Inference and learning in causal models: Lecture 2

Tsinghua Logic Summer School

Tadeg Quillien

University of Edinburgh

2026

Inference by sampling

Inference by sampling

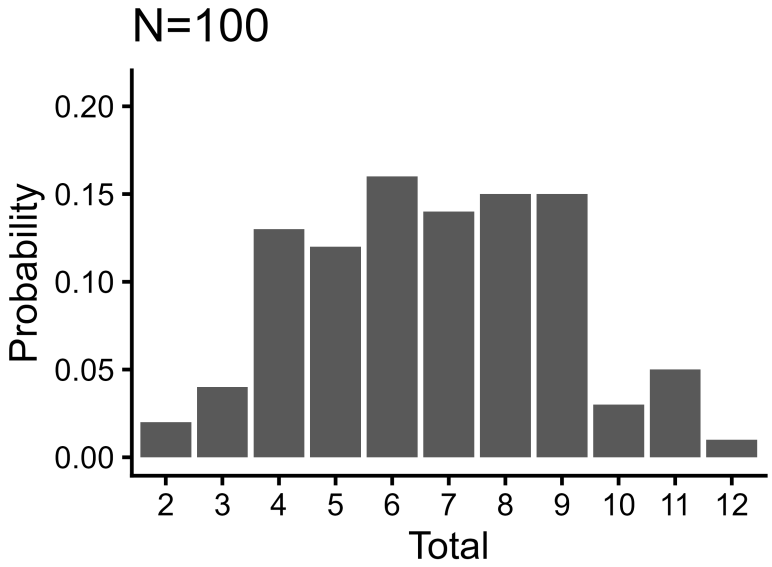
In the previous lecture: we can use the equations of probability theory to compute inferences.

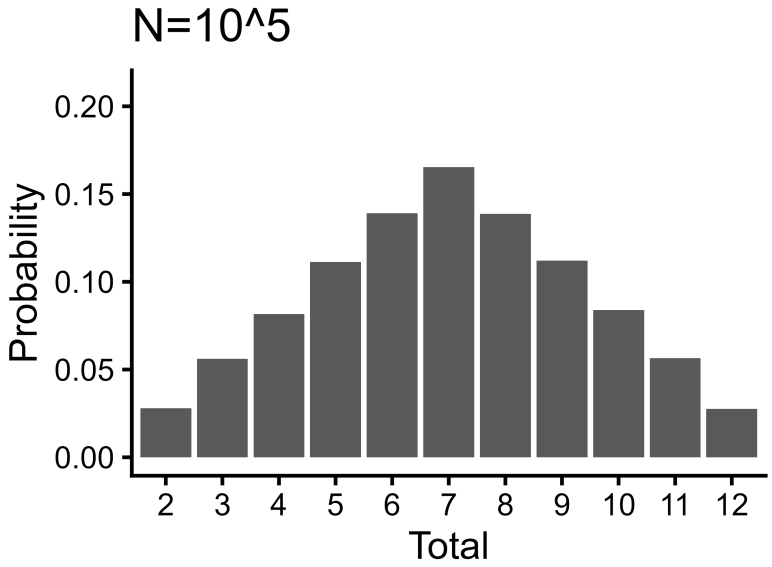
- But these computations can get quite complicated.
- In complex cases, this approach to calculating probabilities will often not work.
- An alternative solution is to use **sampling**.

Sampling: a simple example

What is the probability of getting 6 when rolling two die?

- Simulate many die rolls
- Count the proportion of times where you get 6.



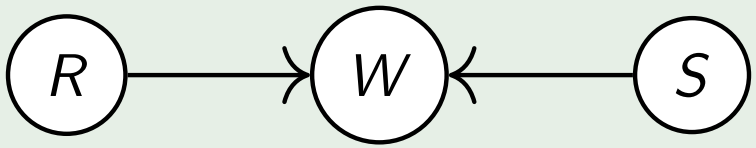


Sampling from a causal model

Each sample is a simulation from the causal model:

- Sample the value of all variables in $\mathcal{U}$,
- Use the structural equations to set the value of variables in $\mathcal{V}$.

Example



$W := R \vee S$; $P(R) = .1$, $P(S) = .5$.

- Sample 1: $R = 0$, $S = 1$, $W = 1$.
- Sample 2: $R = 0$, $S = 0$, $W = 0$,
- etc

Using samples to compute probabilities

A simple scheme is:

- $\hat{P}(X = x)$: the proportion of samples where $X = x$.
- $\hat{P}(X = x|Y = y)$: Among the samples where $Y = y$, the proportion of samples where $X = x$.

Computing conditional probabilities can be expensive

Suppose we have the model $C \to E$

- We want to compute $P(C|E)$
- If E happens 10% of the time, you can only 'keep' 10% of your samples.

In practice, there exists more sophisticated sampling algorithms that are more efficient.

- E.g. $P(E|C)$ can be efficiently estimated by starting every simulation with $\text{Do}(C = 1)$.
- We will see other sampling algorithms later in the course.

(Sampling in psychology)

- The human brain might use sampling to estimate probabilities.
- Sampling is an approximation: small number of samples $\rightarrow$ imprecise estimation.
- This hypothesis can explain why people often make mistakes in probabilistic and causal reasoning.

To find out more see e.g. Zhu, Sanborn & Chater (2020); Davis & Rehder (2020).

From inference *in* causal models to inference *about* causal models

- So far: inference about variable values, given a known causal model.
- In the rest of this lecture (and the next): we relax the assumption that we know the full causal model.

Estimating causal effects

Estimating causal effects

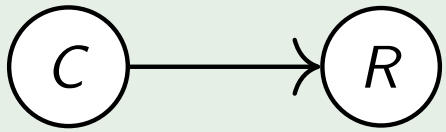
- We assume that you know the structure of the model (i.e. the DAG),
- But not the parameters.
- How can you find out the value of $P(\mathbf{Y}|\text{Do}(\mathbf{X}))$?

How easy this is depends on the data we can collect:

- If we can collect experimental data, estimating $P(\mathbf{Y}|\text{Do}(\mathbf{X}))$ is easy.
- What if we only have observational data?

An easy case

Example



We know from the structure of the network that:

- $P(\mathbf{R}|\text{Do}(\mathbf{C})) = P(\mathbf{R}|\mathbf{C})$,
- $P(\mathbf{C}|\text{Do}(\mathbf{R})) = P(\mathbf{C})$

A more difficult case

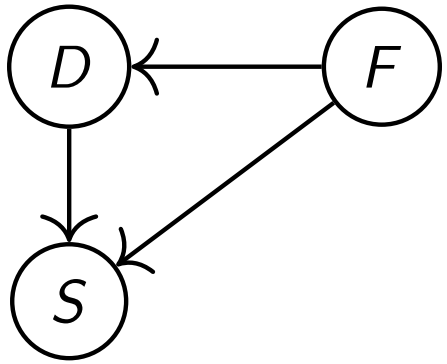
Example

Scientists conducted an experiment to study the effect of a Drug D on a symptom S. They collect non-experimental data. They find that:

- On average, people who take the drug show fewer symptoms (the drug is **good for people**),
- Men who take the drug have more symptoms than men who don't (the drug is **bad for men**),
- Women who take the drug have more symptoms than women who don't (the drug **is bad for women**).

Would you take the drug?

The good-bad-bad drug: an explanation



- Women ($F = 1$) have fewer symptoms on average, but they're also more likely to take the drug for some other reasons.
- $\rightarrow$ This creates an artificially negative correlation between D and S.

Gender is a **confound**: it makes it so that $P(\mathbf{S}|\mathbf{D})$ is not the same as $P(\mathbf{S}|\text{Do}(\mathbf{D}))$.
The good news: we can still use our observational data to estimate $P(\mathbf{S}|\text{Do}(\mathbf{D}))$.

‘Controlling’ for Gender

- On average, people who take the drug show fewer symptoms (the drug is **good for people**),
- Men who take the drug have more symptoms than men who don’t (the drug is **bad for men**),
- Women who take the drug have more symptoms than women who don’t (the drug **is bad for women**).

We have the intuition that the drug is bad because it is bad for women and bad for men.
We can get an unbiased estimate of the causal effect by averaging the effect on men and the effect on women:
$P(S|\text{Do}(D)) = P(S|D,F)P(F) + P(S|D,\neg F)P(\neg F)$.
We say we have **controlled** for Gender.

Control

Definition

The probability of **E** given **C**, **controlling** for Z, is calculated as: $\sum_Z P(\mathbf{E}|\mathbf{C}, Z)P(Z)$

Theorem (informal)

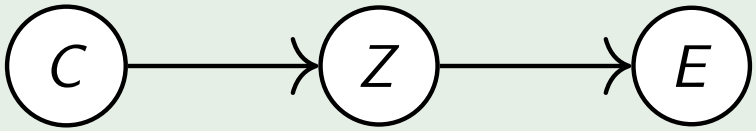
- Under **some** conditions, controlling for Z gives us the causal effect of C on E.
- i.e. $P(\mathbf{E}|\text{Do}(\mathbf{C})) = \sum_Z P(\mathbf{E}|\mathbf{C}, Z)P(Z)$

Warning

Controlling for a variable is not always a good idea! (see next slides)

Controlling is not always a good idea

Example



If we control for Z, we have:

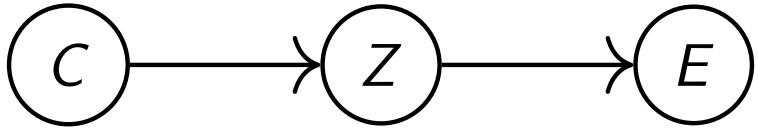
- $P(E|C,Z)P(Z) + P(E|C,\neg Z)P(\neg Z)$.
- Since $P(E|C,Z) = P(E|Z)$, this becomes:
- $P(E|Z)P(Z) + P(E|\neg Z)P(\neg Z) = P(E)$.

So controlling for Z, C has no effect on E.
Here the best strategy to compute $P(E|\text{Do}(C))$ is simply $P(E|\text{Do}(C)) = P(E|C)$.

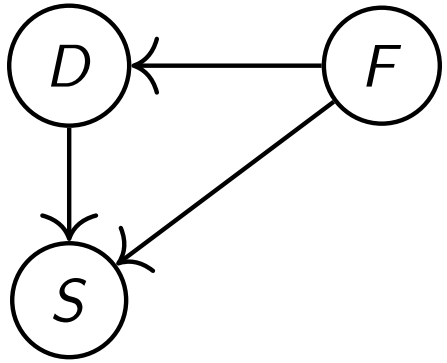
So when should we control and when should we not?

Rules of thumb:

- Don't control for Z if Z is a descendant of C.



- Control for Z if Z is a parent of both C and E.



For a more general theory of when you should or shouldn't control (outside the scope of this course), search for 'back-door criterion' and 'Do-calculus' in Pearl's 2009 book.