Lecture 16

10-20-2021

What did we cover last time?

We introduced observational studies

We made sure to note that observational studies are on a range of plausibility

What are we doing today?

Continue discussing the structure of observational studies

Introduce the Propensity Score

Okay so what if we assume selection on observables?

A common estimand of interest in observational studies is the ATT

$$E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1]$$

This holds under "Strong Ignorability"

$$Y_i(0), Y_i(0) \perp || \perp D|X$$

$$0 < P[D = 1|X] < 1$$

Okay so what if we assume selection on observables?

Given strong ignorability:

$$E[Y_{ij}|D_i = 1, X_i] = E[Y_{ij}|D_i = 0, Xi]$$

In words: once we make an assumption we can continue as if treatment was randomly assigned.

By conditioning on observed covariates, we achieve balance on observables and make the assumption that the only difference between the two groups is the potential outcomes we observe.

Balancing Scores

Under unconfoundedness we remove all bias in comparison between treated and untreated units by adjusting for differences in observed covariates.

In practice, this becomes difficult as the number of covariates to balance on becomes large.

A balance score is a lower dimensional function of the covariates that suffices for removing the bias associated with differences in pretreatment variables.

Balancing Scores

Formally:

$$D_i \perp\!\!\!\perp X_i | b(X_i)$$

In words: a balancing score is a function of covariates such that the probability of receiving treatment given covariates is free of dependence on the covariates given the balancing score.

The Propensity Score is a Balancing Score

Balancing scores have a nice property that if assignment to treatment is unconfounded given the full set of covariates, then assignment is also unconfounded conditioning only on a balancing score.

The propensity score is the conditional probability of treatment (D = 1) given the observed covariates X. We tend to write it as:

$$e(X) = Pr(D = 1|X)$$

The propensity score's balancing property says that:

$$D \perp\!\!\!\perp X|e(X)$$

Propensity Scores and Interpretation

Consider a scenario where we have two units 1, 2 who are assigned to treatment and control. Each has a propensity score of 0.6.

Propensity score methods compare units based on observables who have very similar probabilities of being placed in the treatment group even though those units differed with regard to actual treatment assignment.

If conditional on covariates, two units have the same probability of being treated, then they have similar propensity scores.

Matching and Regression

An attractive feature of matching strategies is that they are typically accompanied by an explicit statement of the CIA needed to give matching estimates a causal interpretation.

A regression is a control strategy too. Both rely on a similar assumption.

> "Any regression model gives the best linear approximation to [your estimate] subject to whatever parameterization you're using. This means that I can't imagine a situation where matching makes sense but regression does not. *Josh Angrist*

Design Phase of Observational Studies

Prior to implementing any estimation strategies we need to go through the design phase of an observational study.

1. Assess Balance
2. Subsample selection if appropriate
3. Assess unconfoundedness

Assess Balance

Assessing balance means comparing the distribution of covariates in the treated and control samples.

Researchers tend to consider the difference in average covariate values by treatment status (usually tested with t-tests)

This is a balance table.

Subsample selection if appropriate

If the basic sample exhibits a substantial amount of imbalance we might do better by constructing a subsample characterized by better balance.

A common procedure to do this is to first estimate the propensity score and then match each treated unit to the closest control unit in terms of estimated propensity score.

Trimming might be appropriate if units have values of the propensity score very close to 0 or very close to 1.

Assessing Unconfoundedness

If our model of the world is true than we're good to go. However, how much do we trust our selection on observables model?

Critics (of which you are one of your own work) will claim that there is some variable that is being omitted and that variable biases your claims.

Sensitivity analysis is a process by which researchers consider how much an unobserved variable would have to affect a result to flip or change the result.

Estimating the Effect of a Treatment on an Unaffected Outcome

One sensitivity test is to estimate our treatment on a variable that we know is not affected by the treatment.

We run the same analysis, but substitute out our outcome variable with the variable not associated with treatment. If there is an effect, we are skeptical that our treatment is well explaining anything.

Note that if we find no effects, this does not imply that our treatment effect is real, just that it is more plausible.

Estimate the Effect of a different treatment

We can also do the same kind of placebo but with a treatment that we know does not cause the outcome.

Suppose we have *two* possible control groups. We can compare estimated treatment effects using the control groups because in both of these groups the treatment indicator is known to be 0.

If we find a significant effect, that suggests that for at least one of the control groups unconfoundedness is violated.

Sensititvity with different pre-treatment variables

Here we again use outcome data for all units, but partition the covariates such that we estimate the effect only a subset to those with the full set.

If there are substantial differences between the restricted and unrestricted models then unconfoundedness either relies critically on all covariates or it does not hold.

What should be reported?

1. We should always start from the raw data and smooth the data slowly, deliberately, and transparently.

2. All our analyses should leverage our design, and by extension the question of interest as much as possible. Theoretical estimands are arguably more important in the observational world.

3. Selection on observables should be regarded skeptically unless compelling evidence to the contrary is provided. Think about the seat belts example.

4. Placebo and sensitivity tests should be conducted wherever possible. We should be very skeptical of studies that did not consider at the design test how to assess their findings.