Causal Inference - HW1

Afek Adler

November 25, 2019

Question 1

Using potential outcomes notation, give an example of a data generating process (a joint distribution) which includes a hidden confounder H, a binary treatment T, and two potential outcomes Y_0 and Y_1, such that:

Ignorability does not hold, and

$$E(Y_1 - Y_0) \neq E[Y|T=1] - E[Y|T=0]$$

where

$$Y = T * Y_1 + (1-T) * T_0$$

Solution

Let

Y_0	Y_1	H	T	Y
1	1	1	1	1
0	1	0	0	0
0	0	1	0	0

$$E[Y|T=1] - E[Y|T=0] = 1 - 0 = 1 \neq E(Y_1 - Y_0) = E(Y_1) - E(Y_0) = 0.666 - 0.333 = 0.333$$

We can see that ignorability does not hold:

$$(Y_0, Y_1) \not\perp\!\!\!\perp T|H \iff p(Y_0, Y_1, T|H) = p(Y_0, Y_1|H) \cdot p(T|H)$$

But:

$$p(Y_0 = Y_1 = T = 0|H = 1) = 0.5 \neq p((Y_0 = Y_1 = 0|H = 1) \cdot p(T = 0|H = 1) = 0.5 * 0.5 = 0.25$$

Question 2

Let $(x_1, t_1, y_1).....(x_n, t_n, y_n)$ be a sample from a randomized con- trolled trial, where for each $i = 1...n, x_i \in R$ is a covariate measured before treatment assignment, $t_i \in \{0, 1\}$ is a binary treatment, and $y_i \in R$ is an outcome measured after the treatment. Let $\pi : R \to \{0, 1\}$ be a policy: a function which for each value of the covariate x assigns a treatment 0 or 1. Let $V(\pi) = E(y|t = \pi(x))$ be the value of the policy $\pi(x)$: what we expect the outcome to be if treatment were assigned according to $\pi(x)$, as opposed to randomly. Give an unbiased estimator of $V(\pi(x))$ for the sample above.

Solution

We know that there was an RCT that "turned" to be an observational study. I will note the the RCT treatment indicator as T. We can look at it from two viewpoints:

(a) Law of total expectation:
Let

$$V(\pi) = E(y|t = \pi(x)) = E(y|\pi(x), T = 1) \cdot p(T = 1) + E(y|\pi(x), T = 0) \cdot p(T = 0)$$

This is already enough as an unbiased estimator since:

$$E(E(y|\pi(x), T = 1)\cdot p(T = 1)+E(y|\pi(x), T = 0)\cdot p(T = 0)) = E(y|\pi(x), T = 1)\cdot p(T = 1)+E(y|\pi(x), T = 0)\cdot p(T = 0$$

Becuase it's an RCT we can assume $p(T = 1) = p(T = 0) = 0.5$
If needed, can further develop:

$$E(y|\pi(x), T = 1)\cdot p(T = 1) = E(y|\pi(x) = 0, T = 1)\cdot p(T = 1, \pi(x) = 0)+E(y|\pi(x) = 1, T = 1)\cdot p(T = 1, \pi(x) = 1)$$

And vice versa for $E(y|\pi(x), T = 1) \cdot p(T = 0)$

WE can interperet it as that the unbiased RCT turned to observational study, with the probability of agreement or disagreement as the weighted scheme.

(b) And IPS:

$$V(\pi) = \frac{1}{n} \sum_{t_i=1} \frac{y_i}{\hat{p}\,(t_i = 0|x_i)} - \frac{1}{n} \sum_{is.tt_i=0} \frac{y_i}{\hat{p}\,(t_i = 0|x_i)} \tag{1}$$

Note - when ϕ is deterministic we have a problem because $\hat{p}\,(t_i = 0|x_i) = 1$ - lack of common support.

Question 3

Give an example of a dataset with features X and one or more observed outcome variables $Y_1....Y_k$. For this dataset give:

1. Two examples of interesting causal questions relating one of the features and one of the outcomes. Explain what would be the treatment and what would be the potential outcomes in this case.

2. Two examples of interesting prediction questions which do not require causal reasoning.

Examples can come from the fields of politics, biology, sports, economics, entertainment, medicine, transportation and so on - use your imagination.

Solution

1. Given a dataset of a music provider (e.g, Spotify) that contains all the history of millions of users. An interesting question to ask is who is which artist is 'better' for the genre. For example, the treatment can be to recommend songs of Kanye or Jay Z and the potential outcome will be how many hip hop songs they listened to after a month. Another interesting question that use the same treatment is how the recommendation will affect the overall music session (the time that the user listens to music in the app since the recommendation until he closes the app).

2. Prediction - time-series forecasting of which band will explode (become famous) in the next t months. How many hours will a specific person listen to tomorrow. How much my taste in music is different than my friends. Creating a tinder app based on music similarity.