Causal Inference
Theory and Applications in Enterprise Computing

Dr. Matthias Uflacker, Johannes Huegle, Christopher Schmidt

April 09, 2019

Agenda
April 9, 2019

- **Lecture Organization**

- **Causal Inference in a Nutshell**
  1. Motivation
  2. A short History
  3. A Paradigm Shift
  4. Causal Graphical Models
  5. The Calculus of Causality
  6. Summary and Outlook
  7. Further Reading
  8. References

Lecture Organization

Lecture Organization
Setup

- Supervisors: Dr. Matthias Uflacker, Johannes Huegle, Christopher Schmidt
- Time: Tuesdays and Wednesday 9.15-10.45 AM
- Location: D.E-9/10, HPI Campus II
- Periods: 4 SWS (6 graded ECTS)
- Lecture Page:
  - https://hpi.de/plattner/teaching/summer-term-2019/causal-inference-theory-and-applications.html
- Enrollment:
  - Sign up for the course until **Fri April 26**

Lecture Organization
Goals

- ***Understand…***
  - opportunities of causal inference
  - the mathematical concepts
  - challenges and problems in the context of real-world applications
  - constraint-based algorithms to derive causal relationships
- ***Do…***
  - work in small teams
  - work on a specific topic in the context of data-driven causal inference
  - implement algorithms and analyze performance results
  - write a scientific report
- ***Improve…***
  - mathematical, analytical, and modeling skills
  - scientific working and writing
  - machine learning techniques

Lecture Organization
Schedule (I/II)

- ***April 8 – April 26: Lecture Period I***
  - Learn theoretical basis of
    - Causal Graphical Models
    - Conditional Independence Testing
    - Causal Structure Learning
- ***April 26 – May 3: Topic Assignments***
  - Form groups and apply for topics until *Fri April 26, 11.59 PM*
  - Assignment and individual meeting on *Tue April 30, 9:00 AM*
- ***May 6 – May 24: Training Period***
  - Get familiar to your topic
  - Tuesday: If needed Q&A + topic related lectures
  - Wednesday: Individual meetings with supervisors
- ***May 27– May 31: Intermediate Presentations***
  - Present your topic, first results, challenges, and solution ideas

Lecture Organization
Schedule (II/II)

- ***June 3 – June 7: Lecture Period II***
  - Learn theoretical basis of causal reasoning on causal graphs
- ***June 10 – June 28: Elaboration Period***
  - Realization and implementation
  - Tuesday: If needed Q&A + topic related lectures
  - Wednesday: Individual meetings with supervisors
- ***July 1 – July 5: Final Presentations***
  - Present your topic, first results, challenges, and solution ideas
- ***July 8 – August 30: Scientific Report***
  - Write scientific report on your findings, incorporate feedback and finalize work
  - Submission: *August 2, 11:59 PM*
  - Peer-Review Submission: *August 16, 11:59 PM*
  - Final Submission: *August 30, 11:59 PM*

Lecture Organization
Grading

- The grading of the seminar works as follows (aka "Leistungserfassungsprozess"):
  - **50%** intermediate and final presentation of implementation results
  - **40%** scientific research article
  - **10%** personal engagement

www.hpi.uni-potsdam.de

- **All individual parts have to be passed** to pass the complete lecture

Causal Inference in a Nutshell

1. Motivation
Causality: An ubiquitous Notion

Teaching Statistics / Volume 22, Issue 2

Storks Deliver Babies (*p*= 0.008)

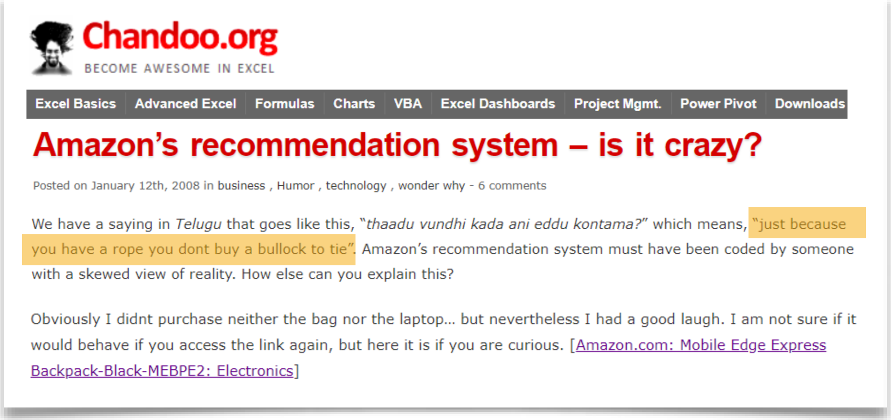
Robert Matthews

First published: 25 December 2001
https://doi.org/10.1111/1467-9639.00013

Abstract

This article shows that a highly statistically significant correlation exists between stork populations and human birth rates across Europe.While storks may not deliver babies, unthinking interpretation of correlation and p-values can certainly deliver unreliable conclusions.

1. Motivation
Causality: An ubiquitous Notion



Chandoo.org
BECOME AWESOME IN EXCEL

Excel Basics | Advanced Excel | Formulas | Charts | VBA | Excel Dashboards | Project Mgmt. | Power Pivot | Downloads

Amazon's recommendation system – is it crazy?

Posted on January 12th, 2008 in business , Humor , technology , wonder why - 6 comments

We have a saying in *Telugu* that goes like this, "*thaadu vundhi kada ani eddu kontama?*" which means, "just because you have a rope you dont buy a bullock to tie". Amazon's recommendation system must have been coded by someone with a skewed view of reality. How else can you explain this?

Obviously I didnt purchase neither the bag nor the laptop… but nevertheless I had a good laugh. I am not sure if it would behave if you access the link again, but here it is if you are curious. [Amazon.com: Mobile Edge Express Backpack-Black-MEBPE2: Electronics]

1. Motivation
Causality: An ubiquitous Notion

Ann Emerg Med. 2017 Jan;69(1):62-72. doi: 10.1016/j.annemergmed.2016.08.007. Epub 2016 Sep 28.

The Effect of Combined Out-of-Hospital Hypotension and Hypoxia on Mortality in Major Traumatic Brain Injury.

Spaite DW[1], Hu C[2], Bobrow BJ[3], Chikani V[4], Barnhart B[5], Gaither JB[6], Denninghoff KR[6], Adelson PD[7], Keim SM[6], Viscusi C[6], Mullins T[8], Sherrill D[9].

Author information

Abstract

STUDY OBJECTIVE: Survival is significantly reduced by either hypotension or hypoxia during the out-of-hospital management of major traumatic brain injury. However, only a handful of small studies have investigated the influence of the combination of both hypotension and hypoxia occurring together. In patients with major traumatic brain injury, we evaluate the associations between mortality and out-of-hospital hypotension and hypoxia separately and in combination.

METHODS: All moderate or severe traumatic brain injury cases in the preimplementation cohort of the Excellence in Prehospital Injury Care study (a statewide, before/after, controlled study of the effect of implementing the out-of-hospital traumatic brain injury treatment guidelines) from January 1, 2007, to March 31, 2014, were evaluated (exclusions: <10 years, out-of-hospital oxygen saturation ≤10%, and out-of-hospital systolic blood pressure <40 or >200 mm Hg). The relationship between mortality and hypotension (systolic blood pressure <90 mm Hg) or hypoxia (saturation <90%) was assessed with multivariable logistic regression, controlling for Injury Severity Score, head region severity, injury type (blunt versus penetrating), age, sex, race, ethnicity, payer, interhospital transfer, and trauma center.

RESULTS: Among the 13,151 patients who met inclusion criteria (median age 45 years; 68.6% men), 11,545 (87.8%) had neither hypotension nor hypoxia, 604 (4.6%) had hypotension only, 790 (6.0%) had hypoxia only, and 212 (1.6%) had both hypotension and hypoxia. Mortality for the 4 study cohorts was 5.6%, 20.7%, 28.1%, and 43.9%, respectively. The crude and adjusted odds ratios for death within the cohorts, using the patients with neither hypotension nor hypoxia as the reference, were 4.4 and 2.5, 6.6 and 3.0, and 13.2 and 6.1, respectively. Evaluation for an interaction between hypotension and hypoxia revealed that the effects were additive on the log odds of death.

CONCLUSION: In this statewide analysis of major traumatic brain injury, combined out-of-hospital hypotension and hypoxia were associated with significantly increased mortality. This effect on survival persisted even after controlling for multiple potential confounders. In fact, the adjusted odds of death for patients with both hypotension and hypoxia were more than 2 times greater than for those with either hypotension or hypoxia alone. These findings seem supportive of the emphasis on aggressive prevention and treatment of hypotension and hypoxia reflected in the current emergency medical services traumatic brain injury treatment guidelines but clearly reveal the need for further study to determine their influence on outcome.

Words Matter: Researchers Should Avoid Implying Causation in Studies of Association

Joshua S. Broder, MD
Duke University School of Medicine, Division of Emergency Medicine, Durham, NC

DOI: https://doi.org/10.1016/j.annemergmed.2017.03.016 | CrossMark

Abstract | Full Text | References

To the Editor:

I read with interest the article by Spaite et al.[1] The authors should be commended for a carefully conducted analysis of the association between survival after major traumatic brain injury and the combination of hypoxia and hypotension. However, even with attempts to statistically isolate hypoxia and hypotension from confounding variables such as injury severity, this nonrandomized cohort study cannot determine causation, which is implied by the title ("The Effect of Combined Out-of-Hospital Hypotension and Hypoxia on Mortality in Major Traumatic Brain Injury"), the abstract, the article, and the Editor's Capsule Summary ("what is the effect on survival of the combination of hypotension and hypoxia compared with either factor alone?").

Although this may appear to be a semantic or minor concern, appropriate terminology in describing research methods, results, and conclusions is of fundamental importance. The history of medicine illustrates the potential harms of misconstruing an association as a causal relationship and acting (with good intentions) to reduce a misperceived effect. For example, estrogen replacement was falsely identified as the cause of improved women's cardiovascular health according to associations from observational cohort studies; prospective randomized controlled trials demonstrated harms from this therapy.[2-5]

1. Motivation
Causality: What is it?

Causality is central notion in science, decision-taking and daily life.

How to reason formally about cause and effect?

Question: How do you define cause and effect?

2. A short History
Causality in Philosophy

The subject of causality has a long history in philosophy.

*"…Thus we remember to have seen that species of object we call flame, and to have felt that species of sensation we call heat. We likewise call to mind their constant conjunction in all past instances. Without any farther ceremony, we call the one **cause** and the other **effect**, and infer the existence of the one from that of the other."*

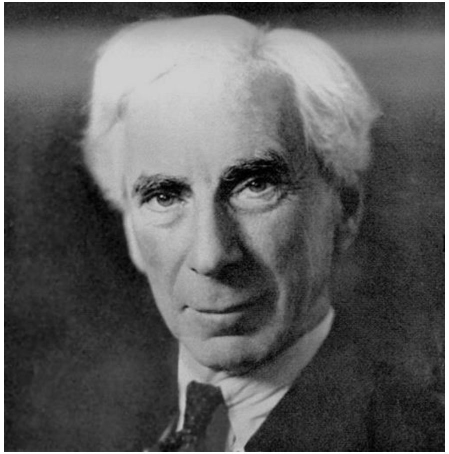
David Hume, A Treatise of Human Nature (1738)

2. A short History
Causality in Philosophy

But: Does the stork really bring babies?

2. A short History
Causality in Philosophy

Some philosophers even proposed to abandon the concept of causality completely.



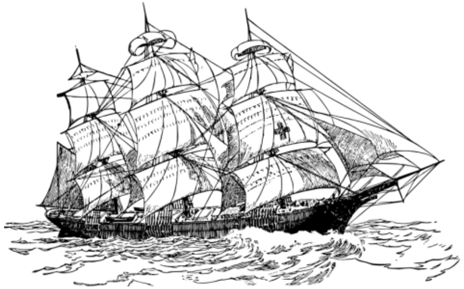
*"All philosophers […] imagine that causation is one of the fundamental axioms or postulates of science, yet, oddly enough, in advanced sciences such as gravitational astronomy, the word `**cause**' never occurs.*
The law of causality, I believe, like much that passes muster among philosophers, is a relict of a bygone age, surviving, like the monarchy, only because it is erroneously sup-posed to do no harm."

Bertrand Russell, On The Notion Of Cause (1912)

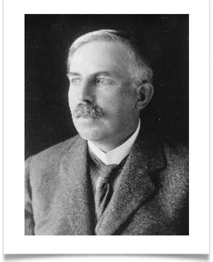
2. A short History
Causality in Statistics

Challenge in Statistics: Draw conclusions from data

- James Lind (1716-1794): **How to treat scurvy?**
  - Scurvy results from a lack of vitamin C
  - 12 scorbutic sailor treated with different acids, e.g. vinegar, cider, lemon
  - Only the condition of the sailor treated by lemon improved



- *"If your experiment needs statistics, you ought to have done a better experiment."*
  Ernest Rutherford (1871-1937)



But: What if you cannot do a *randomized experiment* or receive ambiguous results?

Use statistical tests to validate your hypothesis

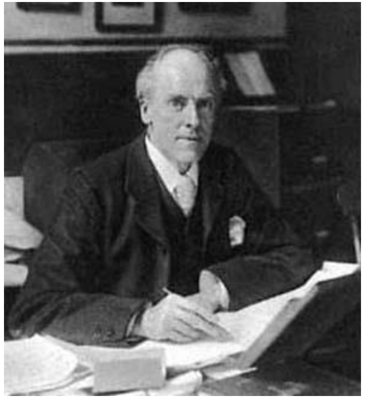
Check whether it is statistically significant that

$$P(recovery \mid lemons) > P(recovery \mid no\ lemons)$$

Or in other words:

"Is there a dependence between recovery and the treatment with lemons?"

2. A short History
Causality in Statistics



"Beyond such discarded fundamentals as `matter' and `force' lies still another fetish amidst the inscrutable arcana of even modern science, namely, the category of cause and effect."

Karl Pearson (1857-1936)

Correlation does not imply causation.

Since then, many statisticians tried to avoid causal reasoning

- *"Considerations of causality should be treated as they have always been in statistics: preferably not at all."* (Terry Speed, 1990)
- *"It would be very healthy if more researchers abandon thinking of and using terms such as cause and effect."* (Bengt Muthen, 1987)

2. A short History
Causality in Statistics

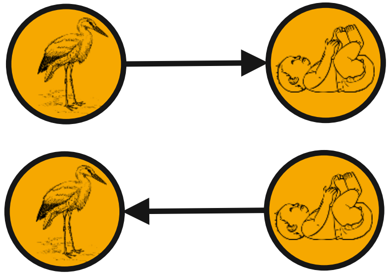
But dependence says us something about causation:

"Common Cause Principle"
Hans Reichenbach (1891-1953)

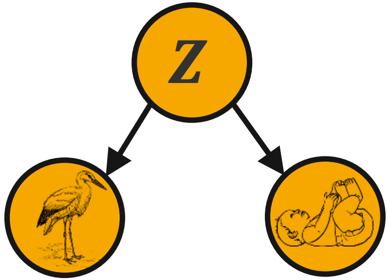
If there is a statistical dependence between variables X and Y, e.g.,

then either

- X causally influences Y (or vise versa), e.g.,

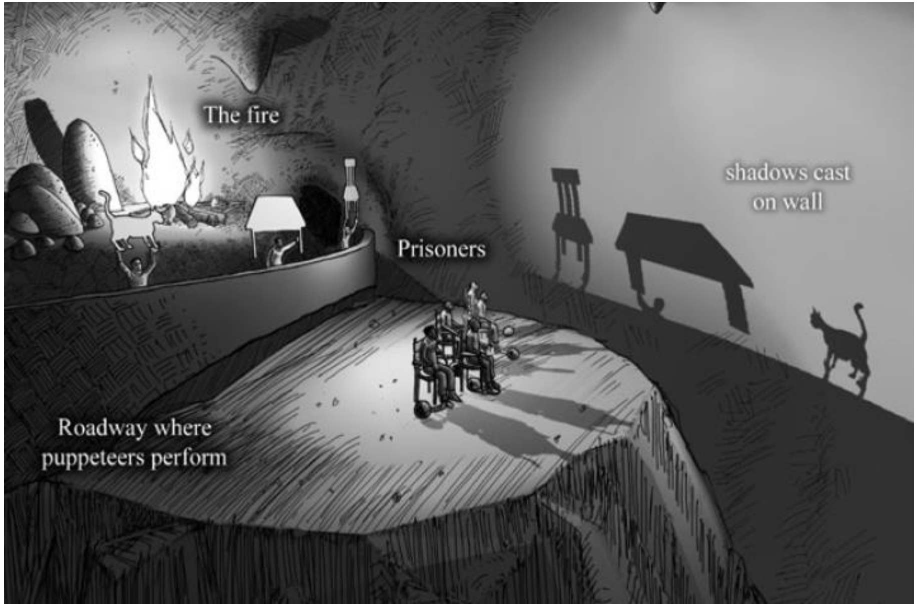


- or there exists Z causally influencing both, e.g.,



3. A Paradigm Shift

The Idea: Plato's Allegory of the Cave



Do not model the distribution of the data but model the mechanisms that generated the data!

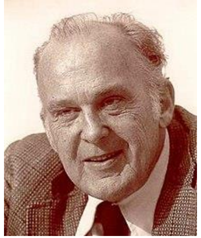
3. A Paradigm Shift
Basic Contributions

- The modeling of the underlying structures provides a language to encode causal relationships – the basis of a causality theory.
- Causality theory helps to decide when, and how, causation can be inferred from domain knowledge and data.

Some people who contributed to causality theories:

Donald Rubin (*1943)

Judea Pearl (*1936)



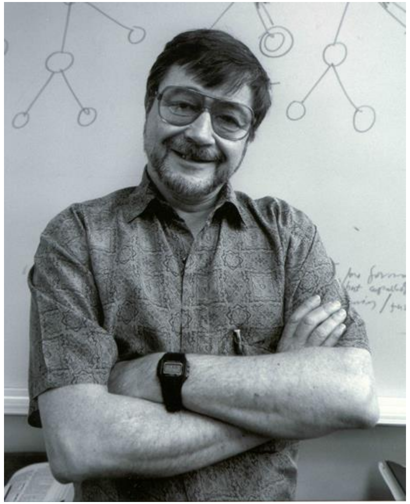
Donald Campbell (1916-1996)

Dawid Philip (*1946)

Clive Granger (1934-2009)

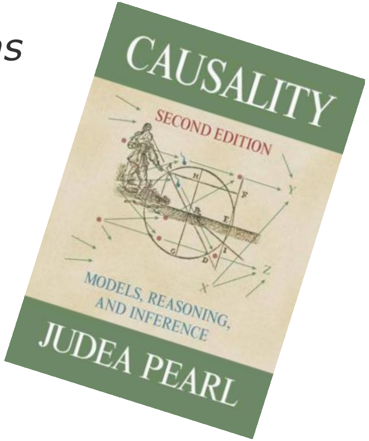
"[…] all approaches to causation are variants or abstractions of […] structural theory […]." Judea Pearl

3. A Paradigm Shift
Structural Causal Models



Judea Pearl
(*1936)

ACM Turing Award 2011:
"For fundamental contributions to artificial intelligence through the development of a calculus for probabilistic and causal reasoning."

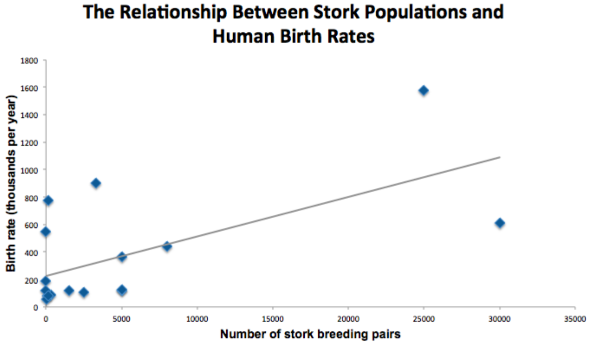


"[…] all approaches to causation are variants or abstractions of […] structural theory […]." Judea Pearl

4. Causal Graphical Models
The Idea in one Example

Do storks deliver babies?

Country	Area (km^2)	Storks (pairs)	Humans (10^6)	Birth rate (10^3/yr)
Albania	28,750	100	3.2	83
Austria	83,860	300	7.6	87
Belgium	30,520	1	9.9	118
Bulgaria	111,000	5000	9.0	117
Denmark	43,100	9	5.1	59
France	544,000	140	56	774
Germany	357,000	3300	78	901
Greece	132,000	2500	10	106
Holland	41,900	4	15	188
Hungary	93,000	5000	11	124
Italy	301,280	5	57	551
Poland	312,680	30,000	38	610
Portugal	92,390	1500	10	120
Romania	237,500	5000	23	367
Spain	504,750	8000	39	439
Switzerland	41,290	150	6.7	82
Turkey	779,450	25,000	56	1576

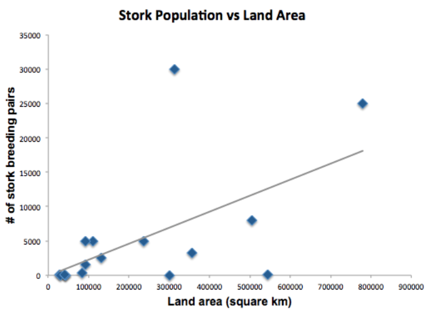


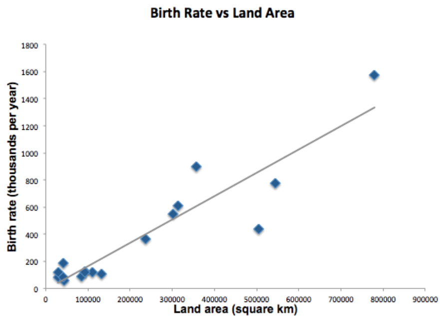
"Highly **statistically significant** degree of correlation between stork populations and birth rates" (or in technical terms, $p = 0.008$)

4. Causal Graphical Models
The Idea in one Example

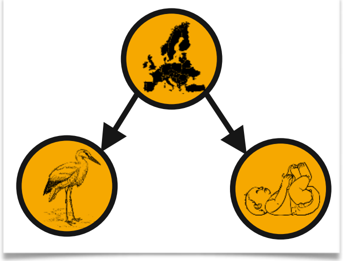
Do storks deliver babies?

Country	Area (km^2)	Storks (pairs)	Humans (10^6)	Birth rate (10^3/yr)
Albania	28,750	100	3.2	83
Austria	83,860	300	7.6	87
Belgium	30,520	1	9.9	118
Bulgaria	111,000	5000	9.0	117
Denmark	43,100	9	5.1	59
France	544,000	140	56	774
Germany	357,000	3300	78	901
Greece	132,000	2500	10	106
Holland	41,900	4	15	188
Hungary	93,000	5000	11	124
Italy	301,280	5	57	551
Poland	312,680	30,000	38	610
Portugal	92,390	1500	10	120
Romania	237,500	5000	23	367
Spain	504,750	8000	39	439
Switzerland	41,290	150	6.7	82
Turkey	779,450	25,000	56	1576





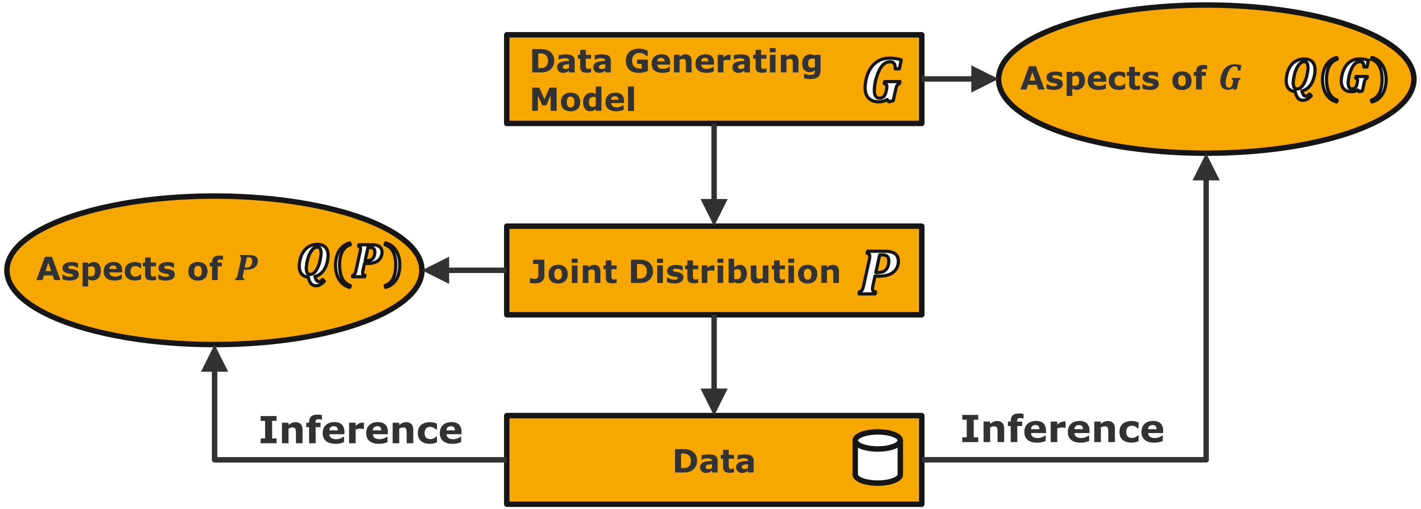
But a **simple variable that affects both** the birth rate and the stork population is the size of each country.



4. Causal Graphical Models
The Concept

Traditional Statistical Inference Paradigm

Paradigm of Structural Causal Models



5. The Calculus of Causality
Causal Inference: How to build a formal Theory?

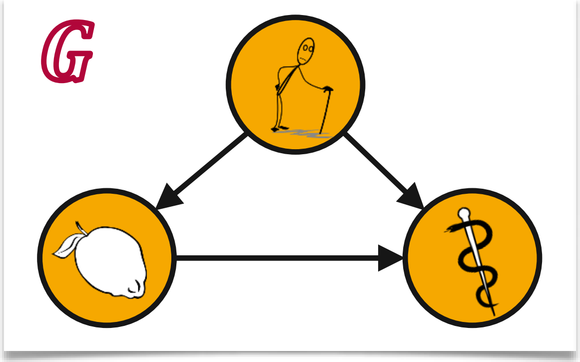
“Causality, although widely used, does not seem to be well-defined” (Lindley and Novick, 1981)

Problem: Probability theory has an associational, and not a causal nature.

To see this:

- Recap the scurvy experiment
- Assume that the data is generated by model $\mathbb{G}$.
  - The **recovery** of the scurvy is causally influenced by the **treatment with lemons**.
  - But now, both the **recovery** of scurvy as well as the **treatment with lemons** are causally influenced by the **age of the sailors**.
- The question remains:

Should we treat scurvy with lemons?

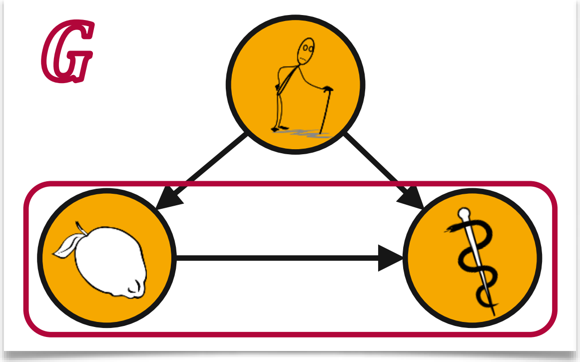


The Associational Nature of Probability Theory

- We run an experiment w.r.t. the model $\mathbb{G}$, i.e., we favor old sailors for treatment with lemons.

- The observed data of all sailors:

Combined	Recovery	No Recovery	Total	**Recovery Rate**
No lemons	20	20	40	50 %
Lemons	16	24	40	40 %
Total	36	44	80	



- Hence, we see that

$$P(recovery|lemons) < P(recovery|no\ lemons)$$

➡ Should we treat scurvy with lemons?

5. The Calculus of Causality
The Associational Nature of Probability Theory

- The observed data of old sailors:

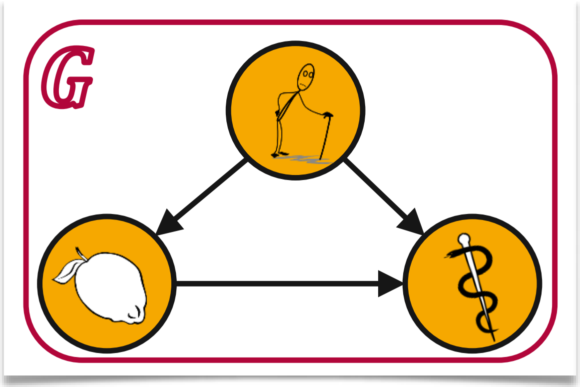
Old	Recovery	No Recovery	Total	Recovery Rate
No lemons	2	8	10	20 %
lemons	9	21	30	30 %
Total	11	29	40	

➡ $P(recovery|lemons, old) > P(recovery|no\ lemons, old)$

- The observed data of young sailors:

Young	Recovery	No Recovery	Total	Recovery Rate
No lemons	18	12	30	60 %
Lemons	7	3	10	70 %
Total	25	15	40	

➡ $P(recovery|lemons, young) > P(recovery|no\ lemons, young)$

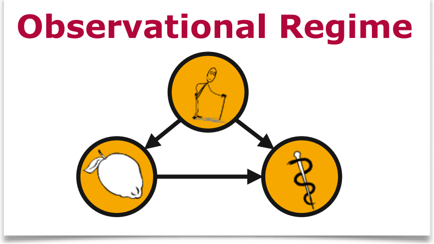


➡ Should we treat scurvy with lemons?

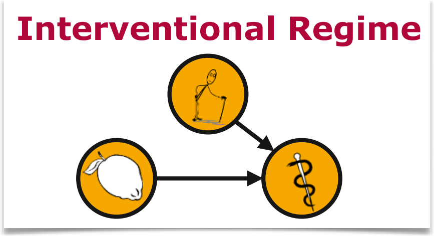
5. The Calculus of Causality
Pearl's contribution: the do-operator

- This reversal of the association between two variables after considering the third variable is called **Simpson's Paradox**.

➡ How to resolute the paradox and find an answer?



vs.



- In an interventional regime, all influences stemming from "natural causes" of the exposure variable are removed (e.g., see randomized experiments).
- Pearl extends probability calculus by introducing a new operator for describing interventions, the **do-operator**.

Example:

$P(lung\ cancer|smoke)$
Probability somebody gets lung cancer, **given that he smokes**.

$P(lung\ cancer|do(smoke))$
Probability somebody gets lung cancer, **if we force the person to smoke**.

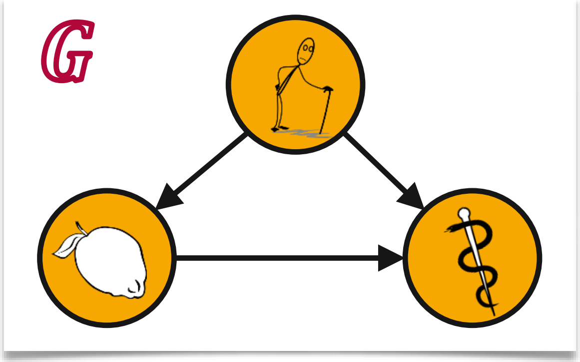
5. The Calculus of Causality
Application of the do-operator

Resolution of the Simpson's paradox

- Simpson's paradox is only paradoxical if we misinterpret

  $P(recovery|lemons)$ as $P(recovery|do(lemons))$
- We should treat scurvy with lemons if

  $P(recovery|do(lemons)) > P(recovery|do(no\ lemons))$



Derivation of the do-operator

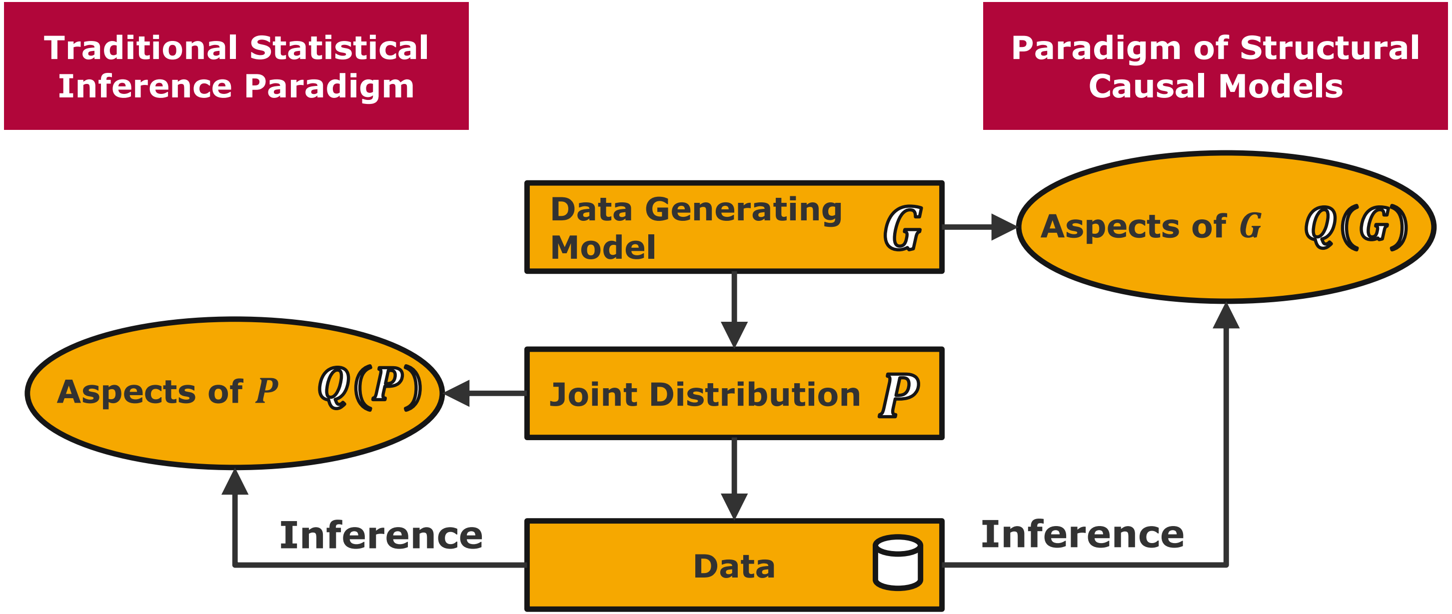
- If *identifiable*,
  $P(\cdot\,|do(\cdot))$ can be calculated from G and observational Data
- In our example, we have

$$P(recovery|do(lemons)) = \sum_{age} P(age)\, P(recovery|age, lemons) = 0.5$$

$$P(recovery|do(no\ lemons)) = \sum_{age} P(age)\, P(recovery|age, no\ lemons) = 0.4$$

➡ We should treat scurvy with lemons!

6. Summary and Outlook
Concept



E.g., what is the sailors' probability of recovery when **we see** a treatment with lemons?

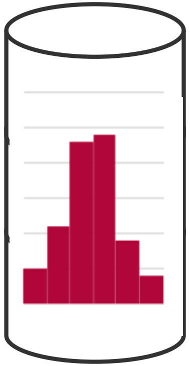
$$Q(P) = P(recovery|lemons)$$

E.g., what is the sailors' probability of recovery if **we do** treat them with lemons?

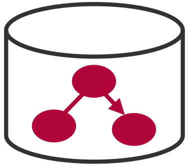
$$Q(G) = P(recovery|do(lemons))$$

6. Summary and Outlook
Inference Procedure

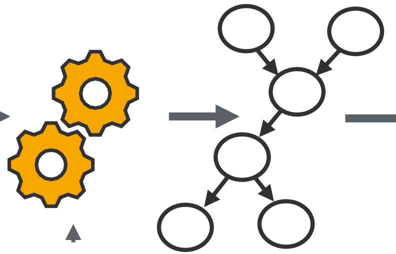
HPI Hasso Plattner Institut



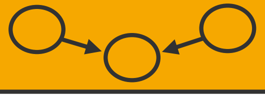
Observational Data



Background Knowledge



Causal Relationships



Causal Structure: *"What are the causal relationships in the system?"*

"How is lung cancer related to smoking and genetics?"

Probabilistic Inference

$P(X_3 | X_1 = x_1, X_2 = x_2)$

$P(X_4 | X_2 = x_2\ )$

Association: *"What is a certain probability if we find the system how it is?"*

"How likely do smoking people get lung cancer?"

Causal Inference

$P(X_3 | do(X_1 = x_1), do(X_2 = x_2))$

$P(X_4 | do(X_2 = x_2))$

Intervention: *"What is a certain probability if we manipulate the system?"*

"What if we ban cigarettes?

Functional Systems

$f_1(x_1, x_2) = e^{\alpha x_1} + \beta x_2 + \gamma$

$f_2(x_3, x_4) = \quad \ldots$

Counterfactuals: *"What if the system would have been different?"*

"What if I had not been smoking the past 2 years?"

Data | **Causal Structure Learning** | **Opportunities** | **Examples**

6. Summary and Outlook
Summary

Traditional statistics, machine learning, etc.

- About associations
- Model the distribution of the data
- Predict given observations

Causal Inference

- About causation
- Model the mechanism that generates the data
- Predict results of interventions

The following questions remain

- What are causal graphical models?
- How to recover these models from data?
- How to do causal inference in this model?

In order to answer these questions, we will learn about

- Causal Graphical Models $\mathbb{G}$
- Conditional Independence Testing $\mathbb{P}$
- Constraint-Based Causal Structure Learning $\rightarrow \mathbb{G}$
- Causal Inference on Causal Graphs $Q(\mathbb{G})$

Uflacker, Huegle, Schmidt

7. Further Reading

Literature

- Pearl, J. (2009). *Causal inference in statistics: An overview*. Statistics Surveys, 3:96-146.
- Pearl, J. (2009). *Causality: Models, Reasoning, and Inference.* Cambridge University Press.
- Pearl, J. (2011). *Simpson's paradox: An anatomy.* Department of Statistics, UCLA
- Spirtes, P., Glymour, C., and Scheines, R. (2000). Causation, Prediction, and Search. The MIT Press.

Lecture

- Judea Pearl's Turing Award Lecture: https://amturing.acm.org/vp/pearl_2658896.cfm

8. References
List of Figures

Thank you
for your attention!