More collider bias

William Lowe

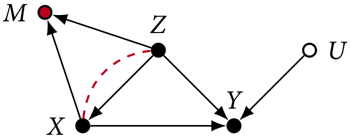
Hertie School

2nd November 2020

Plan

- → Matching (as shown in the lecture)
- → Explaining away in the Towers example (with all the detail worked out)

Matching



In balanced matching, $X \longrightarrow Z$ is exactly offset by X- - - - Z.

We have manufactured 'unfaithfulness' to the graph

Z is gender (50% male, 50% female) and X is treatment.

$$\begin{aligned} P(X = 1 \mid Z = \text{male}) &= 3/4 \\ P(X = 1 \mid Z = \text{female}) &= 1/2 \end{aligned}$$

So

$$P(Z = \text{male} \mid X = 1) = 5/8$$

But now we match each X=1 case with a X=0 of the same gender and set M=1 for each

Now, by careful construction

$$P(Z = \text{male} \mid X = 1, M = 1) = 1/2$$

Selection and control

Lots of relationships have this 'unfaithful' offsetting structure

- → Equilibrium: equal and opposite forces are offset, e.g. in game theory, deterrence, etc.
- → Control: 'Keep this quantity constant!', e.g. driving, policing, planned economies
- → Selection processes: e.g. in biology

It looks like there is no relationship because several have cancelled each other.

This is, as yet, understudied in causal inference (see Imbens, 2020, on modelling supply-demand curves)

Towers

First, let C (CIA conspiracy) and T (terrorist plot) be independently distributed 0/1 variables

We'll model the conditional probability of D (tower destruction) as a 'noisy or'

$$P(D = 1 \mid C, T) = 1 - (1 - c)^{C} (1 - t)^{T}$$

with c and t as success parameters (also probabilities).

For concreteness we'll set

$$\begin{aligned} P(C = 1) &= 0.3 \\ P(T = 1) &= 0.2 \\ c &= 0.7 \\ t &= 0.9 \end{aligned}$$

though the exact choices won't matter for the effect

Towers

Here are the complete prior $P(C, T)$, conditional $P(D \mid C, T)$, and joint $P(D, C, T)$ probability distributions

C	T	D	$P(C, T)$	$P(D \mid C, T)$	$P(D, C, T)$
0	0	0	0.56	1.00	0.56
1	0	0	0.24	0.10	0.02
0	1	0	0.14	0.30	0.04
1	1	0	0.06	0.03	0.00
0	0	1	0.56	0.00	0.00
1	0	1	0.24	0.90	0.22
0	1	1	0.14	0.70	0.10
1	1	1	0.06	0.97	0.06

Notice that $C = 1$ & $T = 1$ is more likely to result in $D = 1$ than either C or T alone, as we'd expect

Towers

We can get the posterior probability distribution $P(C, T \mid D = 1)$ various ways, but simply selecting D=1 cases from the joint distribution and renormalizing is the easiest

C	T	$P(C, T \mid D = 1)$
0	0	0.00
1	0	0.58
0	1	0.26
1	1	0.16

Notice that now both $C = 1$ & $T = 0$ and $C = 0$ & $T = 1$ are more likely than $C = 1$ & $T = 1$.

→ That is the 'explaining away' effect.

References

Imbens, G. W. (2020, March 22). *Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics* (arXiv No. 1907.07271).