UNIVERSITÄT ZU LÜBECK

Automated Planning and Acting

Intro to Causality

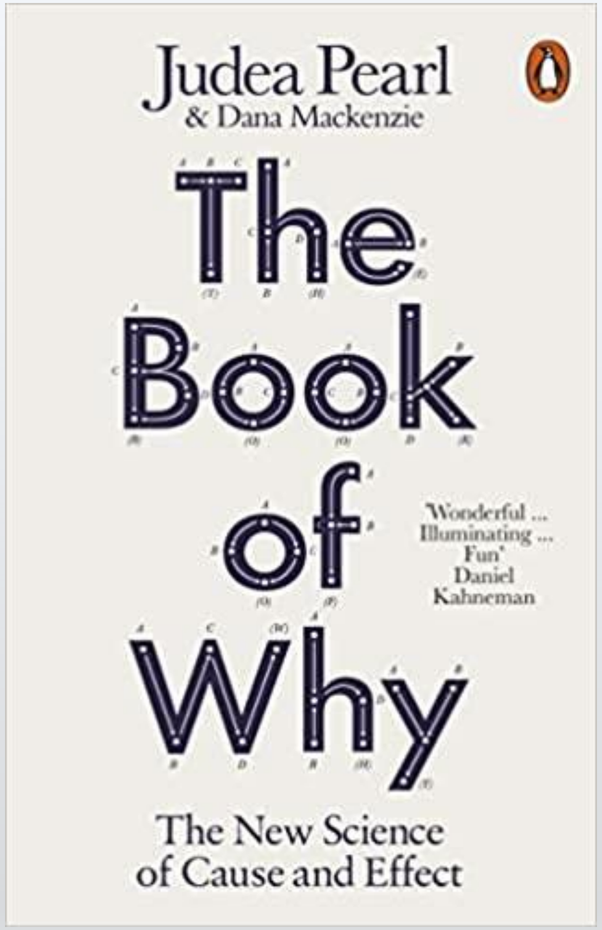
Institute of Information Systems

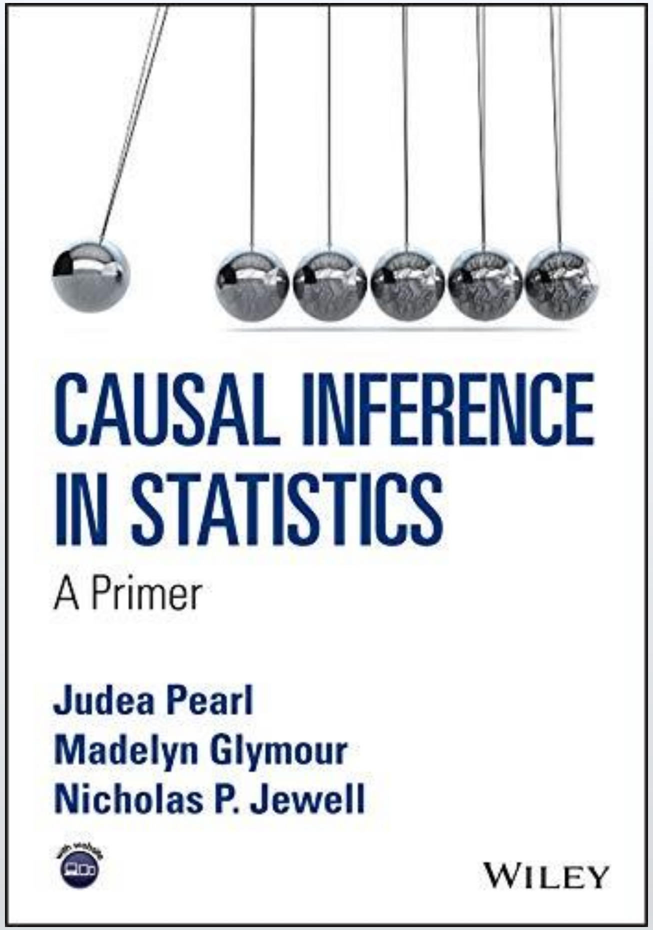
Dr. Mattis Hartwig

Content

1. Planning and Acting with **Deterministic** Models
2. Planning and Acting with **Refinement** Methods
3. Planning and Acting with **Temporal** Models
4. Planning and Acting with **Nondeterministic** Models
5. **Standard** Decision Making
6. Planning and Acting with **Probabilistic** Models
7. **Advanced** Decision Making
8. **Human-aware** Planning
9. Intro to Causality
10. Causal Planning

Two recommended books on causality

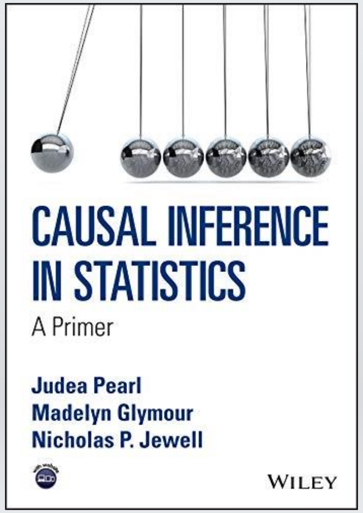




Acknowledgements

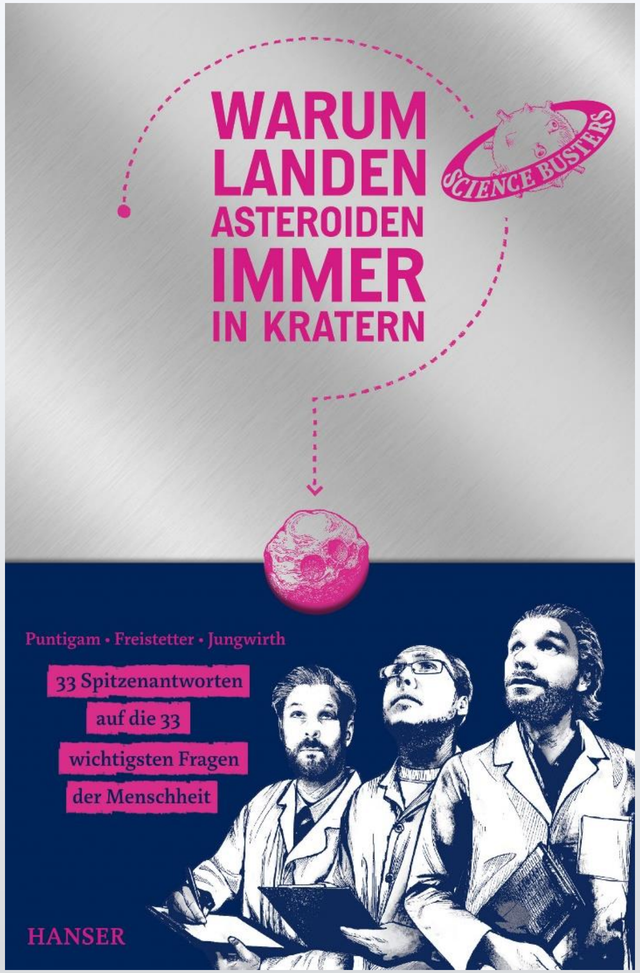
Inspired by Slides from Prof. Dr. Ralf Möller and Dr. Özgür Özçep

Based on "Causal Inference in Statistics: A Primer".



Motivation

- Usual warning: „Correlation is not causation"
- But sometimes (if not very often) one needs causation to understand statistical data



Simpsons Paradox

- Record recovery rates of 700 patients given access to a drug

	Recovery rate with drug	**Recovery rate without drug**
Men	81/87 (93%)	234/270 (87%)
Women	192/263 (73%)	55/80 (69%)
Combined	273/350 (78%)	289/350 (83%)

- Paradox:
  - For men, taking drugs has benefit
  - For women, taking drugs has benefit, too.
  - But: for all persons taking drugs has no benefit

Quiz

How can the data be explained? What is the correct implication?

	Recovery rate with drug	**Recovery rate without drug**
Men	81/87 (93%)	234/270 (87%)
Women	192/263 (73%)	55/80 (69%)
Combined	273/350 (78%)	289/350 (83%)

Resolving the Paradox (Informally)

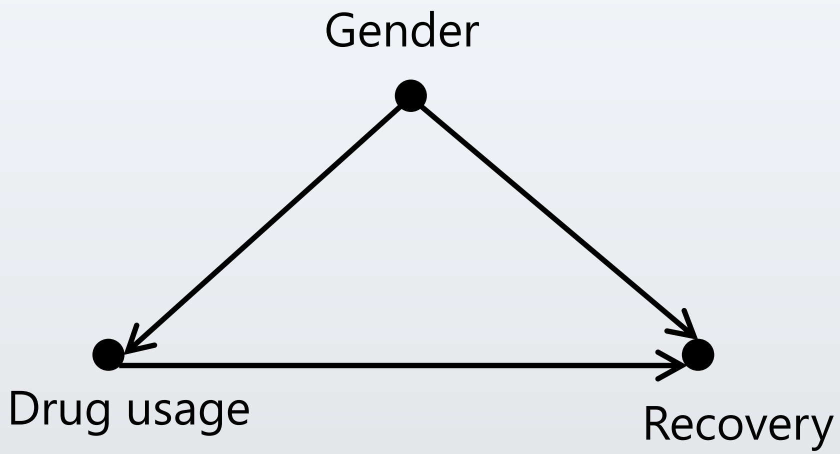
- We have to understand the causal mechanisms that lead to the data in order to resolve the paradox

- In drug example
  - Why has taking drug less benefit for women?
    - Answer: Estrogen has negative effect on recovery
  - Data: Women more likely to take drug than men
  - Choosing randomly any person taking drugs will rather give a woman – and for these recovery is less beneficial

- In this case: Have to consider segregated data
    - (not aggregated data)

Resolving the Paradox Formally

- We have to understand the causal mechanisms that lead to the data in order to resolve the paradox



- Drug usage and recovery have common cause
- Gender is a confounder

Simpson's Paradox (Again)

- Record recovery rates of 700 patients given access to a drug w.r.t. blood pressure (BP) segregation

	Recovery rate without drug	**Recovery rate with drug**
Low BP	81/87 (93%)	234/270 (87%)
High BP	192/263 (73%)	55/80 (69%)
Combined	273/350 (78%)	289/350 (83%)

- BP recorded at end of experiment
- This time segregated data recommend not using drug whereas aggregated data does

Resolving the Paradox (Informally)

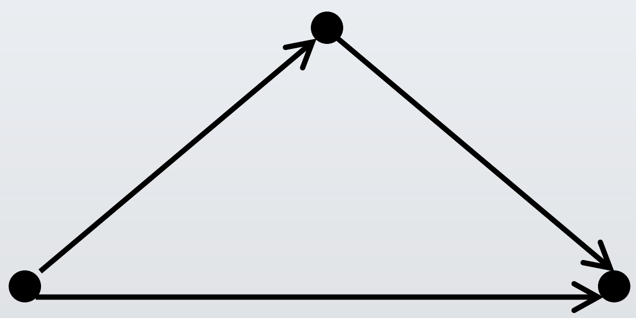
- We have to understand the causal mechanisms that lead to the data in order to resolve the paradox

- In this example
  - Drug effect is: lowering blood pressure (but may have
    - toxic effects)
  - Hence: In aggregated population drug usage recommended
  - In segregated data one sees only toxic effects

Resolving the Paradox Formally

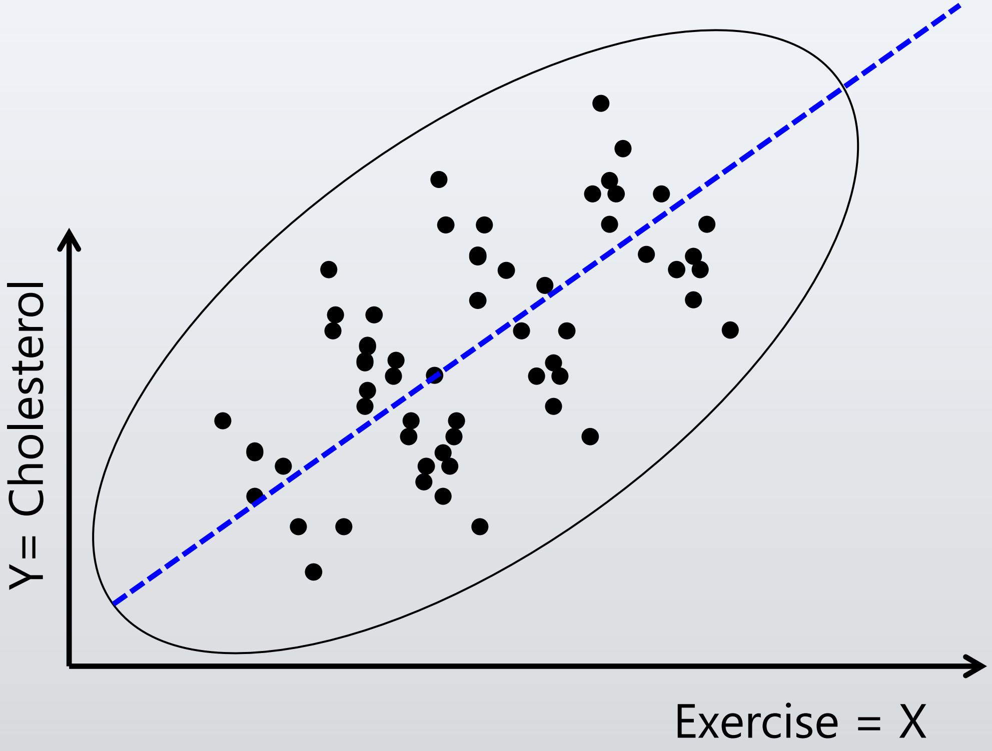
UNIVERSITÄT ZU LÜBECK

- We have to understand the causal mechanisms that lead to the data in order to resolve the paradox

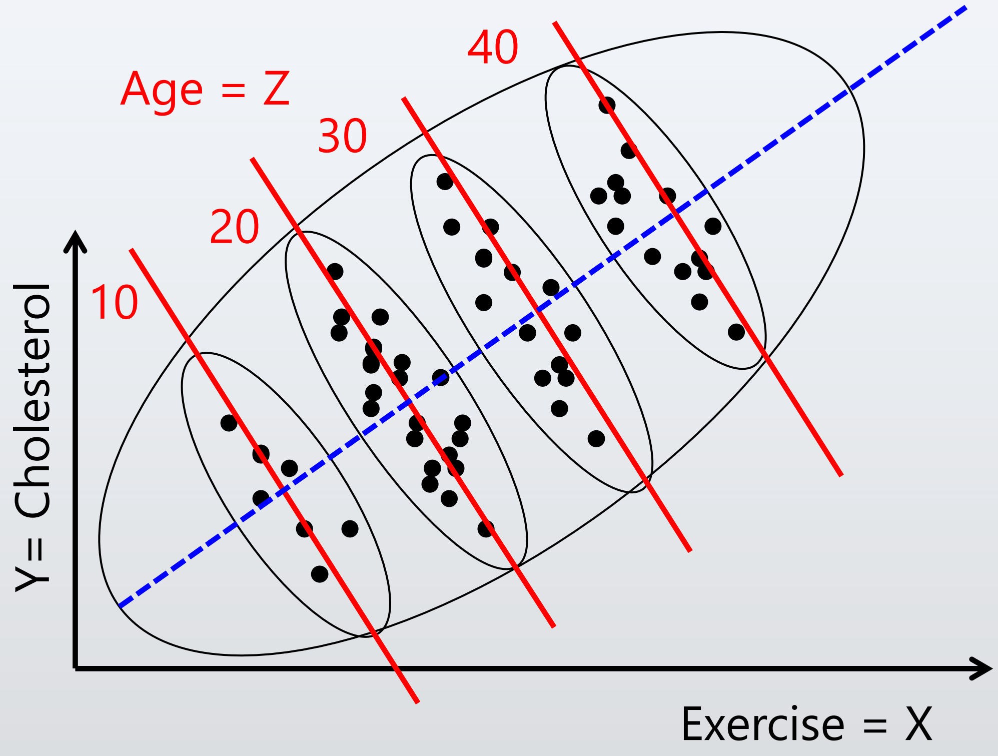


Quiz

- What is happening here? Exercise is bad for Cholesterol levels?



No



Ingredients of a Statistical Theory of Causality

- Working definition of causation
- Method for creating causal models
- Method for linking causal models with features of data
- Method for reasoning over model and data

Working definition of causality

A (random) variable X is a cause of a (random) variable Y if Y - in any way - relies on X for its value

Overview of basic definitions in a probabilistic graph

- Vertices
- Edges
- Adjacency - edge between two nodes
- Complete graph - edge between every pair of nodes
- Directed graph – edge that goes out of one node and into another (arrow)
- Path between X and Y – Sequence of nodes
- Parent
- Child
- Ancestor
- Descendant
- Acyclic

We are working with directed acyclic graphs (DAGs).

Bayesian Networks vs. SCMs

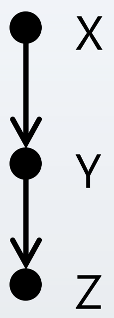
- BNs model statistical dependencies
  - Directed, but not necessarily cause-relation
  - Inherently statistical
  - Default application: discrete variables
- SCMs model causal relations
  - SCMs with random variables (RVs) induce BNs
  - Assumption: There is hidden causal (deterministic) structure behind statistical data
  - More expressive than BNs: Every BN can be modeled by SCMs but not vice versa
  - Default application: continuous variables

Reminder: Conditional Independence

- Event A independent of event B iff $P(A \mid B) = P(A)$
- RV X is independent of RV Y iff
  - $P(X \mid Y) = P(X)$ iff
    - for every x-value of X and for every y-value Y event $X = x$ is independent of event $Y = y$
    - Notation: $(X \perp\!\!\!\perp Y)_P$ or even shorter: $(X \perp\!\!\!\perp Y)$
- X is conditionally independent of Y given Z iff
  - $P(X \mid Y, Z) = P(X \mid Z)$
  - Notation: $(X \perp\!\!\!\perp Y \mid Z)_P$ or even shorter: $(X \perp\!\!\!\perp Y|Z)$

(In)dependences on Chains

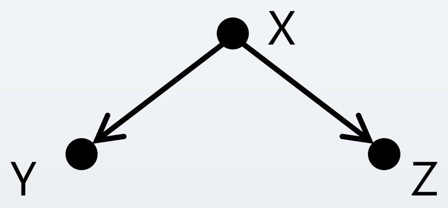
- Z and Y are dependent
- (For some z,y: $P(Z=z \mid Y = y) \neq P(Z = z)$)
- Y and X are dependent
  - **(…)**
- Z and X are likely dependent
- Z and X are independent, conditioned on Y
- (For all x,z,y: $P(Z=z \mid X=x, Y = y) = P(Z = z \mid Y = y)$)



Rule 1 (Conditional Independence in Chains)
Variables X and Z are independent given set of variables Y iff
there is only one path between X and Z and this path is unidirectional and Y intercepts that path

(In)dependences in Forks

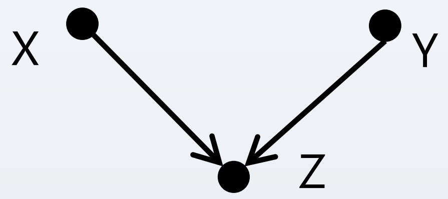
- X and Z are dependent
- ($\exists z,y$: $P(X=x \mid Z = z) \neq P(X = x)$)
- Y and X are dependent
  - …
- Z and Y are likely dependent
- Y and Z are independent, conditioned on X
- ($\forall x,z,y$: $P(Y=y \mid Z=z, X = x) = P(Y = y \mid X = x)$)



Rule 2 (Conditional Independence in Forks)
If variable X is a common cause of variables Y and Z
and there is only one path between Y, Z
then Y and Z are conditionally independent given X.

(In)dependence in Colliders

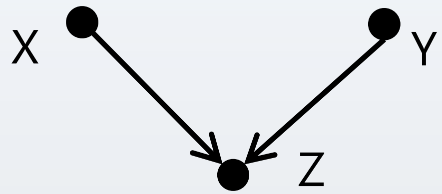
- X and Z are likely dependent
- ($\exists z,y$: $P(X=x \mid Z = z) \neq P(X = x)$)
- Y and Z are likely dependent
- X and Y are independent
- X and Y are likely conditionally dependent, given Z
- ($\exists x,z,y$: $P(X= x \mid Y=y, Z = z) \neq P(X = x \mid Z = z)$)



Rule 3 (Conditional Independence in Colliders)
If a variable Z is the collision node between variables X and Y and there is only one path between X, Y,
then X and Y are unconditionally independent, but are dependent conditional on Z and any descendant of Z

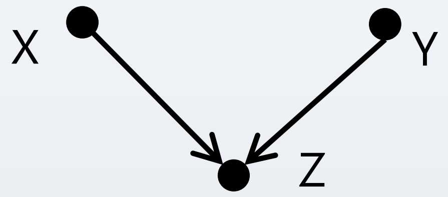
Quiz

- Give an example for a collider and interpret the conditional dependence.



Quiz

- Give an example for a collider and interpret the conditional dependence.



If scholarship received (Z)
but low grade (Y),
then must be musically talented (X)

D-Separation

Property

X independent of Y (conditioned on Z) for all compatible distributions iff
X d-separated from Y by Z in graph

- Z (possibly a set of variables) prohibits the ``flow'' of statistical effects/dependence between X and Y
  - Must block every path
  - Need only one blocking variable for each path

D-Separation definition

Definition (informal)

X is d-separated from Y by Z iff
Z blocks every possible path between X and Y

Definition (formal)

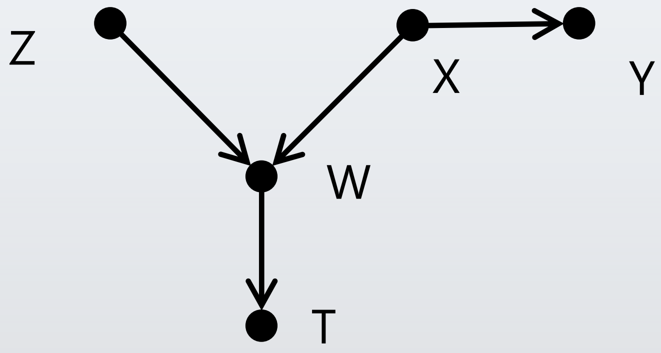
A path p in G (between X and Y) is blocked by Z iff

1. p contains chain $A \rightarrow B \rightarrow C$ or fork $A \leftarrow B \rightarrow C$ s.t. $B \in Z$
   or
2. p contains collider $A \rightarrow B \leftarrow C$ s.t. $B \notin Z$ and all descendants of B are $\notin Z$

If Z blocks every path between X and Y, then X and Y are d-separated conditional on Z, for short: $(X \perp\!\!\!\perp Y \mid Z)_G$

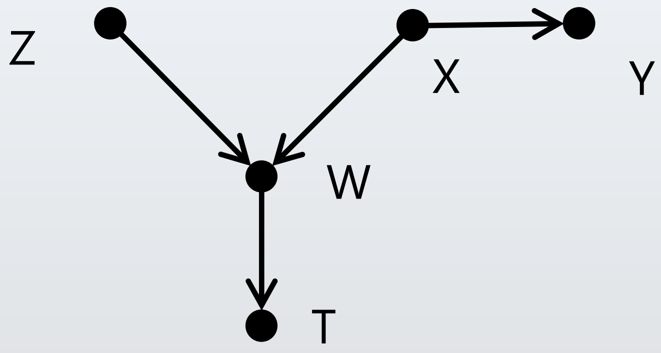
Quiz

- Given an empty conditioning set, are Z and Y dependent?



Quiz

- Given a conditioning set {X}, are Z and Y dependent?
- Given a conditioning set {W}, are Z and Y dependent?
- Given a conditioning set {X,W}, are Z and Y dependent?



Quiz

UNIVERSITÄT ZU LÜBECK

- Is there a situation without R in the conditioning set, where Z and Y are independet?