Analyzing the sensitivity of observational findings using modern prediction methods

Hongseok Namkoong

Decision, Risk, and Operations Division
Columbia Business School
namkoong@gsb.columbia.edu

Bounds on the conditional and average treatment effect with unobserved confounding factors.
Yadlowsky, N., Basu, Duchi, and Tian. Minor revision in Annals of Statistics.
Off-policy policy evaluation for sequential decisions under unobserved confounding.
N., Keramati, Yadlowsky, and Brunskill. NeurIPS 2020.

Causality and decision-making

- Causal understanding is crucial for reliable decision-making

How effective is my ad?

How effective is my new drug?

- Counterfactuals: what would've happened if the person didn't see the ad, or didn't get the drug?
- Today: Use modern prediction models to infer causal effects

Secret to life

Secret to life

The New York Times

Another Benefit to Going to Museums? You May Live Longer

Researchers in Britain found that people who go to museums, the theater and the opera were less likely to die in the study period than those who didn't.

Heavy selection bias based on unobservables (wealth): decision / treatment is intimately connected with health outcomes

The potential outcomes framework

- A feature vector $X \in \mathbb{R}^k$
- Potential outcomes: $Y(1), Y(0)$
- A treatment assignment $Z \in \{0,1\}$
- **Observe $Y := Y(Z)$, never $Y(1-Z)$**

Average Treatment Effect (ATE)

$$ATE = \mathbb{E}[Y(1) - Y(0)]$$

$$= \mathbb{E}_{X \sim P_X}\left[\mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X]\right]$$

$$= \mathbb{E}_{X \sim P_X}\left[\mu_1^\star(X) - \mu_0^\star(X)\right] =: \mathbb{E}_{X \sim P_X}\left[\mu^\star(X)\right]$$

Conditional Average Treatment Effect

- We observe the tuple (X, Y, Z)

Randomized control trials

also called A/B testing, (randomized) experiments

- First try: let's **randomize** treatment assignments

$$Y(1), Y(0) \perp Z$$

- By virtue of randomized assignments, we have

$$\tau = \mathbb{E}[Y(1) - Y(0)]$$

$$= \mathbb{E}[Y(1) \mid Z = 1] - \mathbb{E}[Y(0) \mid Z = 0]$$ ← **observable**

- We can estimate final line from i.i.d. data (Y_i, Z_i)

Recap

- Treatment assignment: Z
- Potential outcome: $Y(0)$, $Y(1)$
- Response $Y := Y(Z)$

Observational studies

- When experimentation is costly, crucial to leverage collected data
- Historically, many important findings from observational data
  - "citrus fruit curing scurvy described in the 1700s or insulin as a treatment for diabetes in the 1920s long preceded the advent of the modern randomized clinical trial."
  - "these methods had in common a reliable method of diagnosis, a predictable clinical course, and a large and obvious effect of the treatment." [Corrigan-Curay et al. 2018]
- These results need to be contextualized and viewed with more skepticism than RCTs

Observational studies

Recap

- Treatment assignment: Z
- Potential outcome: $Y(0),\ Y(1)$
- Response $Y := Y(Z)$

Assumption: $Y(1), Y(0) \perp Z \mid X$ **no unobserved confounding**

- Observed treatment assignments are based on covariate information alone; treatment assignment does not use information about counterfactuals
- Inverse prob. weighting: If you can predict treatment assignments well, estimate

$$\mathbb{E}[Y(1)] = \mathbb{E}\left[\frac{Z}{\mathbb{P}(Z=1 \mid X)} Y\right]$$ using an estimate $\widehat{\mathbb{P}}(Z=1 \mid X)$

- Direct method: If you can predict outcome well, estimate

$\mu_1^\star(X) := \mathbb{E}[Y(1) \mid X] = \mathbb{E}[Y(1) \mid X, Z=1]$, and use $\mathbb{E}[Y(1)] = \mathbb{E}[\mu_1^\star(X)] \approx \frac{1}{n}\sum_{i=1}^{n} \widehat{\mu}_1(X_i)$

Augmented IPW

Recap

- Covariates: X
- Treatment assignment: Z
- Potential outcome: $Y(0),\ Y(1)$
- Response $Y := Y(Z)$

Assumption: $Y(1), Y(0) \perp Z \mid X$

- If you can do either one well, estimate

$$\mathbb{E}[Y(1)] = \mathbb{E}\left[\mu_1^\star(X) + \frac{Z}{\mathbb{P}(Z=1 \mid X)}(Y - \mu_1^\star(X))\right]$$

[Augmented IPW]

- Consistent if you can estimate **either** nuisance parameters consistently: outcomes $\mu_1^\star(X) := \mathbb{E}[Y(1) \mid X]$ or propensity score $e^\star(X) := \mathbb{P}(Z=1 \mid X)$
- Insensitive to errors in nuisance estimates; ***scalable estimator for ATE***

Unobserved confounders

- But there is almost always an unobserved confounders that simultaneously affect potential outcomes and treatment assignments

Judges are more lenient after taking a break, study finds theguardian [Danziger '11]

Overlooked factors in the analysis of parole decisions [Weinshall-Margel '11]

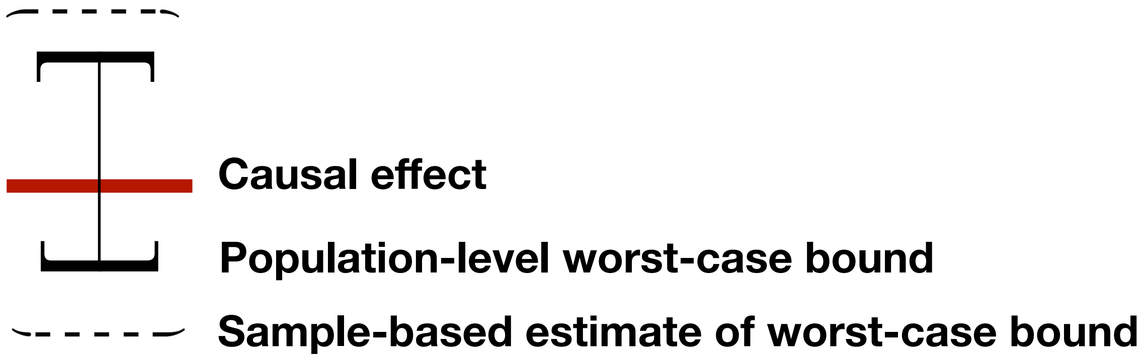
Other examples: Antioxidant vitamin beta carotene [Willett '90, ATBC CPSG '94]

Hormone replacement therapy [Pedersen '03 WHI, Lawlor '04, Rutter '07]

- Visual observations used in clinical decisions (e.g. admission to NICU); drugs preferentially prescribed those who can tolerate them
  - Not properly recorded even at the resolution of large databases

Sensitivity analysis

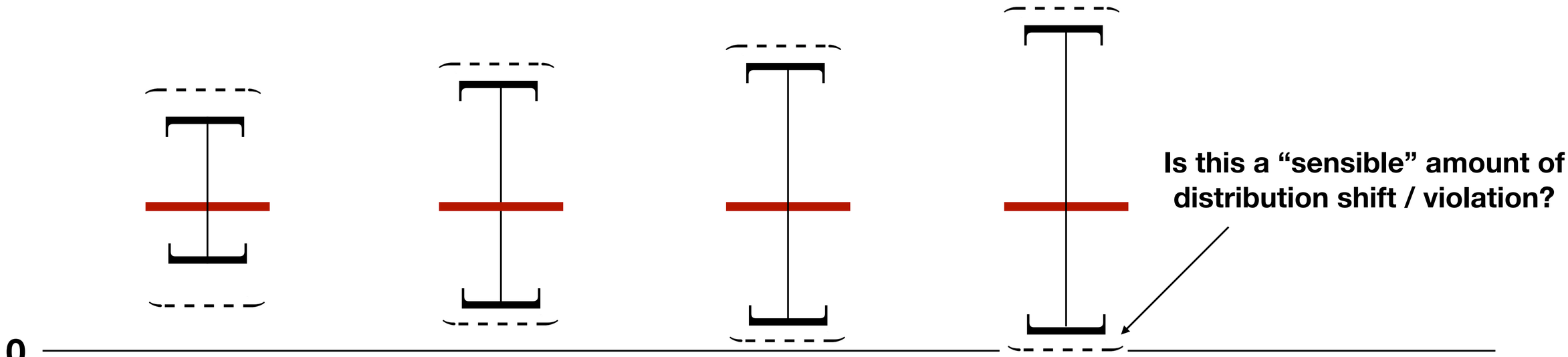
- Analyze sensitivity of CATE and ATE estimates to unobserved confounding
- Sensitivity of a finding: magnitude of violation when endpoint crosses a threshold
- Today: Worst-case bounds on the Doubly Robust / AIPW estimator



0

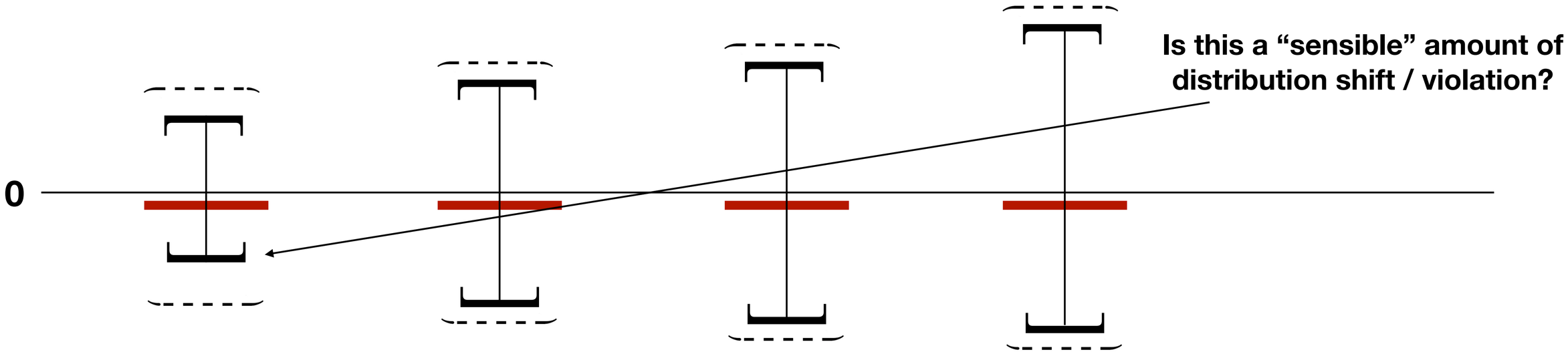
Sensitivity analysis

- Analyze sensitivity of CATE and ATE estimates to unobserved confounding
- Sensitivity of a finding: magnitude of violation when endpoint crosses a threshold
- Today: Worst-case bounds on the Doubly Robust / AIPW estimator



Sensitivity analysis

- Analyze sensitivity of CATE and ATE estimates to unobserved confounding
- Sensitivity of a finding: magnitude of violation when endpoint crosses a threshold
- Today: Worst-case bounds on the Doubly Robust / AIPW estimator



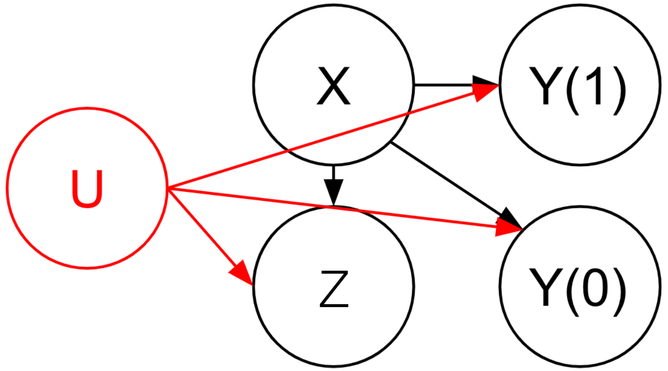
Overview

- First, we analyze the sensitivity of *personalized* treatment effects (CATE)
- Semiparametric methods for worst-case bounds on the ATE
  - ATE is one-dimensional, but there are infinite-dimensional nuisance parameters
  - Use ML methods to estimate nuisance parameters
- Central limit results even when nuisance parameters converge slower

Bounded unobserved confounding

- What if there's a hidden variable U that wasn't observed?

e.g. partial knowledge of counterfactuals



Recap

- Covariates: X
- Treatment assignment: Z
- Potential outcome: $Y(0),\ Y(1)$
- Response $Y := Y(Z)$

Relaxed assumption: Bounded unobserved confounding

There exists $\Gamma > 1$, and U such that $Y(1), Y(0) \perp Z \mid X, U,$

$$u \mapsto \frac{\mathbb{P}(Z = 1 \mid X, U = u)}{\mathbb{P}(Z = 0 \mid X, U = u)}$$ can vary by at most a factor of Γ [Rosenbaum '02]

Bounded unobserved confounding

Relaxed assumption: Bounded unobserved confounding

There exists $\Gamma > 1$, and U such that $Y(1), Y(0) \perp Z \mid X, U,$

$$u \mapsto \frac{\mathbb{P}(Z = 1 \mid X, U = u)}{\mathbb{P}(Z = 0 \mid X, U = u)}$$ can vary by at most a factor of Γ [Rosenbaum '02]

- Equivalent to a logit model: for some function $\kappa(\cdot) \in [0,1]$, $g(\cdot)$,

$$\log \frac{\mathbb{P}(Z = 1 \mid X, U)}{\mathbb{P}(Z = 0 \mid X, U)} = g(X) + \log \Gamma \cdot \kappa(U)$$

FAQs

Relaxed assumption: Bounded unobserved confounding

There exists $\Gamma > 1$, and U such that $Y(1), Y(0) \perp Z \mid X, U,$

$$u \mapsto \frac{\mathbb{P}(Z = 1 \mid X, U = u)}{\mathbb{P}(Z = 0 \mid X, U = u)}$$ can vary by at most a factor of Γ [Rosenbaum '02]

- How do I choose Γ?

➡ Domain expertise (e.g. clinical intuition)

➡ Sensitivity: what would be a clinically significant result? what value of Γ would change its significance?

- Is this the only natural confounding model?

➡ No. We develop modern semiparametric methods under this model, but this framework can be applied to other models.

Literature review

- Extensive literature on sensitivity analysis of matching methods [Rosenbaum ’02a, ‘02b, ’10, 11, 14]
  - Cannot scale to even moderate covariate dimensions
- Parametric sensitivity frameworks in program evaluation (restrictive) [Imbens ’03]
- Some recent works on sensitivity analysis for the IPW under non-standard models of confounding (less interpretable) [Shen et al. ’11, Zhao et al. ’17, Kallus & Zhou ’18, Kallus et al. ‘19]
- Today: **Scalable** sensitivity analysis for the AIPW
  - ML-based approach to analyzing the sensitivity of CATE
  - Modern semiparametric framework

Lower bound for $\mathbb{E}[Y(1) \mid X]$

Recap

- Treatment assignment: Z
- Potential outcome: $Y(0)$, $Y(1)$
- Response $Y := Y(Z)$

• Lower bound unobservables under Γ-bounded unobserved confounding

$$\mathbb{E}[Y(1) \mid X] = \mathbb{P}(Z = 1 \mid X)\mathbb{E}[Y(1) \mid X, Z = 1] + \mathbb{P}(Z = 0 \mid X)\mathbb{E}[Y(1) \mid X, Z = 0]$$

$$= \mathbb{P}(Z = 1 \mid X)\mathbb{E}[Y \mid X, Z = 1] + \mathbb{P}(Z = 0 \mid X)\mathbb{E}[YL_1(Y \mid X) \mid X, Z = 1]$$

where $L_1(\cdot \mid X) := \dfrac{dP(Y(1) \in \cdot \mid X, Z = 1)}{dP(Y(1) \in \cdot \mid X, Z = 0)}$

Lower bound for $\mathbb{E}[Y(1) \mid X]$

unobservable

- Lower bound unobservables under Γ-bounded unobserved confounding

$$\mathbb{E}[Y(1) \mid X] = \mathbb{P}(Z = 1 \mid X)\mathbb{E}[Y(1) \mid X, Z = 1] + \mathbb{P}(Z = 0 \mid X)\mathbb{E}[Y(1) \mid X, Z = 0]$$

$$= \mathbb{P}(Z = 1 \mid X)\mathbb{E}[Y \mid X, Z = 1] + \mathbb{P}(Z = 0 \mid X)\mathbb{E}[YL_1(Y \mid X) \mid X, Z = 1]$$

where $L_1(\cdot \mid X) := \dfrac{dP(Y(1) \in \cdot \mid X, Z = 1)}{dP(Y(1) \in \cdot \mid X, Z = 0)}$

Lower bound for $\mathbb{E}[Y(1) \mid X]$

unobservable

- Lower bound unobservables under Γ-bounded unobserved confounding

$$\mathbb{E}[Y(1) \mid X] = \mathbb{P}(Z = 1 \mid X)\mathbb{E}[Y(1) \mid X, Z = 1] + \mathbb{P}(Z = 0 \mid X)\mathbb{E}[Y(1) \mid X, Z = 0]$$

$$= \mathbb{P}(Z = 1 \mid X)\mathbb{E}[Y \mid X, Z = 1] + \mathbb{P}(Z = 0 \mid X)\mathbb{E}[YL_1(Y \mid X) \mid X, Z = 1]$$

where $L_1(\cdot \mid X) := \dfrac{dP(Y(1) \in \cdot \mid X, Z = 1)}{dP(Y(1) \in \cdot \mid X, Z = 0)}$

observable

Lower bound for $\mathbb{E}[Y(1) \mid X]$

unobservable

- Lower bound unobservables under Γ-bounded unobserved confounding

$$\mathbb{E}[Y(1) \mid X] = \mathbb{P}(Z=1 \mid X)\mathbb{E}[Y(1) \mid X, Z=1] + \mathbb{P}(Z=0 \mid X)\mathbb{E}[Y(1) \mid X, Z=0]$$

$$= \mathbb{P}(Z=1 \mid X)\mathbb{E}[Y \mid X, Z=1] + \mathbb{P}(Z=0 \mid X)\mathbb{E}[YL_1(Y \mid X) \mid X, Z=1]$$

where $L_1(\cdot \mid X) := \dfrac{dP(Y(1) \in \cdot \mid X, Z=1)}{dP(Y(1) \in \cdot \mid X, Z=0)}$

observable

Lemma (YNDBT'18) Under Γ-bounded unobserved confounding, $\dfrac{L_1(y \mid x)}{L_1(y' \mid x)} \leq \Gamma$ a.s.

- Letting $\mathscr{L}_1$ be the set of likelihood ratios s.t. above holds, we have

$$\mathbb{E}[Y(1) \mid X, Z=0] \geq \inf_{L \in \mathscr{L}_1} \mathbb{E}[YL(Y \mid X) \mid X, Z=1] =: \theta_1^\star(X)$$

Bound is tight

Recap

- Treatment assignment: Z
- Potential outcome: $Y(0),\ Y(1)$
- Response $Y := Y(Z)$

Convex Duality

Lemma (YNDBT'18) Under Γ-bounded unobserved confounding, $\dfrac{L_1(y \mid x)}{L_1(y' \mid x)} \leq \Gamma$ a.s.

- Letting $\mathscr{L}_1$ be the set of likelihood ratios s.t. above holds, we have

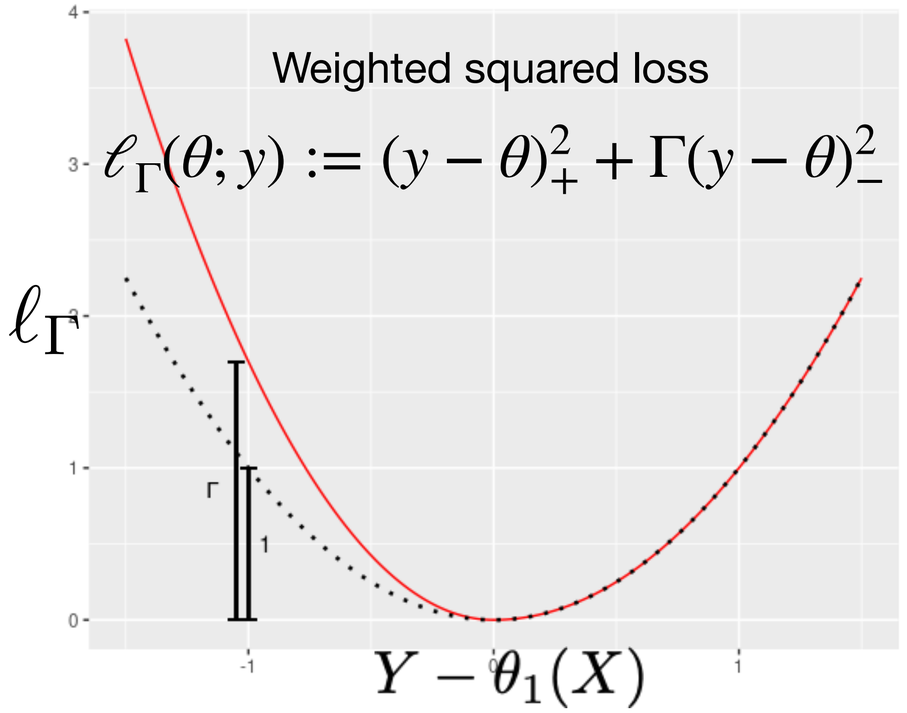
$$\mathbb{E}[Y(1) \mid X, Z = 0] \geq \inf_{L \in \mathscr{L}_1} \mathbb{E}[YL(Y \mid X) \mid X, Z = 1] =: \theta_1^\star(X)$$

- One-dimensional dual for each X

Lemma (YNDBT'18) $\theta_1^\star(X) = \sup\left\{\mu : \mathbb{E}[(Y(1) - \mu)_+ - \Gamma(Y(1) - \mu)_- \mid X, Z = 1] \geq 0\right\}$

What can ML do?

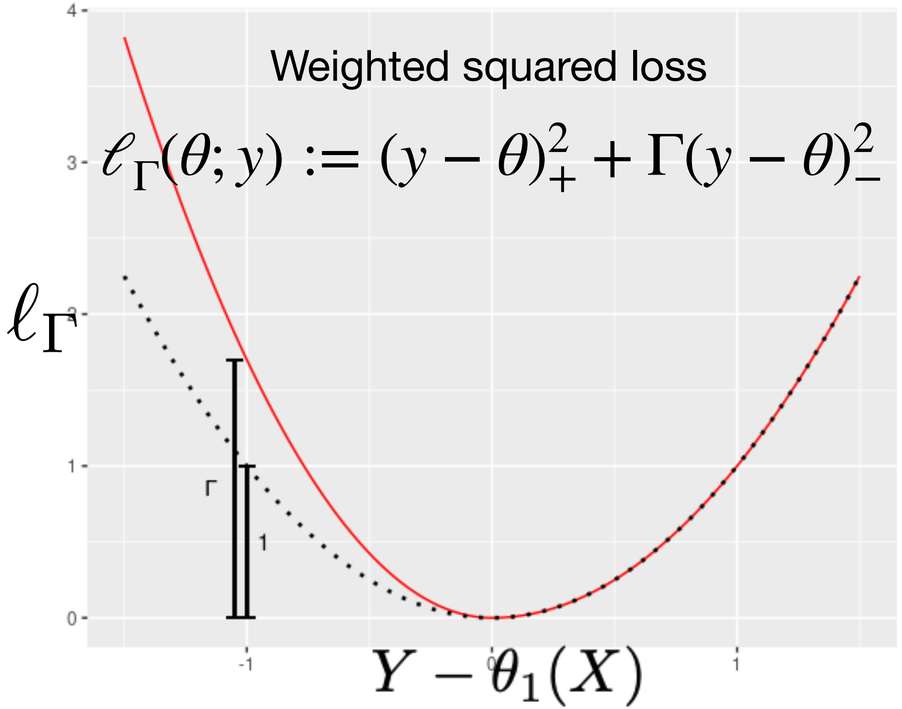
- Tremendous empirical success is **curve-fitting tools** in high-dimensions, under noisy data
- Key ingredients: stochastic optimization & model selection



Sensitivity of CATE via loss minimization

- $\mathbb{E}[Y(1) \mid X, Z = 0] \geq \theta_1^\star(X) = \sup\left\{\mu : \mathbb{E}[(Y(1) - \mu)_+ - \Gamma(Y(1) - \mu)_- \mid X, Z = 1] \geq 0\right\}$
- We want to estimate lower bound $\theta_1^\star$ using flexible ML models ⇒ loss min!

Main result 1: $\theta_1^\star$ is the unique solution to **minimize**$_{\theta(\cdot)}$ $\mathbb{E}[\ell_\Gamma(\theta(X); Y(1)) \mid Z = 1]$



- Solve weighted regression problem using any black-box ML approach
- e.g., random forests, boosted trees, NNs

Lower bound for $\mathbb{E}[Y(1)]$

Recap

- Treatment assignment: Z
- Potential outcome: $Y(0)$, $Y(1)$
- Response $Y := Y(Z)$

- From previous bound, $\mu_1^- = \mathbb{E}[ZY(1) + (1-Z)\theta_1^\star(X)] \leq \mathbb{E}[Y(1)]$
- Use ML methods to estimate nuisance parameters (complex ftns of X)

$$\widehat{\theta}_1(X) = \text{argmin}_{\theta \in \Theta_n} \ \mathbb{E}_n \left[\ell_\Gamma(\theta(X); Y(1)) \mid Z = 1 \right]$$

$$\widehat{e}(X) \approx \mathbb{P}(Z = 1 \mid X)$$

$$\widehat{\nu}_1(X) \approx 1 + (\Gamma - 1)\mathbb{P}(Y(1) \leq \theta_1^\star(X) \mid X, Z = 1)$$

- **Today**: Estimator of μ_1^- insensitive to error in nuisance estimates

Estimation Approach

Recap

- Treatment assignment: Z
- Potential outcome: $Y(0)$, $Y(1)$
- Response $Y := Y(Z)$

- Estimate $\mu_1^- = \mathbb{E}[ZY(1) + (1 - Z)\theta_1^\star(X)] \leq \mathbb{E}[Y(1)]$
- Neyman orthogonality [Neyman '59, Chernozhukov et al. '18]
  - Directional derivative of functional wrt nuisance parameters near true values is zero
  - Ensures that a little perturbation in nuisance parameters near the truth values does not affect functional
  - Central limit rate $O_p(n^{-1/2})$ for the worst-case bound, even when you estimate nuisance parameters at slower rates

Orthogonal Estimator

Recap

- Treatment assignment: Z
- Potential outcome: $Y(0)$, $Y(1)$
- Response $Y := Y(Z)$

- Augmented form

$$\mu_1^- = \mathbb{E}\left[ZY(1) + (1-Z)\theta_1(X) + \frac{Z}{e(X)\nu_1(X)}\left((Y(1)-\theta_1(X))_+ + \Gamma(Y(1)-\theta_1(X))_-\right)\right]$$

where $\theta_1(X) = \text{argmin}_{\theta\in\Theta}\ \mathbb{E}\left[\ell_\Gamma(\theta(X); Y(1)) \mid Z=1\right]$, $e(X) = \mathbb{P}(Z=1 \mid X)$,

$$\nu_1(X) = 1 + (\Gamma - 1)\mathbb{P}(Y(1) \le \theta_1^\star(X) \mid X, Z=1)$$

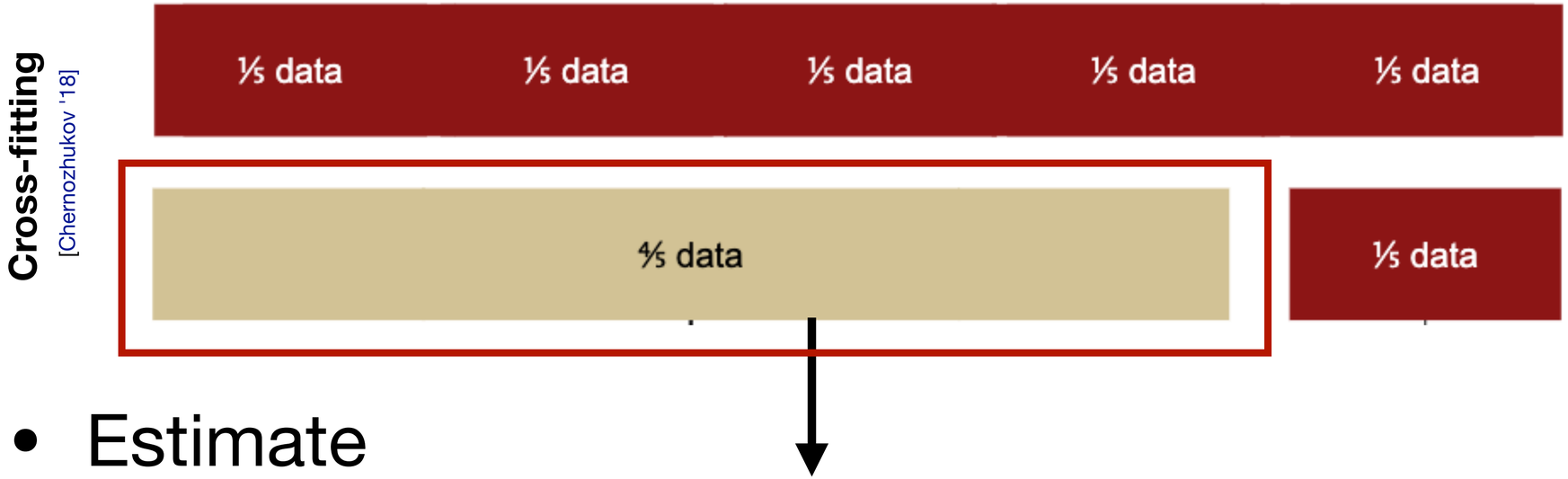
- Bounds the AIPW, one of the most popular estimator for the ATE; equal when $\Gamma = 1$

Insights from the estimator

- AIPW is close to the worst-case bound μ_1^- (a.k.a. robust) when residuals $Y - \hat{\theta}_1(X)$ small
- $\hat{\theta}_1(X)$ trained to predict outcomes based on generic ML approaches
- **Value of features**: Incorporating rich set of covariates can help in terms of robustness to unobserved confounding
- Predicting outcomes accurately pays off—we can utilize existing ML best practices on feature engineering and model selection

Cross-fitting

- First, split data into K-fold partition. Then for each partition…



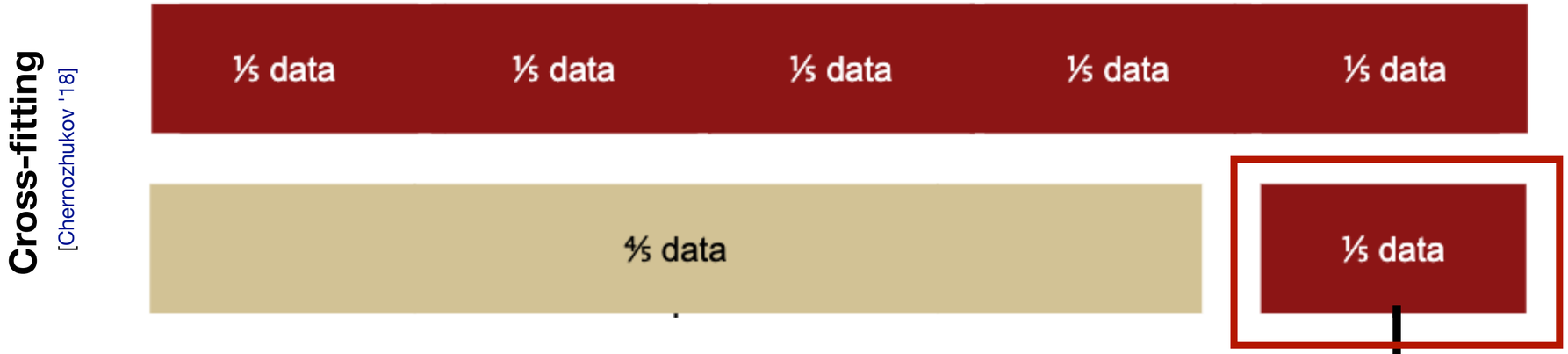
- Estimate nuisance parameters on the auxiliary sample

$$\widehat{\theta}_1(X) = \text{argmin}_{\theta \in \Theta_n} \; \mathbb{E}_n \left[\ell_\Gamma(\theta(X); Y(1)) \mid Z = 1 \right]$$

$$\widehat{e}(X) \approx \mathbb{P}(Z = 1 \mid X)$$

$$\widehat{\nu}_1(X) \approx 1 + (\Gamma - 1)\mathbb{P}(Y(1) \leq \theta_1^\star(X) \mid X, Z = 1)$$

Cross-fitting



Recap

$\mu_z^\star(X) = \mathbb{E}[Y(z) \mid X = x],\ z \in \{0,1\}$

$e^\star(X) = \mathbb{P}(Z = 1 \mid X)$

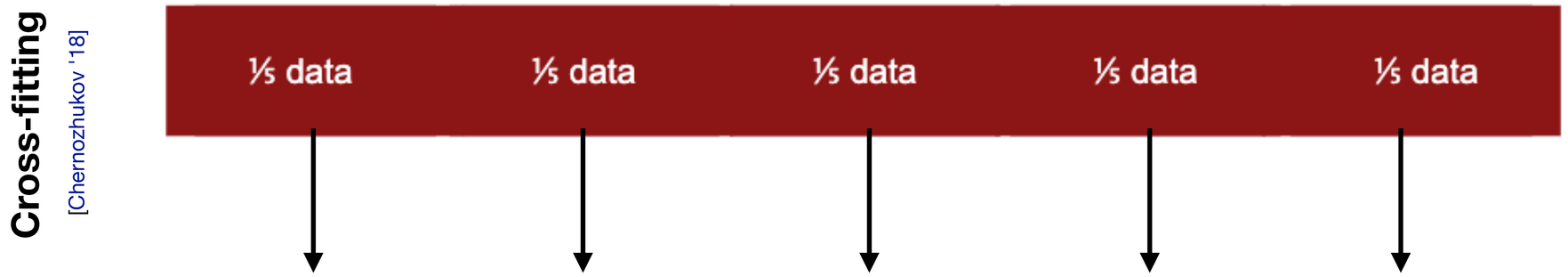
$\nu_1(X) \approx 1 + (\Gamma - 1)\mathbb{P}(Y(1) \leq \theta_1^\star(X) \mid X, Z = 1)$

- Estimate $\bar{\mu}_1$ on the main sample I by plugging in nuisance estimates

$$\widehat{\bar{\mu}}_{1,1} := \frac{1}{|I|}\sum_{i\in I}^{n} Z_i Y_i + (1 - Z_i)\widehat{\theta}_1(X_i) + \frac{Z_i}{\widehat{e}(X_i)}\frac{(Y_i - \widehat{\theta}_1(X_i))_+ + \Gamma(Y_i - \widehat{\theta}_1(X_i))_-}{\widehat{\nu}_1(X_i)}$$

Cross-fitting

- After getting estimates of the lower bound $\bar{\mu}_1$ for each split, average them to get final estimate.



$$\hat{\bar{\mu}}_1 = \frac{1}{5}\left(\hat{\bar{\mu}}_{1,1} + \hat{\bar{\mu}}_{1,2} + \hat{\bar{\mu}}_{1,3} + \hat{\bar{\mu}}_{1,4} + \hat{\bar{\mu}}_{1,5} \right)$$

Main Result II

Recap

$$\widehat{\theta}_1(X) = \text{argmin}_{\theta \in \Theta_n} \; \mathbb{E}_n \left[\ell_\Gamma(\theta(X); Y(1)) \mid Z = 1 \right]$$
$$\widehat{e}(X) \approx \mathbb{P}(Z = 1 \mid X)$$
$$\widehat{\nu}_1(X) \approx 1 + (\Gamma - 1)\mathbb{P}(Y(1) \leq \theta_1^\star(X) \mid X, Z = 1)$$

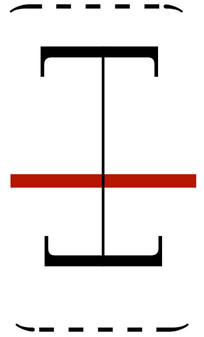
Standard; required for identification and estimation of ATE

- Overlap: $\exists \eta > 0,\ \mathbb{P}(Z = 1 \mid X) \in [\eta, 1 - \eta]$ a.s.
- SUTVA: single version of treatment, no interference between units

Theorem (YNDBT'21)

Let nuisance estimates $\widehat{\theta}_1, \widehat{\nu}, \widehat{e}$ converge at rate $o_p(n^{-1/4})$ to their population counterparts. Under regularity conditions,

$$\frac{\sqrt{n}}{\widehat{\sigma}_n^-}(\widehat{\mu}_1^- - \mu_1^-) \Rightarrow N(0,1)$$ for a known standard deviation estimate $\widehat{\sigma}_n^-$



Main Result II

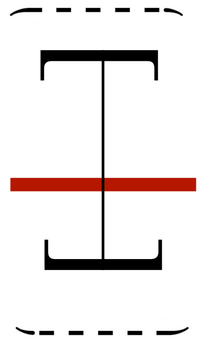
Recap

$$\widehat{\theta}_1(X) = \text{argmin}_{\theta \in \Theta_n} \; \mathbb{E}_n \left[\ell_\Gamma(\theta(X); Y(1)) \mid Z = 1 \right]$$
$$\widehat{e}(X) \approx \mathbb{P}(Z = 1 \mid X)$$
$$\widehat{\nu}_1(X) \approx 1 + (\Gamma - 1)\mathbb{P}(Y(1) \leq \theta_1^\star(X) \mid X, Z = 1)$$

Theorem (YNDBT'21)

Let nuisance estimates $\widehat{\theta}_1, \widehat{\nu}, \widehat{e}$ converge at rate $o_p(n^{-1/4})$ to their population counterparts. Under regularity conditions,

$$\frac{\sqrt{n}}{\widehat{\sigma}_n^-}(\widehat{\mu}_1^- - \mu_1^-) \Rightarrow N(0,1)$$ for a known standard deviation estimate $\widehat{\sigma}_n^-$



- Central limit rates even when ML-based nuisance estimates converge slower
- Can combine with an upper bound on $\mathbb{E}[Y(0)]$ to get a final lower bound on the ATE $\mathbb{E}[Y(1) - Y(0)]$

Extensions

- Similar ideas can be extended to multi-action, sequential decision-making problems when a single decision suffers from unobserved confounding
- Assess sensitivity of the proposed decision policy, compared with the status quo
- Same ingredients: flexible model fitting based on past states and actions

N., Yadlowsky, Keramati, and Brunskill (2020)

https://arxiv.org/abs/2003.05623

Sepsis management in the ICU

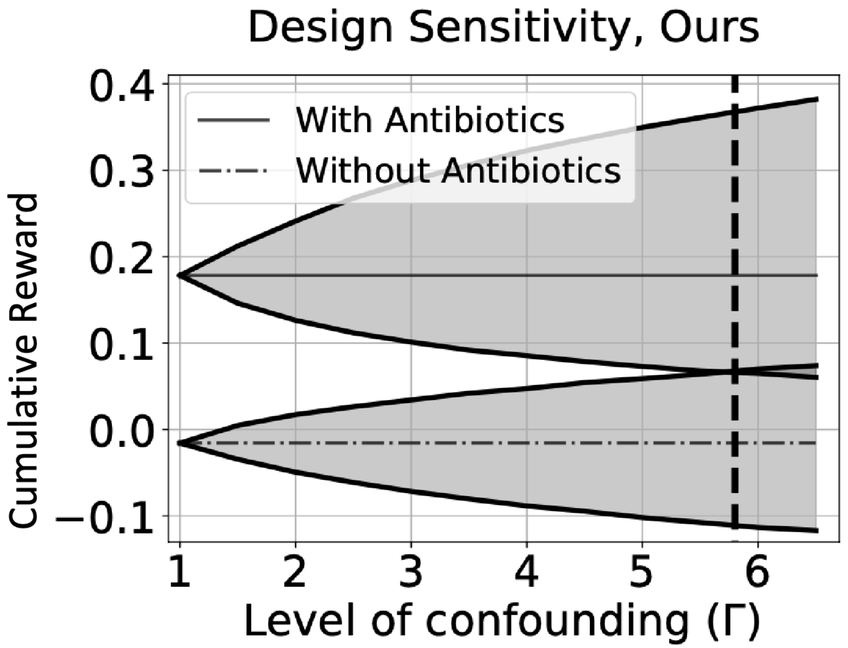
- Sepsis in ICU patients accounts for 1/3 of deaths in hospitals [Howell and Davis '17]
- Automated approaches can manage important medication for sepsis
  - Several AI-based policies have recently been proposed [Futoma '18; Komorowski 18; Raghu 17]
- Due to safety concerns, new treatment policies need to be evaluated offline before thorough online clinical validation
- Today: use our extension to policy evaluation to compare proposals

Sepsis management in the ICU

- ICU data suffers from unobserved confounders
  - Often no pre-existing record of patients in emergency departments (ED)
  - Leaves important patient information unrecorded in subsequent decisions (e.g. comorbidities)
- Stanford ED physician: “initial treatment of antibiotics at admission to the hospital are often confounded by unrecorded factors that affect the eventual outcome (death or discharge from the ICU).”

Proof of concept

- Whether to quickly begin antibiotic treatment is a topic of much discussion: balance early treatment vs. risks of over-prescription [Seymour '17; Sterling '15]
- Two policies: with or without antibiotics in the first step
- We use simulator developed by Oberst and Sontag (2019)



Our ML-based approach allows certifying robustness under realistic values of confounding

Conclusion

- Formulated worst-case bounds on the causal effect
- Flexible ML-based sensitivity framework for the CATE
- Modern semiparametric sensitivity framework for the AIPW => **scalable!**
- Central limit rates even when nuisance estimates converge slower

Bounds on the conditional and average treatment effect with unobserved confounding factors.
Yadlowsky, N., Basu, Duchi, and Tian. Minor revision in Annals of Statistics.
Off-policy policy evaluation for sequential decisions under unobserved confounding.
N., Keramati, Yadlowsky, and Brunskill. NeurIPS 2020.