Introduction to Causal Inference

Felipe José Bravo Márquez

October 6, 2022

Motivation

- Our starting point is the difference between an observation and an intervention (or action).
- We can answer many questions from passive observation alone.
- For example: do 16 year-old drivers have a higher incidence rate of traffic accidents than 18 year-old drivers?
- The answer corresponds to a difference of conditional probabilities.
- Let random variables I, A correspond to traffic incident rate and driver's age correspondingly:

$$P(I|A = 16) - P(I|A = 18) > 0?$$

- Both conditional probabilities can be estimated from a large enough sample drawn from the distribution.
- The answer to the question we asked is solidly in the realm of observational statistics.
- However, important questions often are not observational in nature.

[1] These slides are mainly based on Chapter 9 of [Hardt and Recht, 2021].

Motivation

- Causal question: Would traffic fatalities decrease if we raised the legal driving age by two years?
- Here we are not asking for the frequency of an event in our manifested world.
- This question asks for the effect of a hypothetical **intervention**.
- As a result, the answer is not so simple.
- Even if older drivers have a lower incidence rate of traffic accidents, this might simply be a consequence of additional driving experience.
- There is no obvious reason why an 18 year old with two months on the road would be any less likely to be involved in an accident than a 16 year-old with the same experience.

Motivation

- We can try to address this problem by holding the number of months of driving experience fixed, while comparing individuals of different ages.
- But we quickly run into subtleties.
- What if 18 year-olds with two months of driving experience predominantly live in regions where traffic conditions differ significantly from those in areas where people feel a greater need to drive at a younger age?
- Causal reasoning is a conceptual and technical framework for addressing questions about the effect of hypothetical actions or interventions.
- Once we understand what the effect of an intervention is, we can turn the question around and ask what action plausibly caused an event.
- This gives us a formal language to talk about cause and effect.

The limitations of observation

- Before we develop any new formalism, it is important to understand why we need it in the first place.
- To see why we turn to the example of graduate admissions at the University of California, Berkeley [Bickel et al., 1975].
- The aggregate admission decisions of the six largest departments at the University was compared between male and female applicants.
- The aggregated acceptance rate for these six departments is 44% for men and 30% for women.
- The difference is statistcally significant.
- This led to an investigation into whether the University was acting in a gender-discriminatory manner.
- Recognizing that departments have autonomy over who to admit, we can look at the gender bias of each department.

The limitations of observation

UC Berkeley admissions data from 1973.

	Men		Women	
Department	Applied	Admitted (%)	Applied	Admitted (%)
A	825	62	108	**82**
B	520	60	25	**68**
C	325	**37**	593	34
D	417	33	375	**35**
E	191	**28**	393	24
F	373	6	341	**7**

- Four of the six largest departments show a higher acceptance ratio among women.
- The two other departments with higher acceptance rate for men cannot account for the large difference in acceptance rates that we observed in aggregate.
- It appears that the higher acceptance rate for men that we observed in aggregate seems to have reversed at the department level.

The limitations of observation

- Such reversals are sometimes called Simpson's paradox.
- Even though mathematically they are no surprise.
- It's a fact of conditional probability that there can be events Y (acceptance), A (gender) and a random variable Z (department choice) such that:
  1. $\mathbb{P}(Y|A) < \mathbb{P}(Y|\neg A)$
  2. $\mathbb{P}(Y|A, Z = z) > \mathbb{P}(Y|\neg A, Z = z)$ for all values z that the random variable Z assumes.
- Simpson's paradox nonetheless causes discomfort to some.
- Intuition suggests that a trend which holds for all subpopulations should also hold at the population level.

References I

Bickel, P. J., Hammel, E. A., and O'Connell, J. W. (1975).
Sex bias in graduate admissions: Data from berkeley: Measuring bias is harder than is usually assumed, and the evidence is sometimes contrary to expectation.
Science, 187(4175):398–404.

Hardt, M. and Recht, B. (2021).
Patterns, predictions, and actions: A story about machine learning.
arXiv preprint arXiv:2102.05242.