Statistical Causal Discovery: Exploring Causal Hypotheses

Empirical sciences aim to study causal mechanisms

- Hypothesis A vs. Hypothesis B

Mathematical Theories → Computational Algorithms

Compare different hypotheses using data

Exploring causal hypotheses based on observational data

- Hypotheses necessary to perform experiments
- Experiments often ethically impossible or too costly

Find interesting hypotheses before performing experiments!

Major challenges

Example:
Significant correlation between chocolate consumption and the number of Nobel laureates (23 countries: 2002–2011)
(Messerli, 2012, New England Journal of Medicine)

Chocolate consumption is positively correlated with the number of Nobel laureates

Eating more chocolate increases the number of Nobel laureates??

No, not necessarily!

Different causal relations provide the same correlation

Chocolate consumption (kg/yr/capita)

Chocolate consumption (kg/yr/capita)

Our key contributions

Novel methods capable of identifying causal relations even in the presence of hidden common causes


- Linearity
- Acyclicity
- Non-Gaussianity

Ongoing works and future directions:
- Extensions to more general cases
- Methods to evaluate model assumptions
- Empirical applications

Cases involving more than two variable cases:
- Apply to every pair of variables and integrate the results

(Does not require specifying the number of hidden common causes)

Cases involving more than two variable cases:
- Apply to every pair of variables and integrate the results

Manage this gap!

Cope with hidden common causes!

(Mathematical Theories → Computational Algorithms)

Application areas

Epidemiology

Sleep problems → Depression
or
Sleep problems → Depression

(Moneta et al., 2012, Oxford Bulletin of Economics and Statistics)

Economics

Chocolate consumption (kg/yr/capita)

Chocolate consumption (kg/yr/capita)

Neuroscience

Different information flows in different framing conditions

(Mills-Finnerty et al., 2014, NeuroImage)

Brain activity measurements using functional MRI:
To what extent does which brain region become active?

VMPFC
Ant. Cing

DLPFC
Caudate

VMPFC
Ant. Cing

DLPFC
Caudate

Eating more chocolate increases the number of Nobel laureates??

No, not necessarily!

Different causal relations provide the same correlation

(Mathematical Theories → Computational Algorithms)

Different causal relations provide the same correlation

(Mathematical Theories → Computational Algorithms)