 1/2\}$.
2. $\{(\mathrm{WIPW}(0) - \mathrm{WIPW}(1))/\hat{\sigma}\} \times \{m = 0,\ m = 1/2\}$.

Numerical simulation

Data generation details:

- Potential outcome distributions:

$$Y(0) \sim N(\theta,1),\ Y(1) \sim N(0,0.25);$$

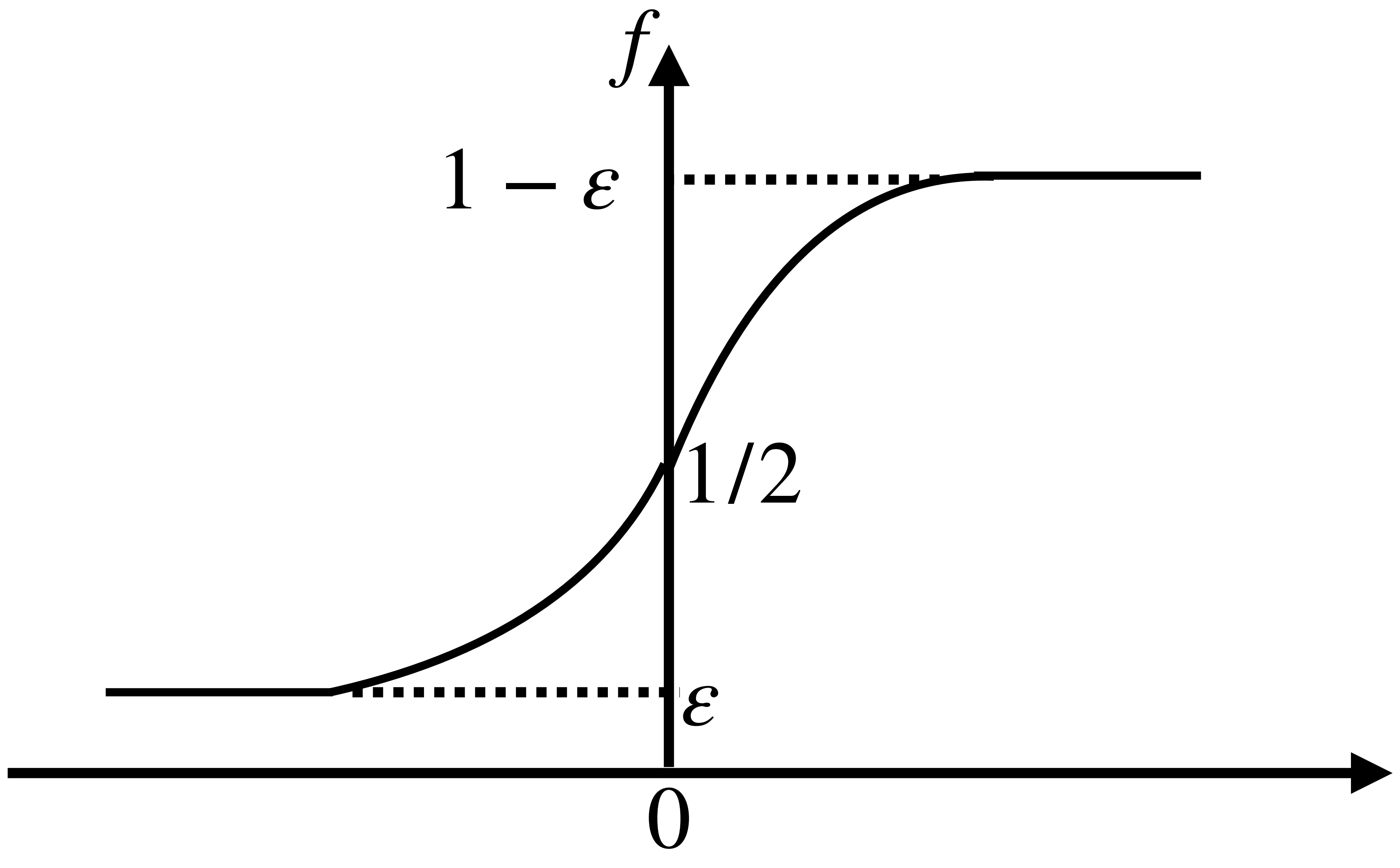
$$Y(0) \sim \mathrm{Ber}(\theta + 0.5),\ Y(1) \sim \mathrm{Ber}(0.5).$$

- Balanced first stage sampling:

Equal probability sampling.

- Sampling algorithm:

Modified Thompson sampling with clipping ε.

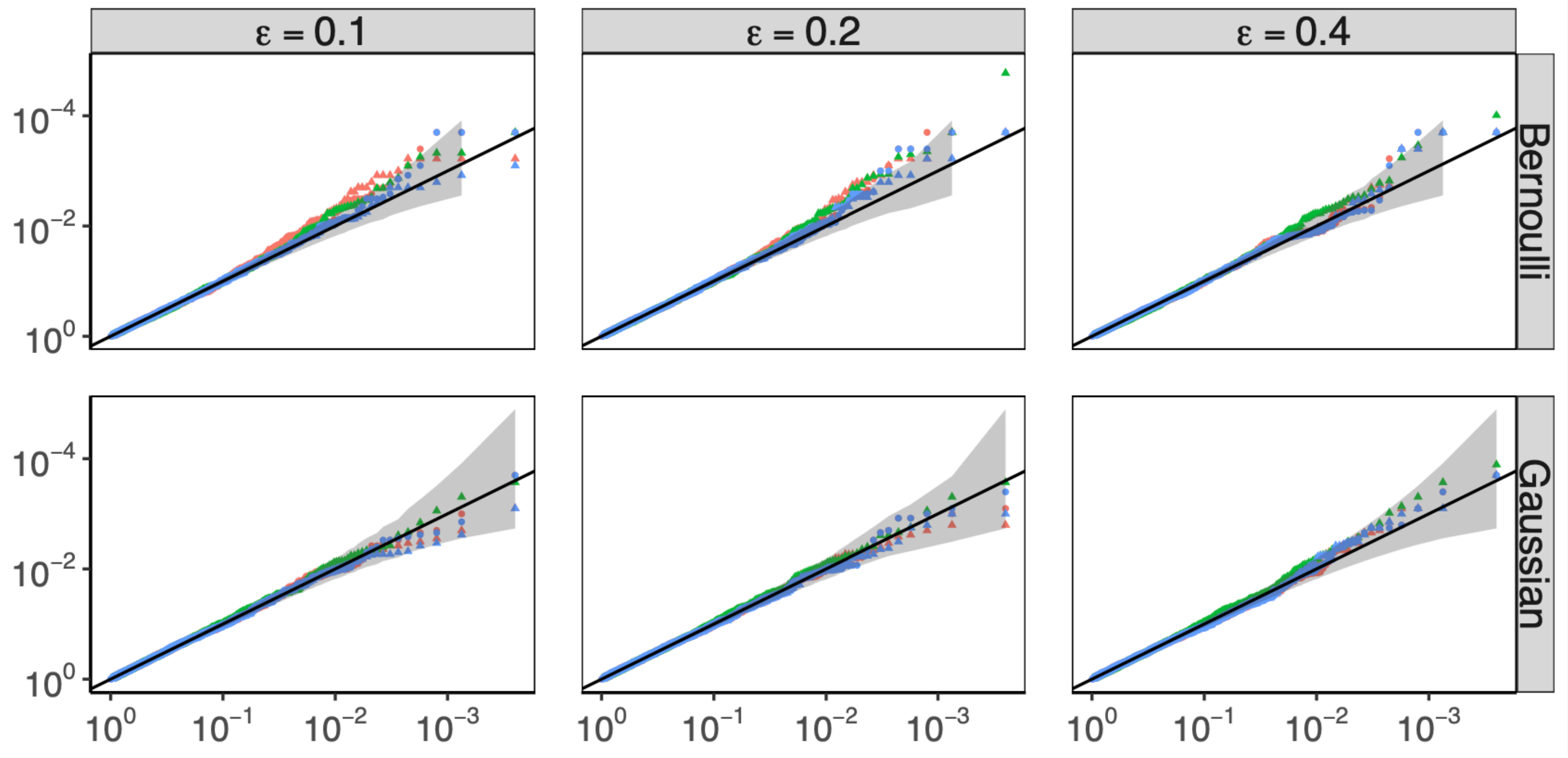


Test statistic choices:

1. $\{(\mathrm{WIPW}(0) - \mathrm{WIPW}(1))\} \times \{m = 0,\ m = 1/2\}$.
2. $\{(\mathrm{WIPW}(0) - \mathrm{WIPW}(1))/\hat{\sigma}\} \times \{m = 0,\ m = 1/2\}$.
3. Sample-splitting with IPW estimator.

Numerical simulation

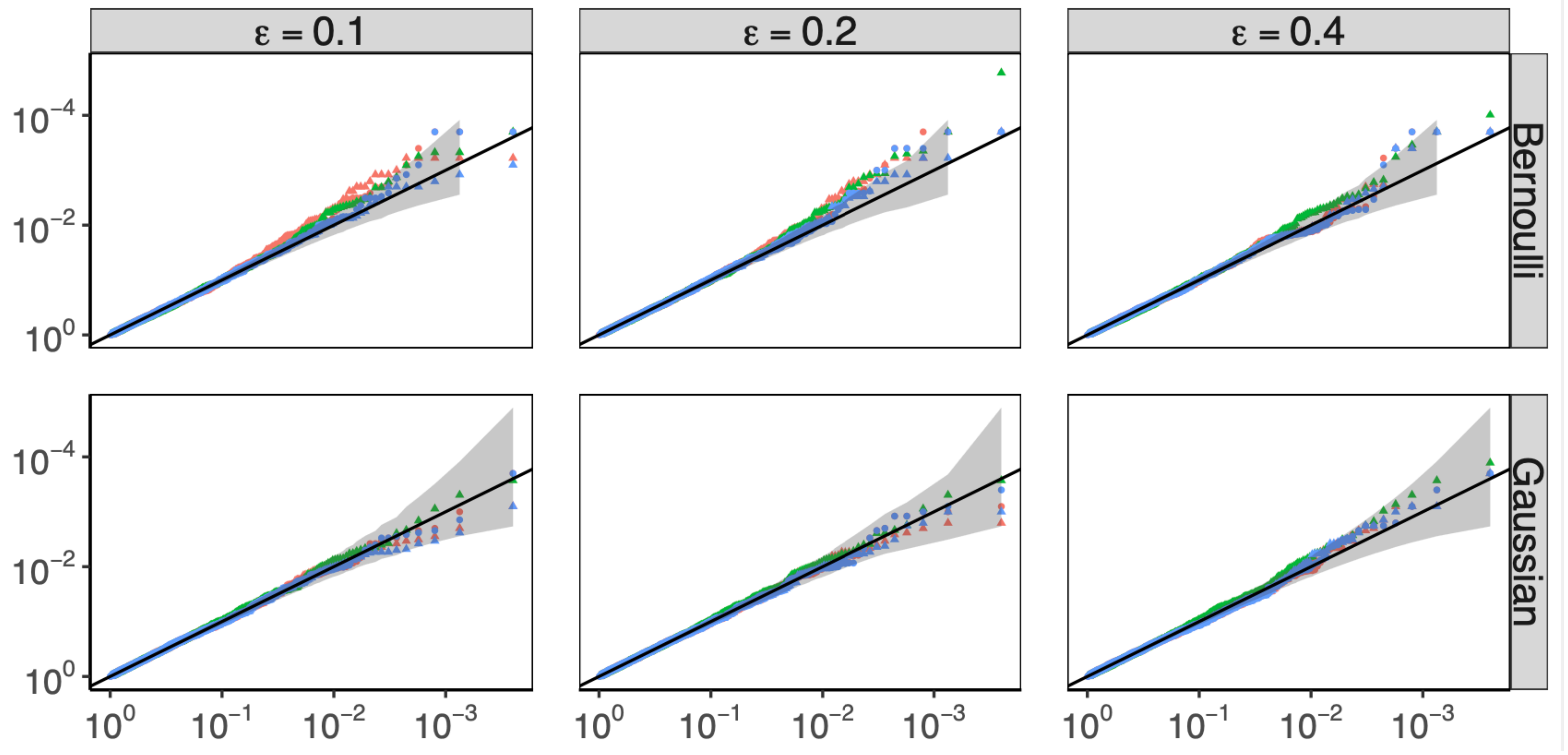
Numerical simulation



$\theta = 0$: null hypothesis

Numerical simulation

Main messages:

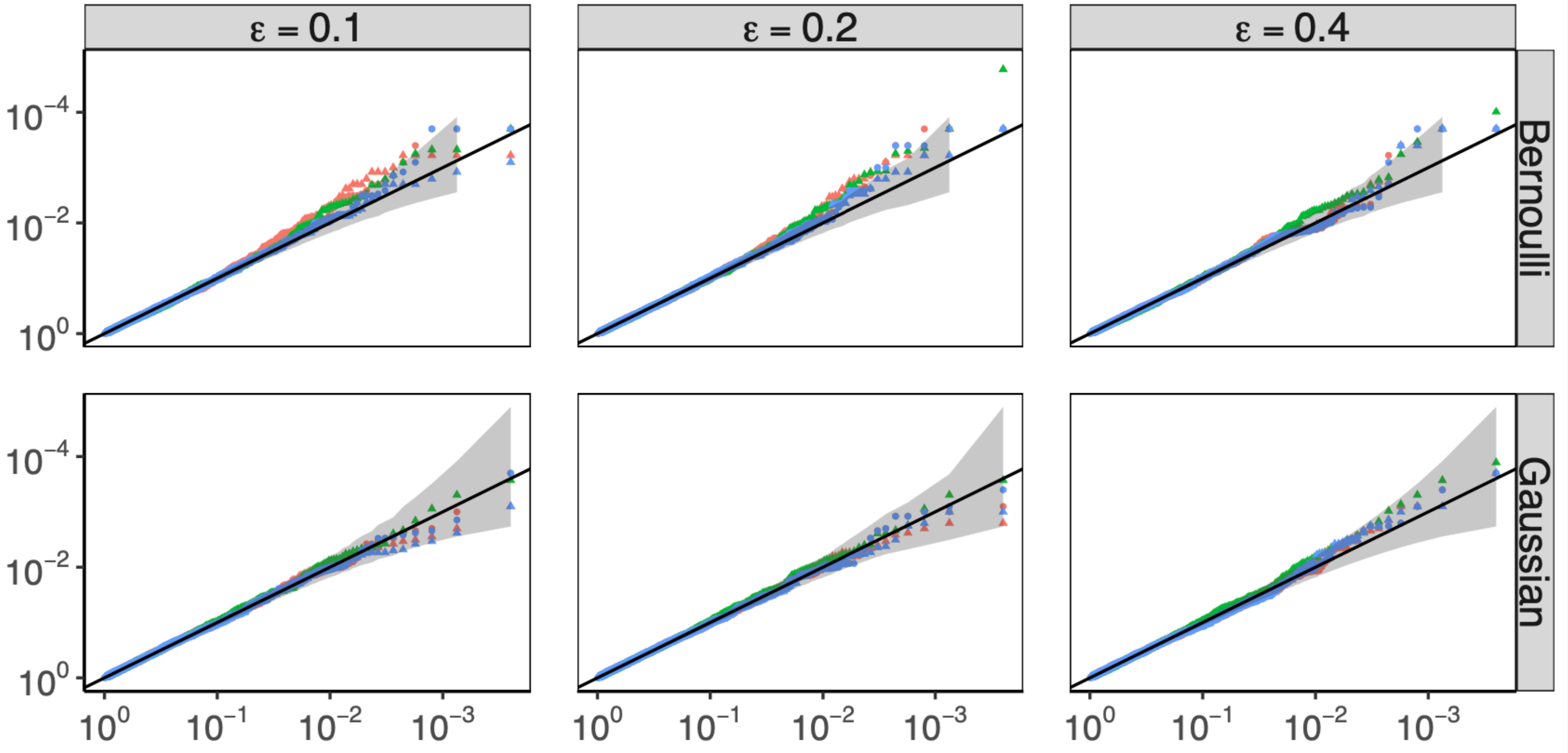


$\theta = 0$: null hypothesis

Numerical simulation

Main messages:

- Type-I error control: all the tests.



$\theta = 0$: null hypothesis

Numerical simulation

Main messages:

- Type-I error control: all the tests.

Numerical simulation

Normalized
Unnormalized
Sample-splitting

Main messages:

- Type-I error control: all the tests.

Numerical simulation

Normalized
Unnormalized
Sample-splitting

$m = 0$ (IPW)
$m = 1/2$ (variance-stablizing IPW)

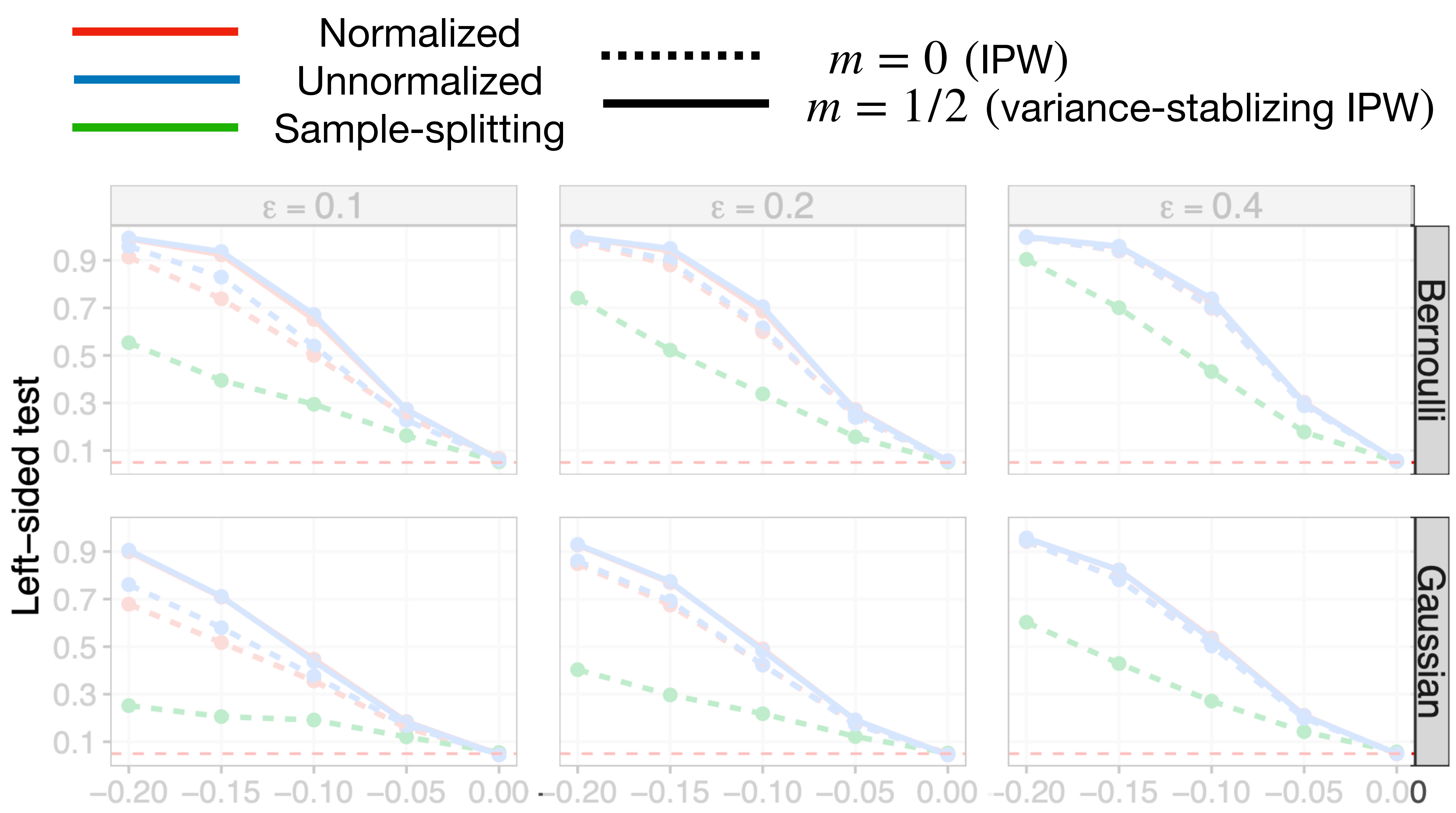
Main messages:

- Type-I error control: all the tests.

Numerical simulation

Main messages:

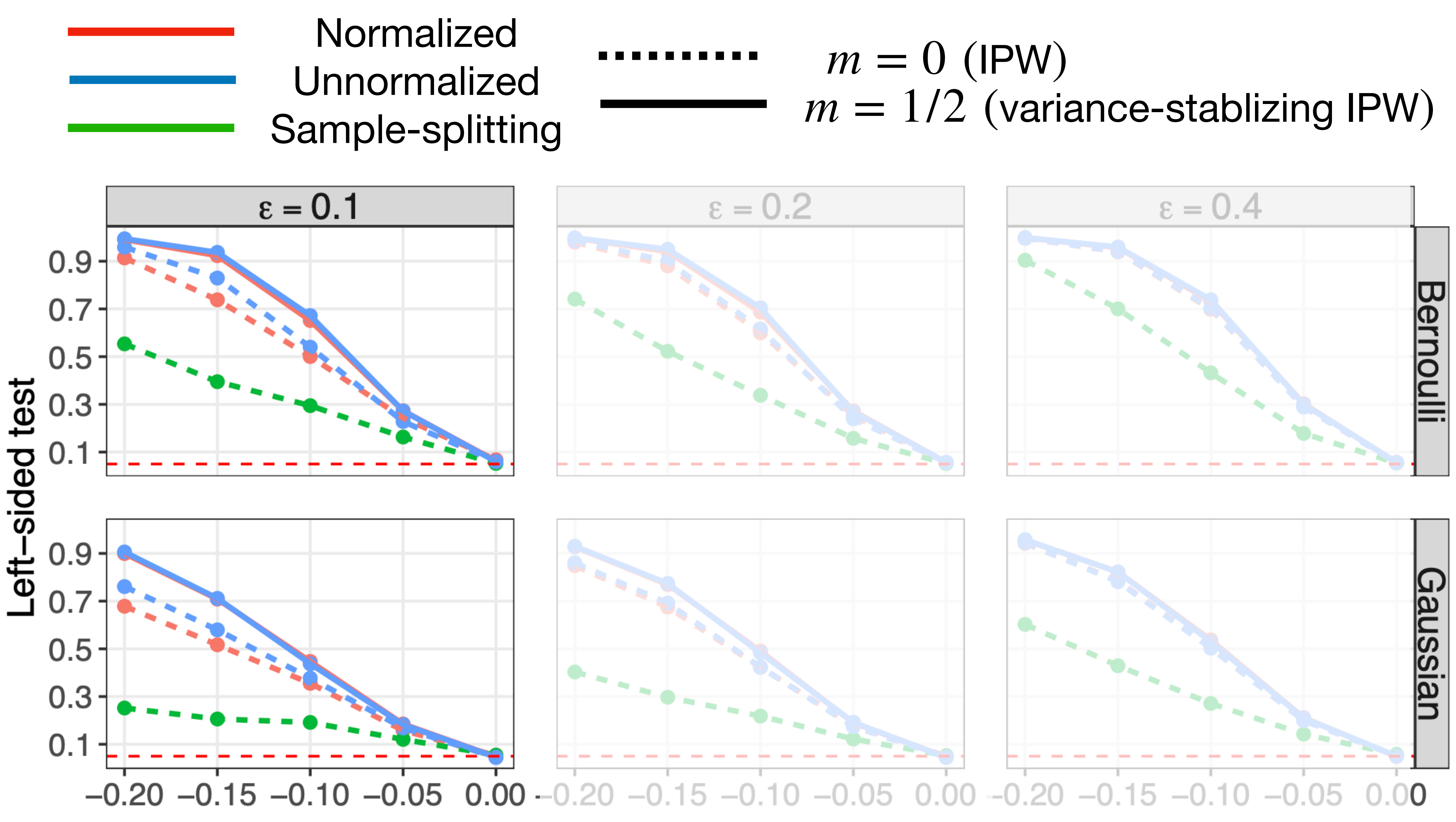
- Type-I error control: all the tests.



Numerical simulation

Main messages:

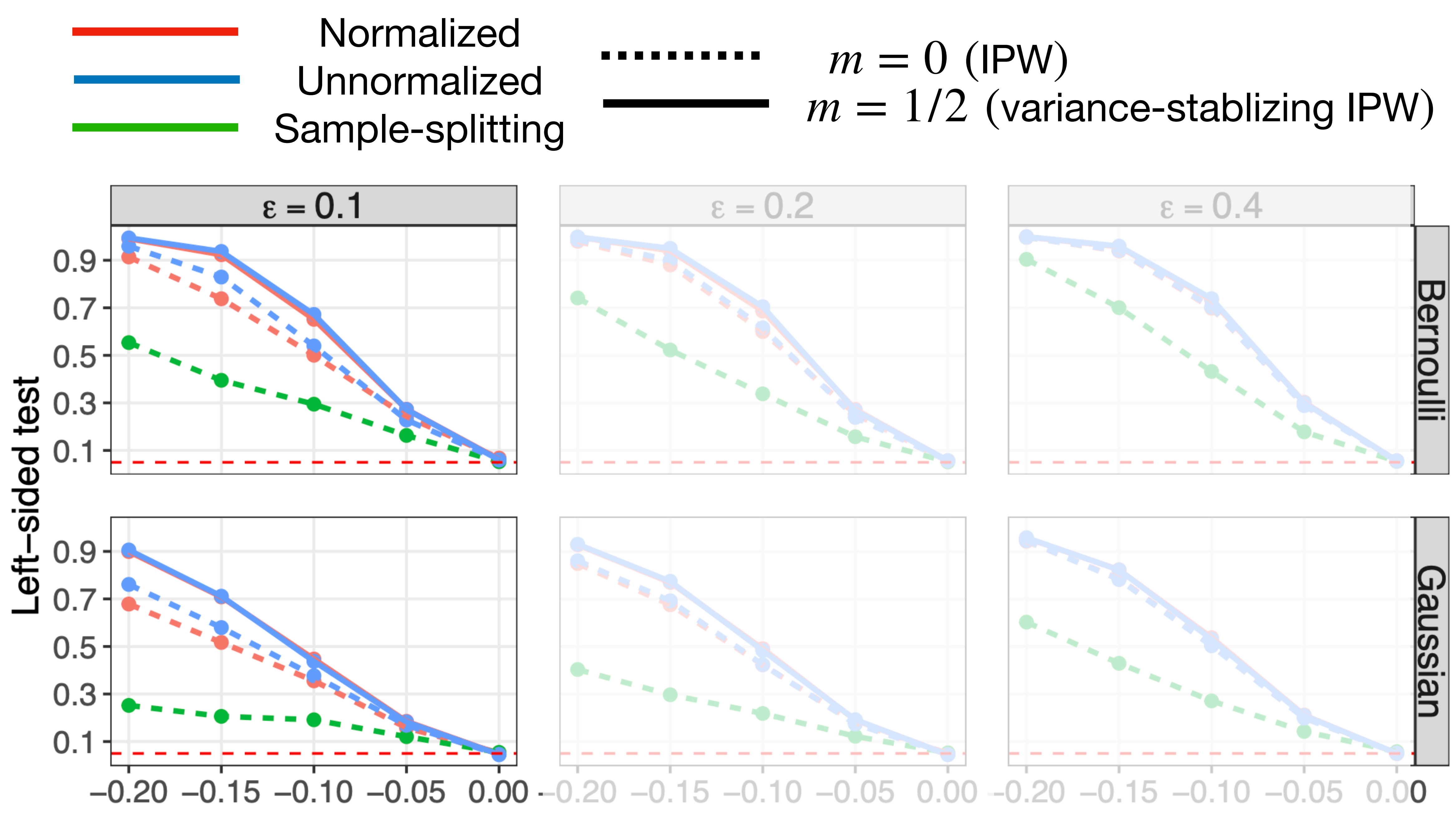
- Type-I error control: all the tests.



Numerical simulation

Main messages:

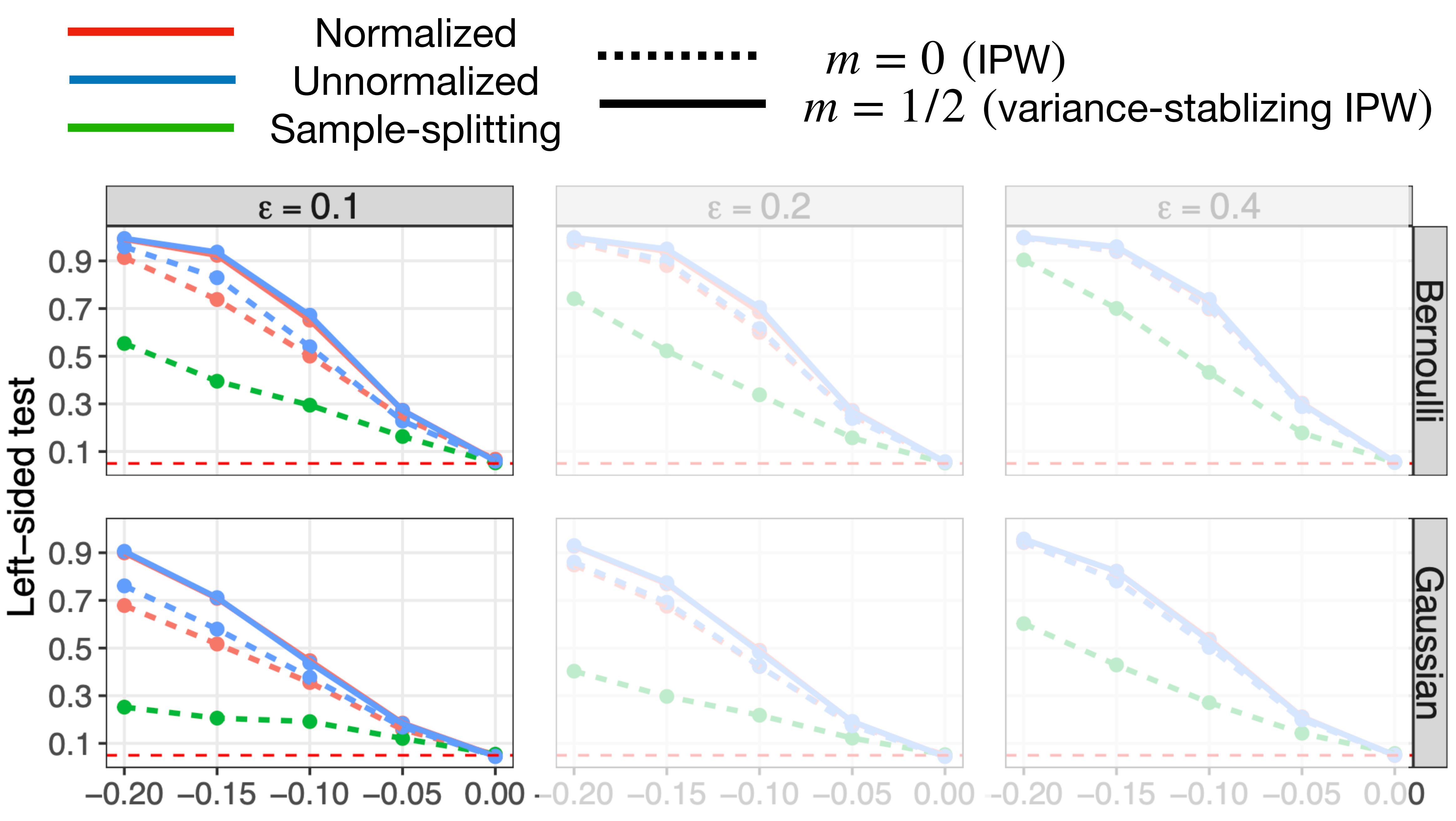
- Type-I error control: all the tests.
- Power: $m = 1/2 > m = 0 >$ sample-splitting.



Numerical simulation

Main messages:

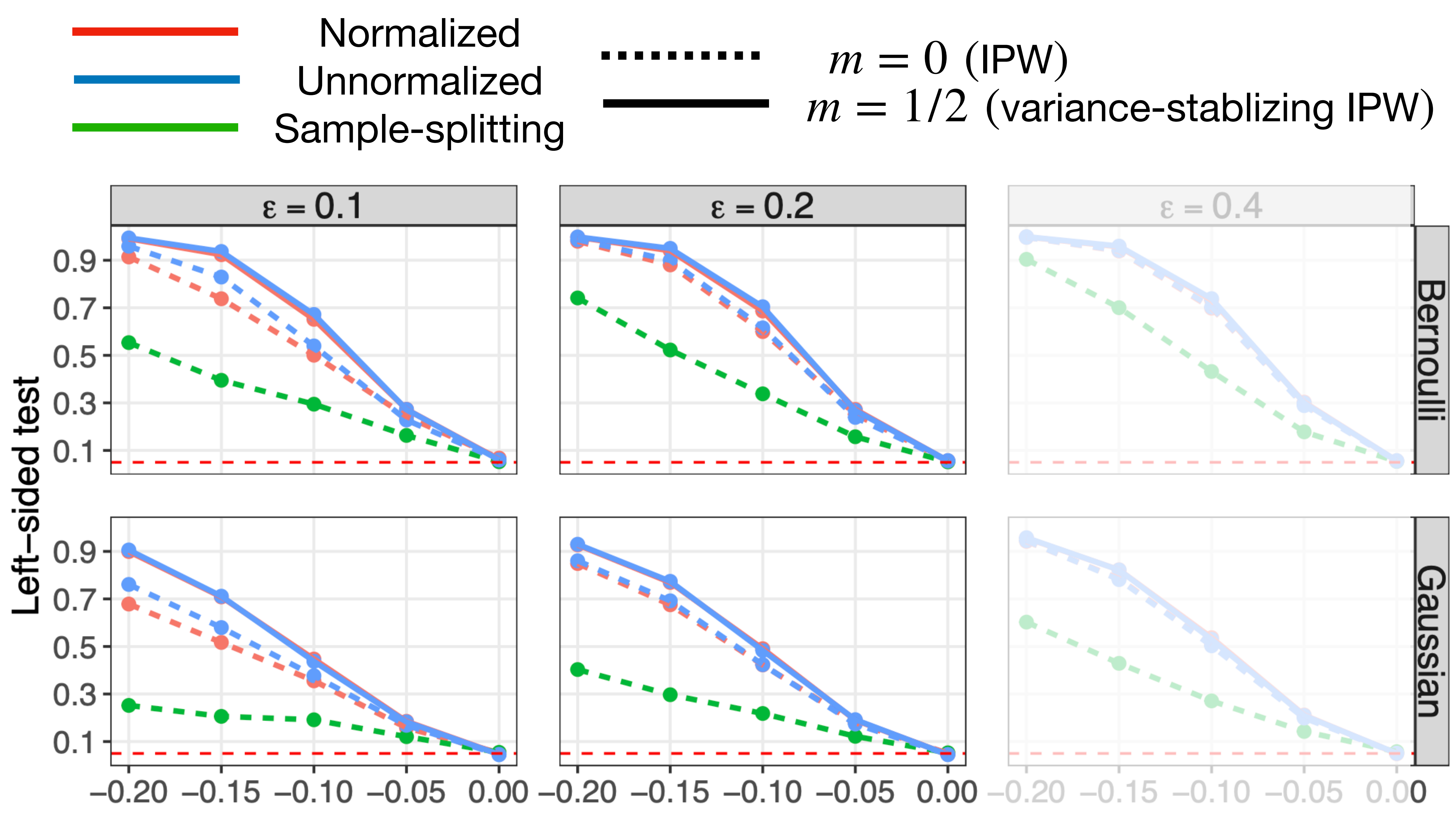
- Type-I error control: all the tests.
- Power: $m = 1/2 > m = 0 >$ sample-splitting.
- Non-equivalence of two normalization schemes.



Numerical simulation

Main messages:

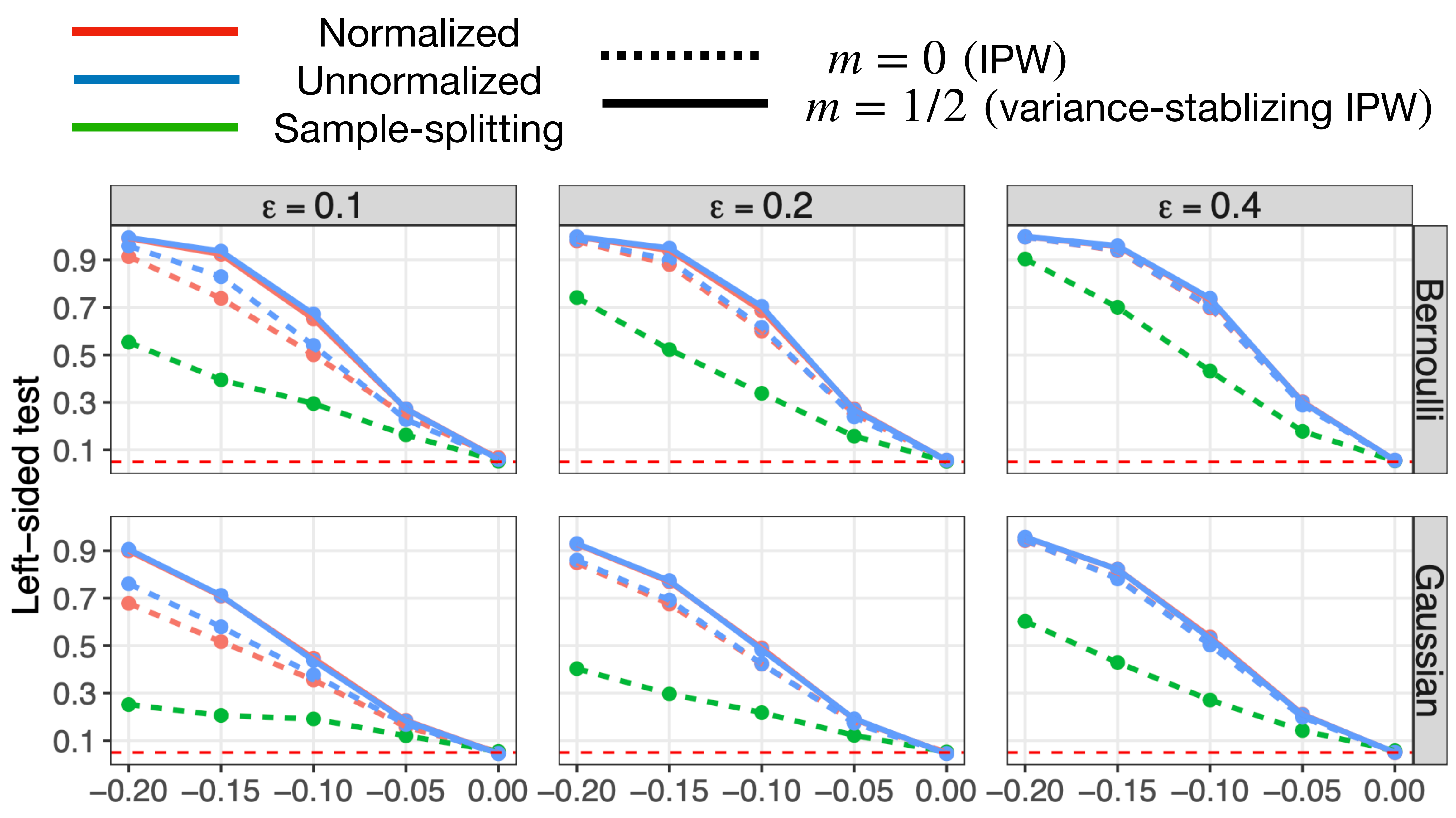
- Type-I error control: all the tests.
- Power: $m = 1/2 > m = 0 >$ sample-splitting.
- Non-equivalence of two normalization schemes.



Numerical simulation

Main messages:

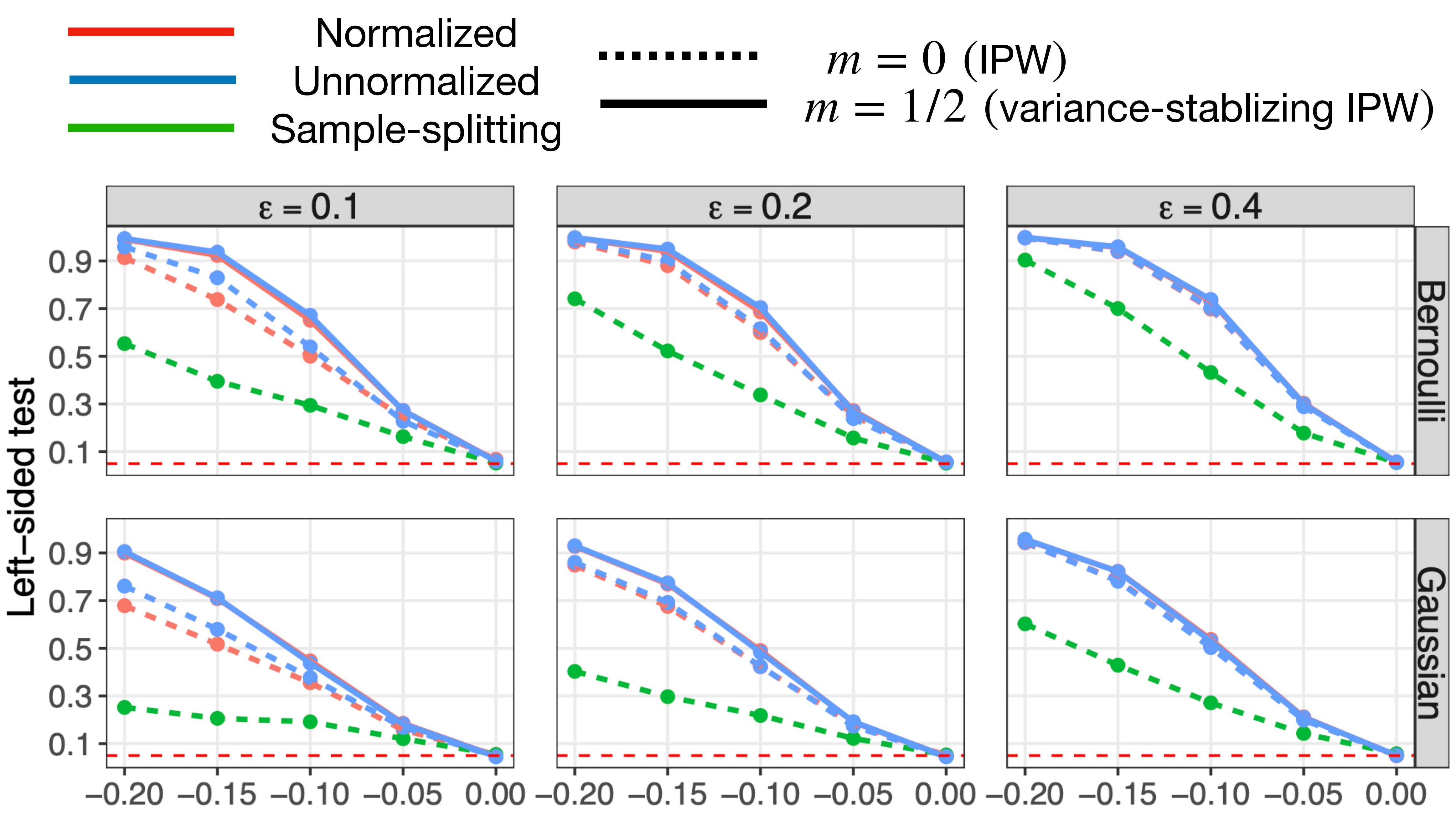
- Type-I error control: all the tests.
- Power: $m = 1/2 > m = 0 >$ sample-splitting.
- Non-equivalence of two normalization schemes.



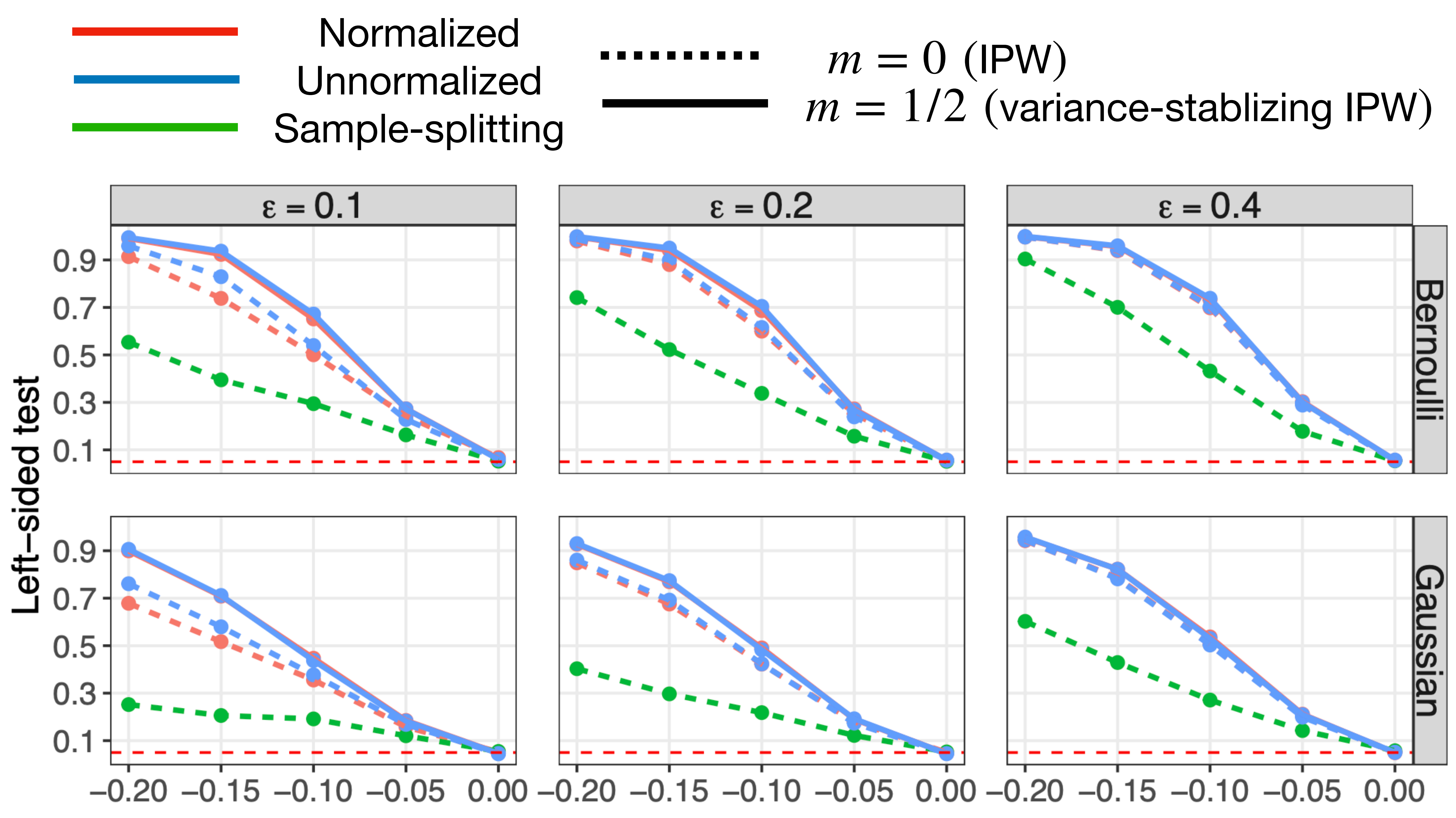
Numerical simulation

Main messages:

- Type-I error control: all the tests.
- Power: $m = 1/2 > m = 0 >$ sample-splitting.
- Non-equivalence of two normalization schemes.
- $m = 1/2$ (solid line) is robust to overlap decay.



Numerical simulation



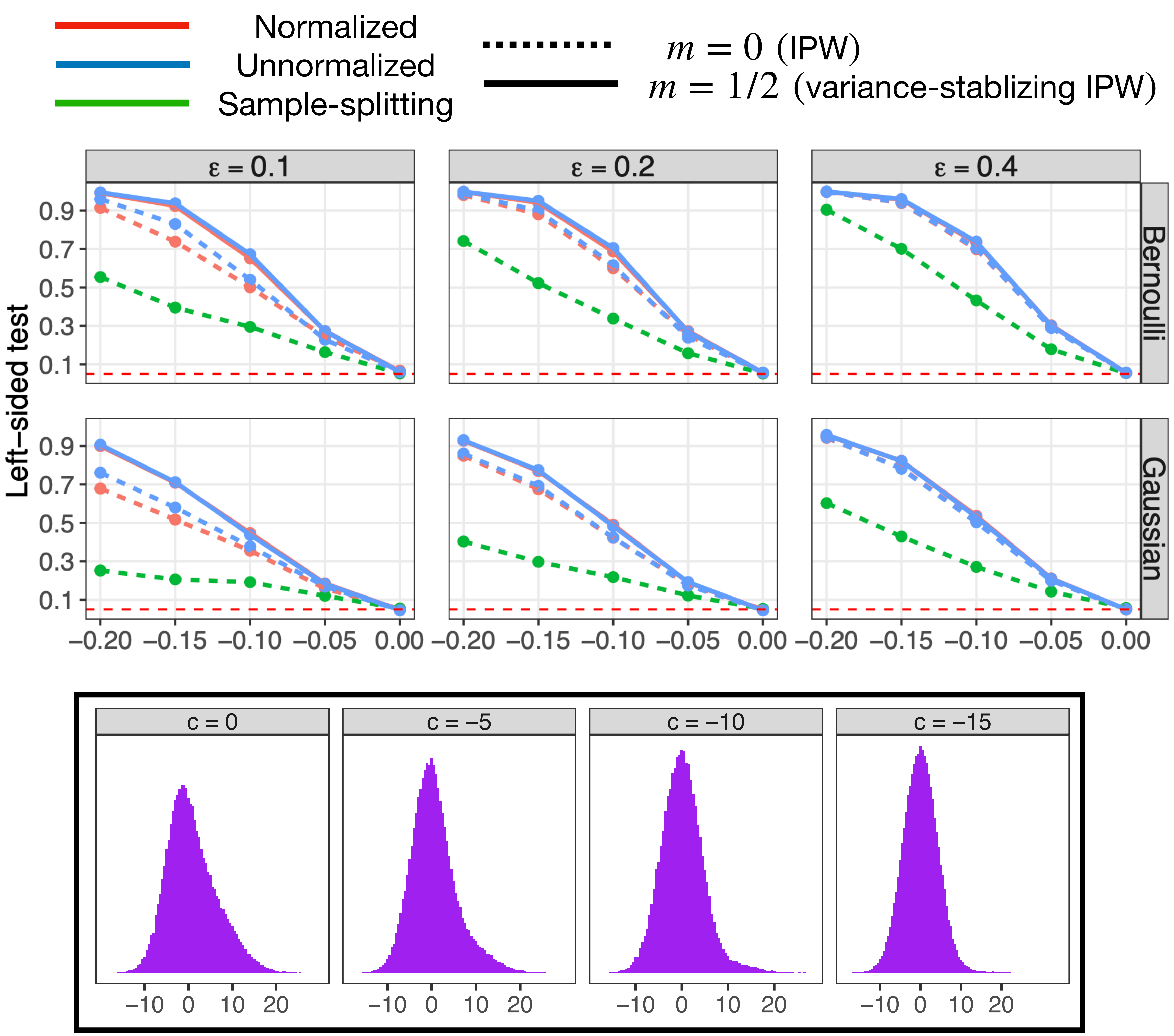
Main messages:

- Type-I error control: all the tests.
- Power: $m = 1/2 > m = 0 >$ sample-splitting.
- Non-equivalence of two normalization schemes.
- $m = 1/2$ (solid line) is robust to overlap decay.
- Left-sided and right-sided tests have asymmetric power curves.

Numerical simulation

Main messages:

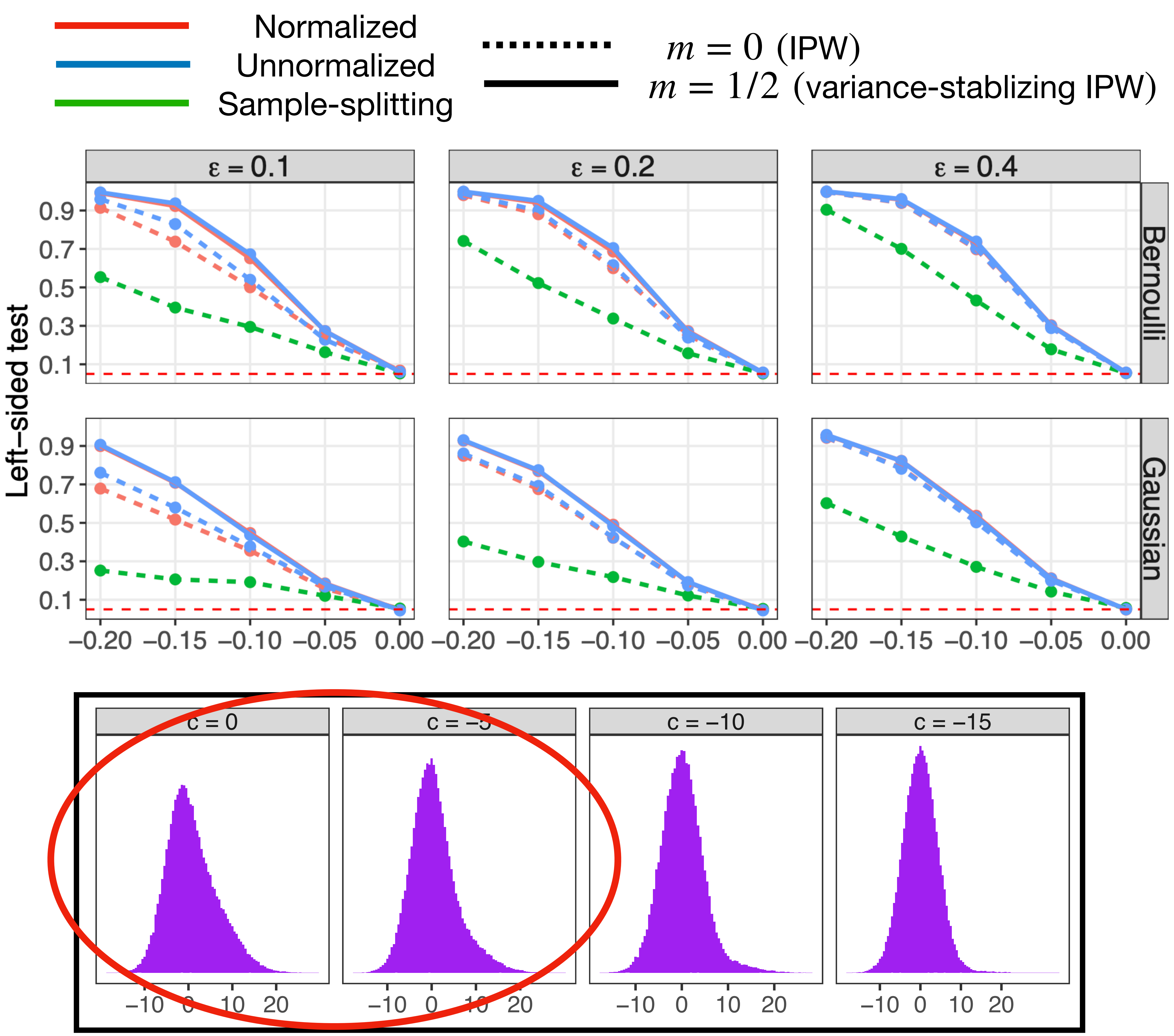
- Type-I error control: all the tests.
- Power: $m = 1/2 > m = 0 >$ sample-splitting.
- Non-equivalence of two normalization schemes.
- $m = 1/2$ (solid line) is robust to overlap decay.
- Left-sided and right-sided tests have asymmetric power curves.



Numerical simulation

Main messages:

- Type-I error control: all the tests.
- Power: $m = 1/2 > m = 0 >$ sample-splitting.
- Non-equivalence of two normalization schemes.
- $m = 1/2$ (solid line) is robust to overlap decay.
- Left-sided and right-sided tests have asymmetric power curves.



Discussion

Discussion

Takeaways

Discussion

Takeaways

- We establish full-spectrum weak limits of WIPW statistics under two-stage response-adaptive experiments under minimal assumptions.

Discussion

Takeaways

- We establish full-spectrum weak limits of WIPW statistics under two-stage response-adaptive experiments under minimal assumptions.
- We propose a fast bootstrap procedure to perform practical hypothesis testing.

Discussion

Takeaways

- We establish full-spectrum weak limits of WIPW statistics under two-stage response-adaptive experiments under minimal assumptions.
- We propose a fast bootstrap procedure to perform practical hypothesis testing.
- We sketch out several extensions on sampling algorithms, involving:

Discussion

Takeaways

- We establish full-spectrum weak limits of WIPW statistics under two-stage response-adaptive experiments under minimal assumptions.
- We propose a fast bootstrap procedure to perform practical hypothesis testing.
- We sketch out several extensions on sampling algorithms, involving:
  - Stopping time.

Discussion

Takeaways

- We establish full-spectrum weak limits of WIPW statistics under two-stage response-adaptive experiments under minimal assumptions.
- We propose a fast bootstrap procedure to perform practical hypothesis testing.
- We sketch out several extensions on sampling algorithms, involving:
  - Stopping time.
  - Early dropping.

Discussion

Takeaways

- We establish full-spectrum weak limits of WIPW statistics under two-stage response-adaptive experiments under minimal assumptions.
- We propose a fast bootstrap procedure to perform practical hypothesis testing.
- We sketch out several extensions on sampling algorithms, involving:
  - Stopping time.
  - Early dropping.

Limitations and future work

Discussion

Takeaways

- We establish full-spectrum weak limits of WIPW statistics under two-stage response-adaptive experiments under minimal assumptions.
- We propose a fast bootstrap procedure to perform practical hypothesis testing.
- We sketch out several extensions on sampling algorithms, involving:
  - Stopping time.
  - Early dropping.

Limitations and future work

- Help plan experiment to maximize statistical power.

Discussion

Takeaways

- We establish full-spectrum weak limits of WIPW statistics under two-stage response-adaptive experiments under minimal assumptions.
- We propose a fast bootstrap procedure to perform practical hypothesis testing.
- We sketch out several extensions on sampling algorithms, involving:
  - Stopping time.
  - Early dropping.

Limitations and future work

- Help plan experiment to maximize statistical power.
- Extend beyond two-stage experiments: fully adaptive sampling.

Discussion

Takeaways

- We establish full-spectrum weak limits of WIPW statistics under two-stage response-adaptive experiments under minimal assumptions.
- We propose a fast bootstrap procedure to perform practical hypothesis testing.
- We sketch out several extensions on sampling algorithms, involving:
  - Stopping time.
  - Early dropping.

Limitations and future work

- Help plan experiment to maximize statistical power.
- Extend beyond two-stage experiments: fully adaptive sampling.
- Explore the optimality within the class of WAIPW test statistics.

Reference

Assumption-lean weak limits and tests for two-stage adaptive experiments. In arXiv, 2025.