Chapter 22. Sensitivity Analysis and Bounds

SeongSik Choi

August 24, 2022

Seoul National University

1. Introduction

Relaxing the unconfoundedness assumption.

$$[W_i \perp (Y_i(0), Y_i(1))] \mid X_i$$

To satisfy this, we should observe all confounders to make the dependence between W_i and $(Y_i(0), Y_i(1))$ 'zero'.(very strong)

$\Rightarrow$ we observe confounders to make the dependence between W_i and $(Y_i(0), Y_i(1))$ 'small'.

How to assess the magnitude of violations from unconfoundedness?
How to estimate ATE under this relaxed assumption?

Two approaches : Manski, Rosenbaum-Rubin

1. Introduction

Instead of focusing on obtaining point estimates of the causal estimands of interest, we end up with ranges of plausible values for these estimands.

Recall IRS lottery data (see 14.6.2 in the textbook).

Before : The estimated [causal effects of ‘winning the lottery’ on ‘annual labor income averaged over the first six years after playing the lottery’] is -5.34 and the p-value is < 0.001 under unconfoundedness.

After : The estimated range of the [causal ($\sim$)] is $[-8.24, -2.44]$ and the p-value is 0.03 under some relaxed assumptions.

1. Introduction

In this section, we restrict the discussion to the case with binary assignment, binary outcomes and a simple case with no observed covariates.

$(W_i, Y_i(0), Y_i(1) \in \{0, 1\}$ and there isn't any observed X_i.)

2. Manski Bounds Analysis

We only observe (W_i, Y_i). We can estimate the three quantities $(p, \mu_{t,1}, \mu_{c,0})$ without unconfoundedness assumption.

where $p = \mathbb{E}(W_i), \mu_{t,1} = \mathbb{E}\left[Y_i(1)|W_i = 1\right] = \mathbb{E}\left[Y_i|W_i = 1\right]$, $\mu_{c,0} = \mathbb{E}\left[Y_i(0)|W_i = 0\right] = \mathbb{E}\left[Y_i|W_i = 0\right]$

But we don't know the quantities

$\mu_{t,0} = \mathbb{E}\left[Y_i(1)|W_i = 0\right], \mu_{c,1} = \mathbb{E}\left[Y_i(0)|W_i = 1\right]$

2. Manski Bounds Analysis

If we know the triple $(p, \mu_{t,1}, \mu_{c,0})$, we can calculate the bound for ATE τ_{sp} as

$$\tau_{\text{sp}} = \mu_{\text{t}} - \mu_{\text{c}} = (p \cdot \mu_{t,1} + (1-p) \cdot \mu_{t,0}) - (p \cdot \mu_{c,1} + (1-p) \cdot \mu_{c,0})$$
$$\in [p \cdot \mu_{t,1} - p - (1-p) \cdot \mu_{c,0}, p \cdot \mu_{t,1} + (1-p) - (1-p) \cdot \mu_{c,0}]$$

where $\mu_t = \mathbb{E}[Y_i(1)] = p \cdot \mu_{t,1} + (1-p) \cdot \mu_{t,0}$
and $\mu_c = \mathbb{E}[Y_i(0)] = p \cdot \mu_{c,1} + (1-p) \cdot \mu_{c,0}$

2. Manski Bounds Analysis

If the triple $(p, \mu_{t,1}, \mu_{c,0}) = (0.3, 0.6, 0.4)$, the bound is

$[p \cdot \mu_{t,1} - p - (1-p) \cdot \mu_{c,0}, p \cdot \mu_{t,1} + (1-p) - (1-p) \cdot \mu_{c,0}]$
= [-0.4, 0.6]

The bound is too wide.(always contains 0)

It's because we don't rule out the extreme cases

(1) $\mu_{t,0} = 1$ and $\mu_{c,1} = 0$ (upper bound)

(2) $\mu_{t,0} = 0$ and $\mu_{c,1} = 1$ (lower bound)

3. Rosenbaum-Rubin Sensitivity Analysis

We still consider the case with no observed covariates X_i.

Assume unconfoundedness given unobserved covariate $U_i \in \{0, 1\}$.

$$[W_i \perp (Y_i(0), Y_i(1))] \mid U_i$$

Define $q = \Pr(U_i = 1)$.

3. Rosenbaum-Rubin Sensitivity Analysis

We still consider the case with no observed covariates X_i.

Assume unconfoundedness given unobserved covariate $U_i \in \{0, 1\}$.

$$[W_i \perp (Y_i(0), Y_i(1))] \mid U_i$$

We consider following models:

$$\Pr(W_i = 1 \mid U_i = u) = \frac{\exp(\gamma_0 + \gamma_1 \cdot u)}{1 + \exp(\gamma_0 + \gamma_1 \cdot u)}$$
$$\Pr(Y_i(1) = 1 \mid U_i = u) = \frac{\exp(\alpha_0 + \alpha_1 \cdot u)}{1 + \exp(\alpha_0 + \alpha_1 \cdot u)}$$
$$\Pr(Y_i(0) = 1 \mid U_i = u) = \frac{\exp(\beta_0 + \beta_1 \cdot u)}{1 + \exp(\beta_0 + \beta_1 \cdot u)},$$

Note that if $(\alpha_1, \beta_1) = 0$ or $\gamma_1 = 0$, then $W_i \perp (Y_i(0), Y_i(1))$.

3. Rosenbaum-Rubin Sensitivity Analysis

There are seven scalar components of the parameter $\theta = (q, \gamma_1, \alpha_1, \beta_1, \gamma_0, \alpha_0, \beta_0)$, which we partition into two subvectors.

(1) $\theta_s = (q, \gamma_1, \alpha_1, \beta_1)$: the sensitivity parameters. We postulate (ranges of) values for them a priori.

(2) $\theta_{\mathrm{e}} = (\gamma_0, \alpha_0, \beta_0)$: the estimable parameters. We estimate them from the data.

3. Rosenbaum-Rubin Sensitivity Analysis

If $\theta_s = (q, \gamma_1, \alpha_1, \beta_1)$ and observed data captured by the triple $(\hat{p}, \hat{\mu}_{t,1}, \hat{\mu}_{c,0})$ are given, we can estimate $\theta_e = (\gamma_0, \alpha_0, \beta_0)$ by finding solution of these three equalities (see Appendix).

$$p = q \cdot \tfrac{\exp(\gamma_0+\gamma_1)}{1+\exp(\gamma_0+\gamma_1)} + (1-q) \cdot \tfrac{\exp(\gamma_0)}{1+\exp(\gamma_0)}$$

$$\begin{aligned}
\mu_{t,1} =& \frac{q \cdot \frac{\exp(\gamma_0+\gamma_1)}{1+\exp(\gamma_0+\gamma_1)}}{q \cdot \frac{\exp(\gamma_0+\gamma_1)}{1+\exp(\gamma_0+\gamma_1)} + (1-q) \cdot \frac{\exp(\gamma_0)}{1+\exp(\gamma_0)}} \cdot \frac{\exp(\alpha_0+\alpha_1)}{1+\exp(\alpha_0+\alpha_1)} \\
&+ \frac{(1-q) \cdot \frac{\exp(\gamma_0)}{1+\exp(\gamma_0)}}{q \cdot \frac{\exp(\gamma_0+\gamma_1)}{1+\exp(\gamma_0+\gamma_1)} + (1-q) \cdot \frac{\exp(\gamma_0)}{1+\exp(\gamma_0)}} \cdot \frac{\exp(\alpha_0)}{1+\exp(\alpha_0)}
\end{aligned}$$

$$\begin{aligned}
\mu_{c,0} =& \frac{q \cdot \frac{1}{1+\exp(\gamma_0+\gamma_1)}}{q \cdot \frac{1}{1+\exp(\gamma_0+\gamma_1)} + (1-q) \cdot \frac{1}{1+\exp(\gamma_0)}} \cdot \frac{\exp(\beta_0+\beta_1)}{1+\exp(\beta_0+\beta_1)} \\
&+ \frac{(1-q) \cdot \frac{1}{1+\exp(\gamma_0)}}{q \cdot \frac{1}{1+\exp(\gamma_0+\gamma_1)} + (1-q) \cdot \frac{1}{1+\exp(\gamma_0)}} \cdot \frac{\exp(\beta_0)}{1+\exp(\beta_0)}
\end{aligned}$$

These values for the estimable parameters $(\gamma_0, \alpha_0, \beta_0)$ are uniquely exist for all values of $(p, \mu_{t,1}, \mu_{c,0})$, and for all values of $\theta_s = (q, \gamma_1, \alpha_1, \beta_1)$.

And similarly with the equalities, we can express $\mu_{t,0}$ and $\mu_{c,1}$ as the functions of $(q, \gamma_1, \alpha_1, \beta_1, \gamma_0, \alpha_0, \beta_0)$.

Finally, $\tau_{\text{sp}} = \mu_t - \mu_c = p \cdot (\mu_{t,1} - \mu_{c,1}) + (1 - p) \cdot (\mu_{t,0} - \mu_{c,0})$ is a function of $(p, \mu_{t,1}, \mu_{c,0})$ and $\theta_s = (q, \gamma_1, \alpha_1, \beta_1)$.
$\tau_{\text{sp}} = \tau\left(q, \gamma_1, \alpha_1, \beta_1 \mid p, \mu_{t,1}, \mu_{c,0}\right)$.

Flow Chart :

$\tau_{\text{sp}} = \tau\left(q, \gamma_1, \alpha_1, \beta_1 \mid p, \mu_{t,1}, \mu_{c,0}\right)$.

If we know $(p, \mu_{t,1}, \mu_{c,0})$, given a set of values Θ for θ_s,

$$\tau_{\text{sp}} \in \left[\tau_{\text{low}}\ , \tau_{\text{high}}\right].$$

where
$$\tau_{\text{low}} = \inf_{(q,\gamma_1,\alpha_1,\beta_1)\in\Theta} \tau\left(q, \gamma_1, \alpha_1, \beta_1 \mid p, \mu_{t,1}, \mu_{c,0}\right),$$
$$\tau_{\text{high}} = \sup_{(q,\gamma_1,\alpha_1,\beta_1)\in\Theta} \tau\left(q, \gamma_1, \alpha_1, \beta_1 \mid p, \mu_{t,1}, \mu_{c,0}\right),$$

Note that if the components of θ_s are close to 0, it implies U is not an important confounder.

How to set a reasonable Θ?

3. Rosenbaum-Rubin Sensitivity Analysis

How to set a reasonable Θ? : We use observed covariates.

If we observe normalized covariates $X_1, \cdots, X_K$,

We can estimate the parameters of the model

$\Pr(W_i = 1 \mid X_{ik}) = \frac{\exp(\delta_{k0}+\delta_{k1}\cdot X_{ki})}{1+\exp(\delta_{k0}+\delta_{k1}\cdot X_{ki})}$

$\Pr\left(Y_i^{\text{obs}} = 1 \mid W_i, X_{ik}\right) = \frac{\exp(\zeta_{k0}+\zeta_{k1}\cdot X_{ki}+\zeta_{k2}\cdot W_i)}{1+\exp(\zeta_{k0}+\zeta_{k1}\cdot X_{ki}+\zeta_{k2}\cdot W_i)}$

and set a Θ as

$$q \in [0,1], \gamma_1 \in \left[-2 \cdot \max_k \left|\hat{\delta}_{k1}\right|, 2 \cdot \max_k \left|\hat{\delta}_{k1}\right|\right],$$

$$\alpha_1, \beta_1 \in \left[-2 \cdot \max_k \left|\hat{\zeta}_{k1}\right|, 2 \cdot \max_k \left|\hat{\zeta}_{k1}\right|\right]$$

(Note that we said we only consider the case with no observed covariates.)

3. Rosenbaum-Rubin Sensitivity Analysis

The bounds analysis can be viewed as an extreme version of a sensitivity analysis.

$\tau_{sp} = \tau\left(q, \gamma_1, \alpha_1, \beta_1 \mid p, \mu_{t,1}, \mu_{c,0}\right)$.

If we let $q = p$, $\gamma_1 \to \infty$. $\alpha_1 \to -\infty$ and $\beta_1 \to -\infty$, then

$$\tau_{sp} \to p \cdot \mu_{t,1} + (1 - p) - (1 - p) \cdot \mu_{c,0},$$

which equals to the upper limit in the Manski bounds.

Similarly, if $q = p$, $\gamma_1 \to \infty, \alpha_1 \to \infty$, and $\beta_1 \to \infty$, then

$$\tau_{sp} \to p \cdot \mu_{t,1} - p - (1 - p) \cdot \mu_{c,0},$$

4. Rosenbaum-Rubin Sensitivity Analysis for p-value

The calculated bounds are obtained when $p, \mu_{t,1}, \mu_{c,0}$ are given.

But we don't know them exactly because of sampling variation.

If we know propensity score, then we can calculate Fisher p-value.
(unconfoundedness is only used in estimating propensity score.)

For example, in lottery data, the statistic $T^{\text{dif}} = \bar{Y}_{\text{t}}^{\text{obs}} - \bar{Y}_{\text{c}}^{\text{obs}}$ is -0.12. Under bernoulli trial with assignment probability 0.47, the p-value can be calculated as 0.026.

So we can calculate maximum p-value given the bound of propensity score.

4. Rosenbaum-Rubin Sensitivity Analysis for p-value

Denote the estimated propensity score under the assumption of unconfoundedness by $\hat{e}_i$, and the actual treatment probability by p_i. For a pre-specified constant Γ, let us assume that

$$|\text{logit}(\hat{e}_i) - \text{logit}(p_i)| \leqslant \Gamma,$$

holds for all $i = 1, \ldots, N$, which gives the bounds of p_i.

$p_i \in (p_{\text{min,i}},\ p_{\text{max,i}})$, where $p_{\text{min,i}} = \text{logit}^{(-1)}(\text{logit}(\hat{e}_i) - \Gamma)$ and $p_{\text{max,i}} = \text{logit}^{(-1)}(\text{logit}(\hat{e}_i) + \Gamma)$

For each $p_1, \cdots, p_N$, we can calculate p-value.

So, we can get the bounds of p-value $\Rightarrow$ the maximum of p-value.

Appendix

The derivation of second equality (in 10p) is as follows.

$$\begin{aligned}
\mu_{t,1} &= \mathbb{E}\left[Y_i(1)|W_i = 1\right] \\
&= \Pr\left(U_i = 1 \mid W_i = 1\right) \cdot \mathbb{E}\left[Y_i(1) \mid W_i = 1, U_i = 1\right] \\
&\quad + \left(1 - \Pr\left(U_i = 1 \mid W_i = 1\right)\right) \cdot \mathbb{E}\left[Y_i(1) \mid W_i = 1, U_i = 0\right] \\
&= \frac{q \cdot \frac{\exp(\gamma_0+\gamma_1)}{1+\exp(\gamma_0+\gamma_1)}}{q \cdot \frac{\exp(\gamma_0+\gamma_1)}{1+\exp(\gamma_0+\gamma_1)} + (1-q) \cdot \frac{\exp(\gamma_0)}{1+\exp(\gamma_0)}} \cdot \frac{\exp\left(\alpha_0 + \alpha_1\right)}{1 + \exp\left(\alpha_0 + \alpha_1\right)} \\
&\quad + \frac{(1-q) \cdot \frac{\exp(\gamma_0)}{1+\exp(\gamma_0)}}{q \cdot \frac{\exp(\gamma_0+\gamma_1)}{1+\exp(\gamma_0+\gamma_1)} + (1-q) \cdot \frac{\exp(\gamma_0)}{1+\exp(\gamma_0)}} \cdot \frac{\exp\left(\alpha_0\right)}{1 + \exp\left(\alpha_0\right)}
\end{aligned}$$