Probabilistic counterfactuals

With an SCM, we can compute values of variables (including counterfactual ones) for each vector u:

u	$X(u)$	$Y(u)$	$Y_1(u)$	$Y_2(u)$	$Y_3(u)$	$X_1(u)$	$X_2(u)$	$X_3(u)$
1	1	2	2	3	4	1	1	1
2	2	4	3	4	5	2	2	2
3	3	6	4	5	6	3	3	3

u	P(u)
1	1/2
2	1/3
3	1/6

We can now compute things like:

- Probabilities of counterfactuals (e.g. $P(Y_2 = 3)$)
- Joint probabilities of observed/counterfactual variables ($P(Y_2 > 3, Y_1 < 4)$)
- Conditional probabilities ($P(Y_3 > Y \mid Y > 2)$)

Generalizing 'conditional interventions'

Recall that with a *conditional intervention*

$$P(Y = y \mid \mathrm{do}(X = g(Z)))$$

(which we defined in terms of $P(Y \mid \mathrm{do}(X = x), Z = z)$), we could only deal with the case where X is not a cause of Z

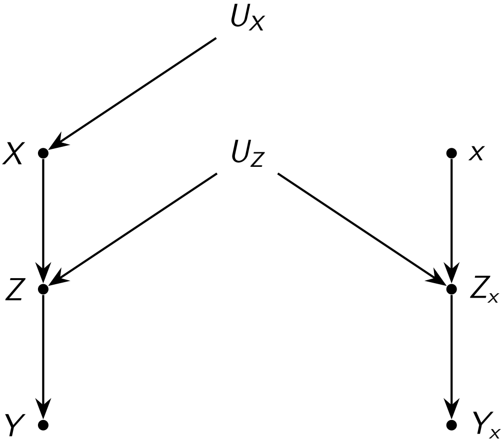
If X *is* a cause of Z, the query becomes counterfactual:

- we need the original Z to determine whether/how to intervene on X;
- under that intervention, we get a new $Z_{X=\ldots}$

(Special case: if X and Z above are the same variable, e.g. additive intervention: see later)

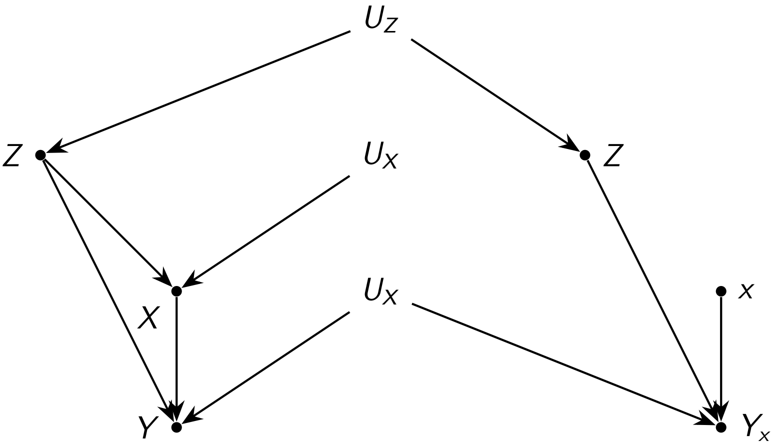
Twin networks

Twin network: draw original and modified model in same graph, with shared exogenous variables:



Conditioning on Z *opens* a d-connecting path between X and Y_x!

Twin networks: closing a path



In this graph, Y_x and X are dependent, but conditioning on Z makes them independent

Backdoor criterion for counterfactuals

Theorem

If a set Z satisfies the backdoor criterion relative to the pair (X, Y) in the unmodified graph, then, for all x, Y_x is conditionally independent of X given Z:

$$P(Y_x \mid X, Z) = P(Y_x \mid Z)$$

Adjustment formula for counterfactuals

$$
\begin{aligned}
P(Y_x = y) &= \sum_z P(Y_x = y \mid Z = z) P(z) \\
&= \sum_z P(Y_x = y \mid Z = z, X = x) P(z) \\
&= \sum_z P(Y = y \mid Z = z, X = x) P(z)
\end{aligned}
$$

(last step: consistency rule)

So we have an adjustment formula for counterfactuals! (Not surprising in this case, because $P(Y_x = y) = P(Y = y \mid \text{do}(X = x)$; more interesting example coming later)

Estimating $E(Y_x)$ from a randomized experiment

What if we don't know the structural equations?

In a randomize controlled trial where X is assigned 0 or 1 at random, we only observe half of the values, but they still give a consistent estimate of $E[Y_x]$ (using adjustment formula with $Z = \varnothing$)

	Predicted potential outcomes		Observed outcomes	
Participant	Y_0	Y_1	Y_0	Y_1
1	1.05	1.95	1.05	■
2	0.44	1.34	■	1.34
3	0.56	1.46	■	1.46
4	0.50	1.40	■	1.40
5	1.22	2.12	1.22	■
6	0.66	1.56	0.66	■
7	0.92	1.82	■	1.82
8	0.44	1.34	0.44	■
9	0.46	1.36	■	1.36
10	0.62	1.52	0.62	■

Nonidentifiability in general

Suppose we don't know the SCM but are considering three possibilities, differing in the structural equation for Y:

Model 1: $Y = f(X, \ldots) + U_Y$ ($\leftarrow$ *additive noise*)

Model 2: $Y = f(X, \ldots) + XU_Y + (1 - X)U'_Y$, where U'_Y is iid with U_Y

Model 3: $Y = f(X, \ldots) + XU_Y - (1 - X)U_Y$ (for U_Y symm. about 0)

The three models give different answers for counterfactual queries about an individual (even if we know u)

But any distribution we can measure (both observational and experimental) will be the same for all three models — so we can never choose one based on data!

Practical uses 1: Recruitment to a program

Example 4.4.1 *A government is funding a job training program aimed at getting jobless people back into the workforce. A pilot randomized experiment shows that the program is effective; a higher percentage of people were hired among those who finished the program than among those who did not go through the program. As a result, the program is approved, and a recruitment effort is launched to encourage enrollment among the unemployed, by offering the job training program to any unemployed person who elects to enroll.*

Lo and behold, enrollment is successful, and the hiring rate among the program's graduates turns out even higher than in the randomized pilot study. The program developers are happy with the results and decide to request additional funding.

Counterargument: the effectiveness now that people can *choose* to enroll might be different from what was measured

Question: What kind of variable could lead to such a difference:

- Place in the graph?
- Role in the story?

Solution

Variables (both binary): X is training, Y is hiring

To settle the debate, we want to compute the effect of treatment on the treated:

$$\begin{aligned} ETT &= E[Y_1 - Y_0 \mid X = 1] \\ &= E[Y_1 \mid X = 1] - E[Y_0 \mid X = 1] \\ &= E[Y \mid X = 1] - \textbf{?} \end{aligned}$$

Adjustment formula: novel case

Suppose set Z satisfies the backdoor criterion. Then

$$\begin{aligned}
P(Y_x \mid X = x') &= \sum_z P(Y_x \mid Z = z, X = x') P(Z = z \mid X = x') \\
&= \sum_z P(Y_x \mid Z = z, X = {\color{red} x}) P(Z = z \mid X = x') \\
&= \sum_z P(Y \mid Z = z, X = x) P(Z = z \mid X = x')
\end{aligned}$$

This is really something new: $P(Y_x \mid X = x')$ does not reduce to $\mathrm{do}(\cdot)$-notation, and the weighting factor $P(Z = z \mid X = x')$ is different than for $P(Y \mid \mathrm{do}(X = x))$

Solution, continued

Using $P(Y_0 \mid X=1) = \sum_z P(Y \mid Z=z, X=0) P(Z=z \mid X=1)$, we can finish the computation of the ETT:

$$
\begin{aligned}
ETT &= E[Y_1 - Y_0 \mid X=1] \\
&= E[Y_1 \mid X=1] - E[Y_0 \mid X=1] \\
&= E[Y \mid X=1] - \sum_y y P(Y_0 = y \mid X=1) \\
&= E[Y \mid X=1] - \sum_y y \sum_z P(y \mid z, X=0) P(z \mid X=1) \\
&= E[Y \mid X=1] - \sum_z P(z \mid X=1) \sum_y y P(y \mid z, X=0) \\
&= E[Y \mid X=1] - \sum_z P(z \mid X=1) E[Y \mid z, X=0]
\end{aligned}
$$

Practical uses 2: Additive interventions

Example 4.4.2 *In many experiments, the external manipulation consists of adding (or subtracting) some amount from a variable X without disabling preexisting causes of X, as required by the do(x) operator. For example, we might give 5 mg/l of insulin to a group of patients with varying levels of insulin already in their systems. Here, the preexisting causes of the manipulated variable continue to exert their influences, and a new quantity is added, allowing for differences among units to continue. Can the effect of such interventions be predicted from observational studies, or from experimental studies in which X was set uniformly to some predetermined value x?*

Question: How to formulate this intervention?

Practical uses 2: Solution

Abbreviate this intervention (increasing X by q) as $add(q)$

Formulation in terms of counterfactuals:

$$E[Y \mid add(q)] - E[Y] = \sum_x E[Y_{x+q} \mid X = x]P(X = x) - E[Y]$$

Suppose Z satisfies the backdoor criterion. Adjustment formula:

$$= \sum_x \sum_z E[Y \mid X = x + q, Z = z]P(Z = z \mid X = x)P(X = x) - E[Y]$$

Practical uses 2: Discussion

Why did we need counterfactuals?

While we could find effect by performing the experiment as described, such an experiment is not of the usual form (and not expressible with do($\cdot$):

- The outcome of such an experiment depends on the original distribution $P(X)$. So unlike do($\cdot$)-type causal effects, the findings are not *invariant*, so less useful

Practical uses 3: Personal decision making

Example 4.4.3 *Ms Jones, a cancer patient, is facing a tough decision between two possible treatments: (i) lumpectomy alone or (ii) lumpectomy plus irradiation. In consultation with her oncologist, she decides on (ii). Ten years later, Ms Jones is alive, and the tumor has not recurred. She speculates: Do I owe my life to irradiation?*

Mrs Smith, on the other hand, had a lumpectomy alone, and her tumor recurred after a year. And she is regretting: I should have gone through irradiation.

Can these speculations ever be substantiated from statistical data? Moreover, what good would it do to confirm Ms Jones's triumph or Mrs Smith's regret?

What are relevant quantities for these two people?

Practical uses 3: Discussion

X and Y binary: $X = 1$ means irradiation; $Y = 1$ means remission

$P(Y \mid \mathrm{do}(X = 1))$: gives answer for population, but not for people in these specific situations

Ms Jones want to know the probability of necessity:

$$PN = P(Y_0 = 0 \mid X = 1, Y = 1)$$

Mrs Smith want to know the probability of sufficiency:

$$PS = P(Y_1 = 1 \mid X = 0, Y = 0)$$

Third quantity (necessary and sufficient):

$$PNS = P(Y_1 = 1, Y_0 = 0)$$

Practical uses 3: Discussion, continued

In general, these probabilities can't be found from observational/experimental data

- We may be able to compute lower and upper bounds for them
- We can find them if we know that for all u, $Y_1(u) \geq Y_0(u)$ ('monotonicity')

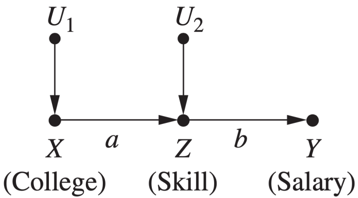
(Details in section 4.5.1; not part of this course)

Practical uses 3: Discussion, continued

Why would we want to know these probabilities?

- Emotional: regret, success
- Legal: blame (see also example 4.4.4 in the book)
- Learning: Gives feedback on our strategy for coming to a decision

Extra question [based on study question 4.3.1]



$$X = U_1$$
$$Z = X + U_2$$
$$Y = Z$$

Assume that U_1 and U_2 are distributed as
$P(U_i = 1) = P(U_i = -1) = \frac{1}{4}$, $P(U_i = 0) = \frac{1}{2}$
Find the expected salary of workers with skill level $Z > 0$ had they received x college education.

Study question 4.4.1 (erratum: (c) should refer to Theorem 4.3.1)

Table 1.1 Results of a study into a new drug, with gender being taken into account

	Drug	No drug
Men	81 out of 87 recovered (93%)	234 out of 270 recovered (87%)
Women	192 out of 263 recovered (73%)	55 out of 80 recovered (69%)
Combined data	273 out of 350 recovered (78%)	289 out of 350 recovered (83%)