Causation and Prediction: Axioms and Explications

Discussants: Kim Luijken & Jeroen Mulder

Causality Reading Group I - Fall 2017

- Example of collider conditioning: nonresponse bias
- Monty Hall

Some Definitions...

- Direct cause
- Indirect cause: screened off by direct cause.
- Boolean variable: events must be sorted.
- Scaled variable: continuous special case of Boolean variable.
- Causal representation convention: a causal link is indeed, really, without a doubt, a causal link. And no interactions are captures in the graph (drawback).
- Causal sufficiency: A set **V** of variables is causally sufficient for a population if and only if in the population every common cause of any two or more variables in **V** is in **V**, or has the same value for all units in the population.

Deterministic Causal Models

- In a deterministic causal structure, the probability distribution over the exogenous variables determines a joint distribution for the entire set of variables in the system.
- Assumption: exogenous variables are jointly independent and *statistical dependence is produced by causal connection*.
- We need a causal sufficient graph(no omitted variables).

... is discussed during the Causality Reading Group.

$$J = (\text{dancing}, \neg\text{dancing}) \tag{1}$$

$$K = (\text{Kim turns on the music}, \neg\text{Kim turns on the music}) \tag{2}$$

$$J = (\text{dancing}, \neg\text{dancing}) \tag{3}$$

$$K = (\text{Kim turns on the music}, \neg\text{Kim turns on the music}) \tag{4}$$

$$A = (\text{Kim and Jeroen drink alcohol}, \neg\text{Kim and Jeroen drink alcohol}) \tag{5}$$

What degree of pseudo-indeterminism is acceptable? Is causality something fundamentally different from an association?

Defining Causality

- "Strictly, therefore, our definitions of causal relations for variables should be relative to a set of possible values for other variables, *but we will ignore this formality and trust to context.*" - Spirtes, Glymour, and Scheines (2000).
- "X is a cause of Y if Y listens to X and decides its value in response to what it hears." - Pearl, Glymour, Jewell (2016)
- Counter factuals.

Group Discussion

- Conditioning on a collider
  - A full paper (Burns & Wieth, 2004) has been written positing that the reason why people find the Monty Hall problem so hard is because of the unintuitive nature of collider conditioning.
- Populations and conditioning
  - Causal structure should be interpreted with respect to a population of interest
  - Often we may be interested in causal relations in a subpopulation, e.g. symptoms of depression among the depressed, relations between income and child support among males
  - By causal sufficiency, we have a sufficient graph if all unobserved common causes are invariant in our population of study
  - By d-sepereation rules, if we condition on an unobserved **endogenous** variable, we introduce bias
  - How should we approach this issue of subpopulations - isn't it still valid to look at subpopulations which are based on something endogenous? In psychology, whether depression is seen as exogenous (latent variable causing symptoms), endogenous (depression is co-presence of symptoms) or neither (depression is an emergent property of a system of symptoms) is based on philosophical outlook.