IDS 702: Module 6.3

Unconfoundedness and overlap

Dr. Olanrewaju Michael Akande

Observational studies

- We will not focus on randomized experiments since most of the data you will have to analyze in practice are actually based on observational studies.
- In observational studies, we do not control or know the assignment mechanism.
- In addition, the presence of measured and unmeasured confounders can create unbalance between the groups.
- Again, to do causal inference, we have to make some structural (often untestable) assumptions, e.g. on the treatment assignment, for identifying causal effects.
- Once we have those general assumptions, we also usually have to make model assumptions to do the actual estimation.

Estimands

Once again, we will focus on the following estimands:

- The average treatment effect (ATE):

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

- The average treatment effect for the treated (ATT):

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0) | W_i = 1].$$

- The average treatment effect for the control (ATC):

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0) | W_i = 0].$$

- For binary outcomes, causal odds ratio (OR) or risk ratio (RR)::

$$\tau = \frac{\mathbb{P}\mathrm{r}[Y_i(1) = 1] / \mathbb{P}\mathrm{r}[Y_i(1) = 0]}{\mathbb{P}\mathrm{r}[Y_i(0) = 1] / \mathbb{P}\mathrm{r}[Y_i(0) = 0]}.$$

Estimands

- The relationship between ATE, ATT and ATC is given by

$$\text{ATE} = \mathbb{P}\text{r}[W_i = 1] \cdot \text{ATT} + \mathbb{P}\text{r}[W_i = 0] \cdot \text{ATC}$$

- In randomized experiments, ATT is equivalent to ATC because treatment and control groups are similar/comparable.
- ATE is then also equivalent to ATT (and ATC).
- In observational studies, ATE is usually different from ATT and ATC.
- The above relation does not hold for ratio estimands.

Assumptions: unconfoundedness

We will need two major assumptions (in addition to SUTVA). The first, we already talked about, that is,

Assumption 1: Unconfoundedness

$$Y_i(0), Y_i(1) \perp W_i | X_i,$$

or using the equivalent form from last class,

$$\mathbb{P}\mathrm{r}[W_i = 1 | X_i, Y_i(0), Y_i(1)] = \mathbb{P}\mathrm{r}[W_i = 1 | X_i]$$

- Assumes that within subgroups defined by values of observed covariates, the treatment assignment is random.
- Rules out unobserved confounders.
- Randomized experiments satisfy unconfoundedness.
- Untestable in most observational studies, but sensitivity can be checked.

Implications of unconfoundedness

- Under unconfoundedness, it turns out that

$$\mathbb{P}\mathrm{r}[Y(w)|X] = \mathbb{P}\mathrm{r}[Y^{\mathrm{obs}}|X, W = w] \quad w = 0, 1.$$

- That is, the observed distribution of Y in treatment arm $W = w$ equals the distribution of the potential outcomes $Y(w)$.

Why does this matter or how does this help us?

- Well, the causal estimands are essentially expectations and probabilities.
- Recall again that ATE is

$$\mathrm{ATE} = \mathbb{E}[Y_i(1) - Y_i(0)].$$

- ATE can then be estimated from the observed data using

$$\mathrm{ATE} = \mathbb{E}_X \left(\mathbb{E}[Y^{\mathrm{obs}}|X, W = 1] - \mathbb{E}[Y^{\mathrm{obs}}|X, W = 0] \right).$$

- Note that we need to average out over the distribution of X since the original formula for ATE does not depend on any X.

Assumptions: overlap

Assumption 2: Overlap (or positivity)

$$0 < \mathbb{P}\mathrm{r}[W_i = 1|X_i] < 1, \quad \text{for all} \quad i.$$

- Notice that this is the probabilistic assignment from last class, that is,

$$0 < \mathbb{P}\mathrm{r}[W_i = 1|X_i, Y_i(0), Y_i(1)] < 1.$$

- However, we can exclude $\{Y_i(0), Y_i(1)\}$ now because of the unconfoundedness assumption.
- $$e(x) = \mathbb{P}\mathrm{r}[W_i = 1|X_i = x]$$

  is usually called the propensity score.

Implications of overlap

- Overlap implies that, in large samples, for all possible values of the covariates, there are both treated and control units.
- This is important within the potential outcomes (or counterfactual) framework both conceptually and operationally (variance inflation).
- Unlike unconfoundedness, overlap can be directly checked from the data often using the estimated propensity scores.
- Unconfoundedness and positivity jointly define the strong ignorability assumption.

Acknowledgements

These slides contain materials adapted from courses taught by Dr. Fan Li.

What's next?

Move on to the readings for the next module!