Recap: causal inference and randomized controlled trials

Rasmus Waagepetersen

October 14, 2025

Causation and association

For a person let Y denote amount of icecream bought a given day, A be an indicator of whether the person wore shorts (1) or not (0), and let W denote outdoor temperature.

We might expect to see a correlation/association between A and Y but this may not reflect a causal relation.

We could estimate $\mathbb{E}[Y|A=1] - \mathbb{E}[Y|A=0]$ but this can not in general be viewed as a causal effect due to confounding by W.

Basic question: how do we define causal effects ?

Essentially two approaches:

- Interventions in structural causal models
- Counterfactuals

Structural causal models (e.g. Pearl et al.: Causal inference in statistics)

Represent relation between variables using equations (structural causal model, SCM):

$$W = f(U_W)$$
$$A = f(W, U_A)$$
$$Y = f(A, W, U_Y)$$

Here I sloppily use f as a generic notation for a function evaluated on one or more random variables. The variables U_W, U_A, U_Y are independent random variables that generate the stochasticity in the model.

The model gives rise to a joint density $p(y, a, w)$ (again I use p as generic notation for a density) which can naturally be factorised as

$$p(y, a, w) = p(y|a, w)p(a|w)p(w)$$

and the model can be visualised in terms of a directed graph (third-last slide)

Intervention

We can now formulate causal effects in terms of interventions which in the simplest form means fixing a variable on a given variable. E.g. fixing $A = 1$. Then A is non-random and the previous SCM reduces to

$$W = f(U_W)$$
$$A = 1$$
$$Y = f(1, W, U_Y)$$

This gives rise to a new joint distribution of Y and W (recall A is now fixed):

$$p_{do(A=1)}(y, w) = p(y|1, w)p(w)$$

("$do(A = 1)$" is Pearl's notation for intervening on a variable. *Not* the same as conditioning on $A = 1$)

Intervention

The SCM after fixing $A = 1$ corresponds to a "world" where A is always 1 but the generating mechanisms for all other variables are left unchanged.

This differs from conditioning on $A = 1$ since unless variables are independent, conditioning on one variable will alter the distributions of the other variables.

The mean of Y under the modified distribution is

$$\mathbb{E}_{do(A=1)}[Y] = \int \mathbb{E}[Y|A=1, W=w]p(w)\mathrm{d}w \qquad (1)$$

which in general differs from

$$\mathbb{E}[Y|A=1] = \int \mathbb{E}[Y|A=1, W=w]p(w|1)\mathrm{d}w$$

where $p(w|1)$ is conditional density of W given $A=1$.

When are the two expectations equal ?

Causal effect of changing A from 0 to 1 is now defined as

$$\mathbb{E}_{do(A=1)}[Y] - \mathbb{E}_{do(A=0)}[Y]$$

Formula (1) goes under various names: backdoor formula, g-formula, standardization,...

Estimation

In practice we may use plug-in estimate (g-computation)

$$\hat{\mathbb{E}}_{do(A=1)}[Y] = \frac{1}{n}\sum_{i=1}^{n}\hat{\mathbb{E}}[Y|A=1, W=w_i]$$

or inverse probability weighting (IPW)

$$\hat{\mathbb{E}}_{do(A=1)}[Y] = \frac{1}{n}\sum_{i=1}^{n}\frac{1[a_i=1]y_i}{\hat{p}(1|w_i)}$$

Counterfactuals (Donald Rubin)

Let $Y(0)$ and $Y(1)$ represent counterfactual outcomes in hypothetical worlds where A is either 0 or 1. In reality we observe Y which is either $Y(0)$ or $Y(1)$. Average causal effect (ATE):

$$\mathbb{E}Y(1) - \mathbb{E}Y(0)$$

Identifying assumptions (cf. Emilie's lecture)

- $(Y(0), Y(1))$ and A conditionally independent given W (conditional unconfoundedness)
- $P(A = a|W = w) > 0$ for all w and a (positivity)
- $Y = Y(A)$ (consistency)

Under these assumptions we saw in Emilie's lecture that

$$\mathbb{E}[Y(a)] = \mathbb{E}[\mathbb{E}[Y|A = a, W]]$$

which precisely equals (1)

Randomized trial

A randomized trial can be represented as

$$W = f(U_W)$$
$$A = U_A$$
$$Y = f(A, W, U_Y)$$

where U_A is a Bernouilli variable. In this case we do have

$$\mathbb{E}_{do(A=a)}[Y] = \mathbb{E}[Y(a)] = \mathbb{E}[Y|A = a]$$

which could simply be estimated by

$$\frac{\sum_{i=1}^{n} y_i 1[a_i = a]}{\sum_{i=1}^{n} 1[a_i = a]} \text{ or } \frac{\sum_{i=1}^{n} y_i 1[a_i = a]}{np(a)}$$

Hence we can evaluate causal effect regardless of whether confounder W is observed or not ! In case of observational studies we must rely on observing all confounders which may or may not be true.

Relation SCM and counterfactuals

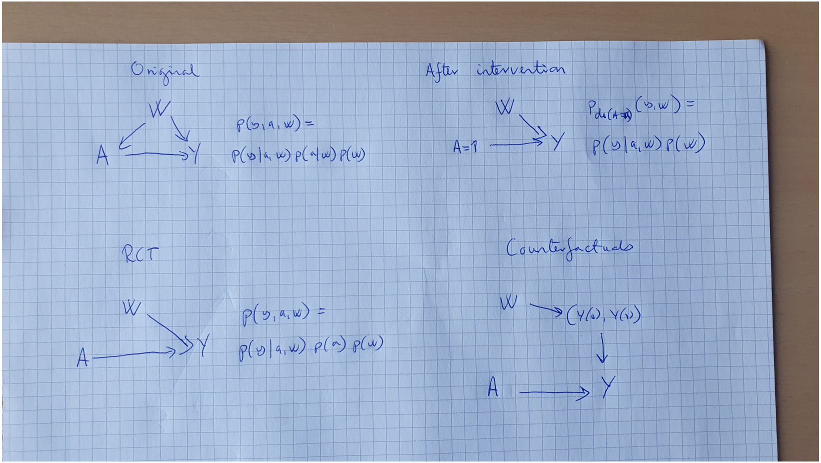
We could modify SCM as follows to explicitly represent counterfactuals:

$$
\begin{aligned}
W &= f(U_W) \\
A &= f(W, U_A) \\
Y(0) &= f(0, W, U_Y) \\
Y(1) &= f(1, W, U_Y) \\
Y &= Y(A)
\end{aligned}
$$

Then A and $(Y(0), Y(1))$ are indeed conditionally independent of A given W and $Y = Y(A)$.

Note: joint distribution of all 5 variables is degenerate since Y is deterministic given specific values of $A, Y(0), Y(1)$.

Graphs



Original
W
A
Y
p(y, a, w) =
p(y|a, w) p(a|w) p(w)
After intervention
W
A=1
Y
P_do(A=1)(y, w) =
p(y|a, w) p(w)
RCT
W
A
Y
p(y, a, w) =
p(y|a, w) p(a) p(w)
Counterfactuals
W
(Y(0), Y(1))
A
Y

We've only considered the very basics

Huge literature on causal inference including books by Pearl et al., Hernan and Robins, Jonas Peters et al. etc.

Exercise

Show that plug-in estimate and IPW estimate are unbiased if true expectations an probabilities are used:

$$\hat{\mathbb{E}}_{do(A=1)}[Y] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[Y|A=1, W=w_i]$$

or inverse probability weighting (IPW)

$$\hat{\mathbb{E}}_{do(A=1)}[Y] = \frac{1}{n} \sum_{i=1}^{n} \frac{1[a_i = a]y_i}{p(a|w_i)}$$