Causality

Jonas Peters
University of Copenhagen

Academic year 2020/2021 – Block 4

UNIVERSITY OF
COPENHAGEN

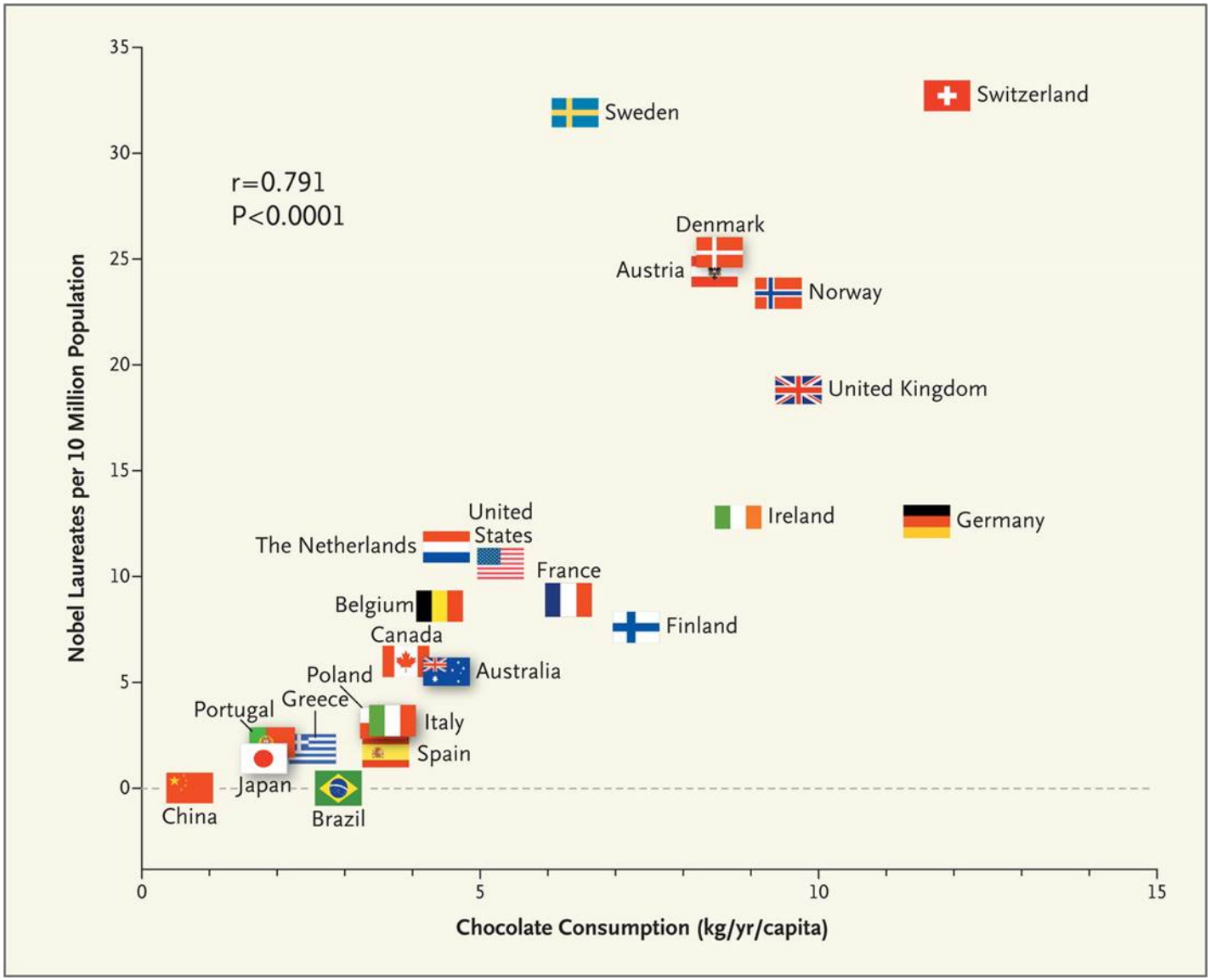
WELCOME!

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 often not cited properly. For references, please see the above mentioned book.

Example: chocolate



F. H. Messerli: *Chocolate Consumption, Cognitive Function, and Nobel Laureates*, N Engl J Med 2012

Example: chocolate

Confectionery news.com

HEADLINES | TRENDS | TECHNOLOGY | PRODUCTS | JOBS | EVENTS | RELATED SITES |

HEADLINES > REGULATION & SAFETY

62 415 10 16

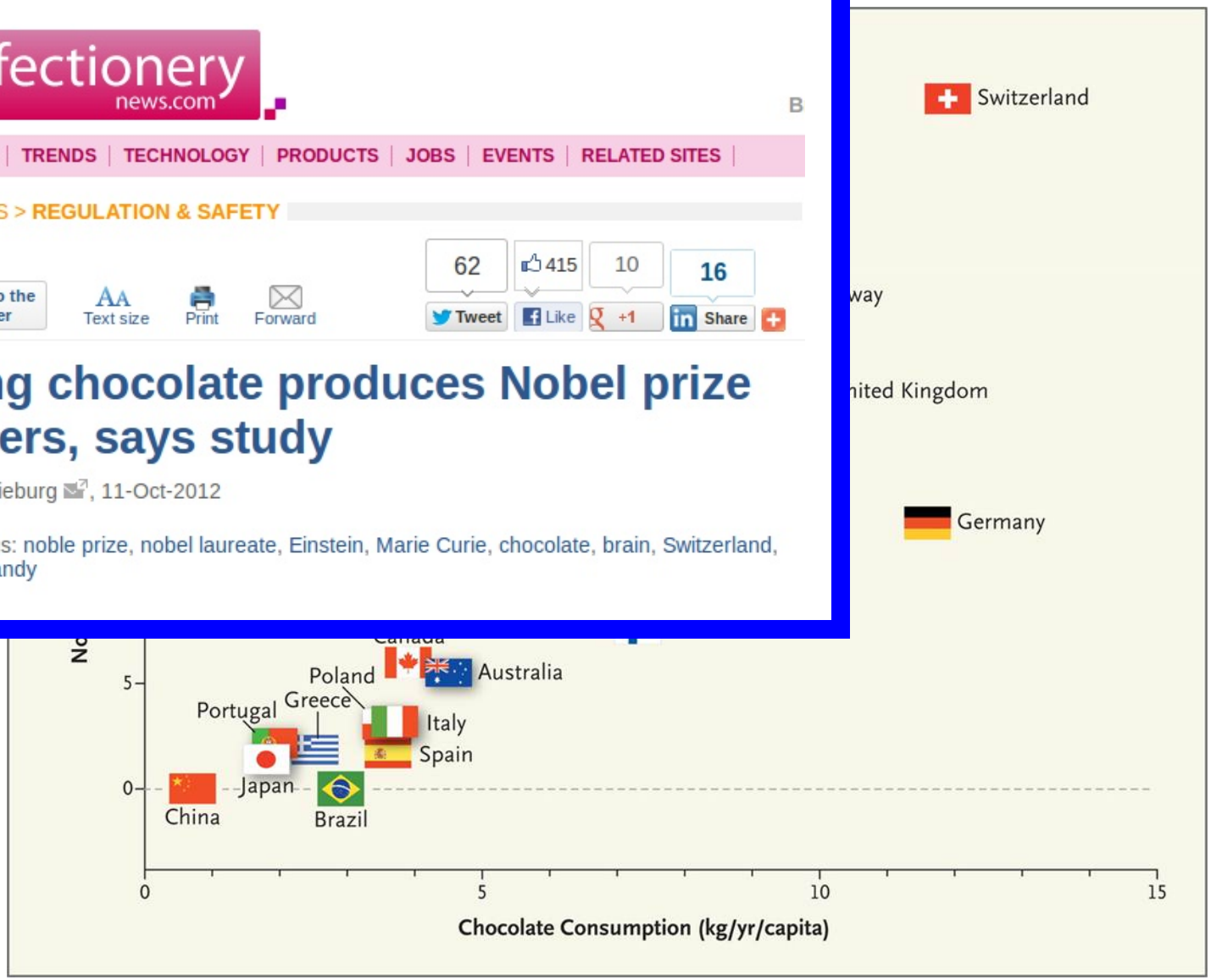
Subscribe to the Newsletter | Text size | Print | Forward

Tweet | Like | +1 | Share

Eating chocolate produces Nobel prize winners, says study

By Oliver Nieburg, 11-Oct-2012

Related tags: noble prize, nobel laureate, Einstein, Marie Curie, chocolate, brain, Switzerland, Sweden, candy

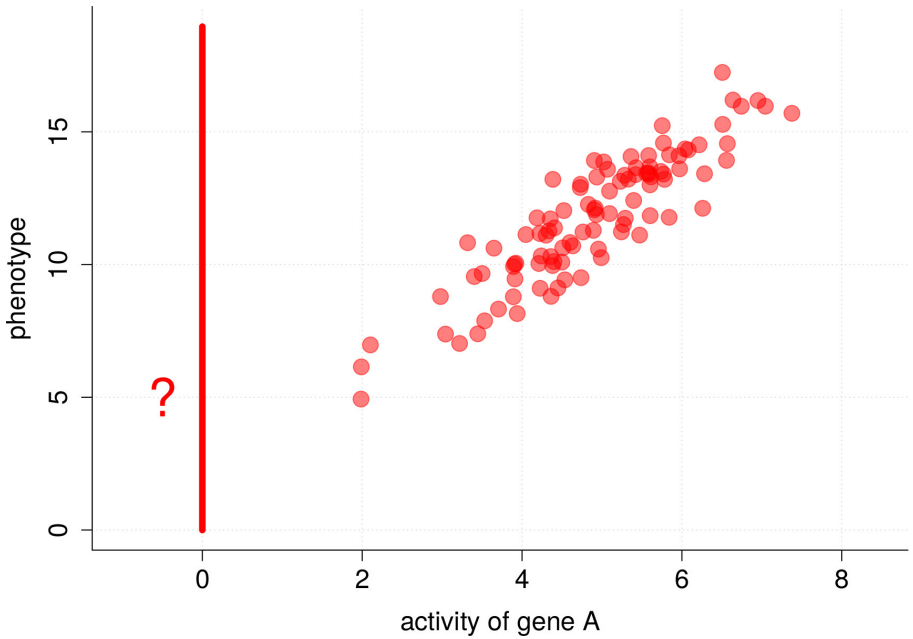


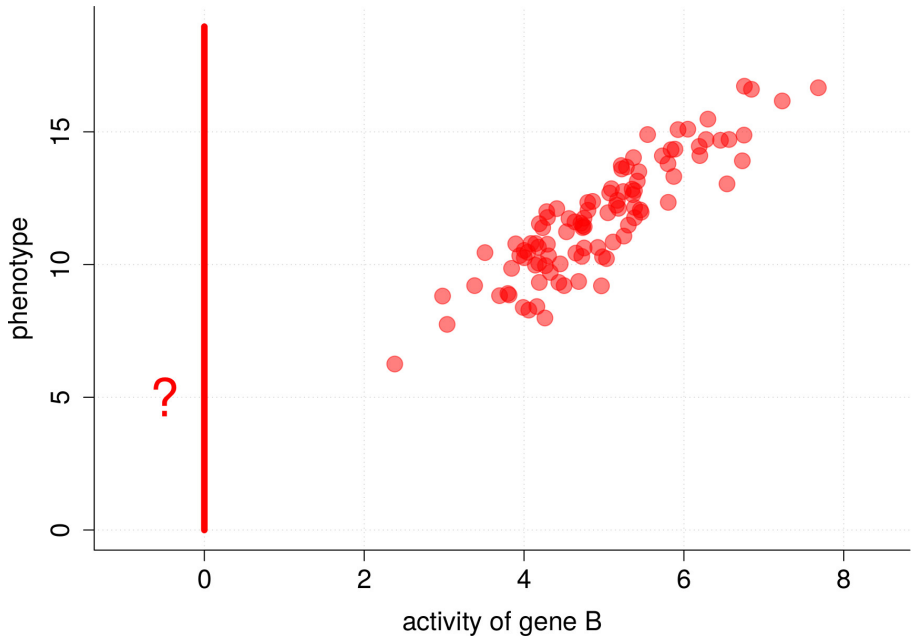
F. H. Messerli: *Chocolate Consumption, Cognitive Function, and Nobel Laureates*, N Engl J Med 2012

Example: chocolate

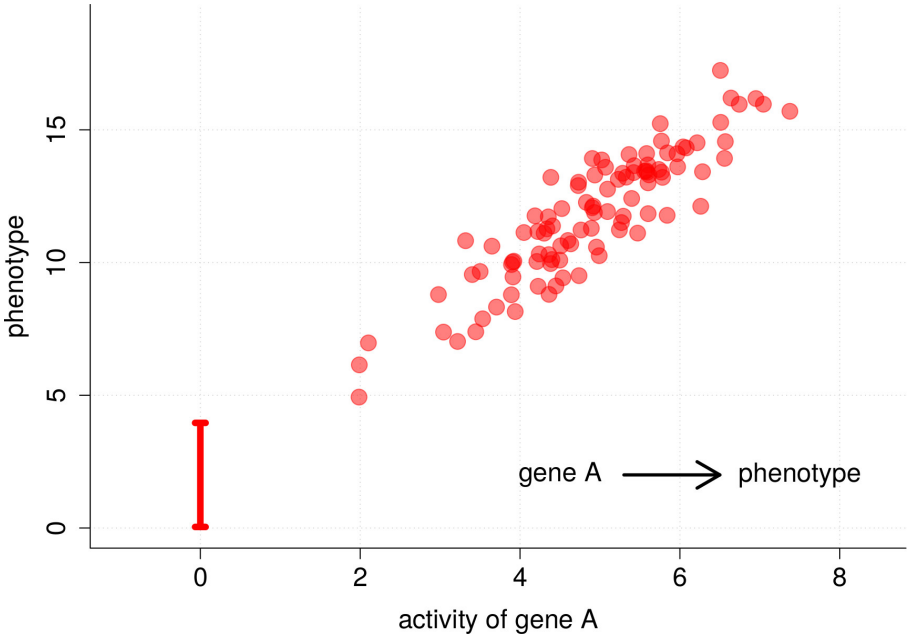
F. H. 12

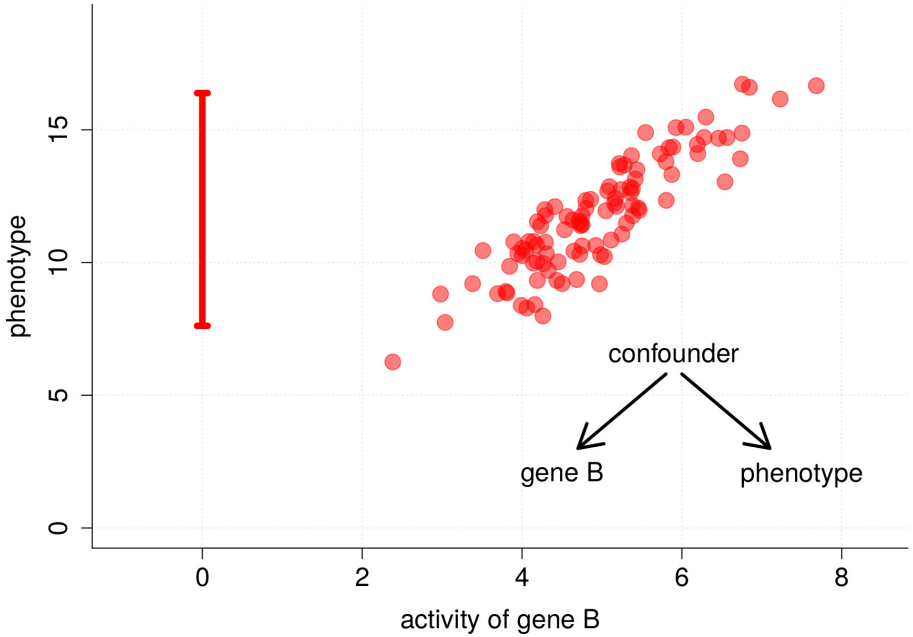
Consider the following problem.





Causality matters!





Example: smoking

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 9,287, or roughly fifteenfold. 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 for-

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%)	$\chi^2=30{\cdot}60$; 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%)	$\chi^2=12{\cdot}97$; 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%)	

Example: smoking

BRITISH MEDICAL JOURNAL

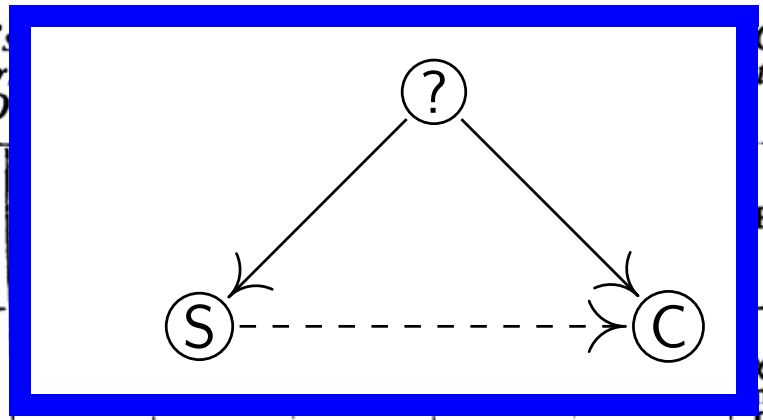


TABLE VII.—*Es* ... *Consumed by Smoker* ... *tients with Diseases O*

Disease Group						Probability Test
Males:						
Lung-carcinoma patients (647)						$\chi^2 = 30{\cdot}60$; n=4; $P < 0{\cdot}001$
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%)	$\chi^2 = 12{\cdot}97$; n=2; $0{\cdot}001 < P < 0{\cdot}01$ (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%)	

...UNG

...uncil

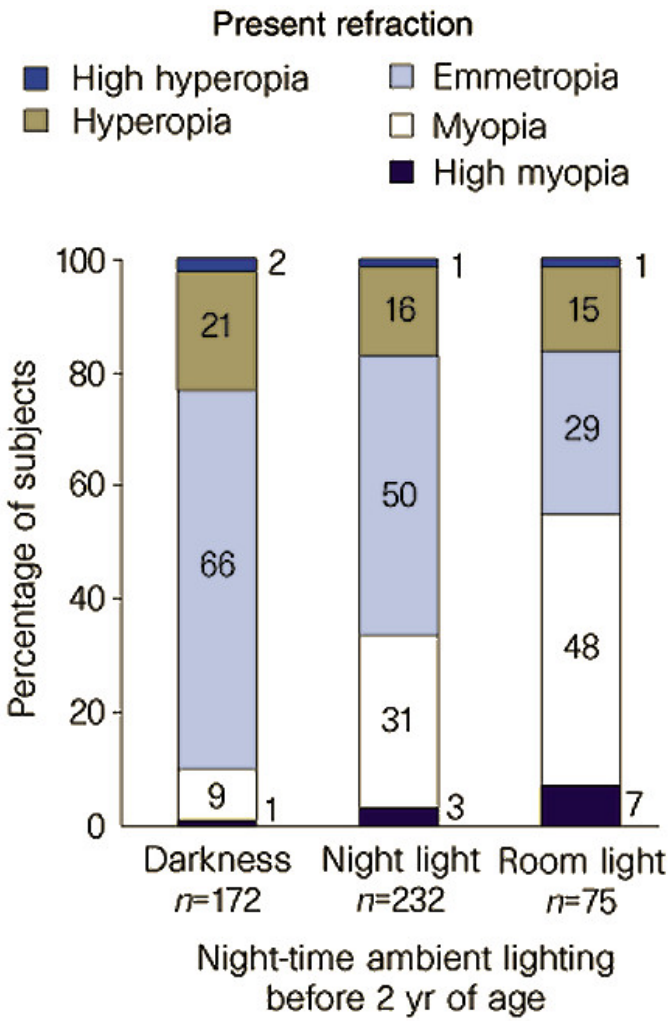
... *Director of the Statistical*

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

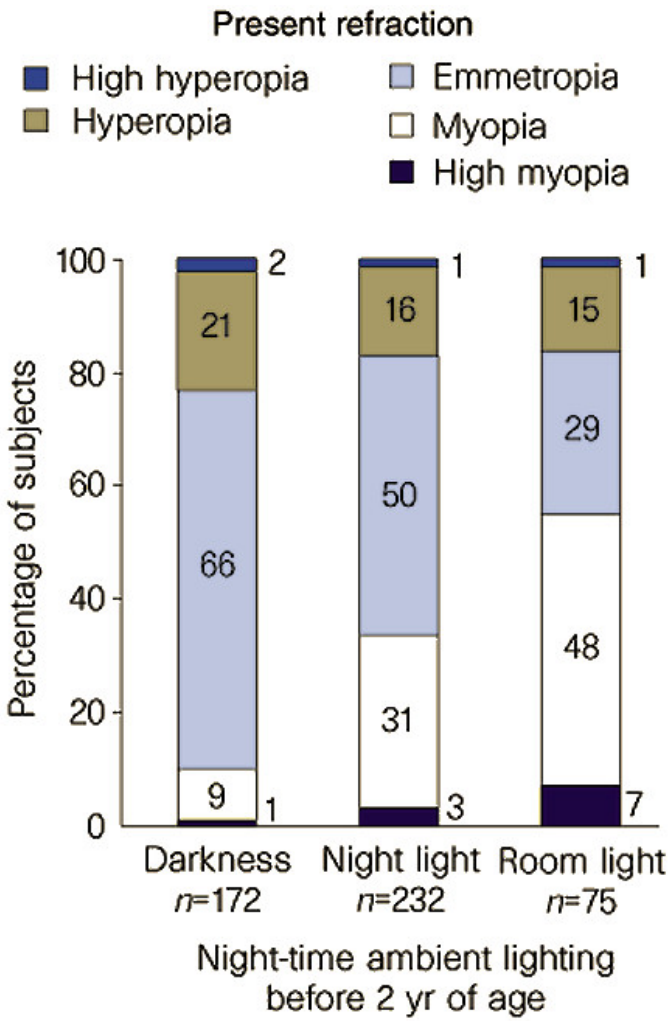
... of the Increase

...m time to time been put for-

Example: 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

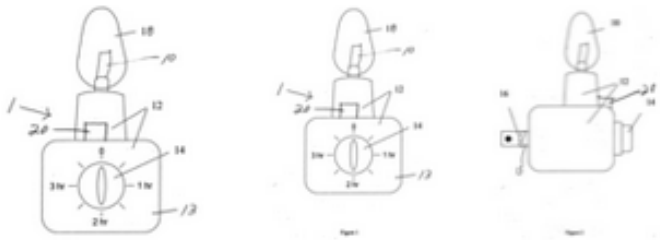
A timer a light and an optional music source is located on or in a housing of a nightlight assembly. When this assembly is plugged into a source of electric power, the timer is set to a selected time for the light and optional music to remain on. After this selected time has elapsed, the light and music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

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

BILDER (3)



BESCHREIBUNG

ANSPRÜCHE (18)

Example: myopia

Patente

Night light with sleep timer

US 20050007889 A1

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

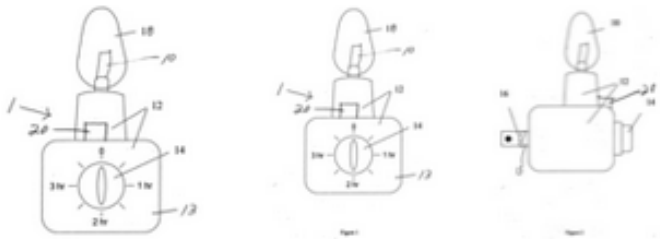
Klassifizierungen (4)

Externe Links: USPTO, USPTO-Zuordnung, Esp

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)

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

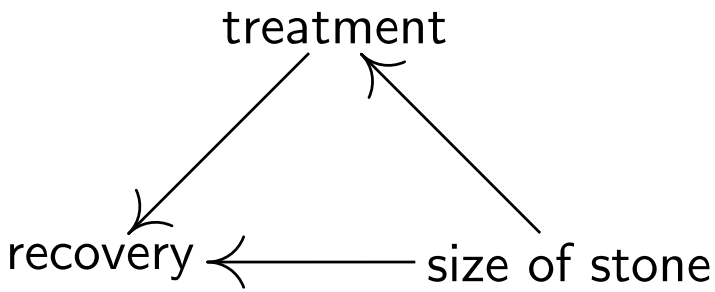
	Treatment A	Treatment B
Small Stones ($\frac{357}{700} = 0.51$)	$\frac{81}{87} = 0.93$	$\frac{234}{270} = 0.87$
Large Stones ($\frac{343}{700} = 0.49$)	$\frac{192}{263} = 0.73$	$\frac{55}{80} = 0.69$
	$\frac{273}{350} = 0.78$	$\frac{289}{350} = 0.83$
	$\frac{562}{700} = 0.80$	

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

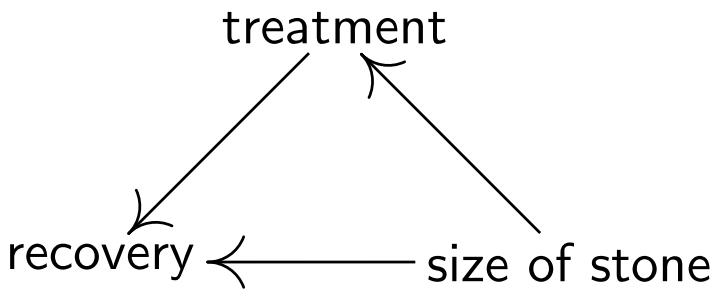
Example: kidney stones

underlying ground truth:



Example: kidney stones

underlying ground truth:

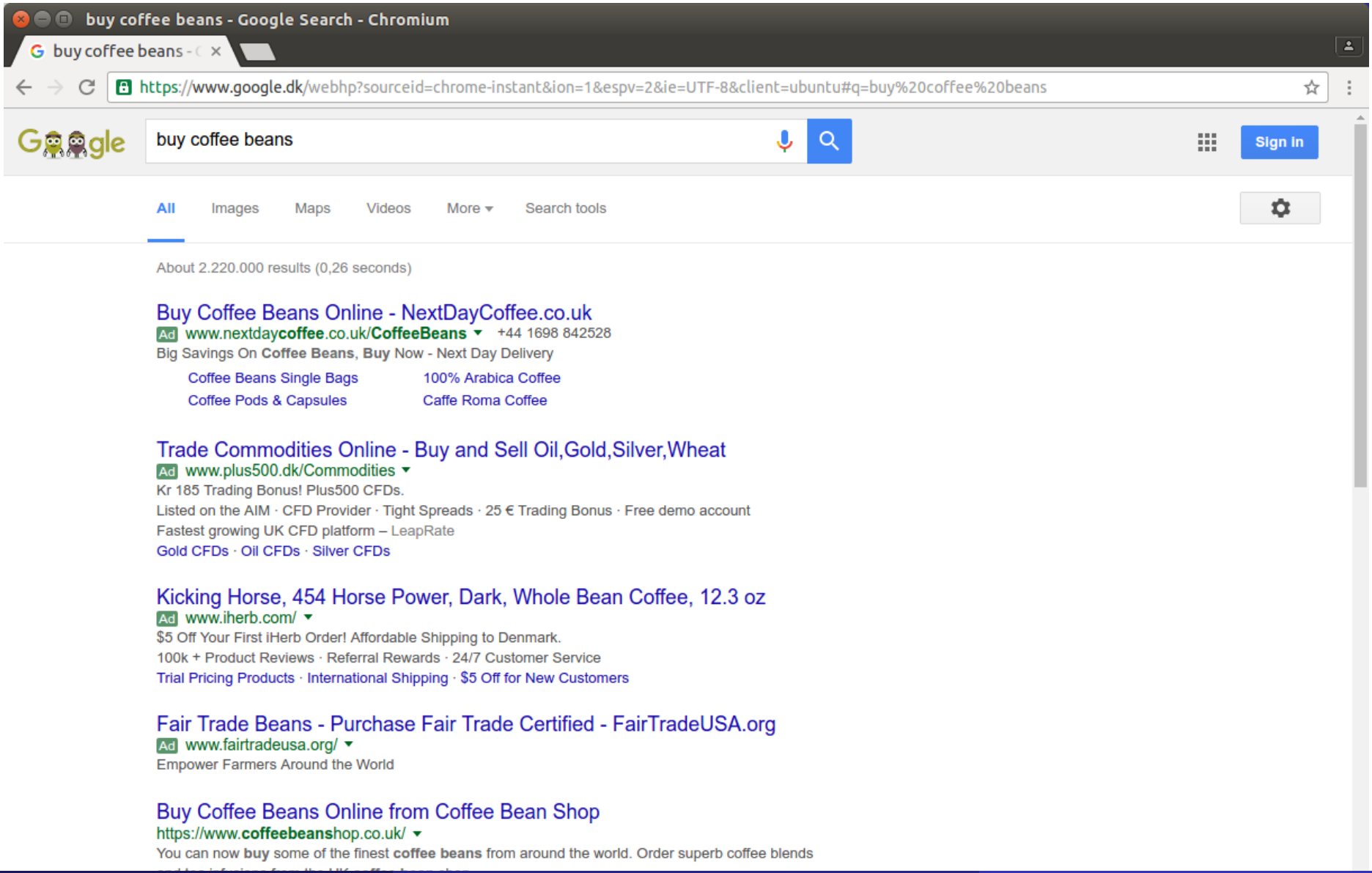


Question: What is the expected recovery if all get treatment B?

(Make treatment independent of size.)

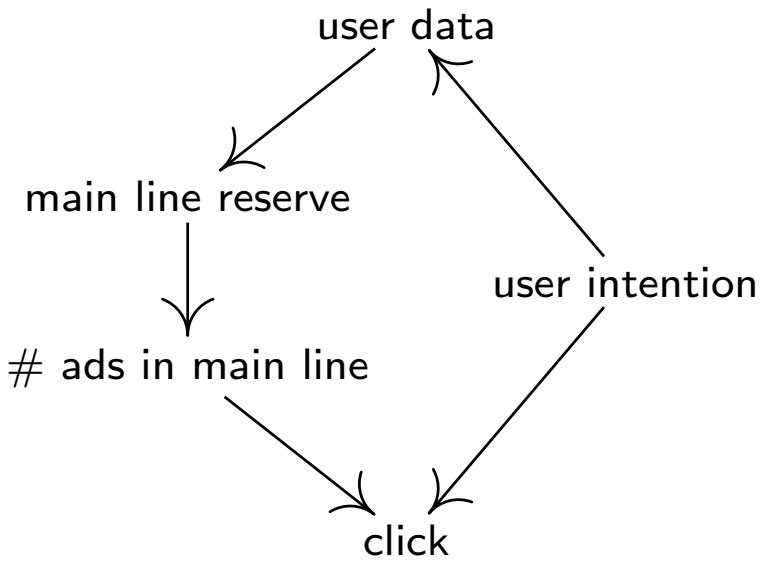
Example: advertisement

Relation to RL



Example: advertisement

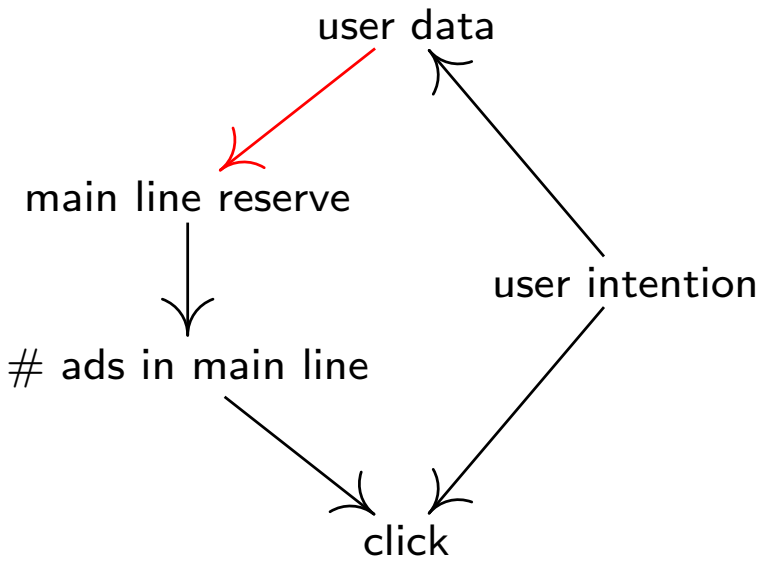
Relation to RL



Bottou et al.: *Counterfactual Reasoning and Learning Systems: The Example of Computational Advertising*, JMLR 2013

Example: advertisement

Relation to RL



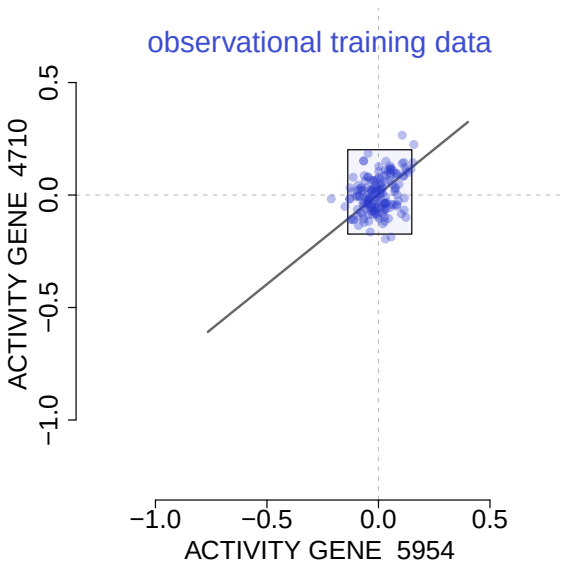
Question: How do we choose an optimal main line reserve?

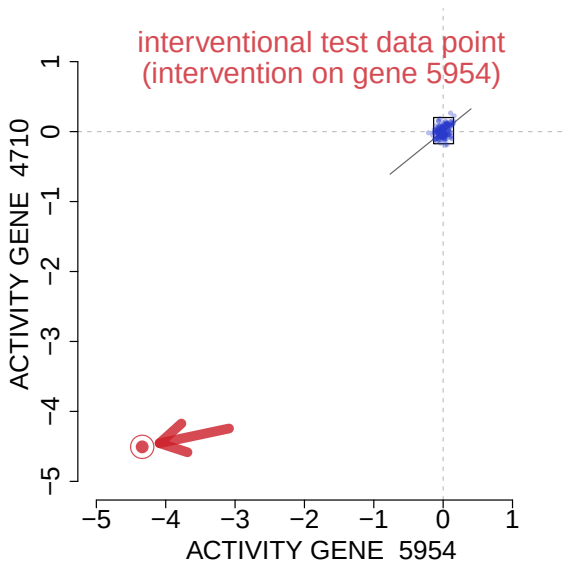
Bottou et al.: *Counterfactual Reasoning and Learning Systems: The Example of Computational Advertising*, JMLR 2013

Example: gene interactions

genetic perturbation experiments for yeast

- $p = 6170$ genes
- $n_{obs} = 160$ wild-types
- $n_{int} = 1479$ gene deletions (targets known)





Example: gene interactions

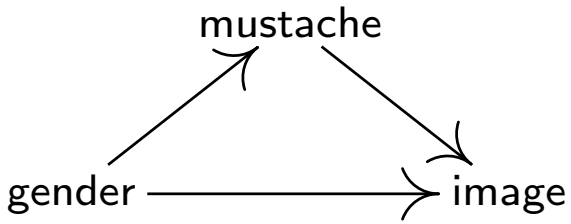
genetic perturbation experiments for yeast

- $p = 6170$ genes
- $n_{obs} = 160$ wild-types
- $n_{int} = 1479$ gene deletions (targets known)

- Causal relationships are often stable!

Kemmeren et al.: Large-scale genetic perturbations reveal reg. networks and an abundance of gene-specific repressors. Cell, 2014

Example: mustache



(a) Intervening vs Conditioning on Mustache, Top: Intervene Mustache=1, Bottom: Condition Mustache=1

Kocaoglu et al: *CausalGAN: Learning Causal Implicit Generative Models with Adversarial Training*, arXiv:1709.02023

- Classical statistics:
  statistical model:

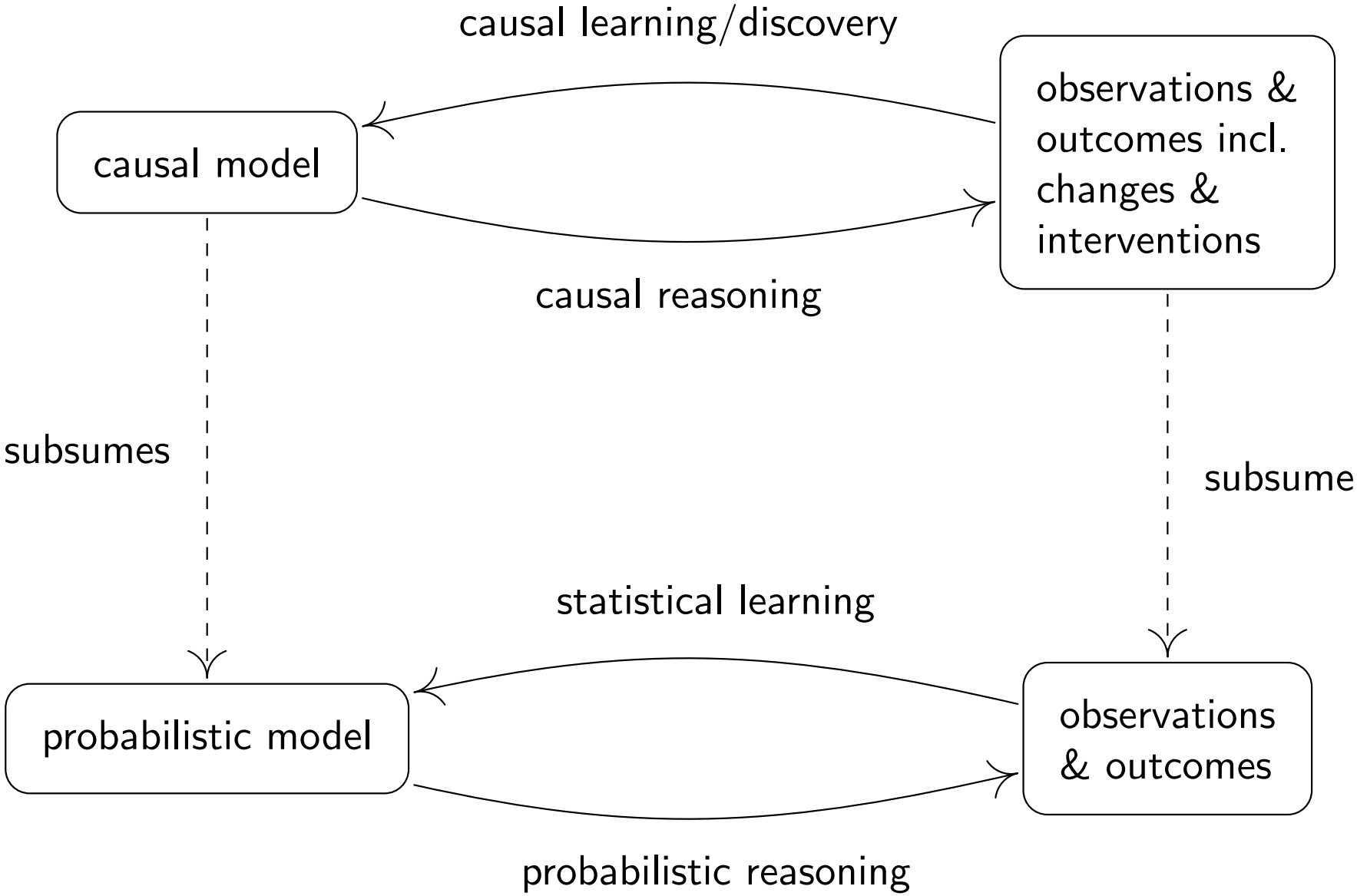
- Classical statistics:
  statistical model: $\{P_\theta, \theta \in \Theta\}$

- 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.


causal learning/discovery
causal model
observations & outcomes incl. changes & interventions
causal reasoning
subsumes
subsume
statistical learning
probabilistic model
observations & outcomes
probabilistic reasoning

- Questions: lectures, TA sessions and padlet (better than emails/absalon messages)
- TA: Sorawit (James) Saengkyongam

- Questions: lectures, TA sessions and padlet (better than emails/absalon messages)
- TA: Sorawit (James) Saengkyongam

Exam Students:

- 6 mandatory assignments (up to two people)
- 5 assignments need to be passed (hand-in in time)
- oral exam (22.6./23.6. via zoom)

- Questions: lectures, TA sessions and padlet (better than emails/absalon messages)
- TA: Sorawit (James) Saengkyongam

Exam Students:

- 6 mandatory assignments (up to two people)
- 5 assignments need to be passed (hand-in in time)
- oral exam (22.6./23.6. via zoom)

Exam PhD Students:

- No assignments (mark exercise for feedback)
- Report at the end (23.6.) about paper, data study or own research problem.
- Cannot contain recycled material.

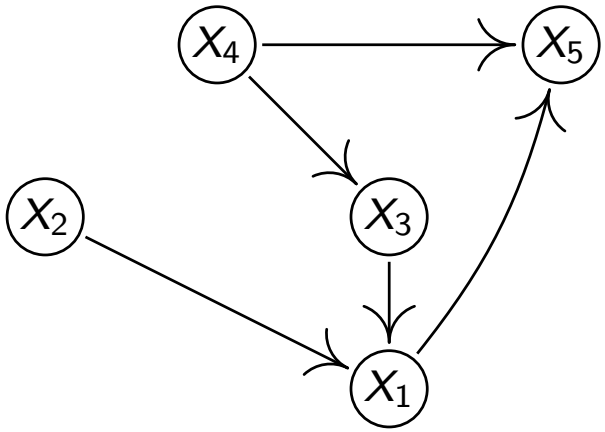
All information on absalon.

Hand-written notes 1

Definition: d-separation

X_i and X_j are d-separated by $\mathcal{S}$ if all paths between X_i and X_j are blocked by $\mathcal{S}$.

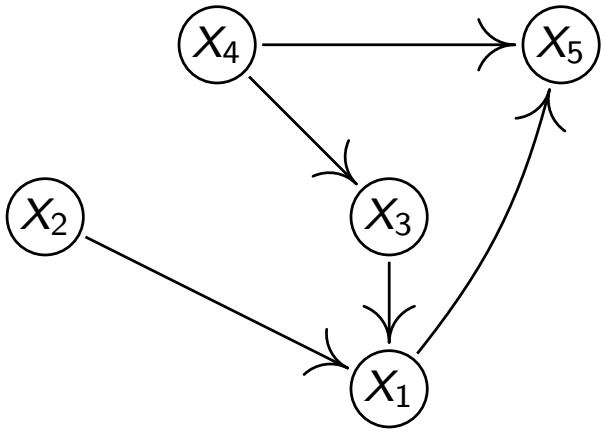
Check, whether all paths blocked!!



Definition: d-separation

X_i and X_j are d-separated by $\mathcal{S}$ if all paths between X_i and X_j are blocked by $\mathcal{S}$.

Check, whether all paths blocked!!



$\circ \cdots \rightarrow \circ \rightarrow \cdots \circ$ blocks a path.

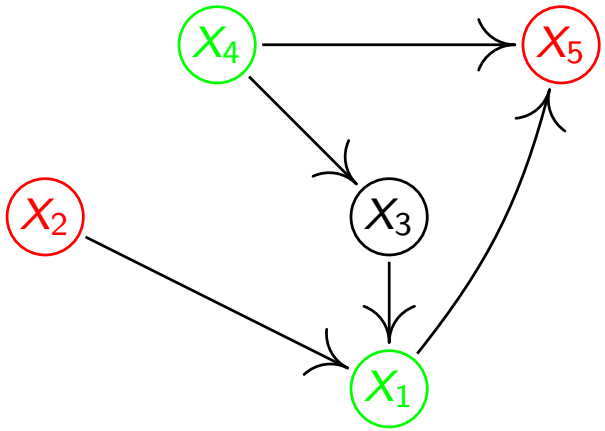
$\circ \cdots \leftarrow \circ \rightarrow \cdots \circ$ blocks a path.

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

Definition: d-separation

X_i and X_j are d-separated by $\mathcal{S}$ if all paths between X_i and X_j are blocked by $\mathcal{S}$.

Check, whether all paths blocked!!



$\circ \cdots \rightarrow \circ \rightarrow \cdots \circ$ blocks a path.

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

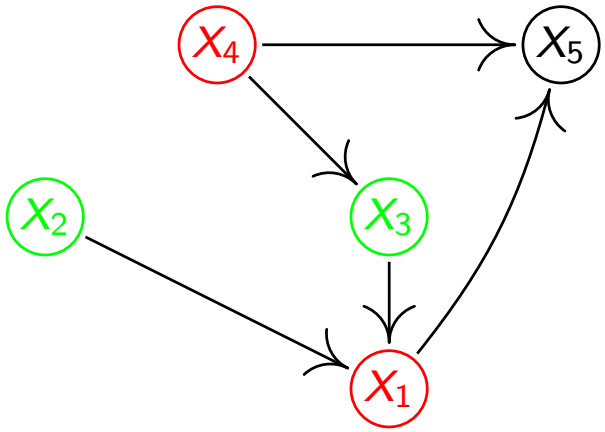
$\circ \cdots \rightarrow \circ \leftarrow \cdots \circ$ blocks a path.

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 $\mathcal{S}$ if all paths between X_i and X_j are blocked by $\mathcal{S}$.

Check, whether all paths blocked!!



$\circ \cdots \rightarrow \circ \rightarrow \cdots \circ$ blocks a path.

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

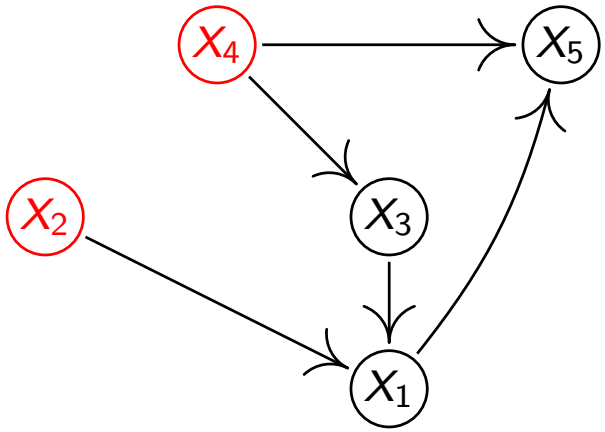
$\circ \cdots \rightarrow \circ \leftarrow \cdots \circ$ blocks a path.

X_2 and X_5 are d-sep. by $\{X_1, X_4\}$
X_4 and X_1 are d-sep. by $\{X_2, X_3\}$

Definition: d-separation

X_i and X_j are d-separated by $\mathcal{S}$ if all paths between X_i and X_j are blocked by $\mathcal{S}$.

Check, whether all paths blocked!!



$\circ \cdots \rightarrow \circ \rightarrow \cdots \circ$ blocks a path.

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

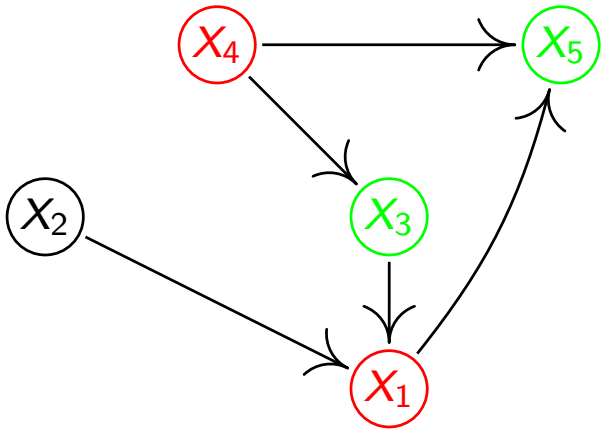
$\circ \cdots \rightarrow \circ \leftarrow \cdots \circ$ blocks a path.

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 $\mathcal{S}$ if all paths between X_i and X_j are blocked by $\mathcal{S}$.

Check, whether all paths blocked!!



$\circ \cdots \rightarrow \circ \rightarrow \cdots \circ$ blocks a path.

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

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

Hand-written notes 2