

Chapters 4 to 7

Vladimir Feinberg

April 25, 2019

Notes on Chapters 4 to 7 of Pearl and Mackenzie's *Book of Why* [?].

4 Confounding

The basic confounder (Fig. 1).

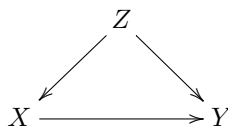
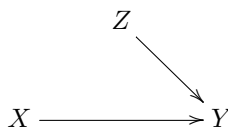


Figure 1

The true causal effect $X \rightarrow Y$ is confounded by the fork centered at Z . Suppose Z positively correlates with X and Y , and X is high. Then by virtue of X being high, Z was probably high, which causes Y to be high; so in this way if we observe (X, Y) pairs alone it's unclear how much X causes increases in Y alone.

Resolving this bias. This can happen in two ways: if we can control application of X , we can perform a randomized controlled trial. Otherwise, we need to control for the confounder Z , and measure the effect at each of the individual values of Z .

In other words, we're interested in what happens when we set X to a particular value of x , denoted $\text{do}(X = x)$. In this scenario, we consider an alternative universe where we choose X , reflected in the following diagram:



In general, performing a do removes all in-arrows to the node(s) being assigned, and then considers a probabilistic model with the remaining connections in the Bayes net.

When controlled treatment is possible, an RCT achieves this by applying treatment randomly, which effectively removes all in-arrows.

4.1 Creating Independence

Recall the rules for independence in isolation of structures in Bayes nets:

1. In a chain $X \rightarrow Y \rightarrow Z$, controlling for Y makes $X \perp\!\!\!\perp Z|Y$.
2. In a fork, $X \leftarrow Y \rightarrow Z$, controlling for Y again has the same effect (confounder example, above).
3. In a collider, $X \rightarrow Y \leftarrow Z$, $X \perp\!\!\!\perp Z$ by default, but conditioning for Y *induces* dependence!

If any link along a chain has independent endpoints, the entire chain has independent endpoints.

As an example of collider bias being induced by control is as follows. Consider the collider $X \rightarrow Y \leftarrow Z$ where X is the sprinklers were turned on, Y is the grass is wet, and Z is it rained yesterday.

If the grass is wet (conditioning/controlling), then the probability that it rained yesterday given that the sprinklers were not turned on is much higher than usual, since something explains why the grass is wet.¹

4.2 The Backdoor Criterion

If there are multiple paths from X to Y , then the backdoor criterion is the condition where all chains starting with an arrow pointing into X are “blocked” by the rules outlined above.

If all such paths are blocked, all that’s left is the causal effect of X on Y .

p159 the “games”

(Working through the examples in book of why ch 4), just go through them all.

What to do when there are multiple paths from X to Y with no colliders? How to compute $\text{do}(X)$ —
only remove all in-arrows. * would be double-counting causal effects * E.g., add arrow $A \rightarrow D$ to game 2

Backdoor is all paths that have in-arrows into X .

Use rules a-d to de-confound.

game 5 instead of having C , what if you had an arrow $Y \rightarrow B$ (cyclic). or $B \rightarrow Y$. then not solvable.

* Do cycles appear in practice, or can you “unwind them” (successful entrepreneur — funding — exit)

5 Smoking

This chapter discusses how causal analysis could resolve the concern over identifying smoking as a leading cause of lung cancer.

At the core of it, it was unclear whether smokers X (an observable property), who were observed to have higher rates of lung cancer Y , did not simply have a smoking gene Z which explained a propensity for smoking and getting lung cancer simultaneously.

¹How does this model behave when there are other possible explanations for the grass being wet, like your neighbor watered it by accident, or similar missing nodes?

In other words, did our world look like Eq. 1 or Eq. 2?



Of course, the smoking industry preferred Eq. 2.

Distinguishing the above two scenarios requires what developed from Cornfield's sensitivity analysis. Nowadays this can probably be done with a variety of tools for [model selection](#).

Worked smoking example

So do you put an arrow diagram 5.2 -i 5.1 or not? Need sensitivity analysis (Cornfield) – how strong would gene smoking need to be in order to explain.

how small does the value found in the arrow have to be before we should ignore it as non-existent?

—

Paradoxes

monty hall bias

simpson's paradox

observational non rct bias -i collider

"if you know you're a man, you shouldn't take the drug, if you know you're a woman, you shouldn't, etc."

but people in general should take drugs.

The problem is we think of every association causally (regardless of the summarization dimension). But different summarizations can conflict in causal prescription. The root problem is you can't go to a causal prescription (take the drug in order to get better).

Actually answering that question requires a causal model. (tbl 6.4)

Lord's paradox = simpsons but numerical. Could be fixed

can we measure the error of the diagram? critical for making trial observation choices.

—

work an example of backdoor adjustment for a binary confounder.

frontdoor adjustment work out ex in 7.1 (for removing mediator effect)

do calculus

3 rules

they are complete, everything that a causal network implies, they'll derive.

What is known * can reduce all reducible do expressions into observational ones (decidable; not every one can be removed for a given study) * can extend this to enabling randomization of possible subgroups of other variables