Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Trang Nguyen

joint work with Liz Stuart, Dan Scharfstein, Betsy Ogburn

arXiv:2303.05032 | trang.nguyen@jhu.edu

CMStatistics 2023-12-17

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

- thank Widemberg da Silva Nobre
- joint work with Liz Stuart, Dan Scharfstein and Betsy Ogburn
- this work is about sens analysis for violation of one of the assumptions used in estimating principal causal effects

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

> Trang Nguyen joint work with Liz Staart. Dan Scharftene. Beter Ochure

> > aXiv2303.05032 | trang.ngxyee@jhu.ndu

CMStatistics 2023-12-17

~

12-1

2023-1

Outline

Background

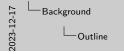
Principal causal effects Principal ignorability (PI) – one identification strategy Sensitivity analysis for PI violation

Mean-based sensitivity analysis: expanding options

Allow different sensitivity parameters Accommodate different estimation methods

Distribution-based sensitivity analysis: a new thing

Method limitations and information use A method using full information (ongoing work) Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects



Background Principal causal effects Principal ignorability (Pt) – one identification strategy Sensitivity analysis for P1 violation

Outline

Mean-based sensitivity analysis: expanding options Allow different sensitivity parameters Accommodate different estimation methods

Nethod limitations and information use A method limitations and information use

i will give a brief introduction to

- principal causal effects
- the principal ignorability assumption
- and sensitivity analysis for its violation

before going into the work that we do on this topic

The noncompliance problem

People may

not take their pill

not attend the training they are assigned to

volunteer less than they are asked to

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Li Background Principal causal effects People may not take their pill not attend the training they are assigned to volunteer less than they are asked to

The noncompliance problem

the study of treatment effects is often complicated by noncompliance

- some people (in drug treatment trials) might not take the pill they are told to take
- some people might not attend the training program they are assigned to
- some people might volunteer fewer hours than they are asked to

The noncompliance problem

People may

not take their pill not attend the training they are assigned to volunteer less than they are asked to

- Z treatment assigned
- Y outcome
- Y(z) potential outcomes
- X baseline covariates
- *S* treatment received

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

ET Background Principal causal effects CO The noncompliance problem

so in addition to the usual variables

- treatment assigned Z (here binary 0/1)
- outcome Y
- potential outcome Y(z) where z is either 1 or 0
- and baseline covariates X

we have a post-treatment assignment variable, generically referred to as treatment received, ${\sf S}$

The noncompliance problem

People may Z treatment adegued not take their pIT vocations not attend the training they are assigned to volunteer less than they are added to X baseline crossinese S treatment received

The noncompliance problem

People may

not take their pill not attend the training they are assigned to volunteer less than they are asked to

Might be intersted in the effect of receiving treatment but those who received \neq those who did not

 $\begin{array}{c} Z & \text{treatment assigned} \\ Y & \text{outcome} \\ \text{ssigned to} & Y(z) & \text{potential outcomes} \\ \text{to} & X & \text{baseline covariates} \\ \end{array}$

- Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects
- Image: Second strain of the second strain of the second strain of the second strain strain

when not all those assigned to treatment receive the treatment

one might be more interested in the effect of receiving the treatment than the effect of being assigned to the treatment

but those who received the treatment (or who volunteered) may not be similar to those who didn't

The noncompliance problem

Might be intersted in the effect of receiving treatment

those who received \neq those who did not

Z treatment assigned

Y(z) potential outcomes X baseline covariates

\$ treatment received

People may

simply comparing their outcomes would be akin to breaking randomization

Principal stratification

People may

not take their pill not attend the training they are assigned to volunteer less than they are asked to

Might be intersted in

the effect of receiving treatment but the groups are not comparable

Principal stratification (Frangakis & Rubin 2002) avoids this problem

by creating a new pre-treatment variable based on potential treatment received and stratifying on it

- Z treatment assigned
- outcome
- Y(z) potential outcomes
- X baseline covariates
- S treatment received
- $\begin{array}{ll} C & \mbox{ principal stratum,} \\ \mbox{ defined based on } \\ \mbox{ potential values } \\ S(1), S(0) \mbox{ of } S \end{array}$

Principal causal effects: $\mathsf{E}[Y(1) - Y(0) \mid C = c]$ Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Background

12-17

2023-1

Principal causal effects
Principal stratification

Principal stratification

the principal stratification framework avoids this problem

by creating a new pre-treatment-assignment variable called principal stratum based on potential values of treatment received

[pointing on RHS] the variable principal stratum (C) here is defined based on S(1) and S(0), which are potential values of S

now we consider effects

– still of being assigned to treatment vs. control, so still the difference between Y(1) and Y(0) – but within each principal stratum

this way we don't break randomization

because C, which is about what type of person someone is (in terms of potential behavior), is a pre-treatment-assignment variable

Principal stratification

People may

not take their pill not attend the training they are assigned to volunteer less than they are asked to

Might be intersted in

the effect of receiving treatment but the groups are not comparable

Principal stratification (Frangakis & Rubin 2002) avoids this problem

by creating a new pre-treatment variable based on potential treatment received and stratifying on it

Two-sided noncompliance: 4 principal strata

S(1)	S(0)	С
1	1	always taker
1	0	complier
0	1	defier
0	0	never taker

- Ζ treatment assigned
 - outcome
- Y(z)potential outcomes
- Х baseline covariates
- S treatment received
- С principal stratum, defined based on potential values *S*(1), *S*(0) of *S*

Principal causal effects: E[Y(1) - Y(0) | C = c] Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

- 12-17 Background Principal causal effects 2023-1
 - Principal stratification

People may	z	treatment assigned
not take their pill	Y	outcome
not attend the training they are assigned to	Y(2)	potential outcomes
volunteer less than they are asked to	×	baseline covariates
Might be intersted in	s	treatment received
the effect of receiving treatment		
but the groups are not comparable	с	principal stratum, defined based on
Principal stratification (Frangahis & Rubin 2002)		potential values
avoids this problem		S(1), S(0) of S
by creating a new pre-treatment variable based on potential treatment received		
	Princip	al causal effects:
and stratifying on it	E[Y	(1) - Y(0) C = c]
Two-sided noncompliance: 4 principal strata		
S(1) S(2) C		
1 1 despirator		

Delected stantification

with binary Z and binary S, there are two settings

in the 2-sided noncompliance setting there are four principal strata based on different combinations of S(1) and S(0) values

Principal stratification

People may

i.e., C = S(1)

not take their pill not attend the training they are assigned to volunteer less than they are asked to

Might be intersted in

the effect of receiving treatment but the groups are not comparable

Principal stratification (Frangakis & Rubin 2002) avoids this problem

by creating a new pre-treatment variable based on potential treatment received and stratifying on it

One-sided noncompliance: 2 principal strata

S(1)	S(0)	С
1	0	complier
0	0	noncomplier

treatment assigned

outcome

Ζ

- Y(z) potential outcomes
- X baseline covariates
- S treatment received
- $\begin{array}{l} C & \mbox{principal stratum,} \\ \mbox{defined based on} \\ \mbox{potential values} \\ S(1), S(0) \mbox{ of } S \end{array}$

Principal causal effects: $\mathsf{E}[Y(1) - Y(0) \mid C = c]$ Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Background Principal causal effects

12-17

2023-1

Principal stratification

Principal stratification People may Z treatment assigned outcome Y(z) potential outcomes baseline covariates treatment received the effect of receiving treatment principal stratum, defined based on potential values S(1), S(0) of S avoids this problem by creating a new pre-treatment variab incipal causal effects: and stratifying on it E[Y(1) - Y(0) | C = c]3(1) 3(2) C 1 0 complex 0 0 remember i.e., C = S(1)

in the 1-sided noncompliance setting, where people assigned to control don't have access to the active treatment

S(0) is always zero

so there are two principal strata based on the two values of S(1)

which we refer to simply as compliers and noncompliers

Examples

Our focus is one-sided noncompliance

target: (non)complier average causal effects (CACE and NACE)

JOBS II for unemployed workers

- Z: week-long training on job search and mental health
- S: attending training

Experience Corps for the elderly

- Z: facilitated program for volunteering to help kids in school
- S: volunteering above a certain number of hours

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Li Background Principal causal effects Our focus is one-sided noncompliance target: (non)complier average causal effects (CACE and NACE)

- Experience Corps for the elderly 2. facilitated program for volumening to help lids in school 5. solunteering above a certain number of hours

our focus is one-sided noncompliance where the target estimands are the effects on people of the two compliance types: the complier average causal effect (CACE) and the noncomplier average causal effect (NACE)

let me give two examples

 $\mathsf{JOBS}\:II$ is a study with unemployed workers where the treatment Z is a week-long training on job search skills and mental health and S, the treatment received variable, is an indicator attending the training or not

here the CACE and NACE are the effects of being assigned to the training vs. not on people who would and people who would not attend the training if they were assigned to it

in Experience Corps

4 / 22

the intervention is a volunteering program for elderly folks as a way to improve their health and wellbeing the investigators define variable S to be volunteering at least a certain number of hours % f(x) = 0

here the CACE and NACE are the effects of the program vs the control condition on people who would and who would not volunteer at that level

Identification challenge: C is not observed under control

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Background

12-17

2023-1

Principal ignorability (PI) – one identification strategy

 \square Identification challenge: C is not observed under control





now the challenge is C (compliance type) is only observed in the intervention arm

in the control arm we have a mixture two types of people and do not know who is who

therefore the CACE and NACE are not identified under standard causal inference assumptions

some additional assumption is required

Two major identification strategies







there are two major strategies for identifying these effects

Two major identification strategies

Sensitivity analysis for principal ignorability violation in estimating

Principal ignorability (PI) – one identification strategy

complier and noncomplier average causal effects

the first one is to assume that for noncompliers the assigned treatment has no effect so the outcome is similar in the two conditions

the second one is to assume that in the control condition, the outcome is similar between compliers and noncompliers, so principal stratum is ignorable, hence the name principal ignorability

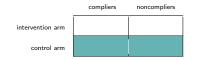
these are very different assumptions and each may be appropriate in some situations but not others

our focus in the current work is PI

12-17

2023-1

Background



$Y \perp\!\!\!\perp C \mid X, Z = 0$

ie within covariate levels, compliers and noncompliers share the same outcome distribution under control

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Image: Background Hereinicipal ignorability (PI) – one identification strategy Image: Background Hereinicipal ignorability (PI)

formally, PI is the assumption of this conditional independence between compliance type and outcome under control

it means that within levels of baseline covariates compliers and noncompliers share the same outcome distribution under control

this assumption may be plausible if we have covariates that are predictive both of compliance and of outcome under control distribution under control

ie within covariate levels, compliant and noncompliant share the same outcome

	compliers	noncompliers
intervention arm	$\mu_{11}(X)$	$\mu_{10}(X)$
control arm	$\mu_{01}(X)$	μ ₀₀ (X)

 $Y \perp\!\!\!\perp C \mid X, Z = 0$

or

 $\underbrace{\mathsf{E}[Y \mid X, Z = 0, C = 1]}_{\mu_{01}(X)} = \underbrace{\mathsf{E}[Y \mid X, Z = 0, C = 0]}_{\mu_{00}(X)}$

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Image: Second strategy Image: Second strate

the important thing it implies is the equality between these two conditional outcome means of compliers and noncompliers (which we label mu-01 and mu-00)

Principal imprability (PI)

 $Y \equiv C \mid X, Z \equiv 0$

 $\underbrace{\mathbb{E}[Y \mid X, Z = 0, C = 1]}_{new[X]} = \underbrace{\mathbb{E}[Y \mid X, Z = 0, C = 0]}_{pen[X]}$

in this picture in the top right here, these two conditional means under control are equal under PI

we also label the two means in the intervention arm as mu-11 and mu-10 these are identified from data

	compliers	noncompliers	
intervention arm	$\mu_{11}(X)$	$\mu_{10}(X)$	
control arm	$\mu_{01}(X)$	μ ₀₀ (X)	

 $Y \perp\!\!\!\perp C \mid X, Z = 0$

$$\underbrace{\mathsf{E}[Y \mid X, Z = 0, C = 1]}_{\mu_{01}(X)} = \underbrace{\mathsf{E}[Y \mid X, Z = 0, C = 0]}_{\mu_{00}(X)}$$

Combined with treatment assignment ignorability, PI identifies CACE, NACE:

$$\mathsf{E}[Y(1) - Y(0) \mid C = c] = \frac{\mathsf{E}\{[\mu_{1c}(X) - \kappa_0(X)]\pi_c(X)\}}{\mathsf{E}[\pi_c(X)]}$$

where

$\kappa_0(X) := E[Y \mid X, Z = 0]$	(mixture outcome mean)
$\pi_c(X) := P(C = c \mid X, Z = 1)$	(principal score)

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Principal ignorability (PI)

where

E[Y | X, Z = 0, C = 1] = E[Y | X, Z = 0, C = 0]

 $s_0(X) := \mathbb{E}[Y \mid X, Z = 0]$ (mixture outcome mean $\pi_c(X) := \mathbb{P}[C = c \mid X, Z = 1]$ (principal score)

Combined with treatment assignment ignorability, PI identifies CACE, NACE

 $E[Y(1) - Y(0) | C = c] = \frac{E[[\mu]_{tr}}{t}$

Background

12-17

2023-1

└──Principal ignorability (PI) – one identification strategy └──Principal ignorability (PI)

the combination of PI with treatment assignment ignorability identifies the two principal causal effects by this formula

which is a weighted average

of a difference in outcome means between the intervention and control conditions

but here, this piece, which is supposed to be the stratum-specific mean mu-0c, is replaced with the mixture outcome mean kappa-0 thanks to PI

and the weight, pi-c(x) here, is the probability of belonging in principal stratum little c given covariate values this is usually called the principal score

	compliers	noncompliers	
intervention arm	$\mu_{11}(X)$	$\mu_{10}(X)$	
control arm	$\mu_{01}(X)$	μ ₀₀ (X)	

 $Y \perp\!\!\!\perp C \mid X, Z = 0$

or

 $\underbrace{\mathsf{E}[Y \mid X, Z = 0, C = 1]}_{\mu_{01}(X)} = \underbrace{\mathsf{E}[Y \mid X, Z = 0, C = 0]}_{\mu_{00}(X)}$

PI is untestable

Need sensitivity analyses

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

└──Background └──Principal ignorability (PI) – one identification strategy

2023-12-17

└──Principal ignorability (PI)

note that PI is an untestable assumption

so we need sensitivity analysis for its violation



Prior sens analysis method that inspired this work

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Background

12-17

2023-1

Sensitivity analysis for PI violation

Prior sens analysis method that inspired this work

Ding and Lu (2017) use a mean ratio sensitivity parameter $\frac{\mu_{\rm H}(X)}{\mu_{\rm H}(X)}=\rho$ and modify a PI-based weighting estimator to incorporate ρ

See also Jurg, Yang and Ding (2022)

Ding and Lu (2017) use a mean ratio sensitivity parameter

 $\frac{\mu_{01}(X)}{\mu_{00}(X)} = \rho$

and modify a PI-based weighting estimator to incorporate ρ

See also Jiang, Yang and Ding (2022)

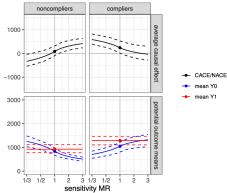
in this prior work, which inspired ours, instead of assuming these two conditional outcome means are equal Ding and Lu used their ratio as a sens param and consider it over a range for the sens analysis

Ding and Lu modified a PI-based weighting estimator to incorporate this param

Example of MR-based sens analysis

earnings in JOBS II

earnings: MR-based sensitivity analysis



Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Background

17

ά. Η.

2023

Sensitivity analysis for PI violation Example of MR-based sens analysis

this is an example of a mean ratio based sens analysis conducted on the outcome earnings at 6 months in the JOBS II study

let me take a moment to orient you to this plot b/c it will appear again

the two columns are for compliers and noncompliers the top panel shows the effect estimates the bottom panel shows the potential outcome means the x-axis shows the sens param

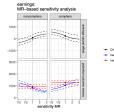
the dots, which are at mean ratio 1, are the PI-based estimates on the right of 1, compliers are assumed to have higher earnings under control than noncompliers with the same covariate values on the left of 1, it's the opposite

we see how the effect estimates change with the sens param

Example of MR-based sens analysis earnings in JOBS II

Example of MR-based sens analysis

earnings in JOBS II



other outcomes for which MR param not ideal

- JOBS II: having a job (binary), depressive symptoms (bounded)
- Experience Corps: generativity (bounded)

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Background

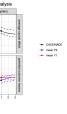
12-17

2023-1

Sensitivity analysis for PI violation Example of MR-based sens analysis other outcomes for which MR param not ideal JORS II: having a jule (kinary), depression symptoms (locanded) 100

while the mean ratio param may be appropriate for earnings

there are other outcomes for which this param is not ideal like binary and bounded outcomes



earnings in JOBS II

Outline

Background

Principal causal effects Principal ignorability (PI) – one identification strategy Sensitivity analysis for PI violation

Mean-based sensitivity analysis: expanding options Allow different sensitivity parameters

Accommodate different estimation methods

Distribution-based sensitivity analysis: a new thing

Method limitations and information use A method using full information (ongoing work) Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Mean-based sensitivity analysis: expanding options

Outline

Background Principal causal effects Principal ignorability (P1) – one identification strategy Sensitivity analysis for P1 violation

Outline

Mean-based sensitivity analysis: expanding options Allow different sensitivity parameters Accommodate different estimation methods

Distribution-based sensitivity analysis: a new thing Method limitations and information use A method using full information (orgoing work)

that motivates this work on mean-based sensitivity analysis which aims to allow different sensitivity parameters suitable to different outcome types

also, we want to accommodate different estimation methods that may be used in practice

2-17

2023-1:

A range of sens assumptions with different sens params

PI: $\mu_{01}(X) = \mu_{00}(X).$

sens-MR: $\frac{\mu_{01}(X)}{\mu_{00}(X)} = \rho,$ sens-OR: $\frac{\mu_{01}(X)/[1 - \mu_{01}(X)]}{\mu_{00}(X)/[1 - \mu_{00}(X)]} = \psi,$ sens-GOR: $\frac{[\mu_{01}(X) - I]/[h - \mu_{01}(X)]}{[\mu_{00}(X) - I]/[h - \mu_{00}(X)]} = \psi$ where *I* and *h* are the outcome lower and upper bounds, sens-SMD: $\frac{\mu_{01}(X) - \mu_{00}(X)}{\sqrt{[\sigma_{01}^{2}(X) + \sigma_{00}^{2}(X)]/2}} = \eta$

where $\sigma_{0c}^2(X) := \operatorname{var}(Y \mid X, Z = 0, C = c)$,

for some range of $\rho,\,\psi$ or η that is considered plausible.

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Provide the sensitivity analysis: expanding options

 Image: Provide the sensitivity parameters

 Image: Provide the sen



A range of sens assumptions with different sens params

for some range of ρ,ψ or η that is considered plausible.

this slide shows a range of sens params

in addition to the mean ratio we consider an odds ratio param, which is suitable for a binary outcome,

a generalized odds ratio, which can be used for an outcome bounded on both ends,

and a standardized mean difference parameter, where the mean difference is divided by a pooled standard deviation. this is suitable for a general continuous variable.

Identification

sens-MR and sens-GOR result in point identification of CACE, NACE

because they help solve the mixture equation

 $\pi_1(X)\mu_{01}(X) + \pi_0(X)\mu_{00}(X) = \kappa_0(X).$

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Provide the sensitivity analysis:
 expanding options

 Image: Provide the sensitivity analysis:
 expanding options

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensity parameters
 Image: Provide th

sens-MR and sens-GCR result in point identification of CACE, NACE because they help salve the mixture equation $v_1(X)v_{01}(X) + v_2(X)v_{02}(X) = v_2(X).$

Identification

both sens-MR and -GOR result in point identification of the CACE and NACE b/c they help solve this mixture equation of the conditional outcome means under control

Identification

sens-MR and sens-GOR result in point identification of CACE, NACE

because they help solve the mixture equation

 $\pi_1(X)\mu_{01}(X) + \pi_0(X)\mu_{00}(X) = \kappa_0(X).$

sens-SMD obtains bounds for CACE, NACE

bounds are narrowed if also assume $1/k \le \frac{\sigma_{01}^2(X)}{\sigma_{00}^2(X)} \le k$ for some k > 1

and reduce to point identification if assume $\sigma_{01}^2(X) = \sigma_{00}^2(X)$ (aka sens-SMDe)

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Provide the sensitivity analysis: expanding options

 Image: Provide the sensitivity provide the sensitivity parameters

 Image: Provide the sensitivity provide the senset the senset the sensitivity provide the sensitivity p

sens-MR and zero-GOR result in point identification of CACE, NACE because they help value the mixture equation $v_1(X)v_{12}(X) + v_2(X)v_{12}(X) = v_2(X).$

seen-SMD obtains bounds for CACE, NACE bounds are narrowed if also assume $1/k \le \frac{\sigma_{2k}^{-1}(3)}{\sigma_{2k}^{-1}(k)} \le k$ for some k > 1and reduce to point identification if assume $\sigma_{2k}^{2}(X) = \sigma_{2k}^{2}(X)$ (also seen-SMD+)

sens-SMD obtains bounds for the effects

these bounds are narrowed if we supplement it with an assumption about the conditional variances and reduce to point identification if we assume equal variance

Identification

sens-MR and sens-GOR result in point identification of CACE, NACE

because they help solve the mixture equation

 $\pi_1(X)\mu_{01}(X) + \pi_0(X)\mu_{00}(X) = \kappa_0(X).$

sens-SMD obtains bounds for CACE, NACE

bounds are narrowed if also assume $1/k \leq \frac{\sigma_{01}^2(X)}{\sigma_{00}^2(X)} \leq k$ for some k > 1and reduce to point identification if assume $\sigma_{01}^2(X) = \sigma_{00}^2(X)$ (aka sens-SMDe)

in all cases, effect identification is via identification of $\mu_{0c}(X)$

by a function of sens param, $\pi_c(X)$, $\kappa_0(X)$ (and $var(Y \mid X, Z = 0)$ w/ sens-SMD)

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Provide the sensitivity analysis:
 expanding options

 Image: Provide the sensitivity analysis:
 expanding options

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensitivity parameters
 Image: Provide the sensitivity parameters

 Image: Provide the sensity parameters
 Image: Provide th

sens-MR and sens-GOR result in point identification of CACE, NACE because they help take the mixture equation $v_1(X)u_2(X) + v_3(X)u_3(X) = v_3(X).$

sens-SMD obtains bounds for CACE, NACE bounds are narround if also assume $1/k \le \frac{p_{0}^{2}(k)}{p_{0}^{2}(k)} \le k$ for some k > 1and reduce to point identification if assume $r_{0}^{2}(X) = r_{0}^{2}(X)$ (also sens-SMD)

in all cases, effect identification is via identification of $\mu w(X)$ by a function of sees param, $v_{\alpha}(X)$, $v_{0}(X)$ (and $var(Y \mid X, Z = 0)$ w/ sees-SME

in all cases effect identification is obtained via identification of mu-Oc(x)

and the result for mu-Oc(x) is a function of the sens param the principal score and mixture outcome mean and for the sens-SMD assumption, also the mixture outcome variance x^{2}

this detail is not so important but will be relevant in the next section

Sens analysis = a modification of main analysis

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

- 12-17 Mean-based sensitivity analysis: expanding options
 - Accommodate different estimation methods
 - -Sens analysis techniques to go with 3 types of PI-based estimators

in practice, a sens analysis often follows and is secondary to a main analysis,

it is thus desirable for the sens analysis to be a hopefully simple modification of the main analysis

people may use different estimators for the main analysis

here we think of 3 types of PI-based estimators for which different sens analysis techniques may apply

2023-1

Sens analysis = a modification of main analysis

- **•** Type A (\approx outcome regression estimators)
 - estimates $\kappa_0(X)$ to first estimate effects conditional on covariates and then aggregates them to estimate CACE/NACE, eg

 $\frac{\sum_{i=1}^{n} \hat{\pi}_{c}(X_{i}) [\hat{\mu}_{1c}(X_{i}) - \hat{\kappa}_{0}(X_{i})]}{\sum_{i=1}^{n} \hat{\pi}_{c}(X_{i})}, \quad \frac{\sum_{i=1}^{n} \frac{Z_{i}}{\hat{\epsilon}(X_{i},Z_{i})} \mathsf{I}(C_{i} = c) [Y_{i} - \hat{\kappa}_{0}(X_{i})]}{\sum_{i=1}^{n} \frac{Z_{i}}{\hat{\epsilon}(X_{i},Z_{i})} \mathsf{I}(C_{i} = c)}$

sens analysis technique: replace $\kappa_0(X)$ by the identification result of $\mu_{0c}(X)$ under the sens assumption

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structu

▶ Type A (c outcome regression estimators) ▶ estimates $q_{ij}(X)$ is form estimate effects conditional on convinces and then aggregates them to estimate CACC/NACC, eg $<math display="block">\frac{\sum_{i=1}^{n} h_i(X) (\lim_{x \in X} |X_i - h_i(X)|)}{\sum_{i=1}^{n} \frac{\sum_{i=1}^{n} \frac{1}{\sqrt{\lambda_i + \lambda_i}} V(G = c) [V_i - h_i(X)]}{\sum_{i=1}^{n} \frac{1}{\sqrt{\lambda_i + \lambda_i}} V(G = c)}$ ▶ ence analytic technique: register $e_q(X)$ by the identifications meant $d_{eq_i}(X)$

The first type consists of outcome regression estimators

Roughly speaking, these as estimators that estimate the $\kappa_0(X)$ to first estimate conditional effects then aggregates the conditional effects to estimate the CACE/NACE

An example is the plug-in estimator << point to it >>.

With this type, the sensitivity analysis technique is to replace $\kappa_0(X)$ with the identification result of $\mu_{0c}(X)$ under the sensitivity assumption

Sens analysis techniques to go with 3 types of PI-based estimators

Sens analysis = a modification of main analysi

Sens analysis = a modification of main analysis

Type B (\approx influence function based estimators)

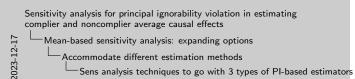
write the CACE/NACE as $\nu_{1c} - \nu_{0c}^{PI}$ π_c

where $\nu_{zc} := \mathsf{E}[\pi_c(X)\mu_{zc}(X)], \quad \nu_{0c}^{\mathsf{PI}} := \mathsf{E}[\pi_c(X)\kappa_0(X)], \quad \pi_c := \mathsf{E}[\pi_c(X)]$

a type B estimator can be expressed as combination of IF-based estimators of π_c , ν_{1c} and ν_{0c}^{Pl}

$$\frac{\hat{\nu}_{1c,\text{if}} - \hat{\nu}_{0c,\text{if}}^{\mathsf{PI}}}{\hat{\delta}_{c,\text{if}}}$$

• sens analysis technique: replace $\hat{\nu}_{0c}^{\text{Pl}}$ with an IF-based estimator of ν_{0c} under the sens assumption



-Sens analysis techniques to go with 3 types of PI-based estimators

Type B includes some influence function based estimators.

We write the effect under PI as this expression that involves three parameters.

A type B estimator can be expressed as a combination of influence function based estimators of these three parameters.

A relevant sensitivity analysis technique then is to replace this component estimator in red with an influence function based estimator of ν_{0c} under the sensitivity assumption

iens analysis techniques to go with 3 types of PI-based estimators ens analysis = a modification of main analysi Type B (n) influence function based estimators

 $\nu_{2e} - \nu_{2e}^{T}$

where $i_{M} := E[t_{T}(X)_{U \in V}[X]], \quad i_{M}^{(2)} := E[t_{T}(X)_{U \in V}[X]], \quad t_{T} := E[t_{T}(X)]$

► a type II estimator can be expressed as combination of IF-based estim

sens analysis technique: replace I an IF-based estimator of a

wite the CACE/NACE as

Sens analysis = a modification of main analysis

- ► Type C (≈ other/weighting estimators)
 - ▶ an example is the pure weighting estimator

 $\frac{\sum_{i=1}^{n} \frac{Z_i}{\hat{\epsilon}(X_i,Z_i)} \mathsf{I}(C_i=c) Y_i}{\sum_{i=1}^{n} \frac{Z_i}{\hat{\epsilon}(X_i,Z_i)} \mathsf{I}(C_i=c)} - \frac{\sum_{i=1}^{n} \frac{1-Z_i}{\hat{\epsilon}(X_i,Z_i)} \hat{\pi}_c(X_i) Y_i}{\sum_{i=1}^{n} \frac{1-Z_i}{\hat{\epsilon}(X_i,Z_i)} \hat{\pi}_c(X_i)}$

no general sens analysis technique

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Provide the second structure
 Provide the second structure

 $\begin{array}{c} \blacktriangleright \mbox{Type C (n other/weighting estimators)$} \\ \hline \mbox{an example is the pure weighting estimators}$$ \\ \hline & \frac{\sum_{i=1}^{n} \frac{1}{2N_{i}^{2} C_{i}^{2}} 1^{i}(C_{i}^{i}=c)Y_{i}}{\sum_{i=1}^{n} \frac{1}{2N_{i}^{2} C_{i}^{2}} 1^{i}(X_{i})}$$ \\ \hline & \frac{\sum_{i=1}^{n} \frac{1}{2N_{i}^{2} C_{i}^{2}} 1^{i}(C_{i}^{i}=c)}{\sum_{i=1}^{n} \frac{1}{2N_{i}^{2} C_{i}^{2}} 1^{i}(X_{i})} $$ \\ \hline & \mbox{an equation set analytic tochologies} $$ \end{array}$

iens analysis techniques to go with 3 types of PI-based estimators

Sens analysis = a modification of main analysis

The other estimators are in type C.

They are not IF-based and do not involve estimating $\kappa_0(X)$.

An example is the pure weighting estimator.

For this type, we do not have a general sensitivity analysis technique in mind, and will need to consider them case by case.

Sens analysis = a modification of main analysis

- ► Type C (≈ other/weighting estimators)
 - ▶ an example is the pure weighting estimator

 $\frac{\sum_{i=1}^{n} \frac{Z_i}{\hat{e}(X_i,Z_i)} \mathsf{I}(C_i=c) Y_i}{\sum_{i=1}^{n} \frac{Z_i}{\hat{e}(X_i,Z_i)} \mathsf{I}(C_i=c)} - \frac{\sum_{i=1}^{n} \frac{1-Z_i}{\hat{e}(X_i,Z_i)} \hat{\pi}_c(X_i) Y_i}{\sum_{i=1}^{n} \frac{1-Z_i}{\hat{e}(X_i,Z_i)} \hat{\pi}_c(X_i)}$

- no general sens analysis technique
- sens-MR: scale Y in control units by a simple function of ρ and $\pi_c(X)$

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structure

 Image: Provide the second structure
 Provide the second structu

 $\label{eq:product} \begin{array}{l} \bullet \quad \mbox{type } C \left(\pi \mbox{ other / weighting estimation} \right) \\ \bullet \quad \mbox{ an example is the pare weighting estimation} \\ & \quad \mbox{ } \frac{\sum_{i=1}^{r} \frac{1}{\pi_i \sum_{i=1}^{r} \sum_{i=1}^{r} \frac{1}{\pi_i \sum_{i=1}^{r}$

Sens analysis techniques to go with 3 types of PI-based estimators

Sens analysis = a modification of main analysi

for the sens-MR case,

there is a simple technique for type C that involves a scaling of the outcome in control units

Sens analysis = a modification of main analysis

- Type A (\approx outcome regression estimators)
 - estimates $\kappa_0(X)$ to first estimate effects conditional on covariates and then aggregates them to estimate CACE/NACE
 - sens analysis technique: replace $\kappa_0(X)$ by the identification result of $\mu_{0c}(X)$ under the sens assumption
- Type B (\approx influence function based estimators)
 - **•** can be expressed as combination of IF-based estimators of π_c , ν_{1c} and ν_{0c}^{PI}
 - sens analysis technique: replace $\hat{\nu}_{0c}^{\text{Pl}}$ with an IF-based estimator of ν_{0c} under the sens assumption
- Type C (\approx other/weighting estimators)
 - no general sens analysis technique
 - sens-MR: scale Y in control units by a simple function of ρ and $\pi_c(X)$

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

- -17 Mean-based sensitivity analysis: expanding options ά ÷ Accommodate different estimation methods 2023
 - -Sens analysis techniques to go with 3 types of PI-based estimators

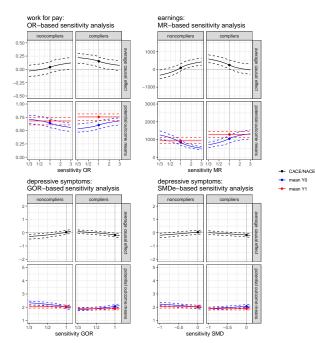
- estimates x₀(X) to first estimate effects conditional on covariates and th aggregates them to estimate CACE/NACE
- aggregates there to estimate CAU.b. (NPC.b. sens analysis technique: replace $n_0(X)$ by the identification result of $\mu_{0n}(X)$ under the sens assumption
- can be expressed as combination of IF-based estimators of multiplication and electronic states and electron sens analysis technique: replace 👯 __ with an IF-based estimator of star
- ► Type C (n: other/weighting estimate no general sens analysis technique anne MIP: acula V in control units he a simula function of a and m.d.Y.

the paper includes more details about different estimators

but these are basically the key ideas

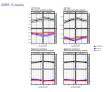
iens analysis techniques to go with 3 types of PI-based estimators Sens analysis = a modification of main analysi

JOBS II results



Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Hean-based sensitivity analysis: expanding options



we applied the different sens analyses with different outcomes in JOBS II

using the OR param for the binary outcome being employed

using the GOR and SMD params for the continuous outcome depressive symptoms

a lot can be said about this specific example, but we don't have time

Some other things we noticed/figured out

- Partial loss of multiple robustness because μ_{0c}(X) is a function of π_c(X) and κ₀(X)
- A pattern of finite-sample bias for the sens analysis where effect estimates are less extreme than should be

because $E[Y_0 | C = c] = \frac{E[\pi_c(X)\mu_{0c}(X)]}{E[\pi_c(X)]}$ is a weighted average where the function being averaged depends on the weight

If IF-based nonparametric estimation: Rate conditions on nuisance estimation for the estimator to be root-n consistent Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Provide the sensitivity analysis:
 expanding options

 Image: Provide the sensititity analysis:
 expanding optio

Some other things we noticed/figured out

 Partial loss of multiple robustness because µ₀₀(X) is a function of v_r(X) and s₀(X)

▶ A pattern of finite-sample bias for the sens analysis where effect estimates are less actreme than should be because $E(h \mid C = c) = \frac{E(x(X))_{H \in V(X)}}{E(n \mid X)}$ is a weighted average where the function bring averaged dependent on the weight

 If If-based nonparametric estimation: Rate conditions on nuisance estimation for the estimator to be root-n consistent

there are several things we noticed or figured out

such as partial loss of multiple robustness for IF-based estimation

finite-sample bias in the sens analysis

and rates conditions for root-n consistent semiparamtric estimation

let's skip this slide, the details are in the paper

Outline

Background

Principal causal effects Principal ignorability (PI) – one identification strategy Sensitivity analysis for PI violation

Mean-based sensitivity analysis: expanding options

Allow different sensitivity parameters Accommodate different estimation methods

Distribution-based sensitivity analysis: a new thing

Method limitations and information use A method using full information (ongoing work) Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Distribution-based sensitivity analysis: a new thing

now I'd like to turn to a different approach

distribution-based sens analysis

which we are working on

Background Principal causal effects Principal ignorability (PI) – one identification strat Sanshido: anabasis for PI addation

Outline

Mean-based sensitivity analysis: expanding options Allow different sensitivity parameters Accommodate different estimation methods

Distribution-based sensitivity analysis: a new thing Method limitations and information use A method using full information (orgoing work)

A general limitation of above methods

is the risk of contradicting the observed data distribution

Assumption	Risk level	Info used from the observed mixture outcome distribution
sens-MR	greatest risk	mean $\kappa_0(X) = E[Y \mid X, Z = 0]$
		lower bound (0)
sens-GOR	less risk	mean
(for nonbinary Y)		both bounds
sens-SMD	less risk	mean
		variance var($Y \mid X, Z = 0$)
sens-OR	no risk	full distribution
(binary Y)		(as mean = probability)

For nonbinary Y, to avoid contradicting the observed data distribution,

sens analysis needs to be fully informed by P(Y | X, Z = 0)

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

2-17 Distribution-based sensitivity analysis: a new thing Η. Method limitations and information use 2023--A general limitation of above methods

Assumption	Rish level	Info used from the observed minime outcome distribution
sens-MR	farming up	mean $ni(X) = E[Y X, Z = 0]$ lower locard (0)
(for moleinary Y)	less risk	malan kath keumik
Grie 3200	Section.	variance $var(Y \mid X, Z = 0)$
Disev Y1	au vich	for man a subshired

For nonbinary Y, to avoid contradicting the observed data distribution sens analysis needs to be fully informed by P(Y | X, Z = 0)

a general limitation of the mean-based methods is the risk of contradicting the observed data distribution

which is due to using some but not all the information in the data

for example, the sens-MR assumption can predict outcome means that exceed the actual outcome range because it is agnostic of the outcome's upper bound

sens-GOR and sens-SMD have lower risk as they use more information, one using both bounds the other using not only the mean but also the variance of the mixture distribution

sens-OR for a binary outcome has no risk, because it actually uses full information

this means it would be ideal for the sens analysis to be fully informed by the observed mixture outcome distribution

A general limitation of above methods

PI: $Y \perp L C \mid X, Z = 0$

(Everything here conditions on X, Z = 0, so this will be implicit.) PE Y # C | X.Z = 0

(Everything here conditions on X, Z = 0, so this will be implicit.)

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Constructing that sens analysis: a new thing

 Image: Constructing that sens analysis: foundation

here's how we construct such a sens analysis

PI says that Y and C are conditionally independent

PI: $Y \perp L C \mid X, Z = 0$

(Everything here conditions on X, Z = 0, so this will be implicit.)

Want a sens assumption

that allows Y and C to be dependent AND helps identify $\mu_{0c}(X)$

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Constructing that sens analysis: a new thing

 Image: Constructing that sens analysis: foundation

we want a sens assumption that allows Y and C to be dependent

Pi: $Y \perp C \mid X, Z = 0$ (for the tension of X, Z = 0

Constructing that sens analysis: foundation

(lowything here conditions on X,Z=0, on this will be implicit.) Want a sense assumption that allows Y and C to be dependent AND helps identify $\mu m(X)$

PI: $Y \perp L C \mid X, Z = 0$

(Everything here conditions on X, Z = 0, so this will be implicit.)

Want a sens assumption

```
that allows Y and C to be dependent AND helps identify \mu_{0c}(X)
```

Principal scores

Recall/recast	$\pi_c(X) = P(C = c \mid X, Z = 0)$	(outcome-agnostic)
Now define	$\tilde{\pi}_c(X,Y) = P(C = c \mid X, Z = 0, Y)$	(outcome-specific)

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Constructing that sens analysis: a new thing

 Image: Constructing that sens analysis: foundation

Constructing that sens analysis: foundation $\label{eq:prod} Pt: \ Y \perp C \mid X, Z = 0$ (For this has sensitive as X, Z = 0

so this will be implicit. Want a sense assumption that allows Y and C to be dependent AND helps identify $\mu m(X)$

ncipal access
$$\begin{split} \text{Recall/recast:} \quad & \pi_{c}(X) = \mathbb{P}(C = c \mid X, Z = 0) \quad (\text{outcome-agnostic}) \\ \text{Now define} \quad & \pi_{c}(X, Y) = \mathbb{P}(C = c \mid X, Z = 0, Y) \quad (\text{outcome-specific}) \end{split}$$

the principal score we have, pi-c(x), does not condition on the outcome

now define pi-tilde-c(x), which additionally conditions on Y

we call this the outcome-specific principal score

PI: $Y \perp L C \mid X, Z = 0$

(Everything here conditions on X, Z = 0, so this will be implicit.)

Want a sens assumption

```
that allows Y and C to be dependent AND helps identify \mu_{0c}(X)
```

Principal scores

Recall/recast	$\pi_c(X) = P(C = c \mid X, Z = 0)$	(outcome-agnostic)
Now define	$\tilde{\pi}_c(X,Y) = P(C = c \mid X, Z = 0, Y)$	(outcome-specific)

 $\tilde{\pi}_c(X, Y)$ is not identified but we know

- $\models \mathsf{E}[\tilde{\pi}_c(X,Y) \mid X, Z=0] = \pi_c(X)$
- ▶ If Y and C are dependent, $\tilde{\pi}_c(X, Y)$ is a function of Y in addition to X

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

A method using full information (ongoing work)

pi-tilde-c is not identified but we know that

its mean conditional on X,Z=0 is pi-c(x)

and that if Y and C are conditionally independent, pi-tilde is a function of Y

 $\perp C \mid X, Z = 0$ (for sything here conditions an X, Z = - so this will be implicit

Constructing that sens analysis: foundation

Wast a sees assumption that allows Y and C to be dependent AND helps identify $\mu n_{*}(X)$

$$\label{eq:rescaled} \begin{split} & \text{Recall/recast} \quad \pi_{c}(X) = \mathbb{P}(C = c \mid X, Z = 0) & (\text{outcome-agreentic}) \\ & \text{Now define} \quad & \tilde{\pi}_{c}(X,Y) = \mathbb{P}(C = c \mid X, Z = 0, Y) & (\text{outcome-specific}) \end{split}$$

$$\begin{split} \mathbb{P}_r(X,Y) & \text{is not identified but we know} \\ \bullet & \mathbb{E}[\mathbb{D}_r(X,Y) \mid X, Z = 0] = m_r(X) \\ \bullet & \text{ If } Y \text{ and } C \text{ are dependent, } \mathbb{P}_r(X,Y) \text{ is a function of } Y \text{ in addition to } X \end{split}$$

PI: $Y \perp L C \mid X, Z = 0$

(Everything here conditions on X, Z = 0, so this will be implicit.)

Want a sens assumption

```
that allows Y and C to be dependent AND helps identify \mu_{0c}(X)
```

Principal scores

Recall/recast	$\pi_c(X) = P(C = c \mid X, Z = 0)$	(outcome-agnostic)
Now define	$ ilde{\pi}_c(X,Y) = P(C=c \mid X,Z=0,Y)$	(outcome-specific)

 $\tilde{\pi}_c(X, Y)$ is not identified but we know

- $\models \mathsf{E}[\tilde{\pi}_c(X,Y) \mid X, Z=0] = \pi_c(X)$
- ▶ If Y and C are dependent, $\tilde{\pi}_c(X, Y)$ is a function of Y in addition to X

(Side note: Very different from the unobserved confounding problem b/c of the observed mixture. Hence exponential tilting doesn't work, except for a binary outcome, in which case = sens-OR.) Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

A method using full information (ongoing work)

let's skip this

Constructing that sens analysis: foundation

 $\mathbf{z}_{-}=\mathbf{u}$ (Everything here conditions as X,Z=0, so this will be implicit.)

Wast a sees assumption that allows Y and C to be dependent AND helps identify $\mu m(X)$

 $\label{eq:response} \begin{array}{ll} \mbox{since} r_{c}(X) = \mathbb{P}(C=c \mid X, Z=0) & (\mbox{outcome-agreentic}) \\ \mbox{Now define} & \ensuremath{\mathbb{P}}_{c}(X, \mathcal{V}) = \mathbb{P}(C=c \mid X, Z=0, \mathcal{V}) & (\mbox{outcome-agreentic}) \end{array}$

 $l_{T}(X, Y)$ is not identified but we know = E[l_{T}(X, Y) | X, Z = 0] = m(X) = If Y and C are dependent, $l_{n}(X, Y)$ is a function of Y in addition to X

(Side rate: Very different from the undeterved confounding problem h/c of the observed minime. Hence represential titling dome's work, enough for a binary solutions, in which scate = sees OR.)

Use shorthand $\tilde{\pi}_1$ for $\tilde{\pi}_1(X, Y)$

Step 1: Assume a distribution for $\tilde{\pi}_1$ with mean $\pi_1(X)$ that allows $\tilde{\pi}_1$ to vary, indexed by a dispersion param

Step 2: Connect *Y* to $\tilde{\pi}_1$ to induce *Y*-*C* dependence

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Constructing that sense analysis: a new thing

 Image: Constructing that sense analysis: 2 steps

Step 1: Assume a distribution for P_1 with mean $\pi_1(X)$ that allows P_1 to vary, indexed by a dispersion param

Step 2: Connect Y to T1 to induce Y-C dependence

With those two known facts, we construct the sens analysis in two steps

[read slide]

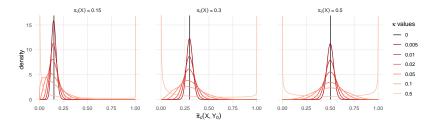
Constructing that sens analysis: 2 steps Use shorthand Its for Ita(X, Y)

Use shorthand $\tilde{\pi}_1$ for $\tilde{\pi}_1(X, Y)$

Step 1: Assume a distribution for $\tilde{\pi}_1$ with mean $\pi_1(X)$ that allows $\tilde{\pi}_1$ to vary, indexed by a dispersion param

use the beta distribution (idea borrowed from Victor Veitch, 2020)

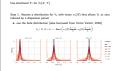
 $ilde{\pi}_1 \mid X, Z = 0 ~\sim~ ext{Beta}\left(\pi_1(X) rac{1-\kappa}{\kappa}, \pi_0(X) rac{1-\kappa}{\kappa}
ight)$



Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Constructing that sens analysis: a new thing

 Image: Constructing that sens analysis: 2 steps



Constructing that sens analysis: 2 steps

for step 1, we assume that conditional on X, Z=0, pi-tilde follows this beta distribution with mean pi and dispersion param kappa

this figure shows for three values of pi (.15, .3 and .5) what different kappa values imply about the distribution of pi-tilde

we see that with kappa as small as 0.02, there is a lot of variation in the probability

Use shorthand $\tilde{\pi}_1$ for $\tilde{\pi}_1(X, Y)$

Step 1: Assume a distribution for $\tilde{\pi}_1$ with mean $\pi_1(X)$ that allows $\tilde{\pi}_1$ to vary, indexed by a dispersion param

use the beta distribution

$$\tilde{\pi}_1 \mid X, Z = 0 \sim \operatorname{Beta}\left(\pi_1(X) \frac{1-\kappa}{\kappa}, \pi_0(X) \frac{1-\kappa}{\kappa}\right)$$

Step 2: Connect *Y* to $\tilde{\pi}_1$ to induce *Y*-*C* dependence

- use quantile-to-quantile mapping between distributions of Y and of $\tilde{\pi}_1$ (given X, Z = 0)
 - same-quantiles: positive Y-C association
 - opposite-quantiles: negative Y-C association

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Constructing that sens analysis: a new thing

 Image: Constructing that sens analysis: 2 steps

Step 2: Assume a distribution for B_2 with mean $\pi_2(X)$ that allows B_1 to vary, indexed by a dispersion param \blacktriangleright use the bota distribution $\kappa_1(X,Z=0 \sim \log_2\left(-\kappa(X)\frac{1-\kappa}{2}, \kappa(X)\frac{1-\kappa}{2}\right)$

Step 2: Connect Y to f_1 to induce Y-C dependence

 use quantile-to-quantile mapping between distributions of Y and of th (given X, Z = 0)
 same-quantiles: postive Y-C association
 opposte-quantiles: negative Y-C association

for step 2

we use quantile-to-quantile mapping between the conditional distributions of Y and of pi-tilde

where same-quantile mapping (ie the 80th percentile in the distribution of Y is connected to the 80th percentile in the distribution of pi-tilde) means positive Y-C association

and opposite-quantile mapping means negative association

Constructing that sens analysis: 2 steps

Use shorthand fis for fis(X, Y)

Use shorthand $\tilde{\pi}_1$ for $\tilde{\pi}_1(X, Y)$

Step 1: Assume a distribution for $\tilde{\pi}_1$ with mean $\pi_1(X)$ that allows $\tilde{\pi}_1$ to vary, indexed by a dispersion param

use the beta distribution

 $\tilde{\pi}_1 \mid X, Z = 0 \sim \operatorname{Beta}\left(\pi_1(X) \frac{1-\kappa}{\kappa}, \pi_0(X) \frac{1-\kappa}{\kappa}\right)$

Step 2: Connect Y to $\tilde{\pi}_1$ to induce Y-C dependence

- use quantile-to-quantile mapping between distributions of Y and of $\tilde{\pi}_1$ (given X, Z = 0)
 - same-quantiles: positive Y-C association
 - opposite-quantiles: negative Y-C association

The sens param: the "signed" κ

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Constructing that sens analysis: a new thing

 Image: Constructing that sens analysis: 2 steps

Step 1: Assume a distribution for b_1 with mean $\pi_2(X)$ that allows b_1 to vary, indexed by a dispersion param **•** use the bata distribution $\approx |X, Z = 0 \sim \operatorname{biss}\left(-z|X|\frac{1-n}{2}, -z|X|\frac{1-n}{2}\right)$

Step 2: Connect Y to R1 to induce Y-C dependence

use quartile-to-quartile mapping between distributions of Y and of the (given X, Z = 0)
 tame-quartile: positive Y-C association
 consolite-supartile: resultive Y-C association

The sens param: the "signed" x

Constructing that sens analysis: 2 steps

Use shorthand th for th(X, Y

the sens param is the signed kappa

where the sign is the sign of the Y-C association

Identification and estimation

Identification:

$$\mu_{0c}(X) = \frac{\mathsf{E}[\tilde{\pi}_{c}Y \mid X, Z=0]}{\mathsf{E}[\tilde{\pi}_{c} \mid X, Z=0]},$$

where $\tilde{\pi}_c$ and Y are quantile-to-quantile connected

Estimation:

- estimate P(Y | X, Z = 0) so can compute quantiles
- then estimate $\mu_{0c}(X)$ using numerical integration

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Construction of the second second

Identification: $\mu_{In}(X) = \frac{E[t_xY \mid X, Z = 0]}{E[t_x \mid X, Z = 0]}$

where 2+ and Y are quantile-to-quantile connected

Estimation: • estimate P(Y | X, Z = 0) so can compute quantiles • then estimate $\mu_{loc}(X)$ using numerical integration

under this assumption about Y-C dependence we have this identification result

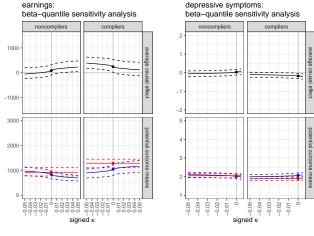
for estimation we first estimate this mixture outcome distribution so that we can compute its quantiles

and then we compute mu-Oc(x) estimates using numerical integration

this method is harder than our earlier methods because it requires estimating this distribution (or its quantile)

but as a result we benefit from more information

Example: JOBS II preliminary results

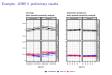


- CACE/NACE - mean Y0 - mean Y1

Results are less extreme than mean-based sens analysis

Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

L Distribution-based sensitivity analysis: a new thing
 A method using full information (ongoing work)
 L earnings: distribution-based (left), mean-based (right)



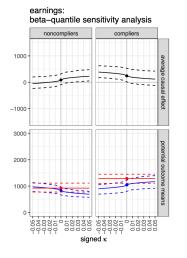
Results are less extreme than mean-based sens analysis

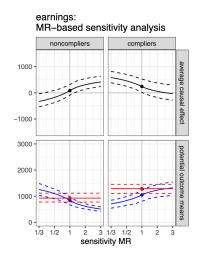
here are results from JOBS II for the two outcomes earnings and depressive symptoms for a rather wide range of signed kappa

what we notice is

these results can be less extreme than what we might get from mean-based sens analysis

earnings: distribution-based (left), mean-based (right)





Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

└─Distribution-based sensitivity analysis: a new thing └─A method using full information (ongoing work) └─depressive symptoms: distribution-based (left),

mean-based (middle and right)

you can see that in this side-by-side shows

of results from the distribution-based and mean-based sens analyses for the outcome earnings

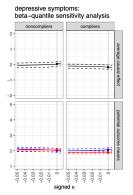


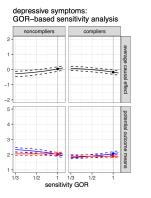
samines: distribution-based (left), mean-based (right)

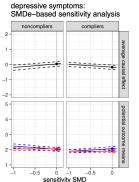
12-17

2023

depressive symptoms: distribution-based (left), mean-based (middle and right)







Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

└─Distribution-based sensitivity analysis: a new thing └─A method using full information (ongoing work) └─depressive symptoms: distribution-based (left),

mean-based (middle and right)

and here is for depressive symptoms the first plot is distribution-based the last two are mean-based

12-17

2023-

and a second sec	warate andre a	anarajan.	angine .
	· · · · · · · · · · · · · · · · · · ·	- Hermitte	H

depressive symptoms: distribution-based (left), mean-based (middle and right)

Outline

Background

Principal causal effects Principal ignorability (PI) – one identification strategy Sensitivity analysis for PI violation

Mean-based sensitivity analysis: expanding options

Allow different sensitivity parameters Accommodate different estimation methods

Distribution-based sensitivity analysis: a new thing

Method limitations and information use A method using full information (ongoing work) Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

 Image: Construction of the second sensitivity analysis: a new thing

 Image: Construction of the second sensitivity analysis: a new thing

 Image: Construction of the second sensitivity analysis: a new thing

 Image: Construction of the second sensitivity analysis: a new thing

 Image: Construction of the second sensitivity analysis: a new thing

 Image: Construction of the second s

Background Principal Caucal effects Principal ignorability (PI) – one identification strategy Sensibility analysis for PI violation

Outline

un-based sensibility analysis: expanding options Allow different sensibility parameters Accommodate different estimation methods

Distribution-based sensitivity analysis: a new thing Method line/tations and information use A method using full information (ongoing work)

to sum up

i have presented our work on sens analysis for principal ignorability violation in the estimation of complier and noncomplier average causal effects

with two approaches a mean-based approach and a distribution-based approach

THANK YOU

and i am happy to take questions