Fundamentals of causal inference: part 5

Justin Sheen

September 1, 2024

1 Introduction

The following is adapted from module five of Prof. Jason Roy's online Coursera course on causal inference https://www.coursera.org/learn/crash-course-in-causality.[1] I wanted to review the basics of causal inference for myself. This is part five of five.

2 Instrumental variables

Imagine we have the usual confounder that annoyingly gets in the way of a causal relationship: $Z \to Y$ and $Z \leftarrow C \to Y$. We now have tools to address this, namely, to randomization to get rid of $Z \leftarrow C$. But what if we also have some unmeasured confounder $Z \leftarrow U \to Y$? Or what about a situation where it would be unethical to randomize treatment Z? For example, if Z is whether pregnant mothers smoke and Y is the birth weight of the baby, we would not be able to randomize pregnant mothers to smoke or not smoke. In these cases we cannot marginalize over all confounders via matching, IPTW, or other methods.

This is where the idea of *instrumental variables* comes in. We can introduce some instrumental variable V that affects the treatment $V \to Z$ (above relationships still assumed). We can think of V as an encouragement; we can encourage pregnant mothers who do smoke to either receive encouragement to stop smoking, or to just receive the usual care they would normally receive without the encouragement. Under an *intention to treat* analysis we can then estimate $E[Y(V = 1)] - E[Y(V = 0)]$ which would be of interest as it is the causal effect of encouragement. Thus, we are estimating an effect based on randomization of the encouragement.

Note that the unmeasured confounder still exists, but there is some hope of estimating a causal effect since we can have some sort of randomization of treatment. Also note that this is not the causal effect of the smoking treatment Z itself.

Instrumental variables have an *exclusion restriction* assumption, which is that V affects Z but does not directly affect Y. Z also can not affect U. This assumption can be a problem for example if a randomized trial is not blinded: participants' knowledge of their treatment status could affect their outcome. Unfortunately this is an assumption that we can't directly check with data, much like the no unmeasured confounders assumption. There is a degree of faith and subject matter knowledge that go into these assumptions.

We want instrumental variables that are strong i.e., highly predictive of treatment. To do this we would have to estimate the proportion of compliers (discussed below). Weak instrumental variables lead to very high variance (confidence intervals wide) since the effective sample size would be very small.

2.1 Randomized trials with noncompliance

For the above relationships, notice that individuals who receive $V = 1$ may not necessarily receive the treatment Z that corresponds to not smoking, i.e., there is noncompliance. When this occurs, our randomized trial starts to look like an observational study since there may be potential confounders based on treatment received, i.e., U or C in the above may confound the relationship between Z and Y if for example older pregnant women are less likely to take the treatment and also have lower birthweight of babies.

Let's imagine each individual has two potential values of treatment. First, there is the treatment they would receive if $V = 1$ and second there is the treatment they would receive if $V = 0$, which we could denote as $Z^{V=1}$ and $Z^{V=0}$ respectively. We can then define the causal effect of treatment assignment on treatment received $E[Z^1 - Z^0]$. This is the proportion treated if everyone had been assigned treatment minus the proportion treated if no one had been assigned treatment. If there is perfect compliance this is just equal to 1. This is generally estimable from the observed data since we have randomized V, as $E[Z^1] = E[Z|V = 1]$ and $E[Z^0] = E[Z|V = 0]$.

We can also estimate $E[Y(V = 1)]$ - $E[Y(V = 0)]$ since $E[Y(V = 1)] = E[Y|V = 1]$ and $E[Y(V = 0)] = E[Y|V = 0]$. But what about the causal effect of received treatment on outcome?

2.2 Compliance classes

We can also label people depending on their compliace: if Z^1 and Z^0 are both 0, these are never-takers, if $Z^1 = 1$ and $Z^0 = 0$ these are compliers, if $Z^1 = 0$ and $Z^0 = 1$ these are defiers, if $Z^1 = 1$ and $Z^0 = 1$ these are always-takers. Compliers and defiers actually have random assignment, so we can learn something about effect of treatment. This is not the case for never-takers and always-takers.

With instrumental variables we try to estimate a local average treatment effect rather than an average causal effect, such as $E[Y(V = 1)|Z^1 = 1, Z^0 = 0] - E[Y(V = 0)|Z^1 = 1, Z^0 = 0]$, or the causal effect in the same population of compliers of assigning the encouragement V, which is equal to $E[Y(Z = 1)|Z^1 = 1, Z^0 = 0] - E[Y(Z = 0)|Z^1 = 1, Z^0 = 0]$. This is known as the complier average causal effect (CACE). Any time we contrast outcomes in the same subpopulation of people we have a valid causal effect.

In observed data, we unfortunately won't identify a person's compliance class completely: we can narrow it down to two options. For example, if an individual has $V = 0$ and $Z = 0$, then we know that their $Z^{V=0} = 0$, and they can either be never-takers or compliers. Similarly if an individual has $V = 0$ and $Z = 1$ we know their $Z^{V=0} = 1$, but we don't know whether they are always-takers or defiers. These are therefore latent, meaning they are not directly observable. Statisticians usually make a *monotonicity assumption* that there are no defiers. That is, no

one does exactly the opposite of what they are assigned. It is called the monotonicity assumption because the probability of treatment should increase with more encouragement.

3 Causal effect identification and estimation for instrumental variable analyses

Our goal is to estimate $E[Y(Z=1)-Y(Z=0)|compliers]$. We had previously discussed how to identify the ITT or intention-to-treat effect which is $E[Y(V=1)-Y(V=0)] = E[Y|V=1]-E[Y|V=0]$. We can use the old trick of conditioning and averaging among subpopulation for example to obtain $E[Y|V=1] = E[Y|V=1, \text{always takers}]P(\text{always takers}) + E[Y|V=1, \text{always takers}]P(\text{always takers}) + E[Y|V=1, \text{never takers}]P(\text{never takers}) +$ $E[Y|V=1, \text{compliers}]P(\text{compliers})$. Note that among never takers and always takers V does nothing, for example $E[Y|V=1, \text{always takers}] = E[Y|\text{always takers}]$. Also note that we don't include condition on $V=1$ in the probabilities e.g. $P(\text{never takers}|V=1) = (\text{never takers})$ since these are attributes that already exist in the population, and should have nothing to do with the randomization of V.

Given these facts, it turns out that when we take the difference $E[Y|V=1] - E[Y|V=0] = E[Y|V=1, \text{compliers}]P(\text{compliers}) - E[Y|V=0, \text{compliers}]P(\text{compliers})$. If we divide by $P(\text{compliers})$ we end up with $E[Y|V=1, \text{compliers}] - E[Y|V=0, \text{compliers}]$ on the right hand side which equals $E[Y(Z=1)|\text{compliers}] - E[Y(Z=0)|\text{compliers}]$ since V translate to Z directly since these are compliers. Note also for the left hand side that $P(\text{compliers}) = E[Z|V=1] - E[Z|V=0]$ since $E[Z|V=1]$ is the proportion that are always takers or compliers and $E[Z|V=0]$ is the proportion of always takers since there are no defiers. Therefore the CACE (causal average causal effect) is the intention to treat effect divided by the causal effect of treatment assignment on treatment received. Thus, even if we don't know exactly who are the compliers, we can still estimate this effect. Note that the ITT effect is an underestimate of CACE because some people assigned to treatment did not take it if $P(\text{compliers})$ is less than 1 (noncompliance).

3.1 Instrumental variables in observational studies

Imagine V occurs naturally. Let's say we have an observational study with very few variables collected, so it would be difficult to make a no unmeasured confounders assumption and we want to use an instrumental variables analysis.

One possible instrumental variable that has been proposed in the literature is calendar time. Sometimes one treatment, A, is used much more often than another treatment, B within a short period of time. Then $V=0$ could be the pre-period of randomization to treatment B whereas $V=1$ could be the post-period of randomization to treatment A.

But does calendar time violate the exclusion restriction? This is why the calendar time period where the change occurs should be short, since other factors can change the outcome over time e.g., general treatments for the condition the treatments were meant to treat would become better over time (hopefully).

Another popular instrumental variable is distance. For example V can be the differential time difference from a NICU to a regular hospital, Z can be treatment at a NICU and Y

can be mortality. For patients, V is then an encouragement either to be treated at a NICU or a regular hospital.

Note that for these observational studies we still think about compliance.

4 Two stage least squares estimation

Recall ordinary least squares (OLS). Suppose we have some model $Y_i = \beta_0 + \beta_1 Z_i + \epsilon_i$. The usual assumption is that treatment Z_i and the error term ϵ_i are independent. However, if there is confounding, for example people assigned to treatment have a higher error term, then they will be correlated and OLS would fail.

In the first stage of Two stage least squares (2SLS) we first regress treatment received Z on the instrumental variable V: $Z_i = \alpha_0 + \alpha_1 V_i + \epsilon_i$. We then obtain $\widehat{Z}_i = \widehat{\alpha_0} + \widehat{\alpha_1} V_i$ which is the predicted value of Z given V. So instead of the actual treatment received of an individual, we are getting, based solely on the instrumental variable, what we predict their treatment would be.

In the second stage we regress outcome on fitted value from stage one, $\widehat{Z}_i$: $Y_i = \beta_0 + \beta_1 \widehat{Z}_i + \epsilon_i$. By the exclusion restriction V is independent of Y given Z, since V only affects Y through Z. So β_1 is the estimate of the causal effect since $\beta_1 = E[Y(\widehat{Z} = 1)] - E[Y(\widehat{Z} = 0)]$ and the first term on the right hand side is equal to $\beta_0 + \beta_1$ and the second term is equal to β_0.

To interpret β_1 lets first recall that the two values of $\widehat{Z}$ are $\widehat{\alpha_0}$ or $\widehat{\alpha_0} + \widehat{\alpha_1}$, which if there is some noncompliance are not equal to 0 or 1. Going from the former term to the latter is essentially what occurs when we go from $V = 0$ to $V = 1$. Thus, the mean difference in the intention to treat effect $\widehat{E}[Y(V = 1)] - \widehat{E}[Y(V = 0)]$ occurs with a $\widehat{\alpha_1}$ change in $\widehat{Z}$, in other words if V changes from 0 to 1 then $\widehat{Z}$ changes by $\widehat{\alpha_1}$ units.

If we see a change in the mean of Y for a $\widehat{\alpha_1}$ unit change of $\widehat{Z}$ of the ITT effect, then for a whole unit change of $\widehat{Z}$ we should see an effect of $\frac{\widehat{E}[Y(V=1)] - \widehat{E}[Y(V=0)]}{\widehat{\alpha_1}}$. Then where does that leave β_1? This is equal to CACE $= \frac{E[Y|V=1] - E[Y|V=0]}{E[Z|V=1] - E[Z|V=0]}$ since the denominator is exactly $\widehat{\alpha_1}$.

2SLS is applicable for situations with many covariates or non-binary data such as a continuous instrumental variable. All we would have to do is in the first stage regress Z on V and covariates X, obtaining fitted values of $\widehat{Z}$, then in the second stage regress Y on $\widehat{Z}$ and X.

Sensitivity analyses on the exclusion restriction assumption and monotonicity assumption should be done to ask, if V affects Y by some amount, would conclusion change, or if there are some percentage of defiers, what percentage would there have to be for conclusions to change, respectively.

References

[1] J. Roy. A crash course in causality: Inferring causal effects from observational data. `https://www.coursera.org/learn/crash-course-in-causality`.