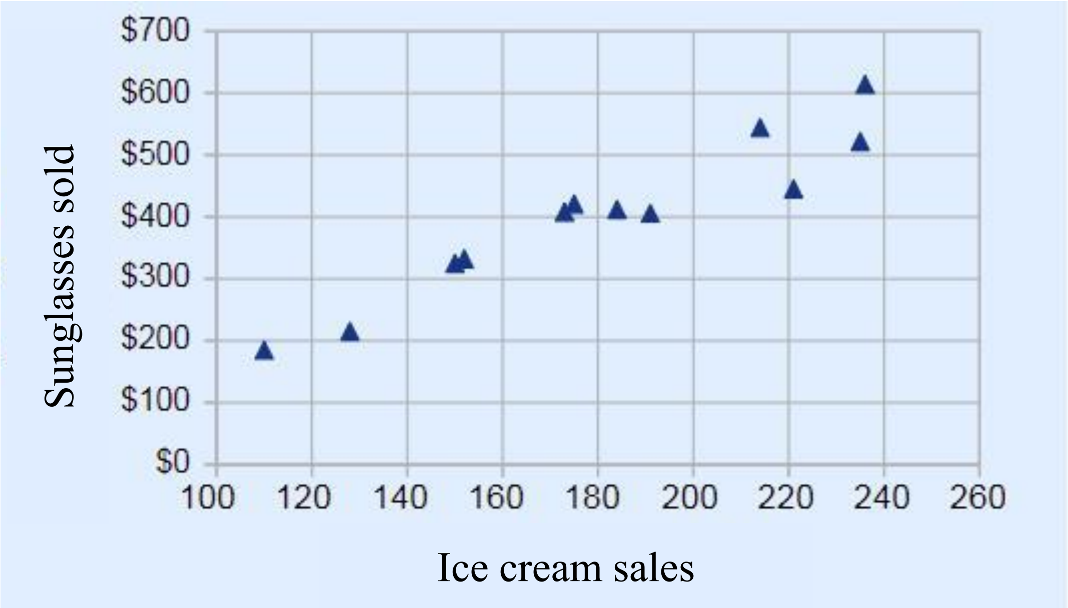
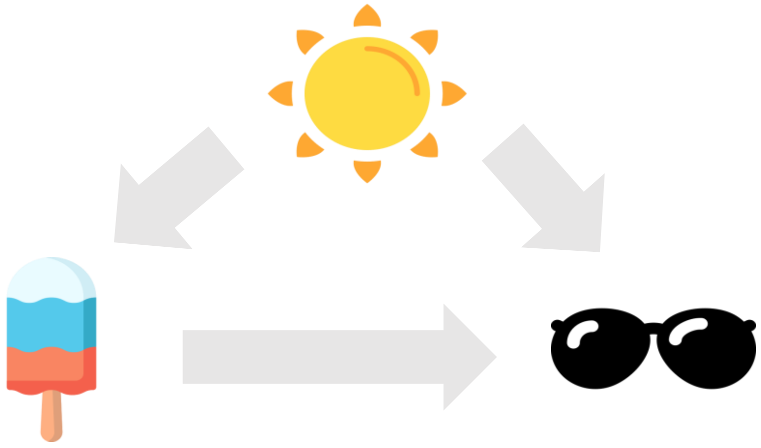
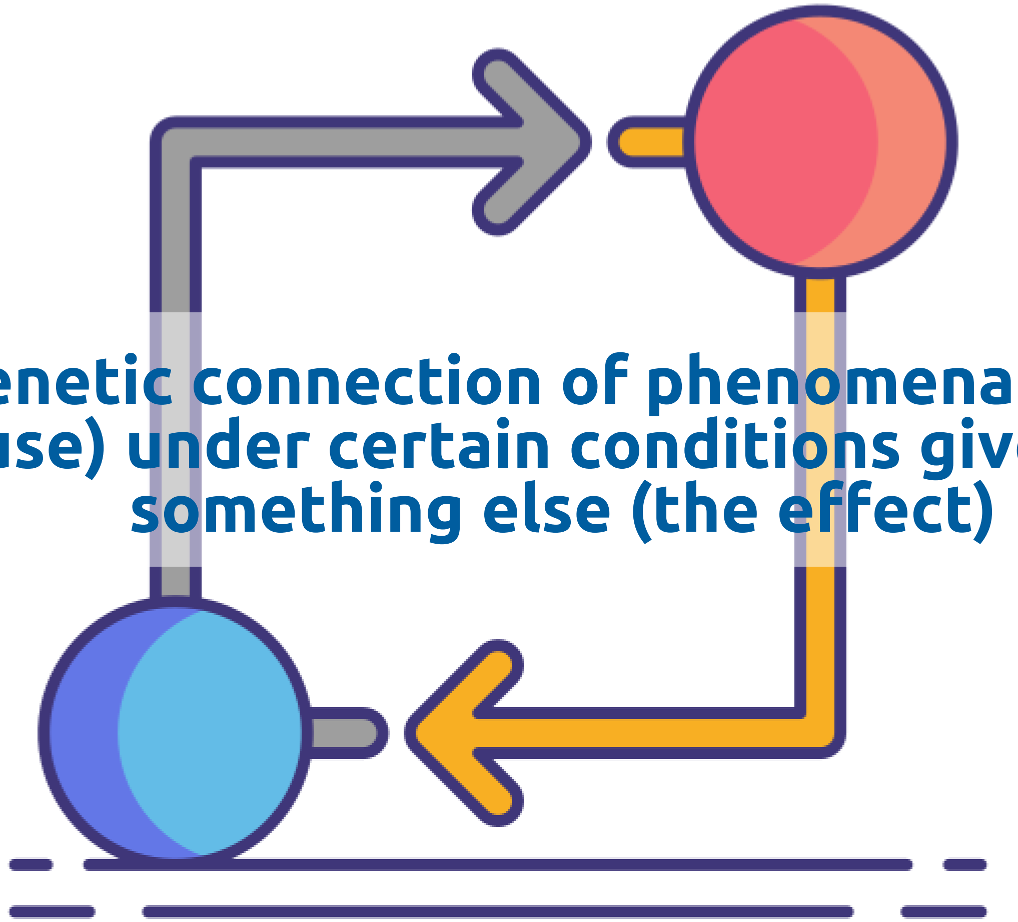
Introduction to Causality

Animesh Danayak, Gayathri Anil
8 June 2020

Does Correlation mean Causation?

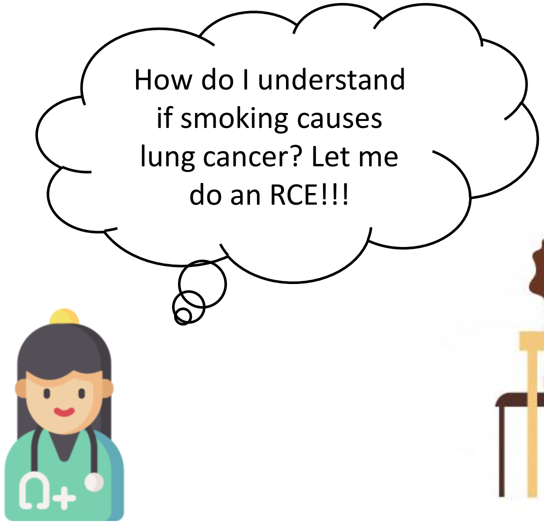






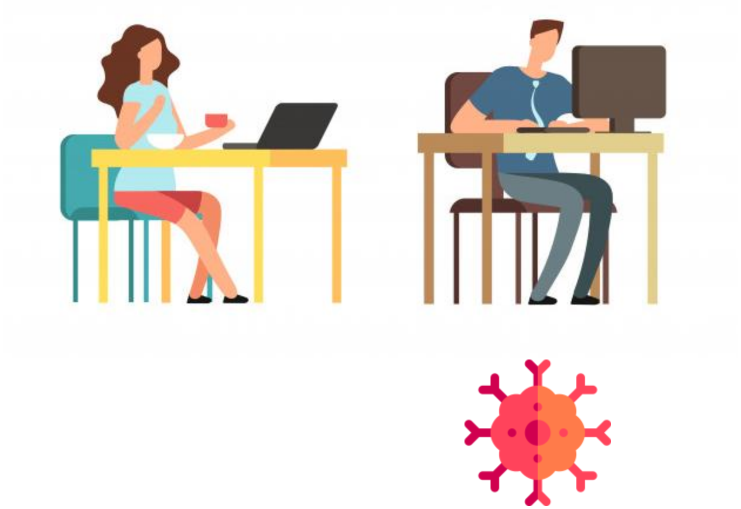
Causality is a genetic connection of phenomena through which one thing (the cause) under certain conditions gives rise to, causes something else (the effect)

A causal relationship is determined via Randomized Controlled Experiments



Test Sample

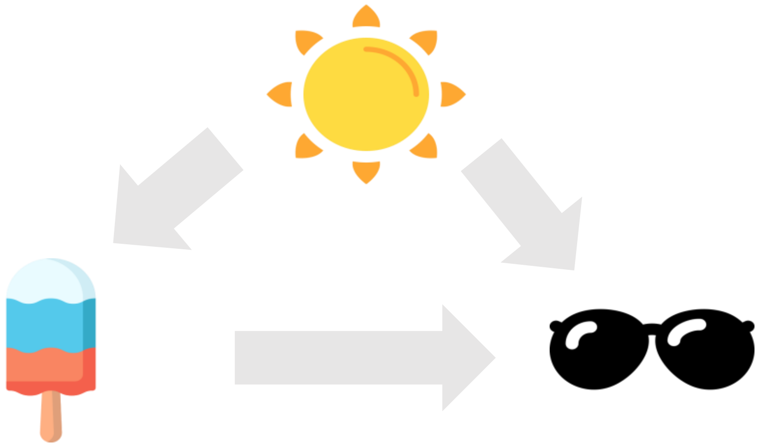
Control Sample

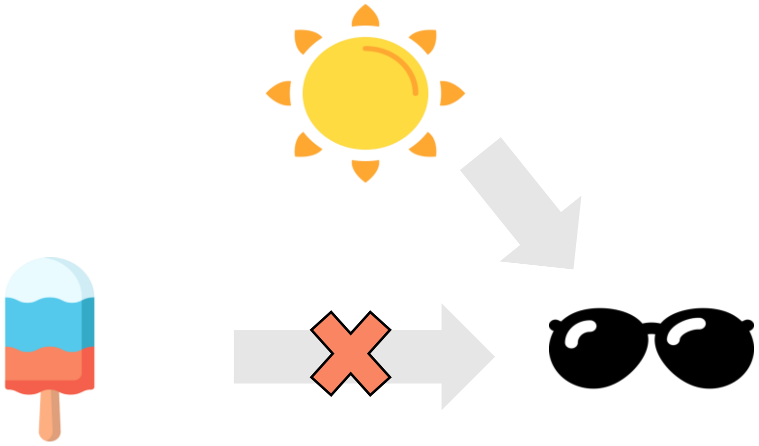


In situations where RCEs are expensive and not feasible, Do - Calculus becomes a data – driven proxy to identify causality

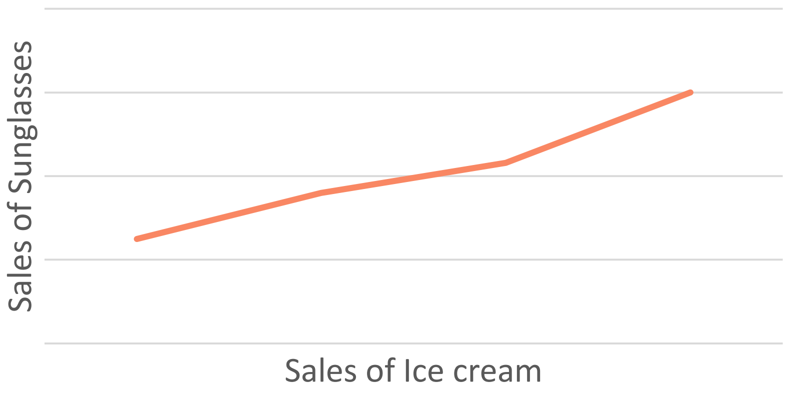
Observational distribution		Interventional distribution
p(Y | X)	→	**p(Y | do(X = x))**
Distribution of Y given that I **observe** variable X takes value X		Distribution of Y if I were to **set** the value of X to x.

do–Calculus helps us discredit non-causal relationships in data



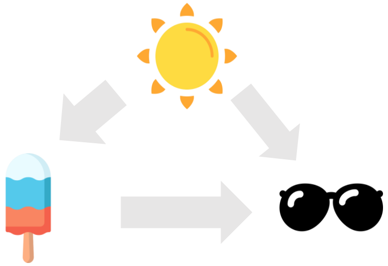


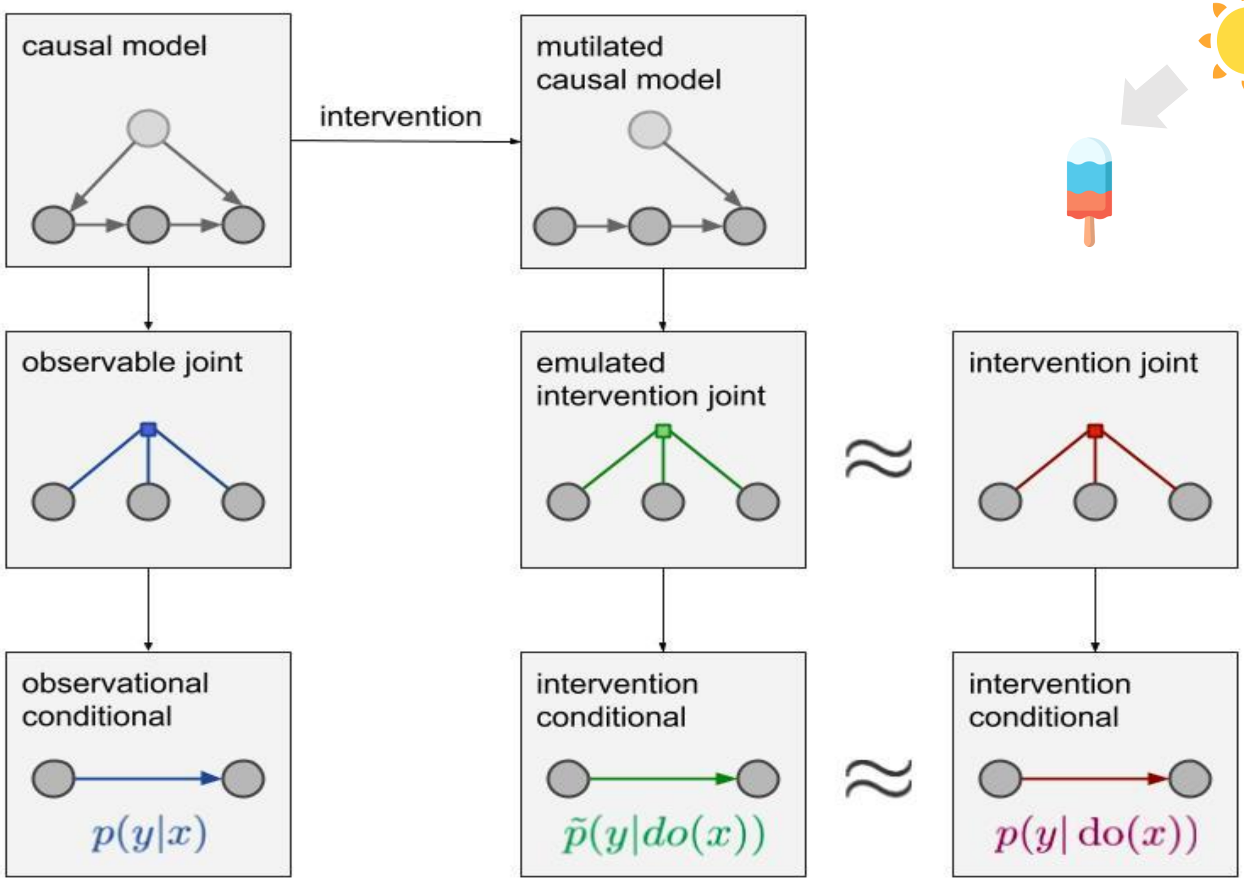
Observational Distribution

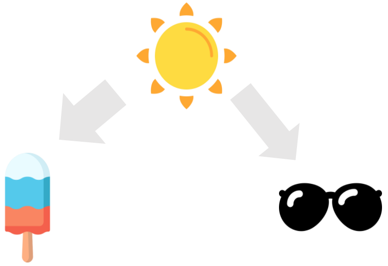


Interventional Distribution

Moving from conventional ML models to causal models







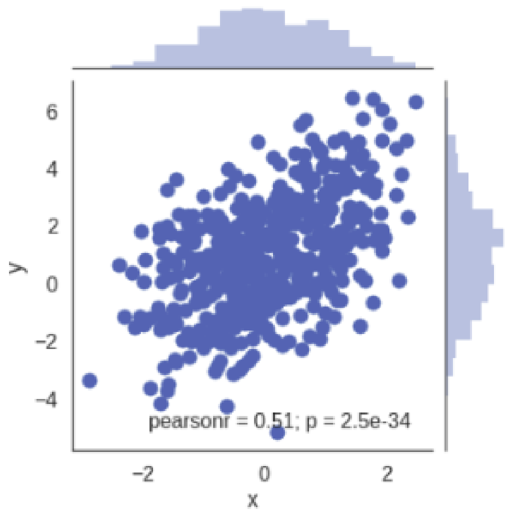
The need of Causal Diagrams for simulating Interventions

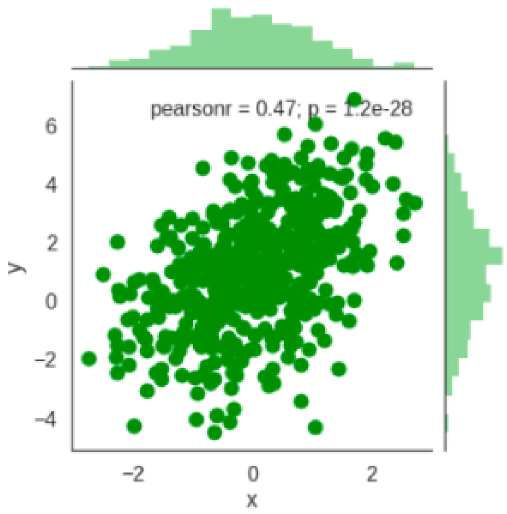
Three scripts:

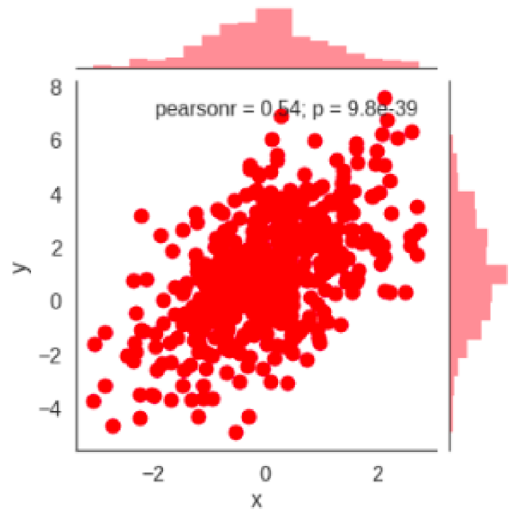
```
x = randn()
y = x + 1 + sqrt(3)*randn()
```

```
y = 1 + 2*randn()
x = (y-1)/4 + sqrt(3)*randn()/2
```

```
z = randn()
y = z + 1 + sqrt(3)*randn()
x = z
```







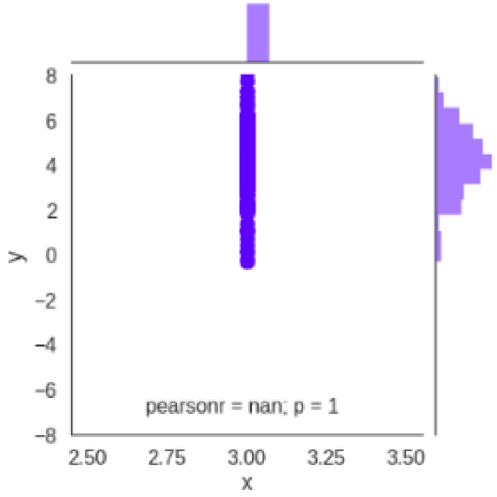
The need of Causal Diagrams for simulating Interventions

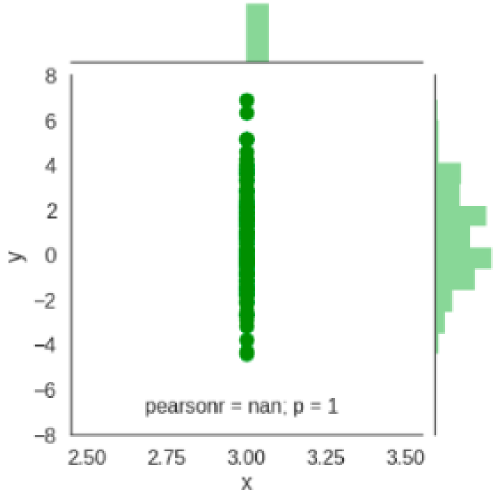
Interventions:

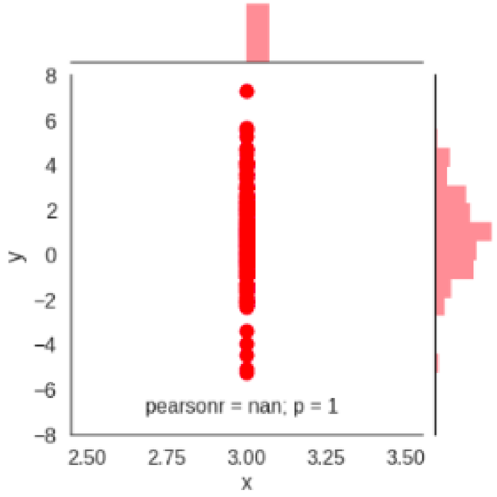
x = randn()
x = 3
y = x + 1 + sqrt(3)*randn()
x = 3

y = 1 + 2*randn()
x = 3
x = (y-1)/4 + sqrt(3)*randn()/2
x = 3

z = randn()
x = 3
x = z
x = 3
y = z + 1 + sqrt(3)*randn()
x = 3

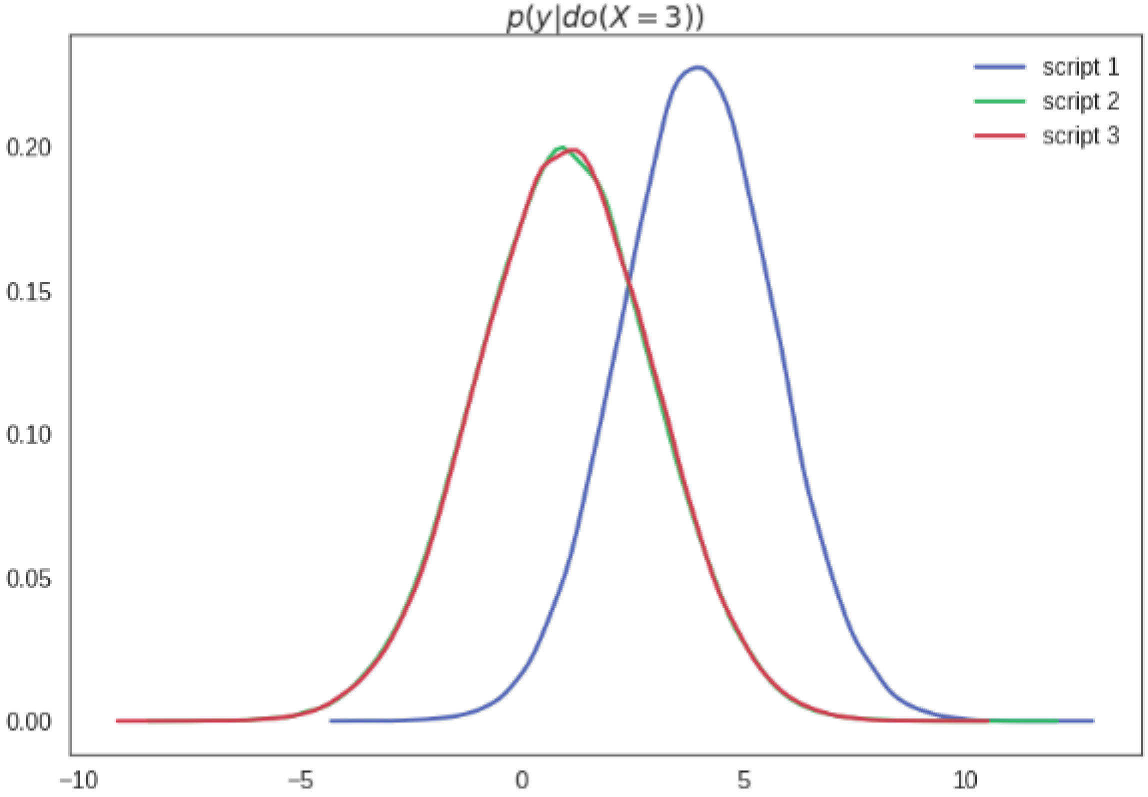






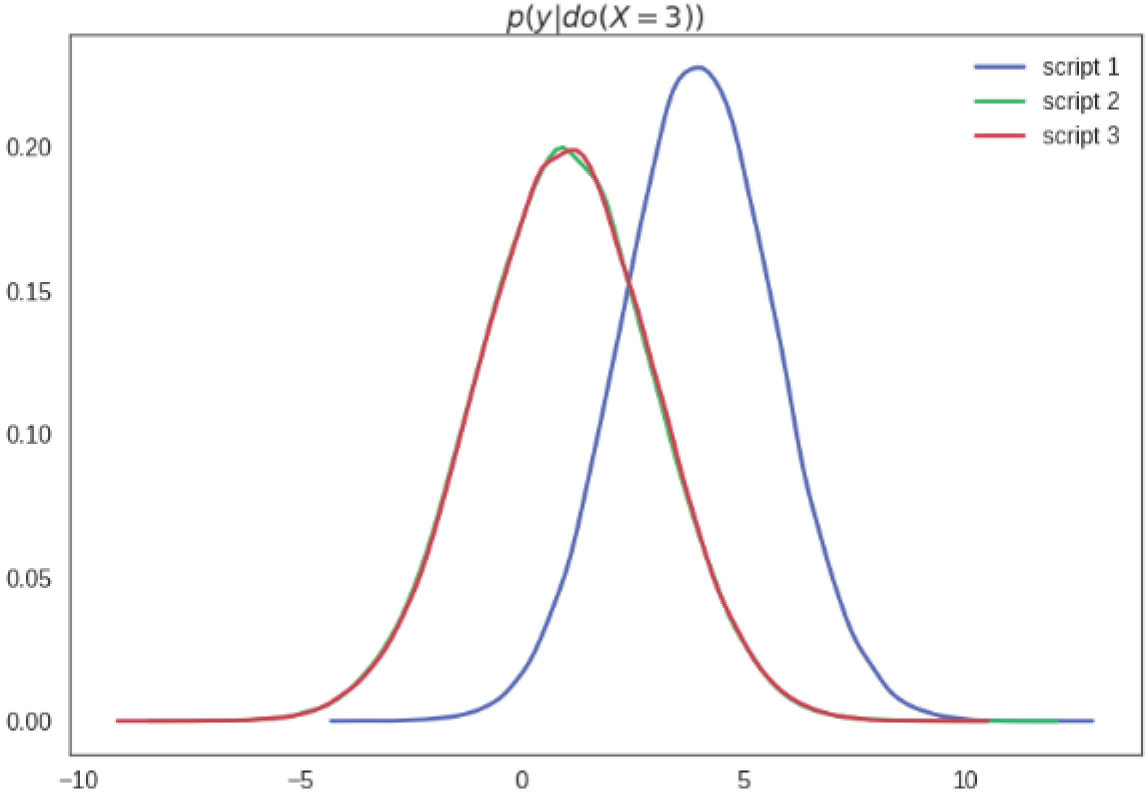
The need of Causal Diagrams for simulating Interventions

Result of this intervention:



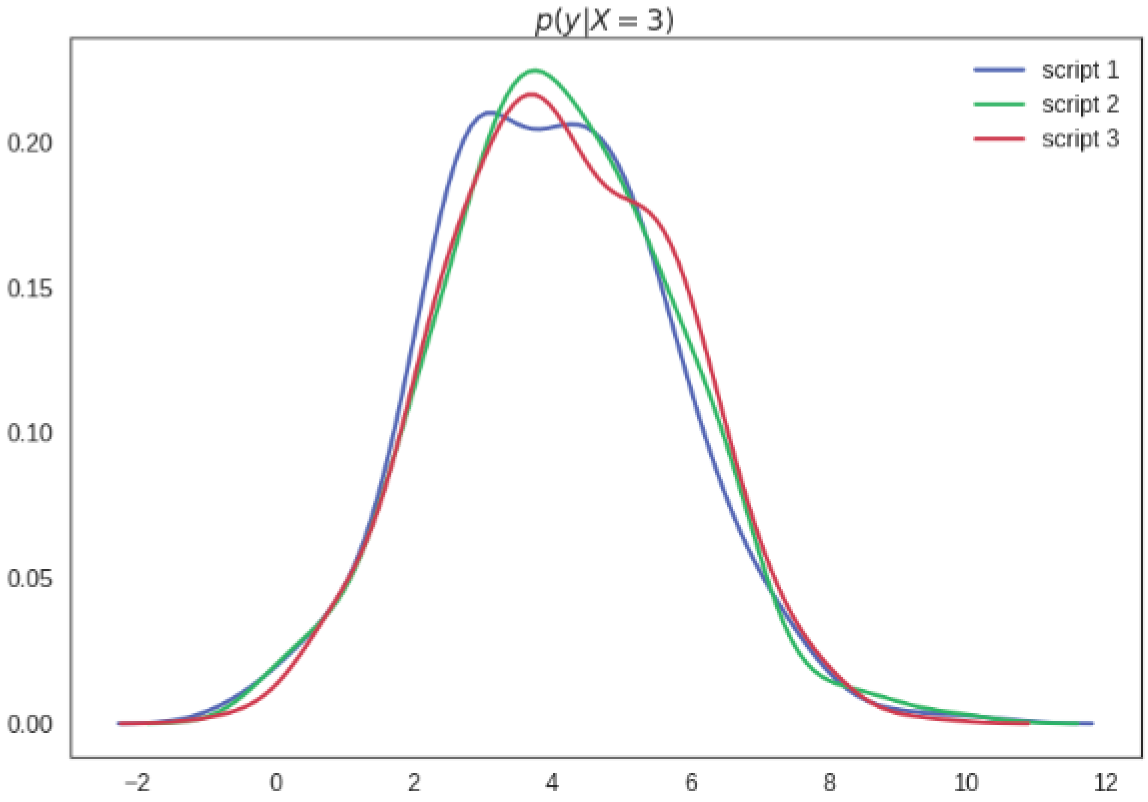
The need of Causal Diagrams for simulating Interventions

Result of this intervention:



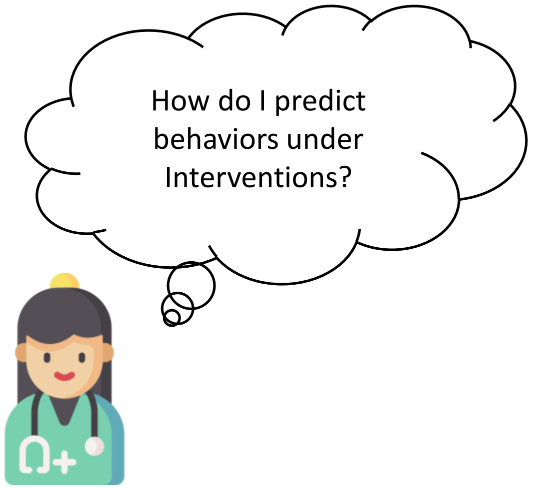
The need of Causal Diagrams for simulating Interventions

If the intervention was not made:



The need of Causal Diagrams for simulating Interventions

- Based on the Joint Distribution of data, the 3 scripts are INDISTINGUISHABLE.
- The scripts behave differently under INTERVENTION
- Consequently, Joint Distribution of data alone is insufficient to predict behaviors under an intervention.



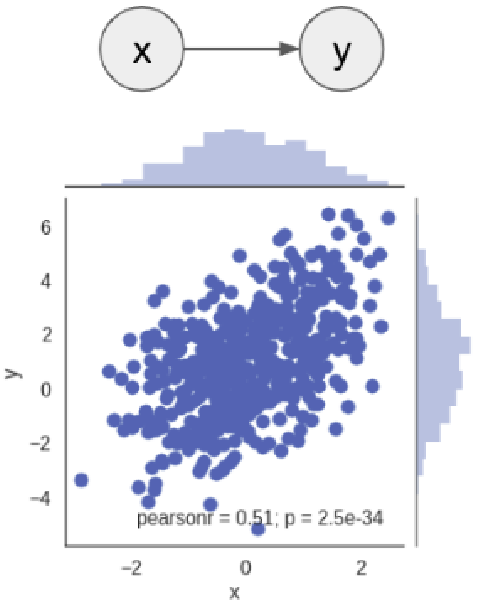
The need of Causal Diagrams for simulating Interventions

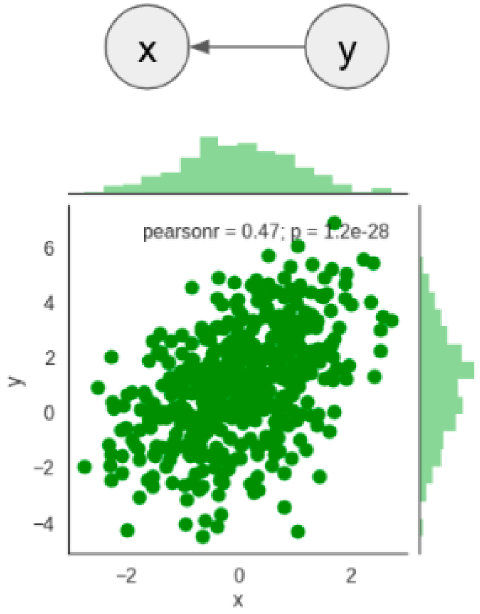
Causal Diagrams:

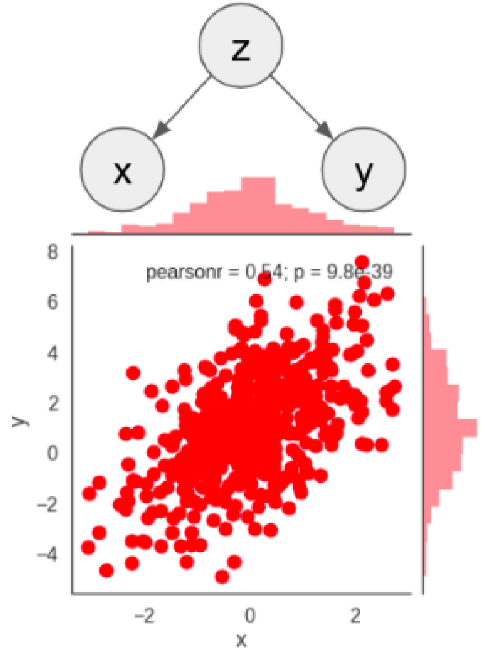
```
x = randn()
y = x + 1 + sqrt(3)*randn()
```

```
y = 1 + 2*randn()
x = (y-1)/4 + sqrt(3)*randn()/2
```

```
z = randn()
y = z + 1 + sqrt(3)*randn()
x = z
```

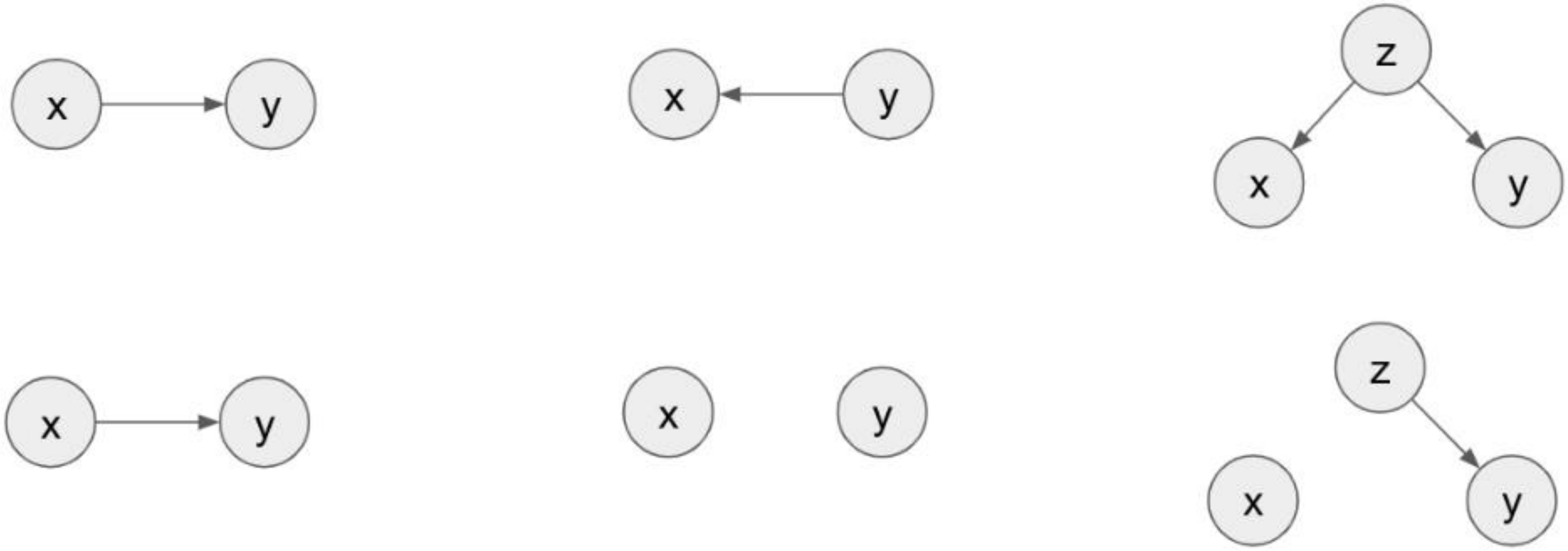






The need of Causal Diagrams for simulating Interventions

Causal Diagrams:



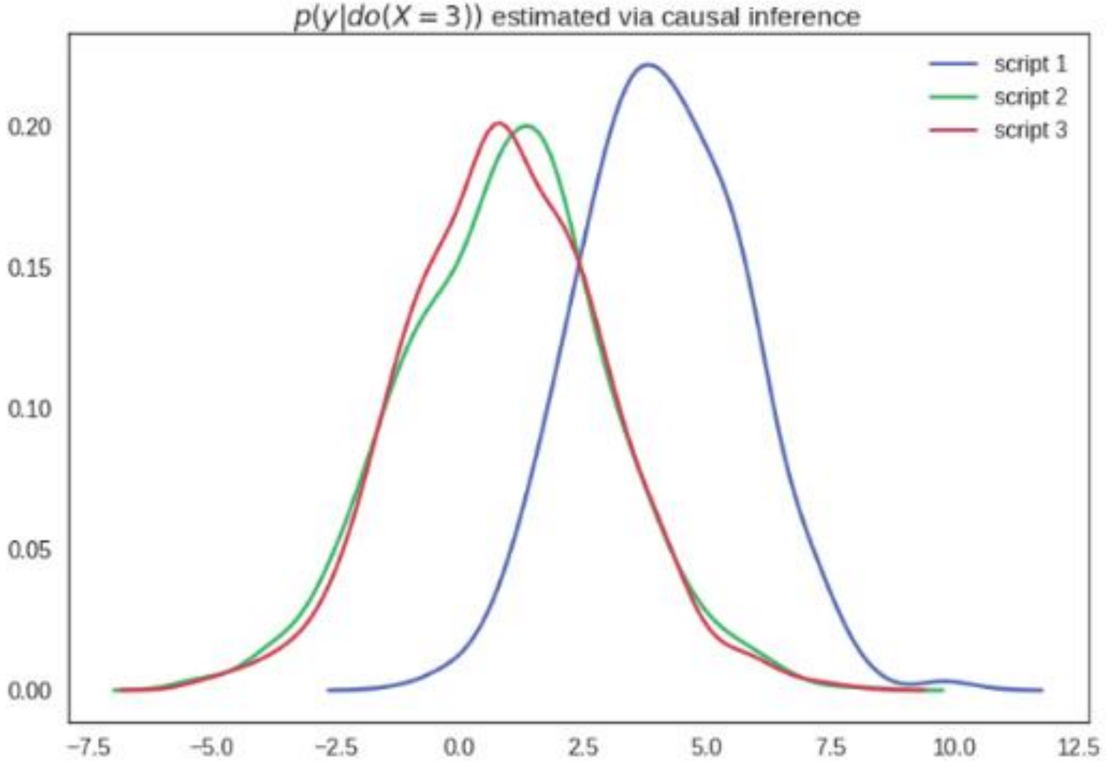
$$P(y|do(X)) = p(y|x)$$

$$P(y|do(X)) = p(y)$$

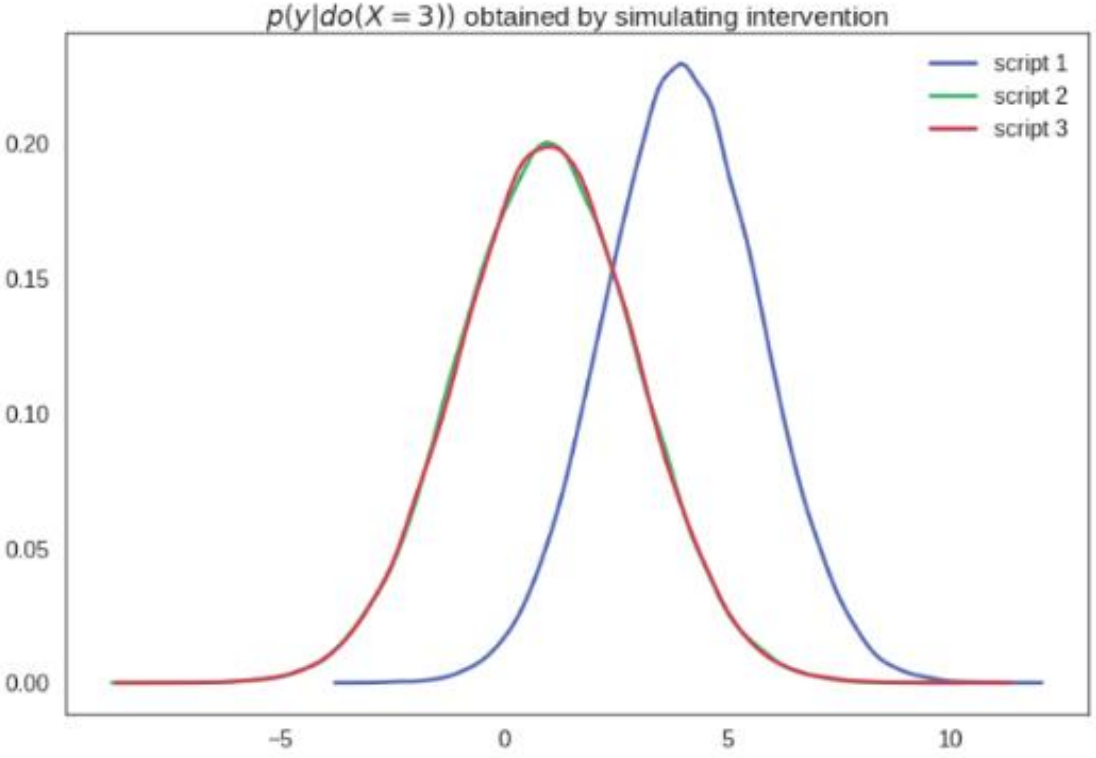
$$P(y|do(X)) = p(y)$$

The need of Causal Diagrams for simulating Interventions

Estimated using Joint Dist.
And Causal Diagram



Running the experiment



The need of Causal Diagrams for simulating Interventions

- The causal diagram allows us to predict how the models will behave under intervention, without carrying out the intervention.
- We can estimate the distribution of ***y*** observed during the intervention experiment using only samples from the script under the normal (non-intervention) situation.
- This is called ***Causal Inference*** *from observational data.*

[Introductory document to Causality]

[Causality brick]

Thank you!

Appendix

References and EoC materials

References

- 8 pillars of Causal wisdom: https://ftp.cs.ucla.edu/pub/stat_ser/r470.pdf
- Understanding Judea Pearl's do-calculus and interventions: https://www.inference.vc/untitled/#:~:text=Causal%20inference%20and%20do%2Dcalculus%20allows%20you%20to%20understand%20a,estimate%20that%20thing%20from%20data.
- Example of how we can use Judea Pearl's do-calculus to prove conditional independence: https://medium.com/@akelleh/introducing-the-do-sampler-for-causal-inference-a3296ea9e78d
- Paper to understand the math behind Judea Pearl's do-calculus: https://ftp.cs.ucla.edu/pub/stat_ser/r402.pdf
- Rubin's idea of Causality:
  - https://stats.stackexchange.com/questions/178159/is-there-an-intuitive-way-to-understand-the-rubin-causal-model-and-the-potential
  - https://en.wikipedia.org/wiki/Rubin_causal_model
- Rubin Causal Model video https://www.youtube.com/watch?v=LrmrH26EhSo
- Slides contrasting Pearl's vs Rubin's model http://leedokyun.com/obs.pdf

EoC materials

- Introductory document to Causality and the brick: https://eoc.mu-sigma.com/search/app/tree/Foundation/?highlightId=af13e4dc-d257-4af6-a991-08f6d77231f0&expanded=false
- **Causality brick:** https://eoc.mu-sigma.com/search/app/tree/Foundation/?highlightId=9c0fc1c5-8151-4441-a6db-72f720fd28a1&expanded=false

Assumptions of Causal Structure Learning

Causal sufficiency: Refers to the absence of latent variables. They can either be modeled explicitly as nodes in the structural equations, or they can manifest themselves as dependence between the noise terms ($\varepsilon 1, \ldots, \varepsilon p$), where the noise terms are assumed to be independent in the absence of latent confounding variables.

Causal faithfulness: The distribution of X generated from structure causal model equation is Markov with respect to the causal DAG, meaning that if A and B are d-separated by S in the causal DAG, then X_A and X_B are conditionally independent given X_S. The reverse implication is called **causal faithfulness**.

Acyclicity: A graph is said to be acyclic if there is no presence of feedback mechanism.