Lecture 1: Introduction to Predictive Modeling

Professor Ilias Bilionis

Causal models and their graphical representation

PREDICTIVE
SCIENCE LABORATORY

Why worry about causality?

- Models that explicitly encode causal relationships are the only useful models.
- Most physical and engineering models are causal models.
  ↳ just need to learn how to interpret the results
- When we extend them to account for known unknowns, we need to make sure that they remain causal.
- Otherwise, we are merely capturing correlations and the results cannot be trusted.
  ↳ cannot use the model to make decisions
- Structural causal models give us the language we need to formalize causality.

PREDICTIVE
SCIENCE LABORATORY

What is a structural causal model?

Example:

- A set of variables we are interested in: could be grown to include additional factors

    - Y: an individual has asthma or not - binary
    - X: the individual is treated or not - binary
    - Z: air pollution level

  gather all variables/players that have a role

SCM

- A set of functions that describes how the variables are connected:

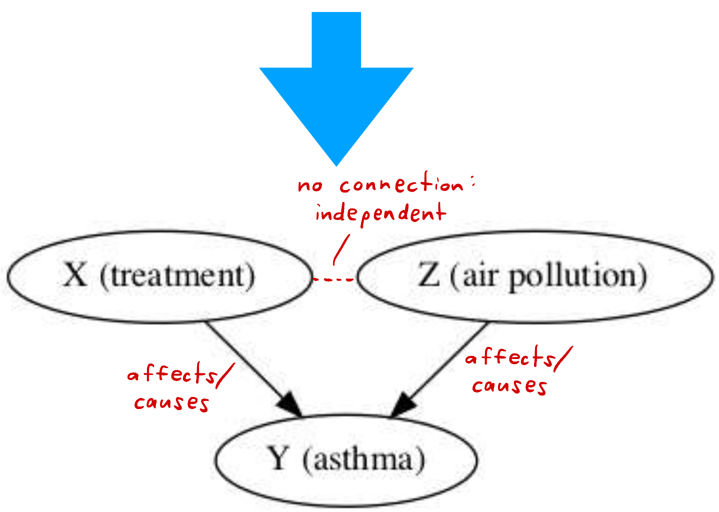
$$Y = f(X, Z) .$$

how X and Z "cause" Y

PREDICTIVE SCIENCE LABORATORY

Graphical representation of causal models

$$Y = f(X, Z)$$



- To each structural causal model there corresponds a directed acyclic graph.
- The edge represents a direct causal link between the parent and the child nodes.

PREDICTIVE
SCIENCE LABORATORY

In Practice:

We typically first build the graph and then work out the function details

$$Y = f(X, Z)$$