CAUSALITY I

Marlene Kretschmer

University of Reading
"CORRELATION DOES NOT IMPLY CAUSATION"

Source: xkcd #552
Including causal (physical) reasoning in the statistical analysis to draw causal conclusions from data
1) Introduction
**Motivation**

We need a *causal understanding* of the world, both for decision-making and for many forms of theory and research.

- What is the effect on global mean temperature if GHG emissions are increasing?
- Will climate change lead to more intense extreme rainfall events in the UK?
- Is El Nino increasing the chance of drought in South Africa?
- What is the effect of melting Arctic sea ice on European climate?
Physics-based methods

As real-world experiments are usually not possible, numerical climate models are used to infer causal relationships of the climate system.

Downside:

Inferences about the real word depend on the realism of the climate model.
Observational data are studied with statistics/data science tools

However, we are usually limited to detect statistical associations (e.g. correlations) but *correlation does not imply causation*

How can we infer *causal* relationships from data?
2) Causal Inference
The concept of *causality* has long been missing in mathematics

Causal inference: the science to extract causal information from data

1. learning causal relationships
2. quantifying causal relationships

In this talk
**The Three Layer Causal Hierarchy**

- **Association:** $P(y | x)$
- **Intervention:** $P(y | do(x))$
- **Counterfactuals:** $P(y_x | x', y')$

Source: The book of why
ASSOCIATION VS. INTERVENTION

What happens if we intervene in $X$?

$X$ causes $Y$?
$Y$ cause $X$?
A common driver $Z$ affects $X$ and $Y$?

Data Doesn’t Speak for Itself!
ASSOCIATION VS. INTERVENTION

Example

X: Pressure  Y: Barometer

Intervening in Y will *not* change X
Intervening in X will change Y
What is the effect on Y if we “do” X=1?

Causal Model  

X --> Y

Estimate causal effect from data to predict the intervention

\[ P(Y \mid do(X) = 1) = P(Y \mid X=1) \]

\[ X = \varepsilon_x \]

\[ Y = 1.5 \times X + \varepsilon_y \]
Association vs. Intervention

To make sense of the data, we need **causal knowledge** about the data-generating mechanisms.

We usually have such “expert knowledge” available ... We should make use of it!
Question: What is the (average) causal effect of X on Y?

1. Use expert knowledge to set a (plausible) causal model

2. Collect data

3. Control for confounders to isolate the causal effect

Confounding is anything that leads to $P(Y|X)$ being different than $P(Y|do(X))$

Linear case:

$$Y = aX + bZ$$
3) Examples from Climate Science
Quantifying the contribution of teleconnections is key to improve our understanding of regional weather and climate variability.

Extracting this information from data is usually difficult!
CORRELATION VS. CAUSATION

American Meteorological Society: “Teleconnection”

A significant [...] correlation in [...] widely separated points.

[...] such correlations suggest that information is propagating [...].

https://glossary.ametsoc.org/wiki/Teleconnection
QUANTIFYING CAUSAL PATHWAYS OF TELECONNECTIONS

How to formally include causal (physical) reasoning in the statistical analysis of teleconnections

Quantifying Causal Pathways of Teleconnections
Marlene Kretschmer, Samantha V. Adams, Alberto Arribas, Rachel Prudden, Niall Robinson, Elena Saggioro, and Theodore G. Shepherd
EXAMPLE 1: COMMON DRIVER

Precipitation in Denmark (DK) and the Mediterranean (MED) are significantly correlated

$$\text{Corr}(\text{DK, MED}) = -0.25$$

Does this reflect a causal relationship?
Example 1: Common Driver

How strong are the causal effects?

\[ \text{DK} = -0.55 \text{ NAO} + \epsilon \]
\[ \text{MED} = +0.42 \text{ NAO} + \epsilon \]

The causal effects explain the correlation

\[ -0.55 \times 0.42 \approx -0.25 \]
**Example 1: Common Driver**

Is our causal model consistent with the data?

DK and MED are independent after regressing out the effect of NAO

→ Corr(DK, MED | NAO) = 0.01

Kretschmer et al. (2021, BAMS)
Example 1: Common Driver

Causal knowledge is needed to interpret both causal and non-causal associations

(JJA mean, NCEP)
EXAMPLE 2: MEDIATOR

What is the effect of ENSO on precipitation in California (CA)?

Kretschmer et al. (2021, BAMS)
What is the effect of ENSO on precipitation in California (CA)?

CA = 0.05 ENSO + 0.79 Jet + ε

We must not interpret this causally!

Kretschmer et al. (2021, BAMS)
**Example 2: Mediator**

What happened here?

CA = 0.05 ENSO + 0.79 Jet + ε

By including “Jet” in the regression model, we blocked the causal pathway from “ENSO” to “Jet”
**EXAMPLE 2: MEDIATOR**

Correct way:

\[ \text{CA} = 0.34 \ \text{ENSO} + \varepsilon \]

Or via product along the pathway:

Jet = 0.37 \ ENSO + \varepsilon

\[ \text{CA} = 0.81 \ \text{Jet} + \varepsilon \]

\[ 0.37 \times 0.81 = 0.30 \]
The Importance of Causal Reasoning

Statistically, example 1 and 2 are undistinguishable

X and Y are correlated
example 1: X = DK and Y = MED
example 2: X = ENSO and Y = CA

X and Y are independent conditional on Z
example 1: Z = NAO
example 2: Z = Jet

The causal interpretation enters through our physical knowledge!
**Example 3: Indirect and Direct Effects**

\[ \text{Jet} = 0.14 \text{ENSO} + \varepsilon \]

**Indirect (stratospheric) pathway:**

\[ \text{Jet} = 0.39 \text{SPV} + 0.04 \text{ENSO} + \varepsilon \]

\[ 0.26 \times 0.39 = 0.10 \]

**Direct (tropospheric) pathway:**

\[ \text{Jet} = 0.04 \text{ENSO} + 0.39 \text{SPV} + \varepsilon \]

**Total effect of ENSO on Jet**

\[ \text{Total} \times 0.14 = 0.14 \]

(OND mean, NCEP)
4) Summary & Outlook
**Basic Causal Structures and Their Implications**

- **Z is a common driver of X and Y**
  - Implication for data:
    - X and Y correlated
    - X and Y are uncorrelated conditional on Z

- **Z is a mediator of X and Y**
  - Implication for data:
    - X and Y uncorrelated
    - X and Y are correlated conditional on Z

- **Z is a common effect of X and Y**
**Question:** What is the (average) causal effect of X on Y?

1. Use expert knowledge to set a (plausible) causal model

```
Z → X
  ↓     ↓
  U     V
  ↓     ↓
Y      Y
```

2. Collect data

```
X
Y
Z
U
V
```

3. Control for confounders to isolate the causal effect

\[
P(Y \mid \text{do}(X)) = P(Y \mid X, Z)
\]

**linear case:**

\[
Y = a\ X + b\ Z
\]
**Take Home Messages**

- Correlations are the result of causal relationships
- We need causal knowledge about the data-generating mechanisms to interpret correlations and to extract the causal effects

- Causal inference gives the formal rules how to achieve this
- Causal knowledge/hypotheses are best expressed using causal networks
- To extract causal effects from data, one needs to control for all confounding factors

Scientific data analysis requires causal reasoning
**Bridging Physics and Statistics**

Physics-based methods

Data-driven methods

Causal Framework

- Inputs
- Hidden Layers
- Outputs
OUTLOOK

Practicals
- See email instructions

Next Lecture
- Conditioning on a common effect
- How to control for the correct pathways to isolate the causal effect of interest
- Beyond linear regression: quantifying non-linear causal relationships
- Outlook: learning causal networks from data