Causal Inference

Chapter 1. Introduction and Framework

1 Introduction.

(1) What are causal questions ?

Examples of problems of interest to causal inference:

(a) Women $>$ 50 yrs: should they be getting regular screening for breast cancer ?

(b) Do citizens of Los Angeles die because of air polution ?

(c) How much mortality and other burden was due to the tobacco industry's miscoduct ?

(d) Women who brake their hip: what prosthesis should they get ?

(e) How does the type of school affect a child's later achievements ?

(f) Does a depression treatment work better if it makes a particular gene have high expression ?

Examples of not clearly defined causal statements:

(a) Are parents more conservative than their children because they are older ?

(b) Is there an effect of gender in this regression ?

cc. There is an important difference between the two: intervention.

cc. Interventions in the examples

A causal question is a problem with a manipulable *intervention*.

cc. The intervention can be hypothetical, but should nevertheless be clearly defined.

(2) Causal questions are important because they instruct policy and decision making.

cc. Policy actions in first group, questions on second group.

(3) What does the new framework do better than others ?

Simpler principles, help better definition, analyses and designs. The benefit is larger when the problems become more demanding.

2 Framework.

Methodologic framework for the course has elements (called Rubin's causal model) as building blocks:

(1) subjects (units), at a particular place and time.

(2) treatments/interventions to compare (e.g., z=0 for standard, z=1 for new).

Fig. 1. treatments.

(3) potential outcomes; e.g., $Y_i(1), Y_i(0)$ are the outcomes that would be observed on the same subjects if assigned new, or, alternatively, if assigned standard treatment.

(4) causal effects (def.): comparisons of potential outcomes for the same subject or same group of subjects.

Fig. 1. two possibilities

- Potential outcomes are not both observed; they are the estimands, i.e., the scientific quantities we will try to estimate.
- Reference: Neyman (1923), Rubin (1974, 1977, 1978)
- Causal effects, in their definition, do not relate to probability distribution for subjects "who got different treatments", or to coefficients of models.
- We do not assume necessarily that the P.O.s are fixed. In fact we could consider an inherent status as random, but that is not important for the scale we discuss here.
- We allow that the potential outomes depend on other factors that relate to that person at the time and place when the treatment is given, and through the follow-up time where the P.O.s would be observed.
- Counterfactual ?
- Not a before-after comparison. Example for why not.

To learn more about causal effects, we need more units.

Fig. 2. Interference

(5) Stability assumptions, to simplify problem (usually, assumed no interference between units, no versions of a single treatment, (SUTVA)).

- Agricultural experiments
- contagious diseases: sutva not applicable for people who interact.
- Relax sutva when not applicable.

Causal effects under sutva.

Fig. 3.

(6) Assignment mechanism.

Fig. 4.

Assignment mechanism is the rule (possibly probabilistic) by which subjects get their actual treatments $\{Z_i\}$. The assigned treatments unmask the potential outcomes $Y_i(Z_i)$ (denoted by Y_i^{obs}), but mask the rest potential outcomes, denoted by Y_i^{mis}. (note: distinguish notation z, Z).

- Assignment mechanism is central, yet not mentioned explicitly in most literature.
- Example: doctor who can give to every patient the best treatment.

  Fig. 5.

- Example: case-crossover design and the "null hypothesis".

  Fig. 6.

The course will formulate problems with this framework

3 Historical review.

Fig. 7.

4 Notes and outline of course.

(a) Other approaches: use comparisons of $pr(Y^{obs}|Z=1,X)$ to $pr(Y^{obs}|Z=0,X)$ as building blocks, or graphs with arrows.

In Rubin's framework, we consider some potential outcomes that we do not see.

Why consider the existence of something we do not see ?

Because such consideration expands the formulation of explanations of observed phenomena.
Example: the fundamental concepts of modern physics are unobservable, only their consequences are observable.

(b) Outline of course, in terms of assignment mechanism.

First need to gain insight of how this framework addresses simple designs, as the randomized design. [Other closely related designs.] Then we consider more demanding settings. We will see that the guidance provided by the standard approach above fails in such settings, and that the problems will be more easily set up and addressed in the new framework.