

Causality

Jonas Peters
University of Copenhagen

Academic year 2021/2022 – Block 4

UNIVERSITY OF
COPENHAGEN



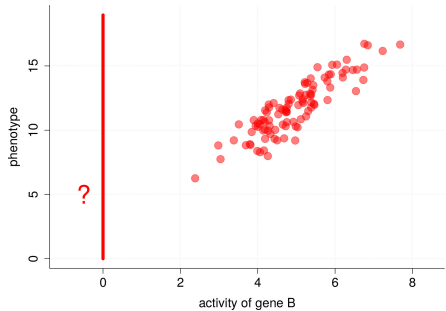
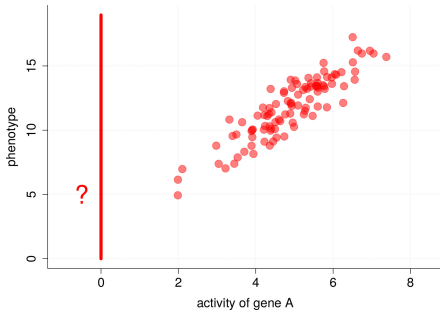
WELCOME!

These slides are used only for visualization. They are not stand-alone material but should be considered as an addition to the reading material, in particular

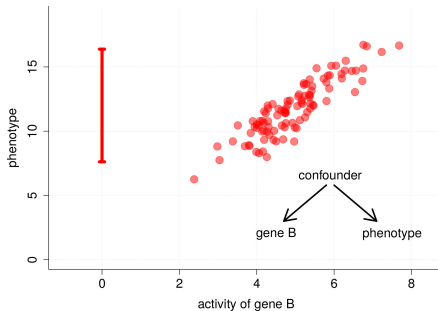
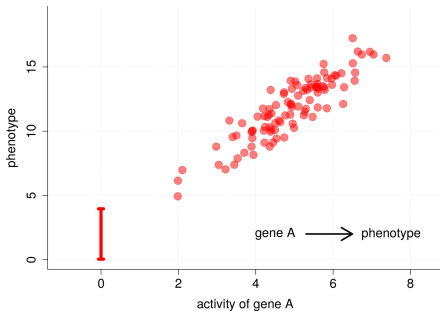
- Book. Peters, Janzing, Schölkopf: Elements of Causal Inference, MIT Press (see also errata).
- Hand-written notes.
- Code examples.

The slides contain many ideas and concepts that are developed by others and these are sometimes not cited properly. For references, please see the above mentioned book.

Consider the following problem.



Causality matters!



BRITISH MEDICAL JOURNAL

LONDON SATURDAY SEPTEMBER 30 1950

SMOKING AND CARCINOMA OF THE LUNG

PRELIMINARY REPORT

BY

RICHARD DOLL, M.D., M.R.C.P.

Member of the Statistical Research Unit of the Medical Research Council

AND

A. BRADFORD HILL, Ph.D., D.Sc.

Professor of Medical Statistics, London School of Hygiene and Tropical Medicine; Honorary Director of the Statistical Research Unit of the Medical Research Council

In England and Wales the phenomenal increase in the number of deaths attributed to cancer of the lung provides one of the most striking changes in the pattern of mortality recorded by the Registrar-General. For example, in the quarter of a century between 1922 and 1947 the annual number of deaths recorded increased from 612 to 8,287, an increase of 13.5 times. This remarkable increase is

whole explanation, although no one would deny that it may well have been contributory. As a corollary, it is right and proper to seek for other causes.

Possible Causes of the Increase

Two main causes have from time to time been put forward. (1) A general increase in the consumption of tobacco, and (2) a general increase in the consumption of alcohol.

Example: smoking

BRITISH MEDICAL JOURNAL

TABLE VII.—*Estimate of Total Amount of Tobacco Ever Consumed by Smokers; Lung-carcinoma Patients and Control Patients with Diseases Other Than Cancer*

Disease Group	No. Who have Smoked Altogether					Probability Test
	365 Cigs.—	50,000 Cigs.—	150,000 Cigs.—	250,000 Cigs.—	500,000 Cigs. +	
Males:						
Lung-carcinoma patients (647)	19 (2.9%)	145 (22.4%)	183 (28.3%)	225 (34.8%)	75 (11.6%)	$\chi^2=30.60$; $n=4$; $P<0.001$
Control patients with diseases other than cancer (622) ..	36 (5.8%)	190 (30.5%)	182 (29.3%)	179 (28.9%)	35 (5.6%)	
Females:						
Lung-carcinoma patients (41) ..	10 (24.4%)	19 (46.3%)	5 (12.2%)	7 (17.1%)	0 (0.0%)	$\chi^2=12.97$; $n=2$; $0.001 < P < 0.01$ (Women smoking 15 or more cigarettes a day grouped together)
Control patients with diseases other than cancer (28) ..	19 (67.9%)	5 (17.9%)	3 (10.7%)	1 (3.6%)	0 (0.0%)	

NG

uncil

Director of the Statistical

no one would deny that it
utory. As a corollary, it is
other causes.

of the Increase

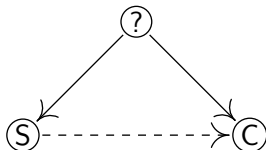
m time to time been put for-

Example: smoking

BRITISH MEDICAL JOURNAL

TABLE VII.—*Effect of Smoking on the Incidence of Lung Cancer*

Disease Group					
Males:					
Lung-carcinoma patients (647)	36 (5.8%)	190 (30.5%)	182 (29.3%)	179 (28.9%)	35 (5.6%)
Control patients with diseases other than cancer (622) ..					
Females:					
Lung-carcinoma patients (41) ..	10 (24.4%)	19 (46.3%)	5 (12.2%)	7 (17.1%)	0 (0.0%)
Control patients with diseases other than cancer (28) ..	19 (67.9%)	5 (17.9%)	3 (10.7%)	1 (3.6%)	0 (0.0%)



Consumed
patients with

Probability
Test

$\chi^2 = 30.60$;
 $n = 4$;
 $P < 0.001$

$\chi^2 = 12.97$;
 $n = 2$;
 $0.001 < P < 0.01$
(Women
smoking 15
or more cig-
arettes a day
grouped to-
gether)

NG

council

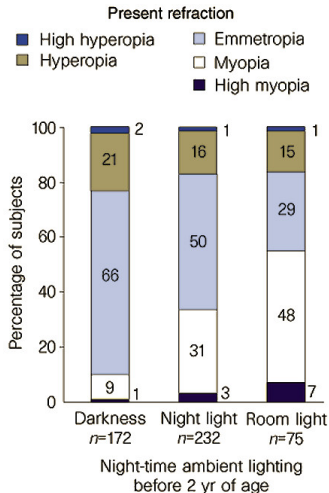
Director of the Statistical

no one would deny that it
utory. As a corollary, it is
other causes.

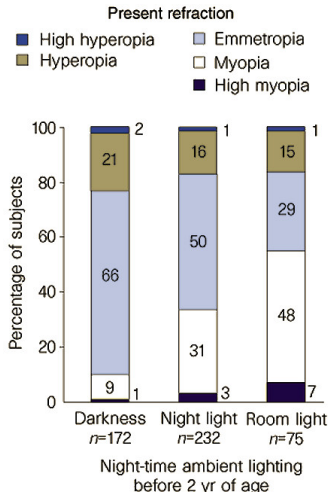
of the Increase

m time to time been put for-

Example: myopia



Example: myopia



“the strength of the association . . . does suggest that the absence of a daily period of darkness during childhood is a potential precipitating factor in the development of myopia”

Quinn, Shin, Maguire, Stone: *Myopia and ambient lighting at night*, Nature 1999

Example: myopia

Patente

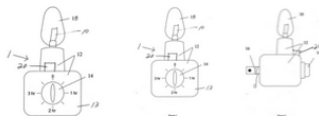
Night light with sleep timer

US 20050007889 A1

ZUSAMMENFASSUNG

A timer a light and an optional music source is located on or in a housing of a nightlight assembly. When this assembly is plugged into a source of electric power, the timer is set to a selected time for the light and optional music to remain on. After this selected time has elapsed, the light and music automatically turns off, allowing for sleep in appropriate darkness and silence.

BILDER (3)



BESCHREIBUNG

ANSPRÜCHE (18)

Veröffentlichungsnummer	US20050007889 A
Publikationstyp	Anmeldung
Anmeldenummer	US 10/614,245
Veröffentlichungsdatum	13. Jan. 2005
Eingetragen	8. Juli 2003
Prioritätsdatum [?]	8. Juli 2003
Erfinder	Karin Peterson
Ursprünglich Bevollmächtigter	Peterson Karin Lyn
Zitat exportieren	BiBTeX , EndNote , F
Klassifizierungen ⁽⁴⁾	
Externe Links:	USPTO , USPTO-Zuordnung , Esp

Example: myopia

Patente

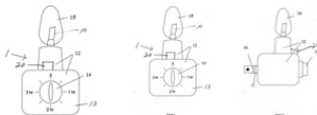
Night light with sleep timer

US 20050007889 A1

ZUSAMMENFASSUNG

A timer a light and an optional music source is located on or in a housing of a nightlight assembly. When this assembly is plugged into a source of electric power, the timer is set to a selected time for the light and optional music to remain on. After this selected time has elapsed, the light and music automatically turns off, allowing for sleep in appropriate darkness and silence.

BILDER (3)



Question: Does the night light with sleep timer help?

BESCHREIBUNG

ANSPRÜCHE (18)

Veröffentlichungsnummer	US20050007889 A
Publikationstyp	Anmeldung
Anmeldenummer	US 10/614,245
Veröffentlichungsdatum	13. Jan. 2005
Eingetragen	8. Juli 2003
Prioritätsdatum [?]	8. Juli 2003
Erfinder	Karin Peterson
Ursprünglich Bevollmächtigter	Peterson Karin Lyn
Zitat exportieren	BiBTeX , EndNote , F
Klassifizierungen	(4)
Externe Links:	USPTO , USPTO-Zuordnung , Esp

Example: kidney stones

	Treatment A	Treatment B
	$\frac{273}{350} = 0.78$	$\frac{289}{350} = 0.83$
	$\frac{562}{700} = 0.80$	

Assume: treatment is chosen only based on size of stones.

Charig et al.: *Comparison of treatment of renal calculi by open surgery, (...)* , British Medical Journal, 1986

Example: kidney stones

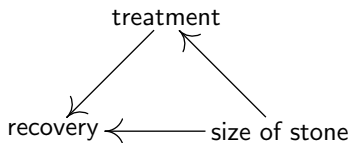
	Treatment A	Treatment B
Small Stones ($\frac{357}{700} = 0.51$)	$\frac{81}{87} = 0.93$	$\frac{234}{270} = 0.87$
Large Stones ($\frac{343}{700} = 0.49$)	$\frac{192}{263} = 0.73$	$\frac{55}{80} = 0.69$
	$\frac{273}{350} = 0.78$	$\frac{289}{350} = 0.83$
	$\frac{562}{700} = 0.80$	

Assume: treatment is chosen only based on size of stones.

Charig et al.: *Comparison of treatment of renal calculi by open surgery, (...)* , British Medical Journal, 1986

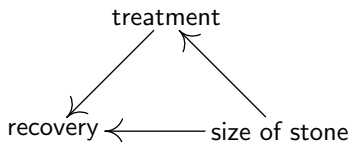
Example: kidney stones

underlying ground truth:



Example: kidney stones

underlying ground truth:



Question: What is the expected recovery if all get treatment B?

(Make treatment independent of size.)

- Classical statistics:
statistical model:

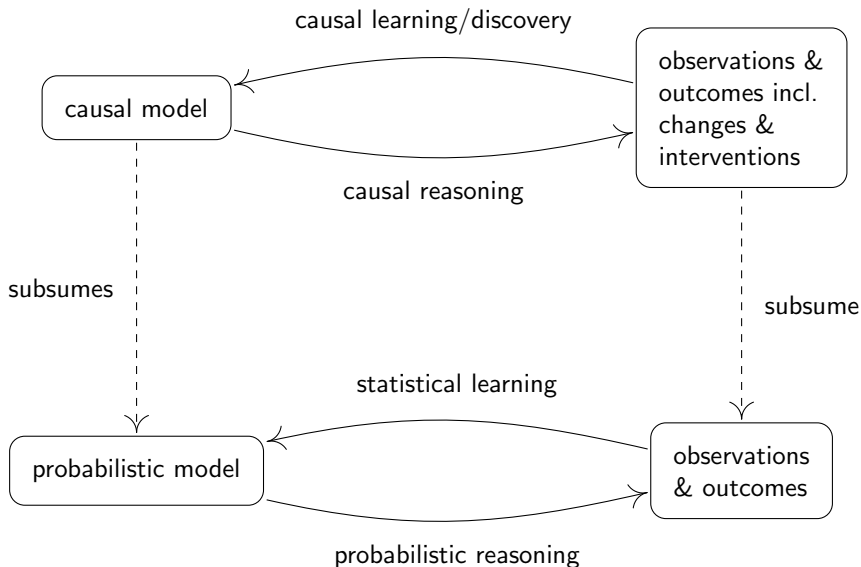
- Classical statistics:
statistical model: $\{P_\theta, \theta \in \Theta\}$

- Classical statistics:
statistical model: $\{P_\theta, \theta \in \Theta\}$
observed data: from P_{θ_0}

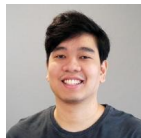
- Classical statistics:
statistical model: $\{P_\theta, \theta \in \Theta\}$
observed data: from P_{θ_0}
inference: investigate θ_0
prediction: use parts of P_{θ_0}

- Classical statistics:
statistical model: $\{P_\theta, \theta \in \Theta\}$
observed data: from P_{θ_0}
inference: investigate θ_0
prediction: use parts of P_{θ_0}
- Causality is often about

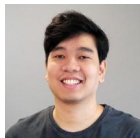
- Classical statistics:
statistical model: $\{P_\theta, \theta \in \Theta\}$
observed data: from P_{θ_0}
inference: investigate θ_0
prediction: use parts of P_{θ_0}
- Causality is often about asking questions about distributions different from the one we have data from.
- We need models relating these distributions.
- We need tools to do causal inference.



- Questions: lectures, TA sessions and padlet (better than emails/absalon messages)
- Format: lecture (1–4), inverted lecture (5–8), ?
- TA: Sorawit (James) Saengkyongam



- Questions: lectures, TA sessions and padlet (better than emails/absalon messages)
- Format: lecture (1–4), inverted lecture (5–8), ?
- TA: Sorawit (James) Saengkyongam



Exam Students:

- 5 mandatory assignments (up to two people)
- 4 assignments need to be passed (choose two exercises, hand-in in time)
- oral exam (21.6./22.6./also: 24.6.)

- Questions: lectures, TA sessions and padlet (better than emails/absalon messages)
- Format: lecture (1–4), inverted lecture (5–8), ?
- TA: Sorawit (James) Saengkyongam



Exam Students:

- 5 mandatory assignments (up to two people)
- 4 assignments need to be passed (choose two exercises, hand-in in time)
- oral exam (21.6./22.6./also: 24.6.)

Exam PhD Students:

- No hand-in of assignments.
- Report at the end (22.6.) about own research problem (alternatively: paper or data study).
- Cannot contain recycled material.

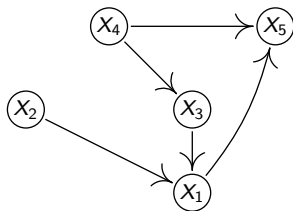
Maybe, there will be a bit of flexibility at the end. What are you interested in?

Hand-written notes 1.

Definition: d -separation

X_i and X_j are d -separated by S if all paths between X_i and X_j are blocked by S .

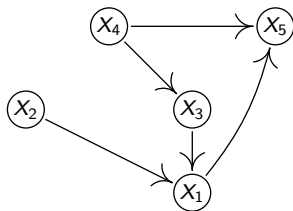
Check, whether all paths blocked!!



Definition: d -separation

X_i and X_j are d -separated by S if all paths between X_i and X_j are blocked by S .

Check, whether all paths blocked!!



$\circ \dots \rightarrow \textcolor{green}{\circ} \rightarrow \dots \circ$ blocks a path.

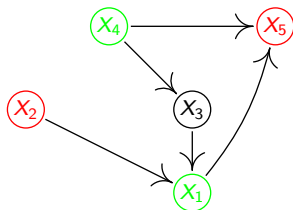
$\circ \dots \leftarrow \textcolor{green}{\circ} \rightarrow \dots \circ$ blocks a path.

$\circ \dots \rightarrow \circ \leftarrow \dots \circ$ blocks a path.

Definition: d -separation

X_i and X_j are d -separated by S if all paths between X_i and X_j are blocked by S .

Check, whether all paths blocked!!



○ ... → ● → ... ○ blocks a path.

○ ... ← ● → ... ○ blocks a path.

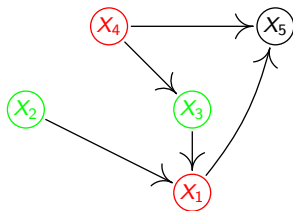
○ ... → ○ ← ... ○ blocks a path.

X_2 and X_5 are d -sep. by $\{X_1, X_4\}$

Definition: d -separation

X_i and X_j are d -separated by S if all paths between X_i and X_j are blocked by S .

Check, whether all paths blocked!!



○ ... → ● → ... ○ blocks a path.

○ ... ← ● → ... ○ blocks a path.

○ ... → ○ ← ... ○ blocks a path.

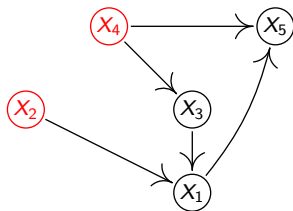
X_2 and X_5 are d -sep. by $\{X_1, X_4\}$

X_4 and X_1 are d -sep. by $\{X_2, X_3\}$

Definition: d -separation

X_i and X_j are d -separated by S if all paths between X_i and X_j are blocked by S .

Check, whether all paths blocked!!



$\circ \dots \rightarrow \text{green} \rightarrow \dots \circ$ blocks a path.

$\circ \dots \leftarrow \text{green} \rightarrow \dots \circ$ blocks a path.

$\circ \dots \rightarrow \circ \leftarrow \dots \circ$ blocks a path.

X_2 and X_5 are d -sep. by $\{X_1, X_4\}$

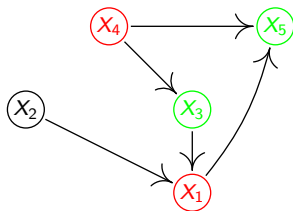
X_4 and X_1 are d -sep. by $\{X_2, X_3\}$

X_2 and X_4 are d -sep. by $\{\}$

Definition: d -separation

X_i and X_j are d -separated by S if all paths between X_i and X_j are blocked by S .

Check, whether all paths blocked!!



○ ... → ○ → ... ○ blocks a path.

○ ... ← ○ → ... ○ blocks a path.

○ ... → ○ ← ... ○ blocks a path.

X_2 and X_5 are d -sep. by $\{X_1, X_4\}$

X_4 and X_1 are d -sep. by $\{X_2, X_3\}$

X_2 and X_4 are d -sep. by $\{\}$

X_4 and X_1 are NOT d -sep. by $\{X_3, X_5\}$

Hand-written notes 2.

p	number of DAGs with p nodes
1	1
2	3
3	25
4	543
5	29281
6	3781503
7	1138779265
8	783702329343
9	1213442454842881
10	4175098976430598143
11	31603459396418917607425
12	521939651343829405020504063
13	18676600744432035186664816926721
14	1439428141044398334941790719839535103
15	237725265553410354992180218286376719253505
16	83756670773733320287699303047996412235223138303
17	62707921196923889899446452602494921906963551482675201
18	99421195322159515895228914592354524516555026878588305014783
19	332771901227107591736177573311261125883583076258421902583546773505
20	2344880451051088988152559855229099188899081192234291298795803236068491263
21	34698768283588750028759328430181088222313944540438601719027559113446586077675521
22	1075822921725761493652956179327624326573727662809185218104090000500559527511693495107583
23	69743329837281492647141549700245804876504274990515985894109106401549811985510951501377122074625

<https://oeis.org/A003024/b003024.txt>

Definition

Given a directed graph (V, E) , a permutation

$$\pi : V \rightarrow V \text{ (bijective)}$$

is called a causal/topological order if

$$j \in DE_i \Rightarrow \pi(i) < \pi(j).$$

(Remember: $\pi^{-1}(1)$ does not have any parents.)

Definition

Given a directed graph (V, E) , a permutation

$$\pi : V \rightarrow V \text{ (bijective)}$$

is called a causal/topological order if

$$j \in DE_i \Rightarrow \pi(i) < \pi(j).$$

(Remember: $\pi^{-1}(1)$ does not have any parents.)

Proposition (Prop. B.2)

For any DAG, there exists a causal order.

Hand-written notes 3.