DECART Summer School 2018:

From Anecdote to Experiment: Causal Inference

Happy the man who has been able to learn the causes of things.

- Virgil

What We'RE UP TO IN OUR DAYS Together

- Introduce and discuss basic concepts
- Hands on approach (get our hands dirty)
- Perspectives from various disciplines
- Exercises in Jupyter and using R
- By day:
  - Day 1 – concepts, fundamentals, R in Jupyter
  - Day 2 – parametric and non-parametric models, more frameworks
  - Day 3- Advanced topics and techniques, paper discussion

"What do you mean 'it just happened'?
Didn't we discuss cause and effect?"

WhAT's A Cause?

- An *explanation* for some event or result (Aristotle)
  - Raises the questions of what is an explanatory object or concept, and what is an effected “thing?”
- An *abstraction*
- *Not* contained in data per se
- More detailed explanations differ across application areas

FIVE MINUTES ON SOME THEORY DEVELOPMENT HIGHPOINTS

- D. Hume (1748)
  - "Counterfactual" as a *false antecedent*
- S. Wright (1921)
  - Path analysis: learn causal relations using statistics
  - Three kinds of relationships:

$$Z \rightarrow X \rightarrow Y \tag{1}$$

$$X \leftarrow Z \rightarrow Y \tag{2}$$

$$X \rightarrow Z \leftarrow Y \tag{3}$$

Developments (cont)

- K. Popper (1959)
  - Three necessary conditions to conclude X is a cause of Y
    - X and Y correlated
    - X precedes Y
    - Alternative explanations for relationship can be ruled out
- D. Lewis(1973)
- Temporal dependence effect on cause:

$$if\ X\ then\ Y,\ if\ \bar{X}\ then\ \bar{Y}$$

MORE RECENTLY, THE "*Causal INFERENCE REVOLUTION*"

- Work in various disciplines has made it clearer in the last few decades that "valid" causal inferences can sometimes be based on observational data
- Important recent innovations include:
  - Potential outcomes framework(s)
  - "Non-parametric" causal model specifications
  - Specifications in graphs
- Seminal work by Hernan, Robins, Pearl, Heckman, Rubin, and many others

CaUSAL MODELLING: GreenlAND & Brumback, 2002

- Four major types of models in health-sciences:
  - causal diagrams/graphs
  - potential outcomes (PO's)
  - structural equation models (SEMs)
  - sufficient-component cause
- Graphs: easily understood qualitative assumptions
- PO's and SEMs: detailed quantitative assumptions

(2002) An overview of relations among causal modeling methods. *Intl. Journal of Epidemiology, 31, 1030-1037*

EffeCT of a DruG on ReCOVERY: Would You PRESCRIBE it?

Gender:	Got Drug:	Didn't Get Drug:
Male	81 of 87 (93%)	234 of 270 (87%)
Female	273 of 350 (78%)	55 of 80 (69%)
Overall	273 of 350 (78%)	289 of 350 (83%)

From Pearl, Glymour and Jewell, 2016

experiments with random assignment To Conditions

- Not a topic we'll actually cover, but it's worth briefly considering why they are the “gold standard” for causal inference
- The basic idea: random assignment unconfounds experimental effects from factors that might be biasing
- Mechanism(s) determining assignment to Tx conditions produce missing values are ignorable, no selection bias(es) affecting estimates
- Potential results for a “unit” S in experimental conditions do not depend on the Tx that S actually receives