Causal Inference
Theory and Applications in Enterprise Computing

Christopher Hagedorn, Johannes Huegle, Dr. Michael Perscheid

May 27, 2020

HPI Hasso Plattner Institut
Digital Engineering · Universität Potsdam

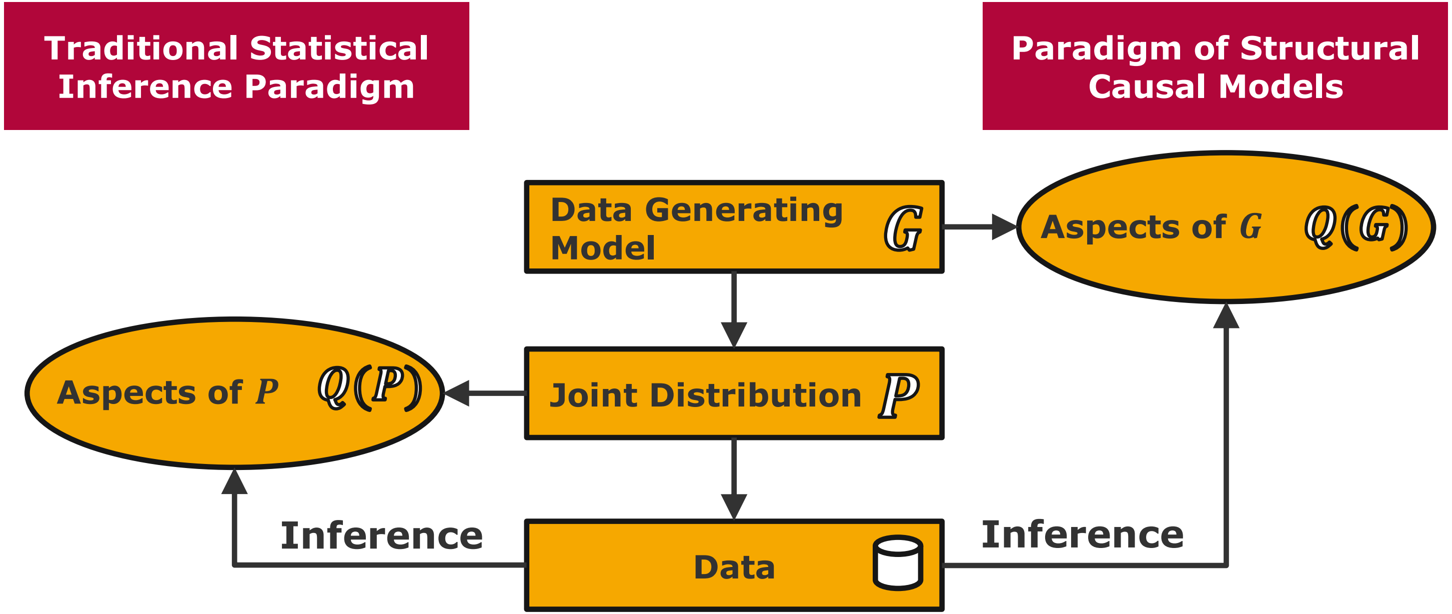
Agenda
May 27, 2020

- **Recap: Causal Inference in a Nutshell**
  - Causal Calculus

- **Jupyter Lab**
  1. Causal Inference in a Nutshell - Cooling House Scenario
  2. Causal Calculus – Exercises

Recap: Causal Inference in a Nutshell

Recap: Causal Inference in a Nutshell
Concept



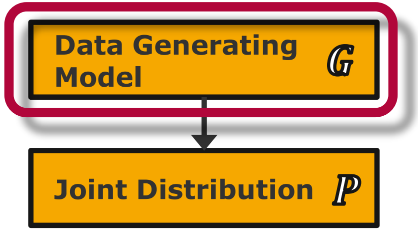
E.g., what is the sailors' probability of recovery when **we see** a treatment with lemons?

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

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

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

Recap: Causal Inference in a Nutshell
Causal Graphical Models



Causal Graphical Model

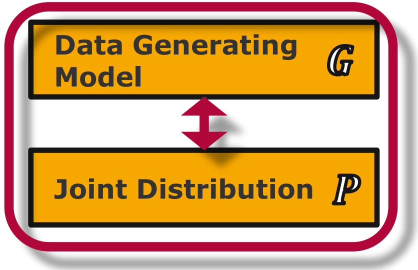
- *Directed Acyclic Graph (DAG)* $G = (V, E)$
  - *Vertices* $V_1, \ldots, V_n$
  - *Directed edges* $E = (V_i, V_j)$, i.e., $V_i \to V_j$
  - *No cycles*
- *Directed Edges* encode direct causes via
  - $V_j = f_j(\mathrm{Pa}(\mathrm{V_j}), \mathrm{N_j})$ with independent noise $N_1, \ldots, N_n$

Causal Sufficiency

- All relevant variables are included in the DAG G

Recap: Causal Inference in a Nutshell

Connecting G and P



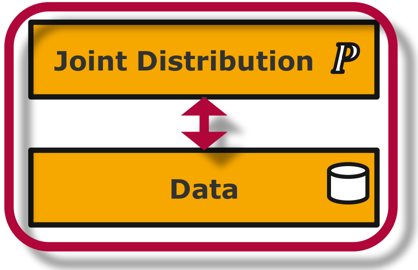
$$(X \perp\!\!\!\perp Y|Z)_G \Rightarrow (X \perp\!\!\!\perp Y|Z)_P$$

- Key postulate: *(Local) Markov Condition*
- Essential mathematical concept: *d-Separation*
  - Idea: *Blocking* of paths
  - Implication: *Global Markov Condition*

$$(X \perp\!\!\!\perp Y|Z)_G \Leftarrow (X \perp\!\!\!\perp Y|Z)_P$$

- Key postulate: *Causal Faithfulness*

Recap: Causal Inference in a Nutshell
Connecting P and



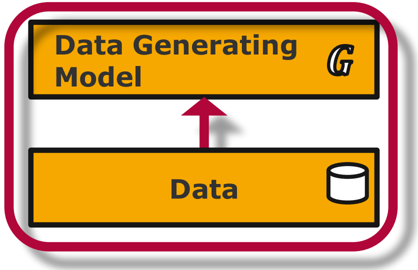
Statistical Inference

- Essential concept: *Point estimator* $\hat{\Theta}$
  - *Statistic* $g(X_1, \ldots, X_n)$ of *random samples* $X_1, \ldots, X_n$ to estimate *population parameter* Θ
- Inference: *Statistical Hypothesis Test*
  - *Null Hypothesis* H_0, claim on a population's property initially assumed to be true
  - *Alternative Hypothesis* H_1, a claim that contradicts H_0
  - Rejection criteria for H_0: c-value $T(x) > \mathrm{c}$ or equivalently p-value $P_{H_0}\big(T(X) > T(x)\big) < \alpha$

$$(\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y} | \boldsymbol{Z})_{\boldsymbol{P}} \Leftarrow$$

- Method: *Conditional Independence Test*
  - Distribution of $\boldsymbol{V} = \{V_1, \ldots, V_N\} \Rightarrow$ dependence measure $T(V_i, V_j, \boldsymbol{S}) \Rightarrow$ hypothesis $H_0: t = 0$

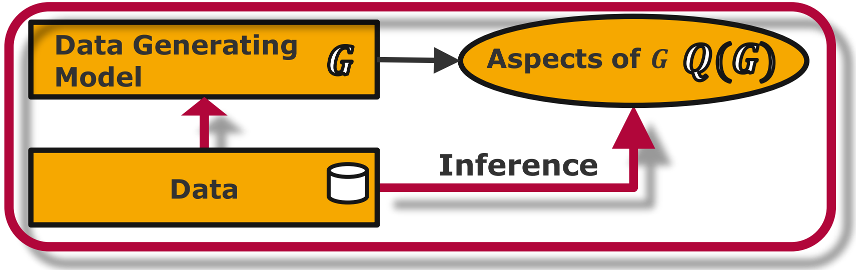
Recap: Causal Inference in a Nutshell
Causal Structure Learning



Causal Structure Learning

- Assumptions: *Causal Sufficiency, Markov Condition, Causal Faithfulness*
- Idea: Accept only those DAG's G for which $(X \perp\!\!\!\perp Y \mid Z)_G \Leftrightarrow (X \perp\!\!\!\perp Y \mid Z)_P$
  - Identifies DAG up to *Markov equivalence class* (i.e., same *skeleton C* and *v-structures*)
  - Markov equivalence class uniquely described by *completed partially directed acyclic graph (CPDAG)*
- Basis: V_i and V_j are linked if and only if there is no $S(V_i, V_j)$ s.t. $\left(V_i \perp\!\!\!\perp V_j \mid S(V_i, V_j)\right)_P$
- Methods:
  - *Constraint-based*: CI testing to derive skeleton together with edge orientation rules
  - *Score-based*: "search-and-score approach"
  - *Hybrid*: Constraint-based skeleton derivation and score-based edge orientation

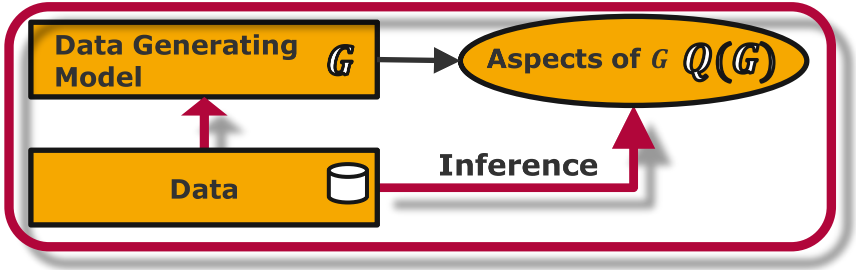
Recap: Causal Inference in a Nutshell
Causal Calculus



Do-Calculus

- Overview: Simpson's paradox is only paradoxical if we misinterpret $P(y|x)$ as $P(y|do(x))$
  - *Bayesian conditioning*, $p(y|x)$, where x is observed variable
  - *Causal conditioning*, $p(y|do(x))$, where we force a specific value x
- Key postulate: *Identifiability*
- Essential mathematical concepts: *Perturbed Graphs*
  - *Back-Door Criterion*, adjustment by conditioning on confounding back-door paths
  - *Front-Door Criterion*, adjustment if conditioning on confounding path is not possible
- Formalism of interventions: *do-Calculus*
  - Rules: *Ignoring observations, action exchange, ignoring actions*
  - Properties: *Calculus is sound and complete*

Recap: Causal Inference in a Nutshell
Adjustment



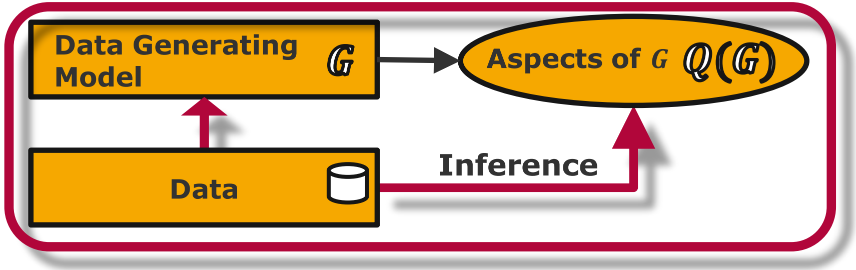
Back-Door Criterion

- A set of variables Z satisfies the *back-door criterion* relative to (V_i, V_j) in a DAG G if:
  - no node in Z is a descendant of V_i; and
  - Z blocks every path between V_i and V_j that contains an arrow to V_i
- *Back-door adjustment*: $P\left(v_j \middle| do(v_i)\right) = \sum_z P\left(v_j | v_i, z\right) P(z)$

Front-Door Criterion

- A set of variables Z satisfies the *front-door criterion* relative to (V_i, V_j) in a DAG G if:
  - Z intercepts all directed paths from V_i to V_j ; and
  - there is no unblocked back-door path from V_i to Z; and
  - all back-door paths from Z to V_j are blocked by V_i
- *Front-door adjustment*: $P\left(v_j \middle| do(v_i)\right) = \sum_z P(z|v_i) \sum_{v_i'} P(v_j | v_i', z) P(v_i')$

Recap: Causal Inference in a Nutshell
do-Calculus



do-Calculus

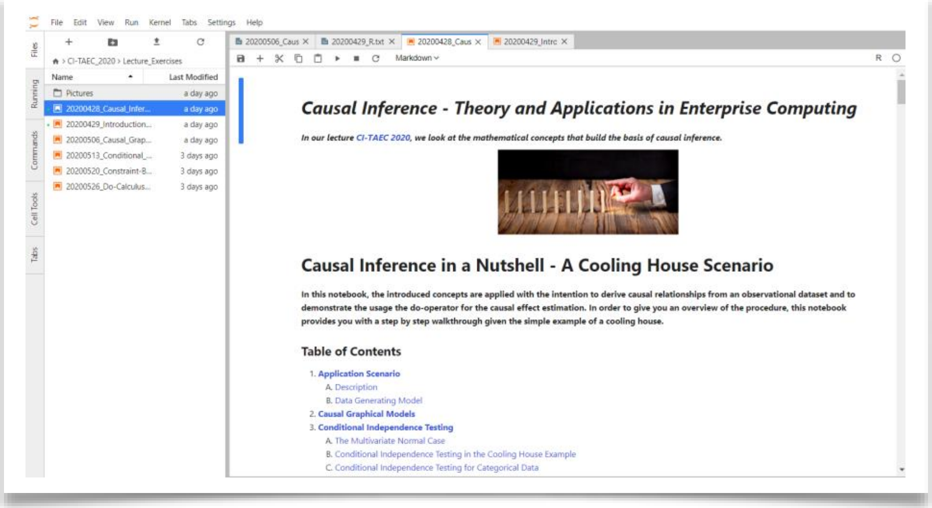
- Let X, Y, Z, and W be arbitrary disjoint sets of nodes in a causal DAG G.
  - *Rule 1: Ignoring observations*
    $$p(y|do(x),z,w) = p(y|do(x),w) \quad if\ (Y \perp\!\!\!\perp Z\,|\,X,W)_{G_{\overline{X}}}$$
  - *Rule 2: Action/Observation exchange (Back-Door)*
    $$p(y|do(x),do(z),w) = p(y|do(x),z,w) \quad if\ (Y \perp\!\!\!\perp Z\,|\,X,W)_{G_{\overline{X},\underline{Z}}}$$
  - *Rule 3: Ignoring actions/interventions*
    $$p(y|do(x),do(z),w) = p(y|do(x),w) \quad if\ (Y \perp\!\!\!\perp Z\,|\,X,W)_{G_{\overline{X},\overline{Z(W)}}}$$

Causal Effects

- *Causal effect of $V_i = v_i$ on V_j:* $P(V_j|do(V_i = v_i))$
- *Causal strength:* Dependent to the causal structures, e.g., *ATE* for binary V_i

Jupyter Lab

Jupyter Lab
Causal Inference in a Nutshell - Cooling House



Topics

- Estimating Causal Effects in the Cooling House Example
- Further Opportunities of Causal Structures

Jupyter Lab

Causal Calculus - Exercises

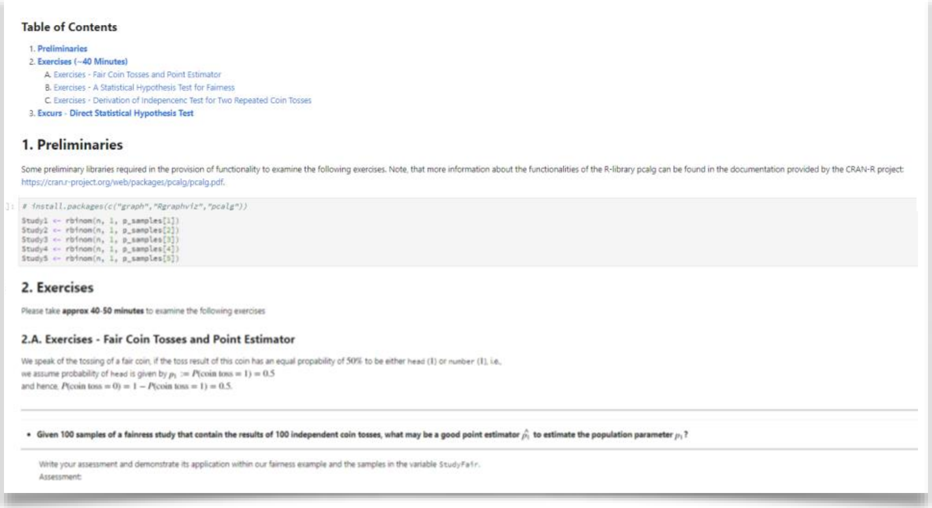


Table of Contents

1. Preliminaries

Some preliminary libraries required in the provision of functionality to examine the following exercises. Note, that more information about the functionalities of the R-library pcalg can be found in the documentation provided by the CRAN-R project: https://cran.r-project.org/web/packages/pcalg/pcalg.pdf.

```
# install.packages(c("graph","Rgraphviz","pcalg"))
Study1 <- rbinom(n, 1, p_samples[1])
Study2 <- rbinom(n, 1, p_samples[2])
Study3 <- rbinom(n, 1, p_samples[3])
Study4 <- rbinom(n, 1, p_samples[4])
Study5 <- rbinom(n, 1, p_samples[5])
```

2. Exercises

Please take **approx 40-50 minutes** to examine the following exercises

2.A. Exercises - Fair Coin Tosses and Point Estimator

We speak of the tossing of a fair coin, if the toss result of this coin has an equal propability of 50% to be either head (1) or number (1), i.e.,
we assume probability of head is given by $p_1 := P(\text{coin toss} = 1) = 0.5$
and hence, $P(\text{coin toss} = 0) = 1 - P(\text{coin toss} = 1) = 0.5$.

- **Given 100 samples of a fairness study that contain the results of 100 independent coin tosses, what may be a good point estimator $\hat{p}_1$ to estimate the population parameter p_1?**

Write your assessment and demonstrate its application within our fairness example and the samples in the variable StudyFair.
Assessment:

Exercises

- Seeing vs. Doing
- Identifiability
- Causal Effects from Observational Data

Thank you
for your attention!