Causal spaces

A mathematical axiomatisation of causality

Junhyung Park, ETH Zürich

UAI Tutorial, 17th August 2026

Mathematical Axiomatisation

Mathematical Axiomatisation

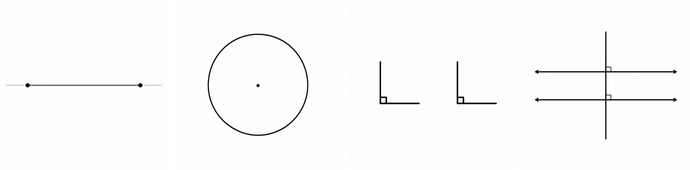
Defining properties that form the foundation of a mathematical theory.

Mathematical Axiomatisation

Defining properties that form the foundation of a mathematical theory.

Euclidean geometry:

- Two points define a straight line.
- A straight line extends infinitely on either side.
- A centre and a radius define a circle.
- All right-angles are equal.
- The parallel postulate.

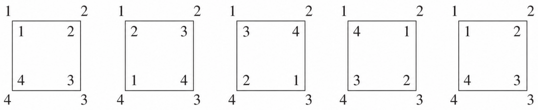


Mathematical Axiomatisation

Defining properties that form the foundation of a mathematical theory.

Group axioms for abstract algebra:

- Closure: $g_1 \circ g_2 \in \mathcal{G}$.
- Identity: $g \circ e = e \circ g = g$.
- Inverse: $g \circ g^{-1} = g^{-1} \circ g = e$.
- Associativity: $(g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)$.

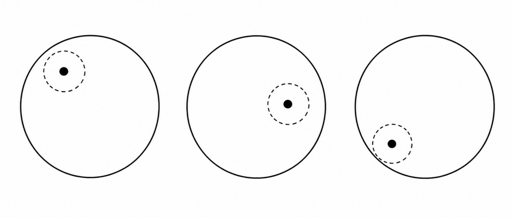


Mathematical Axiomatisation

Defining properties that form the foundation of a mathematical theory.

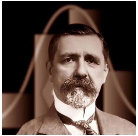
Open set axioms for topology:

- $\emptyset, \Omega$ are open.
- Arbitrary unions of open sets are open.
- Finite intersections of open sets are open.



Mathematical Axiomatisation

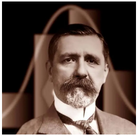
Probability theory

Grundbegriffe der Wahrscheinlichkeitsrechnung, Andrei Kolmogorov, Springer, 1933.

Mathematical Axiomatisation

Probability theory



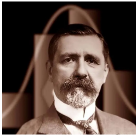
$$(\Omega\ , \mathscr{H}\ , \ \mathbb{P})$$

Set of outcomes $\quad$ σ-algebra of events $\quad$ Probability measure

Grundbegriffe der Wahrscheinlichkeitsrechnung, Andrei Kolmogorov, Springer, 1933.

Mathematical Axiomatisation

Probability theory



$$(\Omega\ ,\mathscr{H},\ \mathbb{P})$$

Set of outcomes

σ-algebra of events

Probability measure

- $\mathbb{P}(\Omega) = 1$;
- $\mathbb{P}(\cup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$.

Grundbegriffe der Wahrscheinlichkeitsrechnung, Andrei Kolmogorov, Springer, 1933.

Mathematical axiomatisation

Desiderata

Mathematical axiomatisation

Desiderata

1. Clearly isolates the essential structural properties of a concept.

Mathematical axiomatisation

Desiderata

1. Clearly isolates the essential structural properties of a concept.
2. Broad enough to include everything we want (minimality).

Mathematical axiomatisation

Desiderata

1. Clearly isolates the essential structural properties of a concept.
2. Broad enough to include everything we want (minimality).
3. Sharp enough to exclude anything we do not want.

Mathematical axiomatisation

Desiderata

1. Clearly isolates the essential structural properties of a concept.
2. Broad enough to include everything we want (minimality).
3. Sharp enough to exclude anything we do not want.
4. Consistent internal logic.

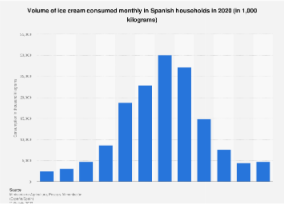
Mathematical axiomatisation

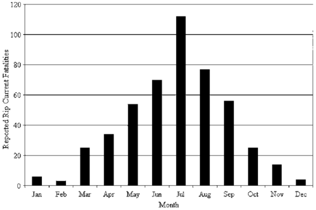
Desiderata

1. Clearly isolates the essential structural properties of a concept.
2. Broad enough to include everything we want (minimality).
3. Sharp enough to exclude anything we do not want.
4. Consistent internal logic.
5. Leads to a rich and fruitful theory.

Is probability theory all we need?

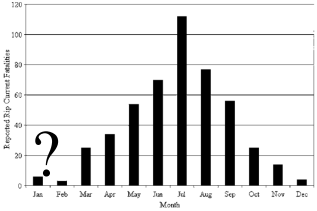
Ice cream sales vs beach accidents





Is probability theory all we need?

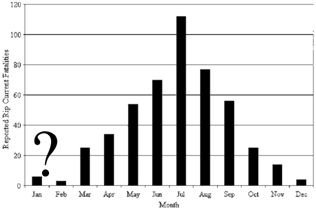
Ice cream sales vs beach accidents



- **Intervention** is at the heart of causality!

Is probability theory all we need?

Ice cream sales vs beach accidents



- **Intervention** is at the heart of causality!
- But the effect of interventions is not deterministic.

Is Probability Theory All We Need?

Is Moderate Drinking Good For Health?

Moderate drinking not better for health than abstaining, analysis suggests

Scientists say flaws in previous research mean health benefits from alcohol were exaggerated

For the regular boozer it is a source of great comfort: the fat pile of studies that say a daily tipple is better for a longer life than avoiding alcohol completely.

Is Probability Theory All We Need?

Is Moderate Drinking Good For Health?

Moderate drinking not better for health than abstaining, analysis suggests

Scientists say flaws in previous research mean health benefits from alcohol were exaggerated

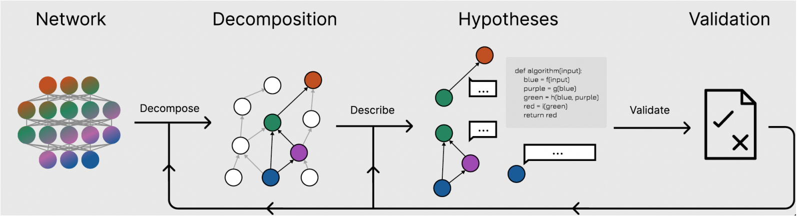
For the regular boozer it is a source of great comfort: the fat pile of studies that say a daily tipple is better for a longer life than avoiding alcohol completely.

But a new analysis challenges the thinking and blames the rosy message on flawed research that compares drinkers with people who are sick and sober.

Scientists in Canada delved into 107 published studies on people's drinking habits and how long they lived. In most cases, they found that drinkers were compared with people who abstained or consumed very little alcohol, without taking into account that some had cut down or quit through ill health.

Modern machine learning

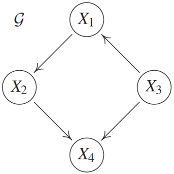
Mechanistic interpretability



Open Problems in Mechanistic Interpretability, Sharkey et al., TMLR, 2025.

Existing frameworks

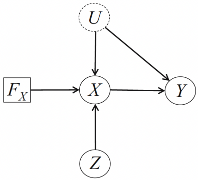
Structural causal models (SCMs)



Potential outcomes

	$Y^{a=0}$	$Y^{a=1}$
Rheia	0	1
Kronos	1	0
Demeter	0	0
Hades	0	0
Hestia	0	0
Poseidon	1	0
Hera	0	0
Zeus	0	1
Artemis	1	1
Apollo	1	0

Decision-theoretic



Causality, Pearl, Cambridge University Press, 2009
Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction, Rubin and Imbens, Cambridge University Press, 2015
Statistical Causality from a Decision-Theoretic Perspective, Dawid, Annual Review of Statistics and its Applications

In the words of Philip Dawid...

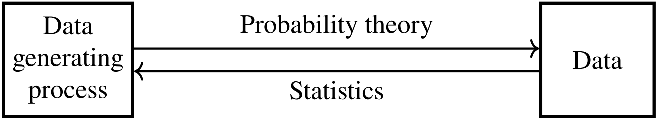
an impressive array of both theoretical and applied developments. As yet, however, there is no one fully accepted foundational basis for this enterprise. Rather, there is a variety of formal and informal frameworks for framing and understanding causal questions, as well as hot discussion about their relationships, merits, and demerits: We might mention, among others, structural equation modeling and path analysis (Wright 1921), potential response models (Rubin 1978), functional models (Pearl 2009), and various forms of graphical representation. This plethora of putative foundations leaves statistical causality in much the same state of confusion as probability theory before Kolmogorov.

This review aims to add to the confusion by describing one particular approach, based on decision-theoretic principles, that its author considers superior to the others (in this of course he is fully aligned with others' attitudes toward their own works). In it, I present the main features

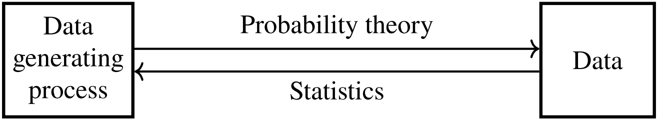
---

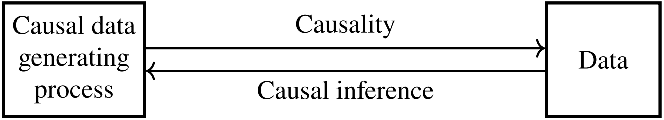
Statistical Causality from a Decision-Theoretic Perspective, Dawid, Annual Review of Statistics and its Applications

Probability vs Statistics

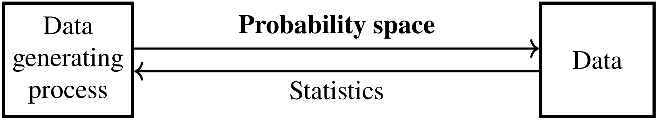


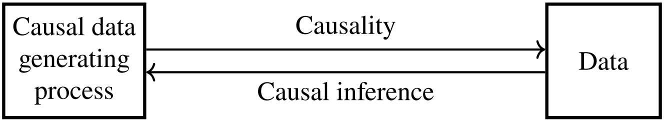
Probability vs Statistics



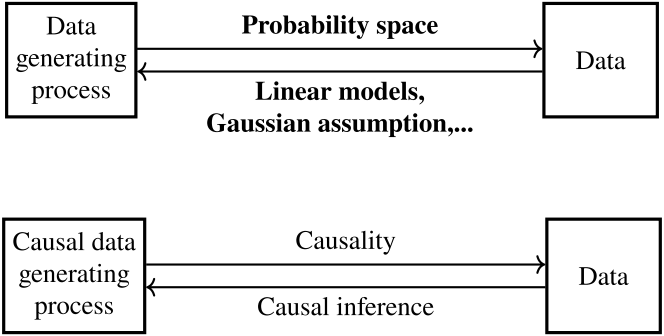


Probability vs Statistics

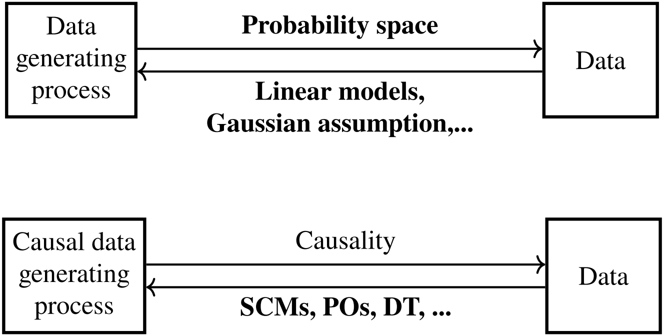




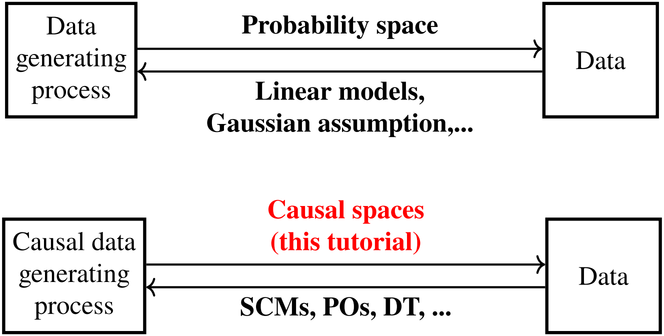
Probability vs Statistics



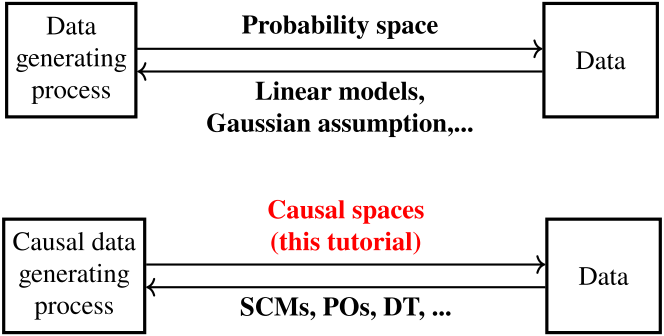
Probability vs Statistics



Probability vs Statistics



Probability vs Statistics



Causality is not yet viewed as a branch of pure mathematics!

Table of Contents

Table of Contents

Notation

- Product measurable space with index set $\mathbb{T}$:

$$(\Omega, \mathscr{H}) = (\times_{t\in\mathbb{T}}\Omega_t, \otimes_{t\in\mathbb{T}}\mathscr{E}_t).$$

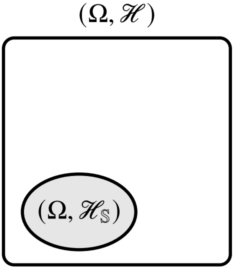
For each $\mathbb{S} \subseteq \mathbb{T}$, a sub-$\sigma$-algebra $\mathscr{H}_{\mathbb{S}}$ of $\mathscr{H}$.

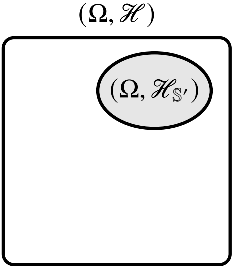
Notation

- Product measurable space with index set $\mathbb{T}$:

  $$(\Omega, \mathscr{H}) = (\times_{t \in \mathbb{T}} \Omega_t, \otimes_{t \in \mathbb{T}} \mathscr{E}_t).$$

  For each $\mathbb{S} \subseteq \mathbb{T}$, a sub-$\sigma$-algebra $\mathscr{H}_\mathbb{S}$ of $\mathscr{H}$.
  Intuition:
  - **$\mathscr{H} = \mathscr{H}_\mathbb{T}$ entire space / all variables.**
  - **$\mathscr{H}_\mathbb{S}$ subspace / subset of the variables.**





Notation

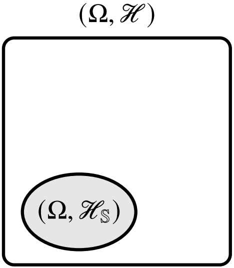
- Product measurable space with index set $\mathbb{T}$:

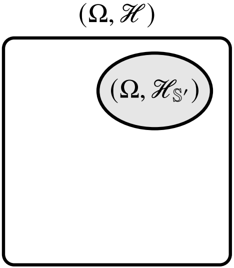
$$(\Omega, \mathscr{H}) = (\times_{t\in\mathbb{T}}\Omega_t, \otimes_{t\in\mathbb{T}}\mathscr{E}_t).$$

  For each $\mathbb{S} \subseteq \mathbb{T}$, a sub-$\sigma$-algebra $\mathscr{H}_{\mathbb{S}}$ of $\mathscr{H}$.
  Intuition:
  - $\mathscr{H} = \mathscr{H}_{\mathbb{T}}$ **entire space / all variables.**
  - $\mathscr{H}_{\mathbb{S}}$ **subspace / subset of the variables.**
- Transition probability kernel $K_{\mathbb{S}} : \Omega_{\mathbb{S}} \times \mathscr{H} \to [0, 1]$

$$K_{\mathbb{S}}(x, \cdot) \to [0, 1].$$

  For each $x \in \Omega_{\mathbb{S}}$, $K_{\mathbb{S}}(x, \cdot)$ is a measure on $(\Omega, \mathscr{H})$.



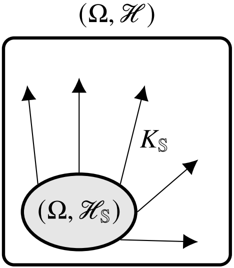


Notation

- Product measurable space with index set $\mathbb{T}$:

$$(\Omega, \mathscr{H}) = (\times_{t \in \mathbb{T}} \Omega_t, \otimes_{t \in \mathbb{T}} \mathscr{E}_t).$$

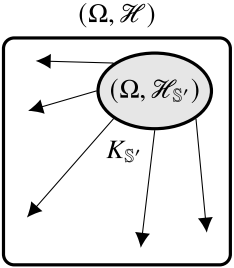
  For each $\mathbb{S} \subseteq \mathbb{T}$, a sub-$\sigma$-algebra $\mathscr{H}_{\mathbb{S}}$ of $\mathscr{H}$.
  Intuition:
  - $\mathscr{H} = \mathscr{H}_{\mathbb{T}}$ **entire space / all variables.**
  - $\mathscr{H}_{\mathbb{S}}$ **subspace / subset of the variables.**



- Transition probability kernel $K_{\mathbb{S}} : \Omega_{\mathbb{S}} \times \mathscr{H} \to [0, 1]$

$$K_{\mathbb{S}}(x, \cdot) \to [0, 1].$$

  For each $x \in \Omega_{\mathbb{S}}$, $K_{\mathbb{S}}(x, \cdot)$ is a measure on $(\Omega, \mathscr{H})$.
  Intuition: conditional distribution.



Causal spaces

A **causal space** is a quadruple $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$, where

Causal spaces

A **causal space** is a quadruple $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$, where

- $(\Omega, \mathscr{H}, \mathbb{P}) = (\times_{t\in\mathbb{T}}\Omega_t, \otimes_{t\in\mathbb{T}}\mathscr{E}_t, \mathbb{P})$ is a probability space;

Causal spaces

A **causal space** is a quadruple $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$, where

- $(\Omega, \mathscr{H}, \mathbb{P}) = (\times_{t\in\mathbb{T}}\Omega_t, \otimes_{t\in\mathbb{T}}\mathscr{E}_t, \mathbb{P})$ is a probability space;
- $\mathbb{K} = \{K_\mathbb{S} : \mathbb{S} \subseteq \mathbb{T}\}, \quad K_\mathbb{S} : \Omega_\mathbb{S} \times \mathscr{H} \to [0, 1]$,

where $K_\mathbb{S}$ are **causal kernels**

Causal spaces

A **causal space** is a quadruple $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$, where

- $(\Omega, \mathscr{H}, \mathbb{P}) = (\times_{t\in\mathbb{T}}\Omega_t, \otimes_{t\in\mathbb{T}}\mathscr{E}_t, \mathbb{P})$ is a probability space;
- $\mathbb{K} = \{K_\mathbb{S} : \mathbb{S} \subseteq \mathbb{T}\}, \;\; K_\mathbb{S} : \Omega_\mathbb{S} \times \mathscr{H} \to [0, 1]$,

where $K_\mathbb{S}$ are **causal kernels**, such that

(i) for all $A \in \mathscr{H}$ and $x \in \Omega$, $K_\emptyset(x; A) = \mathbb{P}(A)$;

(ii) for all $A \in \mathscr{H}_\mathbb{S}$ and $x \in \Omega$, $K_\mathbb{S}(x; A) = \mathbf{1}_A(x)$.

Causal spaces

A **causal space** is a quadruple $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$, where

- $(\Omega, \mathscr{H}, \mathbb{P}) = (\times_{t\in\mathbb{T}}\Omega_t, \otimes_{t\in\mathbb{T}}\mathscr{E}_t, \mathbb{P})$ is a probability space;
- $\mathbb{K} = \{K_\mathbb{S} : \mathbb{S} \subseteq \mathbb{T}\}, \quad K_\mathbb{S} : \Omega_\mathbb{S} \times \mathscr{H} \to [0, 1]$,

where $K_\mathbb{S}$ are **causal kernels**, such that

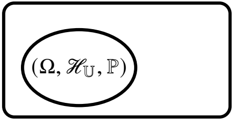
(i) for all $A \in \mathscr{H}$ and $x \in \Omega$, $K_\emptyset(x; A) = \mathbb{P}(A)$;

(ii) for all $A \in \mathscr{H}_\mathbb{S}$ and $x \in \Omega$, $K_\mathbb{S}(x; A) = \mathbf{1}_A(x)$.

$\mathbb{P}$ is the **observational** measure.

Interventions

$(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$

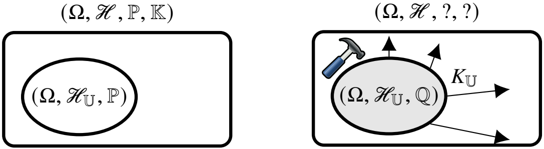


Interventions

An intervention is the process of

(a) choosing a sub-σ-algebra $\mathscr{H}_{\mathbb{U}}$, and

(b) placing any measure $\mathbb{Q}$ on $(\Omega, \mathscr{H}_{\mathbb{U}})$.



Interventions

An intervention is the process of

(a) choosing a sub-σ-algebra $\mathscr{H}_{\mathbb{U}}$, and

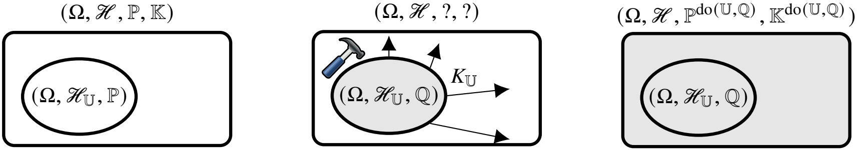
(b) placing any measure $\mathbb{Q}$ on $(\Omega, \mathscr{H}_{\mathbb{U}})$.

New causal space: $(\Omega, \mathscr{H}, \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}, \mathbb{K}^{\mathrm{do}(\mathbb{U},\mathbb{Q})})$, where

$$\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$$

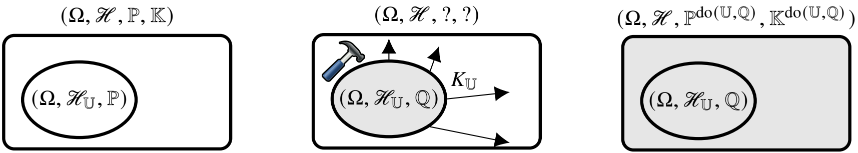
and $\mathbb{K}^{\mathrm{do}(\mathbb{U},\mathbb{Q})} = \{K_{\mathbb{S}}^{\mathrm{do}(\mathbb{U},\mathbb{Q})} : \mathbb{S} \subseteq \mathbb{T}\}$ with

$$K_{\mathbb{S}}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(x, A) = \int \mathbb{Q}(dx'_{\mathbb{U}\setminus\mathbb{S}}) K_{\mathbb{S}\cup\mathbb{U}}\Big((x_{\mathbb{S}}, x'_{\mathbb{U}\setminus\mathbb{S}}); A\Big).$$



Axioms of causal kernels

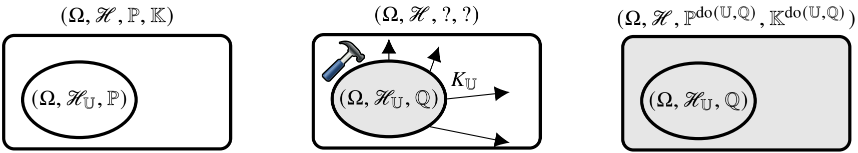
Recall that causal kernels $K_{\mathbb{U}}$ from $(\Omega, \mathscr{H}_{\mathbb{U}})$ into $(\Omega, \mathscr{H})$ satisfy



Axioms of causal kernels

Recall that causal kernels $K_{\mathbb{U}}$ from $(\Omega, \mathscr{H}_{\mathbb{U}})$ into $(\Omega, \mathscr{H})$ satisfy

(i) for all $A \in \mathscr{H}$ and $x \in \Omega$, $K_{\emptyset}(x; A) = \mathbb{P}(A)$;



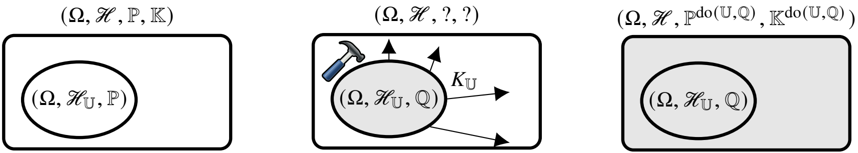
Axioms of causal kernels

Recall that causal kernels $K_{\mathbb{U}}$ from $(\Omega, \mathscr{H}_{\mathbb{U}})$ into $(\Omega, \mathscr{H})$ satisfy

(i) for all $A \in \mathscr{H}$ and $x \in \Omega$, $K_{\emptyset}(x; A) = \mathbb{P}(A)$;

Intuition on the axioms:

(i) $\mathbb{P}^{\mathrm{do}(\emptyset,\mathbb{Q})}(A) = \mathbb{P}(A)$. *(Trivial intervention)*



Axioms of causal kernels

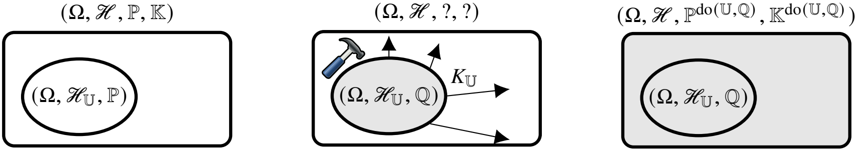
Recall that causal kernels $K_{\mathbb{U}}$ from $(\Omega, \mathscr{H}_{\mathbb{U}})$ into $(\Omega, \mathscr{H})$ satisfy

(i) for all $A \in \mathscr{H}$ and $x \in \Omega$, $K_{\emptyset}(x; A) = \mathbb{P}(A)$;

(ii) for all $A \in \mathscr{H}_{\mathbb{U}}$ and $x \in \Omega$, $K_{\mathbb{U}}(x; A) = \mathbf{1}_A(x)$.

Intuition on the axioms:

(i) $\mathbb{P}^{\mathrm{do}(\emptyset, \mathbb{Q})}(A) = \mathbb{P}(A)$. *(Trivial intervention)*



Axioms of causal kernels

Recall that causal kernels $K_{\mathbb{U}}$ from $(\Omega, \mathscr{H}_{\mathbb{U}})$ into $(\Omega, \mathscr{H})$ satisfy

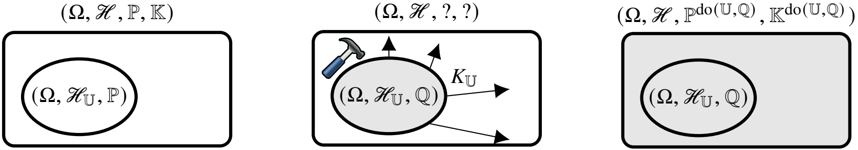
(i) for all $A \in \mathscr{H}$ and $x \in \Omega$, $K_{\emptyset}(x; A) = \mathbb{P}(A)$;

(ii) for all $A \in \mathscr{H}_{\mathbb{U}}$ and $x \in \Omega$, $K_{\mathbb{U}}(x; A) = \mathbf{1}_A(x)$.

Intuition on the axioms:

(i) $\mathbb{P}^{\mathrm{do}(\emptyset, \mathbb{Q})}(A) = \mathbb{P}(A)$. *(Trivial intervention)*

(ii) $\mathbb{P}^{\mathrm{do}(\mathbb{U}, \mathbb{Q})}(A) = \int \mathbb{Q}(dx)\mathbf{1}_A(x) = \mathbb{Q}(A)$. *(Interventional determinism)*



Theorem

The new causal space $(\Omega, \mathscr{H}, \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}, \mathbb{K}^{\mathrm{do}(\mathbb{U},\mathbb{Q})})$ obtained after intervention is indeed a causal space, i.e.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}$ is a probability measure;
- $\mathbb{K}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}$ satisfies the two axioms on the previous page.

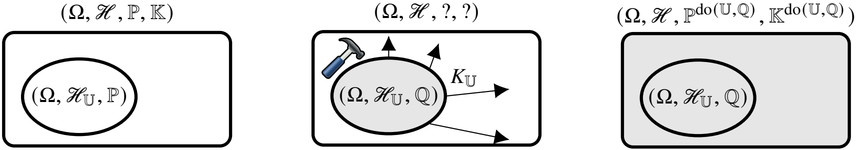
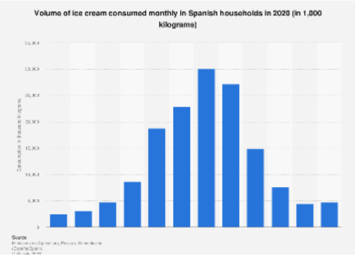
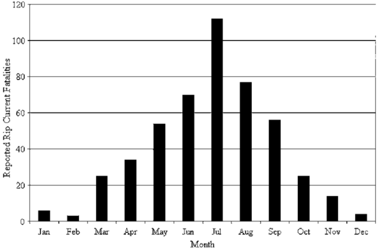


Table of Contents

Example

Ice cream sales and beach accidents

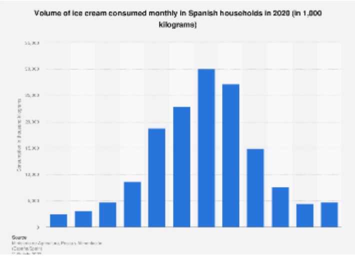


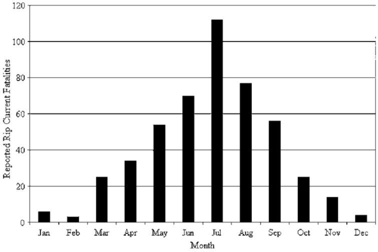


- Causal space: $(\Omega_{\text{ice}} \times \Omega_{\text{acc}}, \mathscr{E}_{\text{ice}} \otimes \mathscr{E}_{\text{acc}}, \mathbb{P}, \mathbb{K})$.

Example

Ice cream sales and beach accidents

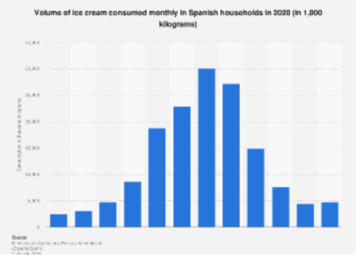


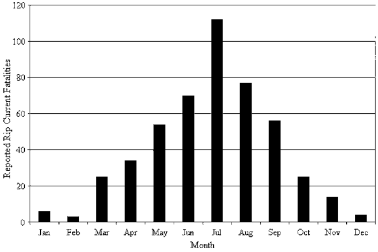


- Causal space: $(\Omega_{\text{ice}} \times \Omega_{\text{acc}}, \mathscr{E}_{\text{ice}} \otimes \mathscr{E}_{\text{acc}}, \mathbb{P}, \mathbb{K})$.
- $\mathbb{P}$ has strong dependence.

Example

Ice cream sales and beach accidents





- Causal space: $(\Omega_{\text{ice}} \times \Omega_{\text{acc}}, \mathscr{E}_{\text{ice}} \otimes \mathscr{E}_{\text{acc}}, \mathbb{P}, \mathbb{K})$.
- $\mathbb{P}$ has strong dependence.
- For causal kernels, let
  - $K_{\text{ice}}(x; A) = \mathbb{P}(A)$ for all $A \in \mathscr{E}_{\text{acc}}$; and
  - $K_{\text{acc}}(y; B) = \mathbb{P}(B)$ for all $B \in \mathscr{E}_{\text{ice}}$.

19 / 46

Example

Ice Cream Sales and Beach Accidents

A Number of beach accidents.

I Ice cream sales.

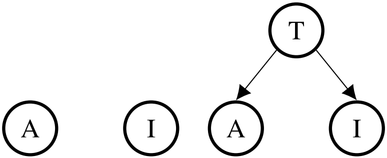
Example

Ice Cream Sales and Beach Accidents

A Number of beach accidents.

I Ice cream sales.

T Temperature.



Example

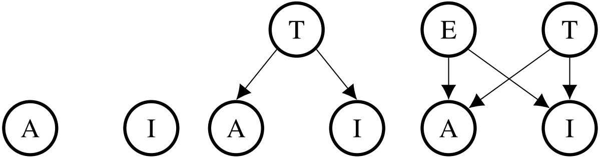
Ice Cream Sales and Beach Accidents

A Number of beach accidents.

I Ice cream sales.

T Temperature.

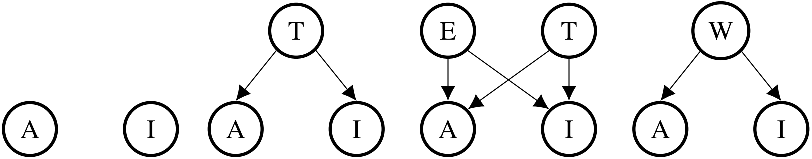
E Economy.



Example

Ice Cream Sales and Beach Accidents

- A Number of beach accidents.
- I Ice cream sales.
- T Temperature.
- E Economy.
- W World.



Example

Ice Cream Sales and Beach Accidents

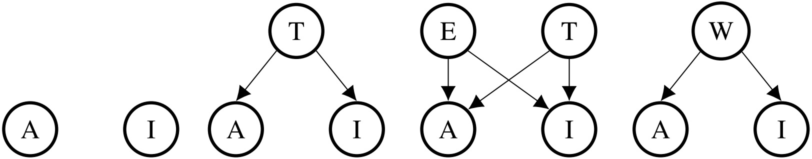
A Number of beach accidents.

I Ice cream sales.

T Temperature.

E Economy.

W World.



Recall:

- Causal space: $(\Omega_{\text{ice}} \times \Omega_{\text{acc}}, \mathscr{E}_{\text{ice}} \otimes \mathscr{E}_{\text{acc}}, \mathbb{P}, \mathbb{K})$.
- $\mathbb{P}$ has strong correlation.
- For causal kernels, let
  - $K_{\text{ice}}(x; A) = \mathbb{P}(A)$ for all $A \in \mathscr{E}_{\text{acc}}$; and
  - $K_{\text{acc}}\big(y; B\big) = \mathbb{P}(B)$ for all $B \in \mathscr{E}_{\text{ice}}$.

Example

Crop Yield and Price

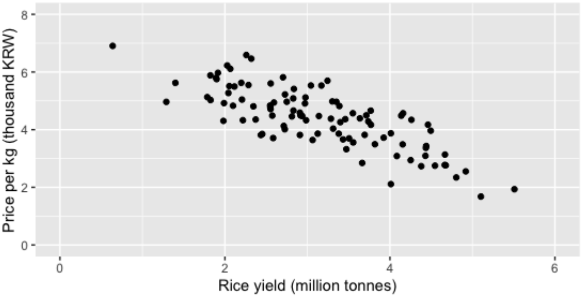
Causal space: $(\Omega_{\text{rice}} \times \Omega_{\text{price}}, \mathscr{E}_{\text{rice}} \otimes \mathscr{E}_{\text{price}}, \mathbb{P}, \mathbb{K})$.

Example

Crop Yield and Price

Causal space: $(\Omega_{\text{rice}} \times \Omega_{\text{price}}, \mathscr{E}_{\text{rice}} \otimes \mathscr{E}_{\text{price}}, \mathbb{P}, \mathbb{K})$.

No Intervention:



Example

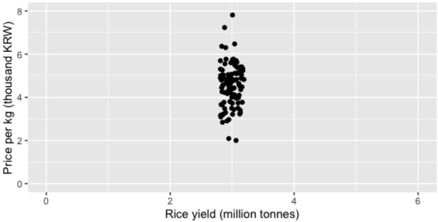
Crop Yield and Price

Causal space: $(\Omega_{\text{rice}} \times \Omega_{\text{price}}, \mathscr{E}_{\text{rice}} \otimes \mathscr{E}_{\text{price}}, \mathbb{P}, \mathbb{K})$.

Intervention to fix rice = 3:

$$K_{\text{rice}}(3; A) = \int_A \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-4.5)^2} dx$$

for $A \in \mathscr{E}_{\text{price}}$.



Example

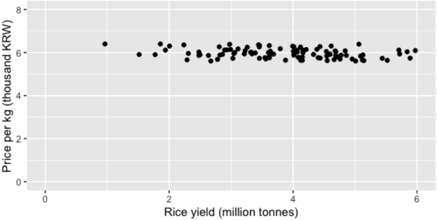
Crop Yield and Price

Causal space: $(\Omega_{\text{rice}} \times \Omega_{\text{price}}, \mathscr{E}_{\text{rice}} \otimes \mathscr{E}_{\text{price}}, \mathbb{P}, \mathbb{K})$.

Intervention to fix price = 6:

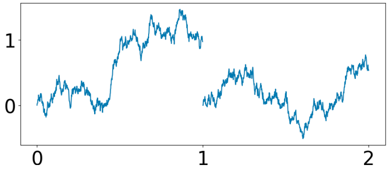
$$K_{\text{price}}(6; B) = \int_B \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-4)^2} dx$$

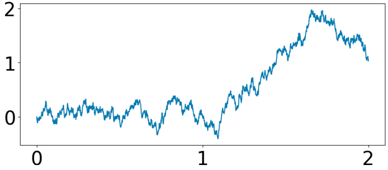
for $B \in \mathscr{E}_{\text{rice}}$.



Example

1-dimensional Brownian motion



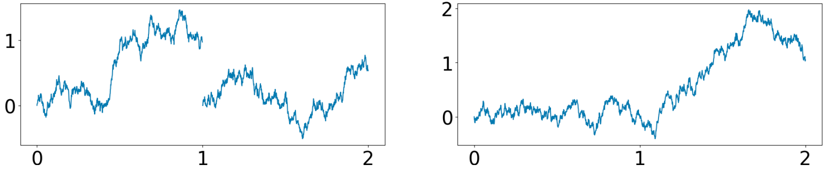


- Causal space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P}, \mathbb{K})$[1].

[1] For each $t \in \mathbb{R}_+$, $\Omega_t = \mathbb{R}$ and $\mathscr{E}_t$ is the Borel σ-algebra.

Example

1-dimensional Brownian motion

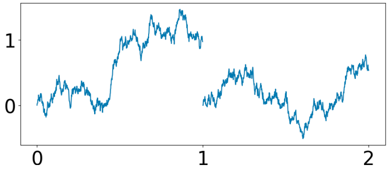


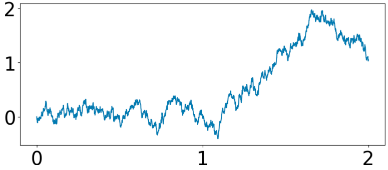
- Causal space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P}, \mathbb{K})$[1].
- $\mathbb{P}$ is the Wiener measure.

[1]For each $t \in \mathbb{R}_+$, $\Omega_t = \mathbb{R}$ and $\mathscr{E}_t$ is the Borel σ-algebra.

Example

1-dimensional Brownian motion





- Causal space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P}, \mathbb{K})$[1].
- $\mathbb{P}$ is the Wiener measure.
- For values $y \in \Omega_t$, the causal kernels are

$$K_{\{1\}}\big(0; dy\big) = \begin{cases} \mathbb{P}(dy), & 0 \le t < 1, \\ \delta_0(dy), & t = 1, \\ \frac{1}{\sqrt{2\pi(t-1)}} e^{-\frac{y^2}{2(t-1)}}\, dy, & t > 1. \end{cases}$$

[1] For each $t \in \mathbb{R}_+$, $\Omega_t = \mathbb{R}$ and $\mathscr{E}_t$ is the Borel σ-algebra.

Table of Contents

Mathematical axiomatisation

1. Clearly isolates the essential structural properties of a concept.
2. Broad enough to include everything we want (minimality).
3. Sharp enough to exclude anything we do not want.
4. Consistent internal logic.
5. **Leads to a rich and fruitful theory.**

Mathematical axiomatisation

Desiderata

1. Clearly isolates the essential structural properties of a concept.
2. Broad enough to include everything we want (minimality).
3. Sharp enough to exclude anything we do not want.
4. Consistent internal logic.
5. **Leads to a rich and fruitful theory.**
   - **Unify and generalise existing theory.**

Mathematical axiomatisation

Desiderata

1. Clearly isolates the essential structural properties of a concept.
2. Broad enough to include everything we want (minimality).
3. Sharp enough to exclude anything we do not want.
4. Consistent internal logic.
5. **Leads to a rich and fruitful theory.**
   - **Unify and generalise existing theory.**
   - Go beyond existing theories.

Recap: Independence

from probability theory

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B), \text{ or equivalently, } \mathbb{P}(A \mid B) = \mathbb{P}(A).$$

Recap: Independence

from probability theory

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B), \text{ or equivalently, } \mathbb{P}(A \mid B) = \mathbb{P}(A).$$

2.1 Independence

Measure theory ends and probability begins with the definition of independence. We begin with what is hopefully a familiar definition and

Probability: Theory and Examples, Durrett, Cambridge University Press, 2019

Recap: Independence

from probability theory

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B), \text{ or equivalently, } \mathbb{P}(A \mid B) = \mathbb{P}(A).$$

2.1 Independence

Measure theory ends and probability begins with the definition of independence. We begin with what is hopefully a familiar definition and

5 Independence

This section is about independence, a truly probabilistic concept. For

Probability: Theory and Examples, Durrett, Cambridge University Press, 2019
Probability and Stochastics, Çinlar, Springer, 2011

Causal effect

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.

Causal effect

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.

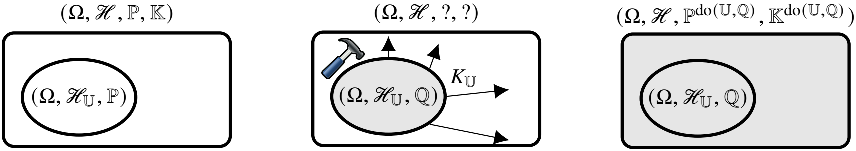
- $K_{\mathbb{U}}$ has *no causal effect* on A if $K_{\mathbb{U}}(x; A) = \mathbb{P}(A)$ for all $x \in \Omega$.

Causal effect

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.

- $K_{\mathbb{U}}$ has *no causal effect* on A if $K_{\mathbb{U}}(x; A) = \mathbb{P}(A)$ for all $x \in \Omega$.

Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.
Suppose $K_{\mathbb{U}}$ has no causal effect on A.



Causal effect

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.

- $K_{\mathbb{U}}$ has *no causal effect* on A if $K_{\mathbb{U}}(x; A) = \mathbb{P}(A)$ for all $x \in \Omega$.

Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.
Suppose $K_{\mathbb{U}}$ has no causal effect on A.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \mathbb{P}(A)$ for any $\mathbb{Q}$.

Causal effect

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.

- $K_{\mathbb{U}}$ has *no causal effect* on A if $K_{\mathbb{U}}(x; A) = \mathbb{P}(A)$ for all $x \in \Omega$.

Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.
Suppose $K_{\mathbb{U}}$ has no causal effect on A.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \mathbb{P}(A)$ for any $\mathbb{Q}$.

  Proof: $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A) = \int \mathbb{Q}(dx) \mathbb{P}(A) = \mathbb{P}(A)$.

Causal effect

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.

- $K_{\mathbb{U}}$ has *no causal effect* on A if $K_{\mathbb{U}}(x; A) = \mathbb{P}(A)$ for all $x \in \Omega$.

Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.
Suppose $K_{\mathbb{U}}$ has no causal effect on A.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \mathbb{P}(A)$ for any $\mathbb{Q}$.

  Proof: $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A) = \int \mathbb{Q}(dx) \mathbb{P}(A) = \mathbb{P}(A)$.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A \cap B) = \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(B)$ for all $B \in \mathscr{H}_{\mathbb{U}}$.

Causal effect

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.

- $K_{\mathbb{U}}$ has *no causal effect* on A if $K_{\mathbb{U}}(x; A) = \mathbb{P}(A)$ for all $x \in \Omega$.

Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.
Suppose $K_{\mathbb{U}}$ has no causal effect on A.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \mathbb{P}(A)$ for any $\mathbb{Q}$.

  Proof: $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A) = \int \mathbb{Q}(dx) \mathbb{P}(A) = \mathbb{P}(A)$.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A \cap B) = \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(B)$ for all $B \in \mathscr{H}_{\mathbb{U}}$.

  Proof:
  $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A \cap B) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A \cap B) =$

Causal effect

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.

- $K_{\mathbb{U}}$ has *no causal effect* on A if $K_{\mathbb{U}}(x; A) = \mathbb{P}(A)$ for all $x \in \Omega$.

Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.
Suppose $K_{\mathbb{U}}$ has no causal effect on A.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \mathbb{P}(A)$ for any $\mathbb{Q}$.

  Proof: $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A) = \int \mathbb{Q}(dx) \mathbb{P}(A) = \mathbb{P}(A)$.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A \cap B) = \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(B)$ for all $B \in \mathscr{H}_{\mathbb{U}}$.

  Proof: Using the second axiom $K_{\mathbb{U}}(x; B) = \mathbf{1}_B(x)$,
  $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A \cap B) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A \cap B) = \int \mathbb{Q}(dx) \mathbf{1}_B(x) K_{\mathbb{U}}(x; A)$

Causal effect

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.

- $K_{\mathbb{U}}$ has *no causal effect* on A if $K_{\mathbb{U}}(x; A) = \mathbb{P}(A)$ for all $x \in \Omega$.

Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.
Suppose $K_{\mathbb{U}}$ has no causal effect on A.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \mathbb{P}(A)$ for any $\mathbb{Q}$.

  Proof: $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A) = \int \mathbb{Q}(dx) \mathbb{P}(A) = \mathbb{P}(A)$.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A \cap B) = \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(B)$ for all $B \in \mathscr{H}_{\mathbb{U}}$.

  Proof: Using the second axiom $K_{\mathbb{U}}(x; B) = \mathbf{1}_B(x)$,
  $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A \cap B) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A \cap B) = \int \mathbb{Q}(dx) \mathbf{1}_B(x) K_{\mathbb{U}}(x; A)$
  $= \mathbb{P}(A) \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; B)$

Causal effect

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.

- $K_{\mathbb{U}}$ has *no causal effect* on A if $K_{\mathbb{U}}(x; A) = \mathbb{P}(A)$ for all $x \in \Omega$.

Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.
Suppose $K_{\mathbb{U}}$ has no causal effect on A.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \mathbb{P}(A)$ for any $\mathbb{Q}$.

  Proof: $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A) = \int \mathbb{Q}(dx) \mathbb{P}(A) = \mathbb{P}(A)$.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A \cap B) = \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(B)$ for all $B \in \mathscr{H}_{\mathbb{U}}$.

  Proof: Using the second axiom $K_{\mathbb{U}}(x; B) = \mathbf{1}_B(x)$,
  $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A \cap B) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A \cap B) = \int \mathbb{Q}(dx) \mathbf{1}_B(x) K_{\mathbb{U}}(x; A)$
  $= \mathbb{P}(A) \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; B) = \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(B)$.

Causal effect

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.

- $K_{\mathbb{U}}$ has *no causal effect* on A if $K_{\mathbb{U}}(x; A) = \mathbb{P}(A)$ for all $x \in \Omega$.

Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.
Suppose $K_{\mathbb{U}}$ has no causal effect on A.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \mathbb{P}(A)$ for any $\mathbb{Q}$.

  Proof: $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A) = \int \mathbb{Q}(dx)\mathbb{P}(A) = \mathbb{P}(A)$.

- $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A \cap B) = \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A)\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(B)$ for all $B \in \mathscr{H}_{\mathbb{U}}$.

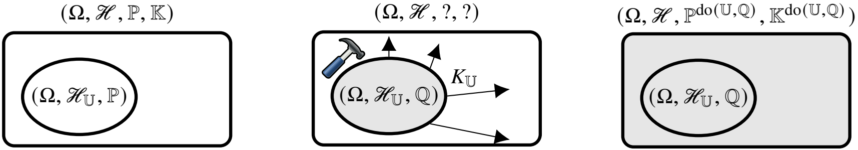
  Proof: Using the second axiom $K_{\mathbb{U}}(x; B) = \mathbf{1}_B(x)$,
  $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A \cap B) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A \cap B) = \int \mathbb{Q}(dx) \mathbf{1}_B(x) K_{\mathbb{U}}(x; A)$
  $= \mathbb{P}(A) \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; B) = \mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A)\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(B)$.

- The converse also holds!

Sources

Identifiability

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.
Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.

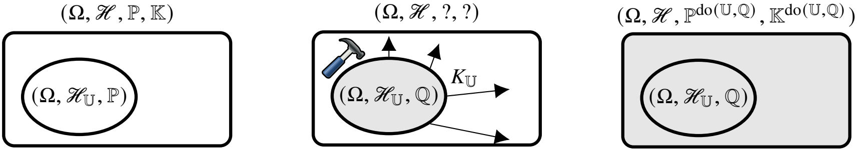


Sources

Identifiability

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.
Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.

- $K_{\mathbb{U}}$ is a *source* of A if $K_{\mathbb{U}}(\cdot; A) = \mathbb{P}(A \mid \mathscr{H}_{\mathbb{U}})$ almost surely.

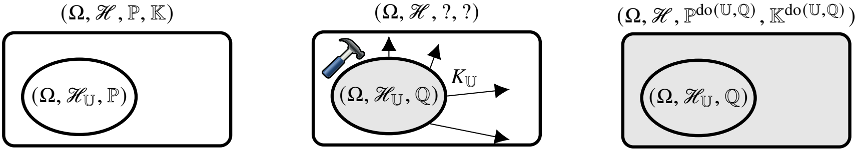


Sources

Identifiability

Recall: causal space $(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$ and causal kernel $K_{\mathbb{U}}(x; \cdot)$.
Recall: interventional measure $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx) K_{\mathbb{U}}(x; A)$.

- $K_{\mathbb{U}}$ is a *source* of A if $K_{\mathbb{U}}(\cdot; A) = \mathbb{P}(A \mid \mathscr{H}_{\mathbb{U}})$ almost surely.
- *Identifiability:* $\mathbb{P}^{\mathrm{do}(\mathbb{U},\mathbb{Q})}(A) = \int \mathbb{Q}(dx)\mathbb{P}(A \mid \mathscr{H}_{\mathbb{U}})$.



Two-out-of-three

Independence–causal effect–source

In summary:

Two-out-of-three

Independence–causal effect–source

In summary:

- **Independence:** $\mathbb{P}(A \mid \mathscr{H}_{\mathbb{U}}) = \mathbb{P}(A)$ (almost surely).

Two-out-of-three

Independence–causal effect–source

In summary:

- **Independence:** $\mathbb{P}(A \mid \mathscr{H}_{\mathbb{U}}) = \mathbb{P}(A)$ (almost surely).
- **No causal effect:** $K_{\mathbb{U}}(\cdot; A) = \mathbb{P}(A)$.

Two-out-of-three

Independence–causal effect–source

In summary:

- **Independence:** $\mathbb{P}(A \mid \mathscr{H}_{\mathbb{U}}) = \mathbb{P}(A)$ (almost surely).
- **No causal effect:** $K_{\mathbb{U}}(\cdot; A) = \mathbb{P}(A)$.
- **Source:** $K_{\mathbb{U}}(\cdot; A) = \mathbb{P}(A \mid \mathscr{H}_{\mathbb{U}})$ (almost surely).

Two-out-of-three

Independence–causal effect–source

In summary:

- **Independence:** $\mathbb{P}(A \mid \mathscr{H}_{\mathbb{U}}) = \mathbb{P}(A)$ (almost surely).
- **No causal effect:** $K_{\mathbb{U}}(\cdot; A) = \mathbb{P}(A)$.
- **Source:** $K_{\mathbb{U}}(\cdot; A) = \mathbb{P}(A \mid \mathscr{H}_{\mathbb{U}})$ (almost surely).

If two of the three holds, then the third also holds!

Do-calculus

Graphical criteria for identification

Causality, Pearl, Cambridge University Press, 2009

Do-calculus

Graphical criteria for identification

Rule 1 (*Insertion/deletion of observations*):

$$P(y \mid \hat{x}, z, w) = P(y \mid \hat{x}, w) \quad \text{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}}}.$$

Causality, Pearl, Cambridge University Press, 2009

Do-calculus

Graphical criteria for identification

Rule 1 (*Insertion/deletion of observations*):

$$P(y \mid \hat{x}, z, w) = P(y \mid \hat{x}, w) \quad \mathit{if}\ (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}}}.$$

Rule 2 (*Action/observation exchange*):

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, z, w) \quad \mathit{if}\ (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}\underline{Z}}}.$$

Causality, Pearl, Cambridge University Press, 2009

Do-calculus

Graphical criteria for identification

Rule 1 (*Insertion/deletion of observations*):

$$P(y \mid \hat{x}, z, w) = P(y \mid \hat{x}, w) \quad \text{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}}}.$$

Rule 2 (*Action/observation exchange*):

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, z, w) \quad \text{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}\underline{Z}}}.$$

Rule 3 (*Insertion/deletion of actions*):

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, w) \text{ if } (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}, \overline{Z(W)}}},$$

where Z(W) is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

Causality, Pearl, Cambridge University Press, 2009

Do-calculus

Graphical criteria for identification

Rule 1 (*Insertion/deletion of observations*):

$$P(y \mid \hat{x}, z, w) = P(y \mid \hat{x}, w) \quad \text{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}}}.$$

Independence

Rule 2 (*Action/observation exchange*):

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, z, w) \quad \text{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}\underline{Z}}}.$$

Rule 3 (*Insertion/deletion of actions*):

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, w) \text{ if } (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}, \overline{Z(W)}}},$$

where $Z(W)$ is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

Causality, Pearl, Cambridge University Press, 2009

Do-calculus

Graphical criteria for identification

Rule 1 (*Insertion/deletion of observations*):

$$P(y \mid \hat{x}, z, w) = P(y \mid \hat{x}, w) \quad \text{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}}}.$$

Independence

Rule 2 (*Action/observation exchange*):

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, z, w) \quad \text{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}\underline{Z}}}.$$

Sources

Rule 3 (*Insertion/deletion of actions*):

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, w) \text{ if } (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}, \overline{Z(W)}}},$$

where Z(W) is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

Causality, Pearl, Cambridge University Press, 2009

Do-calculus

Graphical criteria for identification

Rule 1 (*Insertion/deletion of observations*):

$$P(y \mid \hat{x}, z, w) = P(y \mid \hat{x}, w) \quad \text{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}}}.$$

Independence

Rule 2 (*Action/observation exchange*):

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, z, w) \quad \text{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}\underline{Z}}}.$$

Sources

Rule 3 (*Insertion/deletion of actions*):

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, w) \text{ if } (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}, \overline{Z(W)}}},$$

where $Z(W)$ is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

Causal effects

Causality, Pearl, Cambridge University Press, 2009

Do-calculus

Graphical criteria for identification

Rule 1 (*Insertion/deletion of observations*):

$$P(y \mid \hat{x}, z, w) = P(y \mid \hat{x}, w) \quad \text{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}}}.$$

Independence

Rule 2 (*Action/observation exchange*):

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, z, w) \quad \text{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}\underline{Z}}}.$$

Sources

Rule 3 (*Insertion/deletion of actions*):

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, w) \text{ if } (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}, \overline{Z(W)}}},$$

Causal effects

where $Z(W)$ is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

- Causal spaces: definitions.

Causality, Pearl, Cambridge University Press, 2009

Do-calculus

Graphical criteria for identification

Rule 1 (*Insertion/deletion of observations*): Independence

$$P(y \mid \hat{x}, z, w) = P(y \mid \hat{x}, w) \quad \textit{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}}}.$$

Rule 2 (*Action/observation exchange*): Sources

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, z, w) \quad \textit{if } (Y \perp\!\!\!\perp Z) \mid X, W)_{G_{\overline{X}\underline{Z}}}.$$

Rule 3 (*Insertion/deletion of actions*): Causal effects

$$P(y \mid \hat{x}, \hat{z}, w) = P(y \mid \hat{x}, w) \textit{ if } (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}, \overline{Z(W)}}},$$

where Z(W) is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

- Causal spaces: definitions.
- SCM do-calculus: criteria under modelling assumptions.

Causality, Pearl, Cambridge University Press, 2009

Table of Contents

Mathematical axiomatisation

Desiderata

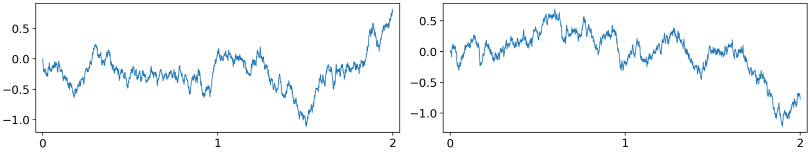
1. Clearly isolates the essential structural properties of a concept.
2. Broad enough to include everything we want (minimality).
3. Sharp enough to exclude anything we do not want.
4. Consistent internal logic.
5. **Leads to a rich and fruitful theory.**
   - Unify and generalise existing theory.
   - **Go beyond existing theories.**

Table of contents

Table of contents

Example: 1-dimensional Brownian motion

Observational

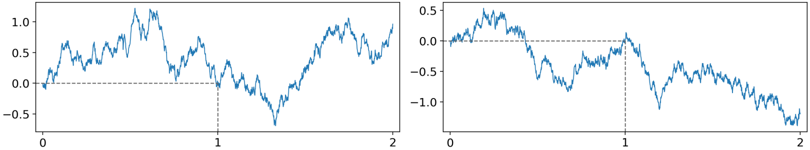


- Probability space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P})$[2],
- $\mathbb{P}$ is the Wiener measure.

[2]For each $t \in \mathbb{R}_+$, $\Omega_t = \mathbb{R}$ and $\mathscr{E}_t$ is the Borel σ-algebra.

Example: 1-dimensional Brownian motion

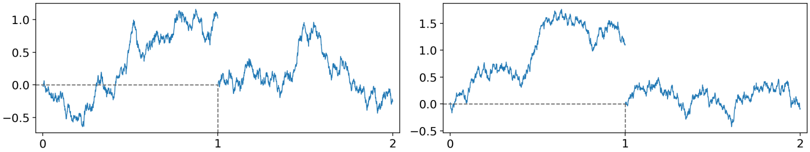
Conditioned



- Probability space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P})$,
- $\mathbb{P}$ is the Brownian bridge measure.

Example: 1-dimensional Brownian motion

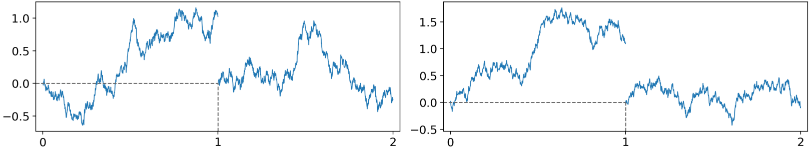
Intervention at a fixed time point



- Causal space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P}, \mathbb{K})$.
- $\mathbb{P}$ is the Wiener measure.

Example: 1-dimensional Brownian motion

Intervention at a fixed time point

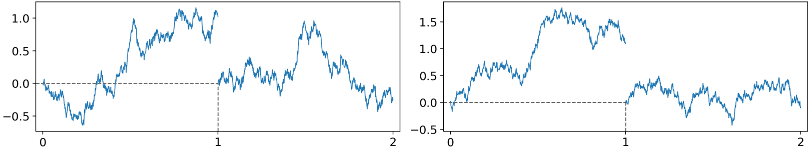


- Causal space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P}, \mathbb{K})$.
- $\mathbb{P}$ is the Wiener measure.
- For values y of $Y \in \Omega_t$, the causal kernel is

$$K_{\{1\}}((\eta,0);dy) = \begin{cases} \delta_{\eta_t}(dy), & 0 \leq t < 1, \\ \delta_0(dy), & t = 1, \\ \frac{1}{\sqrt{2\pi(t-1)}} e^{-\frac{y^2}{2(t-1)}}\, dy, & t > 1. \end{cases}$$

Example: 1-dimensional Brownian motion

Intervention at a fixed time point



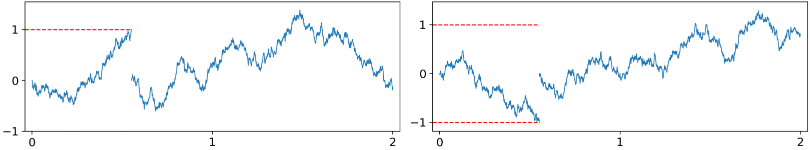
- Causal space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P}, \mathbb{K})$.
- $\mathbb{P}$ is the Wiener measure.
- For values y of $Y \in \Omega_t$, the causal kernel is

$$K_{\{1\}}((\eta,0);dy) = \begin{cases} \delta_{\eta_t}(dy), & 0 \le t < 1, \\ \delta_0(dy), & t = 1, \\ \frac{1}{\sqrt{2\pi(t-1)}} e^{-\frac{y^2}{2(t-1)}}\, dy, & t > 1. \end{cases}$$

- **Future values are affected by the intervention; past values are not.**

Example: 1-dimensional Brownian motion

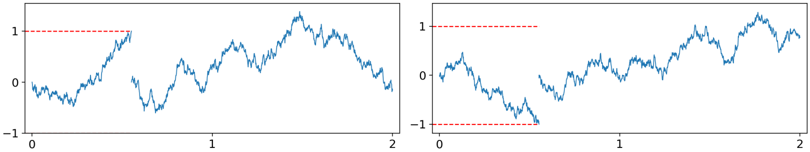
Targeted intervention at a stopping time



- Causal space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P}, \mathbb{K})$.
- $\mathbb{P}$ is the Wiener measure.

Example: 1-dimensional Brownian motion

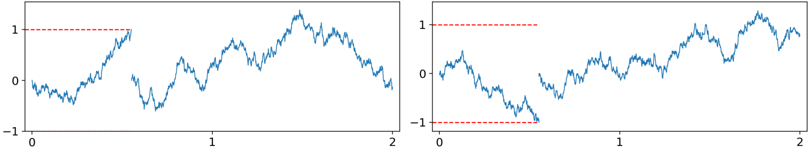
Targeted intervention at a stopping time



- Causal space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P}, \mathbb{K})$.
- $\mathbb{P}$ is the Wiener measure.
- Define the stopping time $\tau(\eta) = \inf\{s \geq 0 : |\eta_s| = 1\}$.

Example: 1-dimensional Brownian motion

Targeted intervention at a stopping time

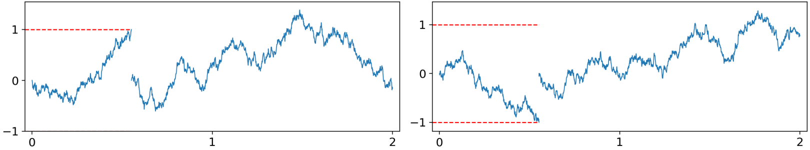


- Causal space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P}, \mathbb{K})$.
- $\mathbb{P}$ is the Wiener measure.
- Define the stopping time $\tau(\eta) = \inf\{s \geq 0 : |\eta_s| = 1\}$.
- Targeted causal kernel satisfies

$$K_{\{\tau(\eta)\}}((\eta,0),dy) = \begin{cases} \delta_{\eta_t}(dy), & 0 \leq t < \tau(\eta), \\ \delta_0(dy), & t = \tau(\eta), \\ \frac{1}{\sqrt{2\pi(t-\tau(\eta))}} e^{-\frac{y^2}{2(t-\tau(\eta))}}\, dy, & t > \tau(\eta). \end{cases}$$

Example: 1-dimensional Brownian motion

Targeted intervention at a stopping time



- Causal space: $(\times_{t\in\mathbb{R}_+}\Omega_t, \otimes_{t\in\mathbb{R}_+}\mathscr{E}_t, \mathbb{P}, \mathbb{K})$.
- $\mathbb{P}$ is the Wiener measure.
- Define the stopping time $\tau(\eta) = \inf\{s \geq 0 : |\eta_s| = 1\}$.
- Targeted causal kernel satisfies

$$K_{\{\tau(\eta)\}}((\eta, 0), dy) = \begin{cases} \delta_{\eta_t}(dy), & 0 \leq t < \tau(\eta), \\ \delta_0(dy), & t = \tau(\eta), \\ \frac{1}{\sqrt{2\pi(t-\tau(\eta))}} e^{-\frac{y^2}{2(t-\tau(\eta))}}\, dy, & t > \tau(\eta). \end{cases}$$

- **The time of intervention is path-dependent.**

Table of Contents

What is counterfactual thinking?

counterfactual *adjective*

coun·ter·fac·tu·al ˌkau̇n-tər-ˈfak-chə-wəl -chəl, -shwəl, -chü-əl

Synonyms of *counterfactual* >

: contrary to fact

counterfactual assumptions

counterfactual

[koun-ter-**fak**-choo-*uhl*] Phonetic (Standard) IPA

noun

Logic.

a conditional statement the first clause of which expresses something contrary to fact, as “If I had known.”

*Meaning of **counterfactual** in English*

counterfactual

adjective • formal

UK /ˌkaʊn.təˈfæk.tʃu.əl/ **US** /ˌkaʊn.t̬ɚˈfæk.tʃu.əl/

Add to word list

thinking about what did not happen but could have happened, or relating to this kind of thinking:

- *Thoughts about how an embarrassing event might have turned out differently are known to psychologists as counterfactual thinking.*

Definition of 'counterfactual'

counterfactual

in British English

(ˌkauntəˈfæktʃʊəl) *logic*

adjective

1. expressing what has not happened but could, would, or might under differing conditions

noun

2. a conditional statement in which the first clause is a past tense subjunctive statement expressing something contrary to fact, as in
if she had hurried she would have caught the bus

What is counterfactual thinking?

Imagine parallel worlds.

What is counterfactual thinking?

Imagine parallel worlds.

How are they related?

Counterfactuals in psychology

- Explain the past and prepare for the future.

Counterfactual Thought, Byrne, Annual Review of Psychology, 2016

Counterfactuals in psychology

- Explain the past and prepare for the future.
- Intentions and decisions.

Counterfactual Thought, Byrne, Annual Review of Psychology, 2016

Counterfactuals in psychology

- Explain the past and prepare for the future.
- Intentions and decisions.
- Modulate emotions such as regret and relief.

Counterfactual Thought, Byrne, Annual Review of Psychology, 2016

Counterfactuals in psychology

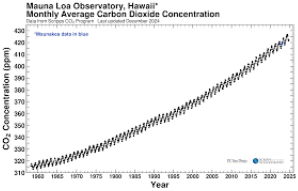
- Explain the past and prepare for the future.
- Intentions and decisions.
- Modulate emotions such as regret and relief.
- Support moral judgments such as blame.

Counterfactual Thought, Byrne, Annual Review of Psychology, 2016

Counterfactuals in other fields

Counterfactuals in other fields

Counterfactuals in other fields



Counterfactual spaces

Ladder of causation?

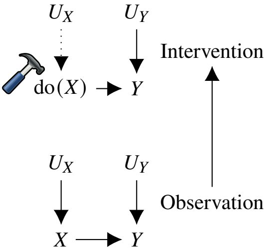
$$\begin{tikzcd} U_X \arrow[d] & U_Y \arrow[d] \\ X \arrow[r] & Y \end{tikzcd}$$

Observation

The Book of Why: The New Science of Cause and Effect, Pearl and Mackenzie, Basic Books, 2018.

Counterfactual spaces

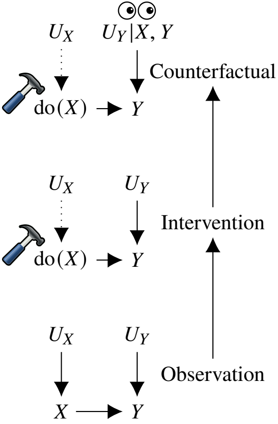
Ladder of causation?



The Book of Why: The New Science of Cause and Effect, Pearl and Mackenzie, Basic Books, 2018.

Counterfactual spaces

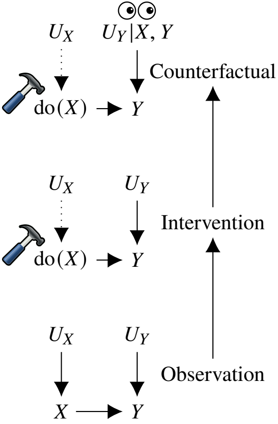
Ladder of causation?



The Book of Why: The New Science of Cause and Effect, Pearl and Mackenzie, Basic Books, 2018.

Counterfactual spaces

Ladder of causation?

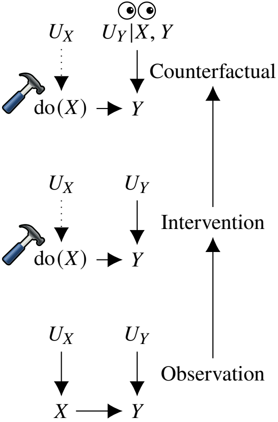


Probability spaces
$(\Omega, \mathscr{H}, \mathbb{P})$

The Book of Why: The New Science of Cause and Effect, Pearl and Mackenzie, Basic Books, 2018.

Counterfactual spaces

Ladder of causation?

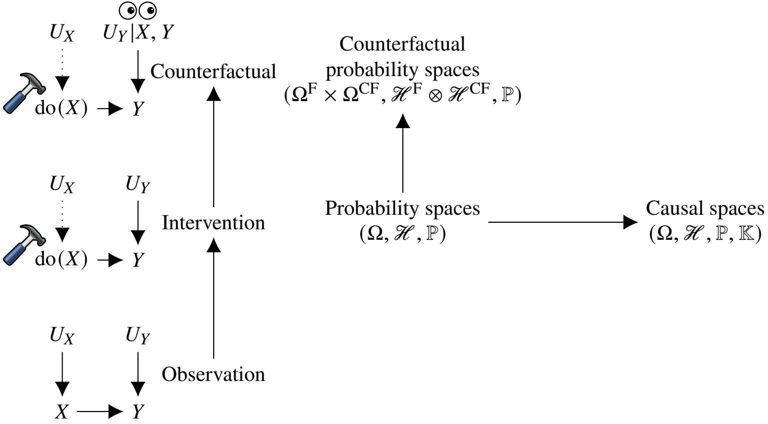


$$\begin{array}{c}\text{Probability spaces}\\(\Omega, \mathscr{H}, \mathbb{P})\end{array} \longrightarrow \begin{array}{c}\text{Causal spaces}\\(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})\end{array}$$

The Book of Why: The New Science of Cause and Effect, Pearl and Mackenzie, Basic Books, 2018.

Counterfactual spaces

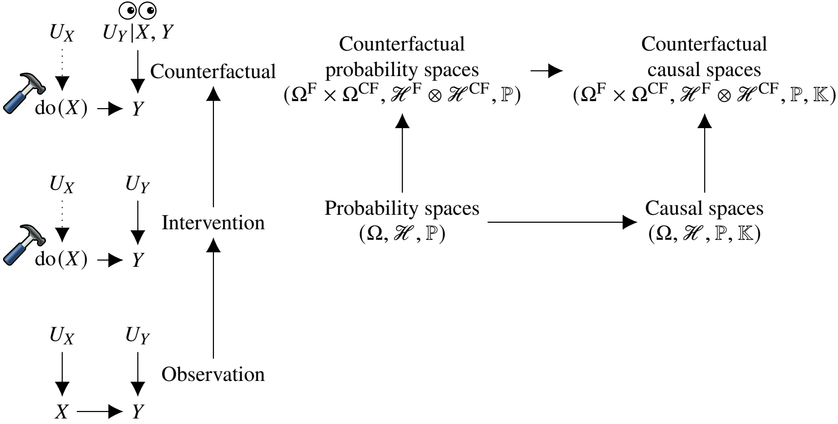
Ladder of causation?



The Book of Why: The New Science of Cause and Effect, Pearl and Mackenzie, Basic Books, 2018.

Counterfactual spaces

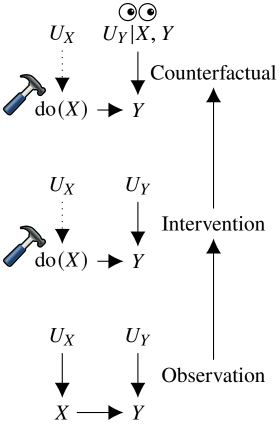
Ladder of causation?



The Book of Why: The New Science of Cause and Effect, Pearl and Mackenzie, Basic Books, 2018.

Counterfactual spaces

Ladder of causation?



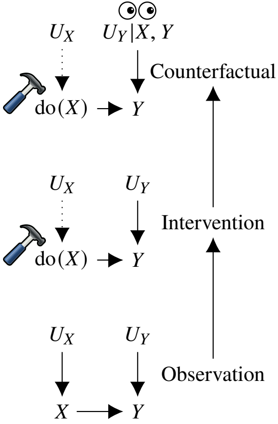
$$\begin{tikzcd}
\begin{array}{c}\text{Counterfactual}\\ \text{probability spaces}\\ (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P})\end{array} \arrow[r] & \begin{array}{c}\text{Counterfactual}\\ \text{causal spaces}\\ (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P}, \mathbb{K})\end{array} \\
\begin{array}{c}\text{Probability spaces}\\ (\Omega, \mathscr{H}, \mathbb{P})\end{array} \arrow[u] \arrow[r] & \begin{array}{c}\text{Causal spaces}\\ (\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})\end{array} \arrow[u]
\end{tikzcd}$$

Interventions and counterfactuals are orthogonal concepts.

The Book of Why: The New Science of Cause and Effect, Pearl and Mackenzie, Basic Books, 2018.

Counterfactual spaces

Ladder of causation?



Counterfactual probability spaces
$(\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P})$

Counterfactual causal spaces
$(\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P}, \mathbb{K})$

Probability spaces
$(\Omega, \mathscr{H}, \mathbb{P})$

Causal spaces
$(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})$

Interventions and counterfactuals are orthogonal concepts.

Shared information between worlds is encoded in $\mathbb{P}$ and $\mathbb{K}$.

The Book of Why: The New Science of Cause and Effect, Pearl and Mackenzie, Basic Books, 2018.

Table of Contents

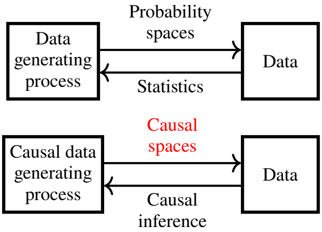
Summary

$$\begin{tikzcd}[column sep=huge]
\boxed{\begin{array}{c}\text{Data}\\\text{generating}\\\text{process}\end{array}} \arrow[r, shift left, "\begin{array}{c}\text{Probability}\\\text{spaces}\end{array}"] & \boxed{\text{Data}} \arrow[l, shift left, "\text{Statistics}"]
\end{tikzcd}$$

$$\begin{tikzcd}[column sep=huge]
\boxed{\begin{array}{c}\text{Causal data}\\\text{generating}\\\text{process}\end{array}} \arrow[r, shift left, "\color{red}\begin{array}{c}\text{Causal}\\\text{spaces}\end{array}"] & \boxed{\text{Data}} \arrow[l, shift left, "\begin{array}{c}\text{Causal}\\\text{inference}\end{array}"]
\end{tikzcd}$$

$$\begin{tikzcd}
\begin{array}{c}\text{Counterfactual}\\\text{probability spaces}\\(\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P})\end{array} \arrow[r] & \begin{array}{c}\text{Counterfactual}\\\text{causal spaces}\\(\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P}, \mathbb{K})\end{array} \\
\begin{array}{c}\text{Probability}\\\text{spaces}\\(\Omega, \mathscr{H}, \mathbb{P})\end{array} \arrow[u] \arrow[r] & \begin{array}{c}\text{Causal}\\\text{spaces}\\(\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})\end{array} \arrow[u]
\end{tikzcd}$$

Summary

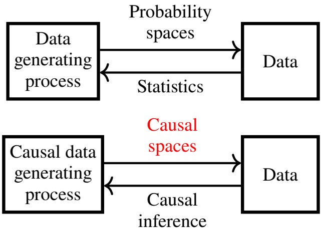


$$
\begin{tikzcd}
\begin{array}{c}\text{Counterfactual}\\ \text{probability spaces}\\ (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P})\end{array} \arrow[r] & \begin{array}{c}\text{Counterfactual}\\ \text{causal spaces}\\ (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P}, \mathbb{K})\end{array} \\
\begin{array}{c}\text{Probability}\\ \text{spaces}\\ (\Omega, \mathscr{H}, \mathbb{P})\end{array} \arrow[u] \arrow[r] & \begin{array}{c}\text{Causal}\\ \text{spaces}\\ (\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})\end{array} \arrow[u]
\end{tikzcd}
$$

Causal spaces

- are **a measure-theoretic axiomatisation** of causality;

Summary

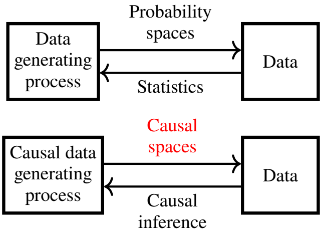


$$\begin{tikzcd}
\text{Counterfactual probability spaces } (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P}) \arrow[r] & \text{Counterfactual causal spaces } (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P}, \mathbb{K}) \\
\text{Probability spaces } (\Omega, \mathscr{H}, \mathbb{P}) \arrow[u] \arrow[r] & \text{Causal spaces } (\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K}) \arrow[u]
\end{tikzcd}$$

Causal spaces

- are **a measure-theoretic axiomatisation** of causality;
- whose axioms say

Summary

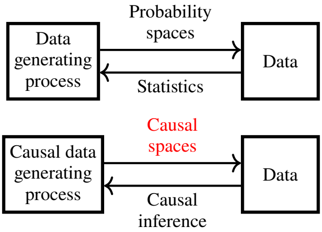


$$\begin{tikzcd}
\substack{\text{Counterfactual} \\ \text{probability spaces} \\ (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P})} \arrow[r] & \substack{\text{Counterfactual} \\ \text{causal spaces} \\ (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P}, \mathbb{K})} \\
\substack{\text{Probability} \\ \text{spaces} \\ (\Omega, \mathscr{H}, \mathbb{P})} \arrow[u] \arrow[r] & \substack{\text{Causal} \\ \text{spaces} \\ (\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})} \arrow[u]
\end{tikzcd}$$

Causal spaces

- are **a measure-theoretic axiomatisation** of causality;
- whose axioms say
  - doing nothing changes nothing (trivial intervention); and

Summary

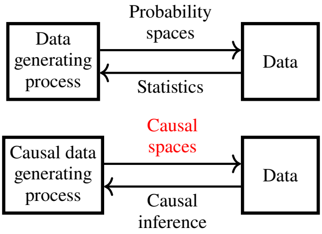


$$
\begin{tikzcd}
\begin{array}{c}\text{Counterfactual} \\ \text{probability spaces} \\ (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P})\end{array} \arrow[r] & \begin{array}{c}\text{Counterfactual} \\ \text{causal spaces} \\ (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P}, \mathbb{K})\end{array} \\
\begin{array}{c}\text{Probability} \\ \text{spaces} \\ (\Omega, \mathscr{H}, \mathbb{P})\end{array} \arrow[u] \arrow[r] & \begin{array}{c}\text{Causal} \\ \text{spaces} \\ (\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})\end{array} \arrow[u]
\end{tikzcd}
$$

Causal spaces

- are **a measure-theoretic axiomatisation** of causality;
- whose axioms say
  - doing nothing changes nothing (trivial intervention); and
  - when giving X value x, X has value x (interventional determinism);

Summary

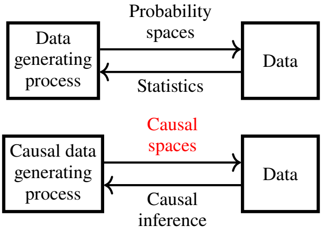


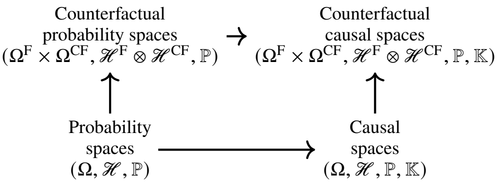
$$
\begin{tikzcd}
\begin{array}{c}\text{Counterfactual}\\ \text{probability spaces}\\ (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P})\end{array} \arrow[r] & \begin{array}{c}\text{Counterfactual}\\ \text{causal spaces}\\ (\Omega^{\mathrm{F}} \times \Omega^{\mathrm{CF}}, \mathscr{H}^{\mathrm{F}} \otimes \mathscr{H}^{\mathrm{CF}}, \mathbb{P}, \mathbb{K})\end{array} \\
\begin{array}{c}\text{Probability}\\ \text{spaces}\\ (\Omega, \mathscr{H}, \mathbb{P})\end{array} \arrow[u] \arrow[r] & \begin{array}{c}\text{Causal}\\ \text{spaces}\\ (\Omega, \mathscr{H}, \mathbb{P}, \mathbb{K})\end{array} \arrow[u]
\end{tikzcd}
$$

Causal spaces

- are **a measure-theoretic axiomatisation** of causality;
- whose axioms say
  - doing nothing changes nothing (trivial intervention); and
  - when giving X value x, X has value x (interventional determinism);

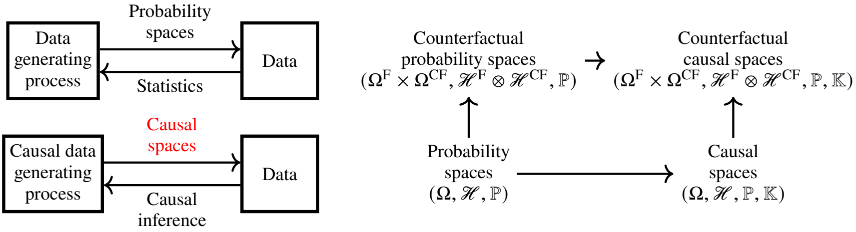
Summary





- **Our hope is that causal spaces play the same role that probability spaces have played for the theory of randomness, for the theory of causality, as the foundation of an area of pure mathematics.**

Summary

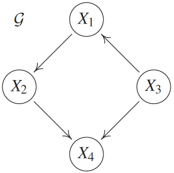


- **Our hope is that causal spaces play the same role that probability spaces have played for the theory of randomness, for the theory of causality, as the foundation of an area of pure mathematics.**
- We view SCMs and potential outcomes are valuable tools to specify (counterfactual) causal spaces.

Thank you.

Existing frameworks

Structural causal models (SCMs)



Potential outcomes

	$Y^{a=0}$	$Y^{a=1}$
Rheia	0	1
Kronos	1	0
Demeter	0	0
Hades	0	0
Hestia	0	0
Poseidon	1	0
Hera	0	0
Zeus	0	1
Artemis	1	1
Apollo	1	0

Causality, Pearl, Cambridge University Press, 2009
Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction, Rubin and Imbens, Cambridge University Press, 2015

Existing frameworks

Desiderata of axioms

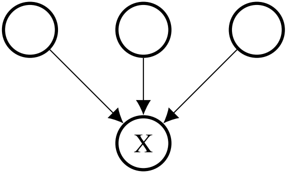
	SCMs	Potential outcomes
Isolation of structural properties	✗	✗
Broadness	✗	✓
Sharpness	✓	✗
Consistent internal logic	✗	✓
Rich theory.	✓	✓

Causality, Pearl, Cambridge University Press, 2009
Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction, Rubin and Imbens, Cambridge University Press, 2015

Table of Contents

How is the causal information encoded?

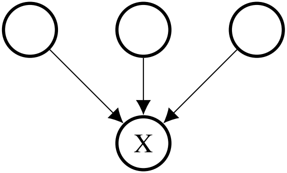
SCMs: $X = f_X(\dots)$



How the system affects a sub-system.

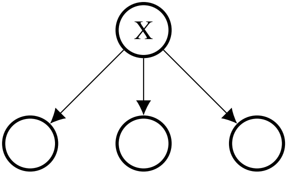
How is the causal information encoded?

SCMs: $X = f_X(...)$



How the system affects a sub-system.

Causal spaces: $K_X(x; \ldots) = \ldots$



How a sub-system affects the system.

What are the primitive objects?

SCMs

- Set of variables $X_1, ..., X_n$.
- Noise distribution.
- Structural equations
  $X_i = f_i(\ldots)$.

What are the primitive objects?

SCMs

- Set of variables $X_1, ..., X_n$.
- Noise distribution.
- Structural equations $X_i = f_i(\ldots)$.
- The system is *deterministic*.
- Observational and interventional distributions are *derived* from these primitive objects.

What are the primitive objects?

SCMs

- Set of variables $X_1, ..., X_n$.
- Noise distribution.
- Structural equations $X_i = f_i(\ldots)$.
- The system is *deterministic*.
- Observational and interventional distributions are *derived* from these primitive objects.

Causal Spaces

- Probability space $(\Omega, \mathscr{H}, \mathbb{P})$.
- Causal kernels K_S.

What are the primitive objects?

SCMs

- Set of variables $X_1, ..., X_n$.
- Noise distribution.
- Structural equations $X_i = f_i(\ldots)$.
- The system is *deterministic*.
- Observational and interventional distributions are *derived* from these primitive objects.

Causal Spaces

- Probability space $(\Omega, \mathscr{H}, \mathbb{P})$.
- Causal kernels K_S.
- The system is *fundamentally stochastic*.
- Observational and interventional distributions *are* the primitive objects.

Relationship between observational and interventional measures

SCMs

- Observational and interventional distributions are *coupled* through the structural equations.

Relationship between observational and interventional measures

SCMs

- Observational and interventional distributions are *coupled* through the structural equations.

Causal Spaces

- Observational and interventional distributions are a priori completely *decoupled*.