

Causality

Evgeny Burnaev

Rodrigo Rivera

Skoltech

References

The course is based on materials from

- Ilya Shpitser, course on Causal Inference, Johns Hopkins
- Kun Zhang, course on Causality and Machine Learning, Carnegie Mellon University
- Jonas Peters, course on Causlity, Univ. of Copenhagen

The main players are

- [UCLA](#): Judea Pearl
- [CMU](#): Peter Spirtes, Clark Glymour, Richard Scheines, Kun Zhang
- [Harvard University](#): Donald Rubin, Jamie Robins
- [ETH Zürich](#): Peter Bühlmann, Nicolai Meinshausen, Stefan Bauer
- [Max-Planck-Institute Tübingen](#): Dominik Janzing, Bernhard Schölkopf, Mateo Rojas-Carulla
- [University of Amsterdam](#): Joris Mooij

Two Quotes on Causality from the 1740s



We may define a cause to be an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second, ... where, if the first object had not been the second never had existed.

David Hume (1748)



Their cases were as similar as I could have them. They all in general had putrid gums, the spots and lassitude, with weakness of the knees. They lay together in one place, being a proper apartment for the sick in the forehold...

James Lind (1747)

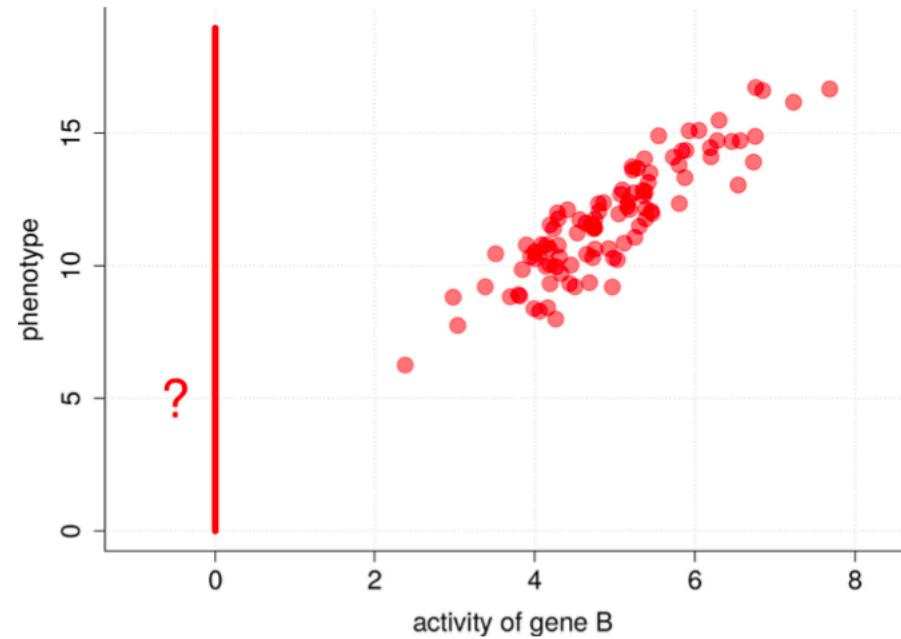
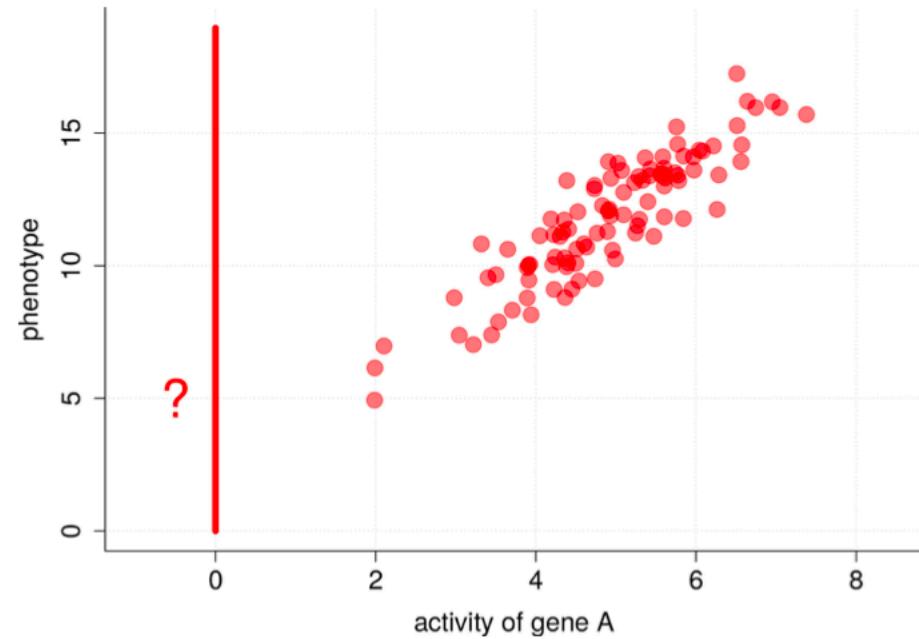
Association vs. Causation

- Most scientific inquiry/data analyses have one of two goals:
 - Association/prediction, i.e., determine predictors or variables associated with the outcome of interest.
 - Causality, i.e., understand factors that cause or have an effect on the outcome of interest.
- We are often told that association is not causation.
- Spurious correlations are very common!

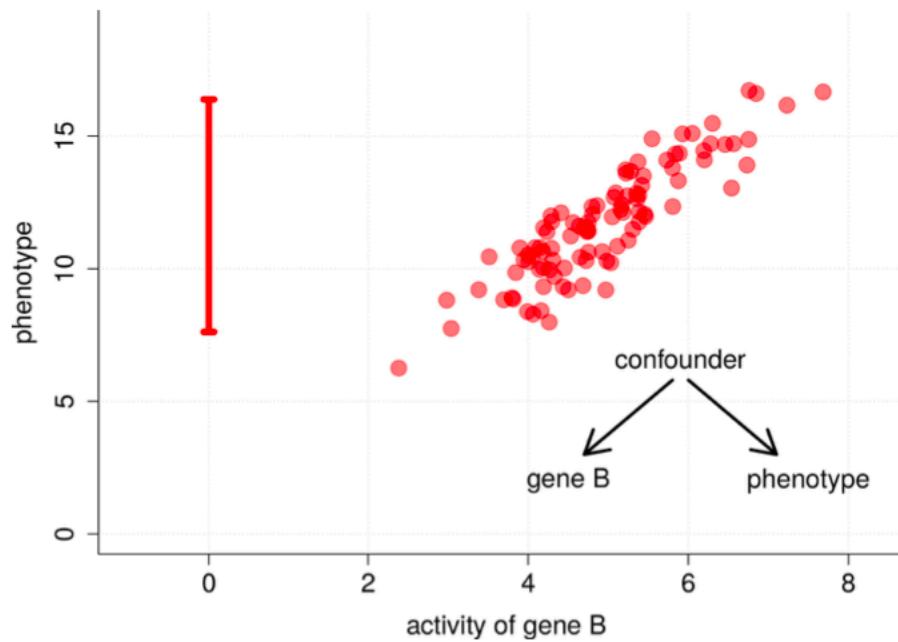
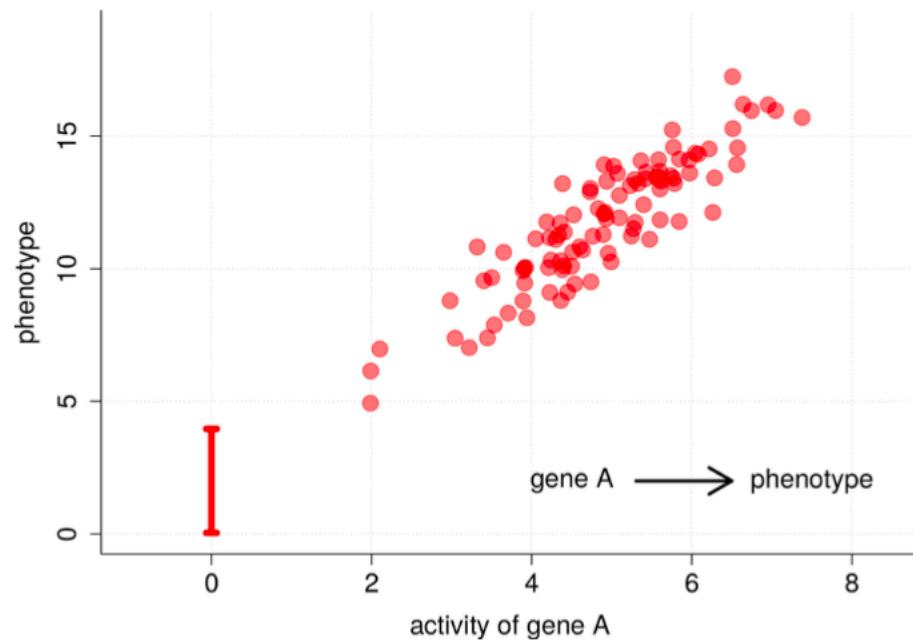
Association vs. Causation

- Most scientific inquiry/data analyses have one of two goals:
 - Association/prediction, i.e., determine predictors or variables associated with the outcome of interest.
 - Causality, i.e., understand factors that cause or have an effect on the outcome of interest.
- We are often told that association is not causation.
- Spurious correlations are very common!
 - When can we claim an association is causation?
 - This is what causal inference is about.

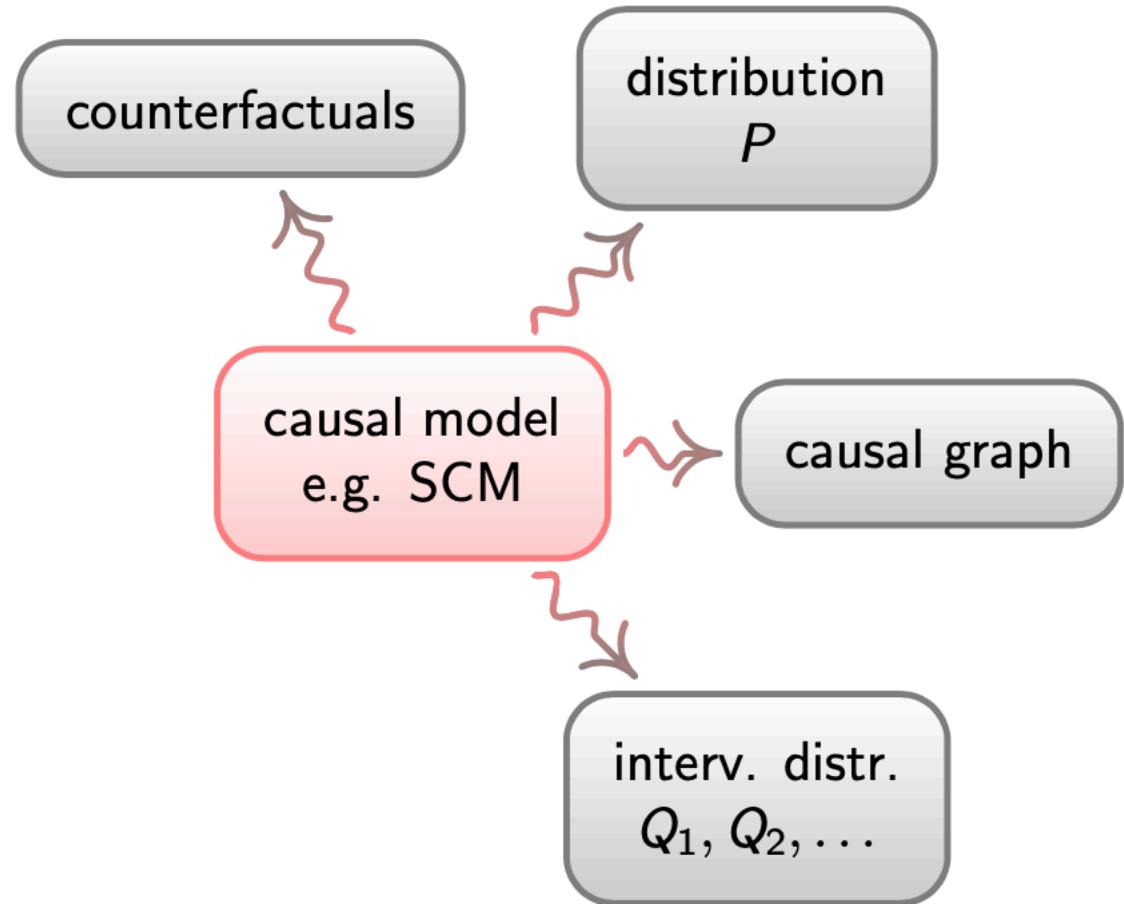
Step 1: Toy problem



Step 2: Causality matters!

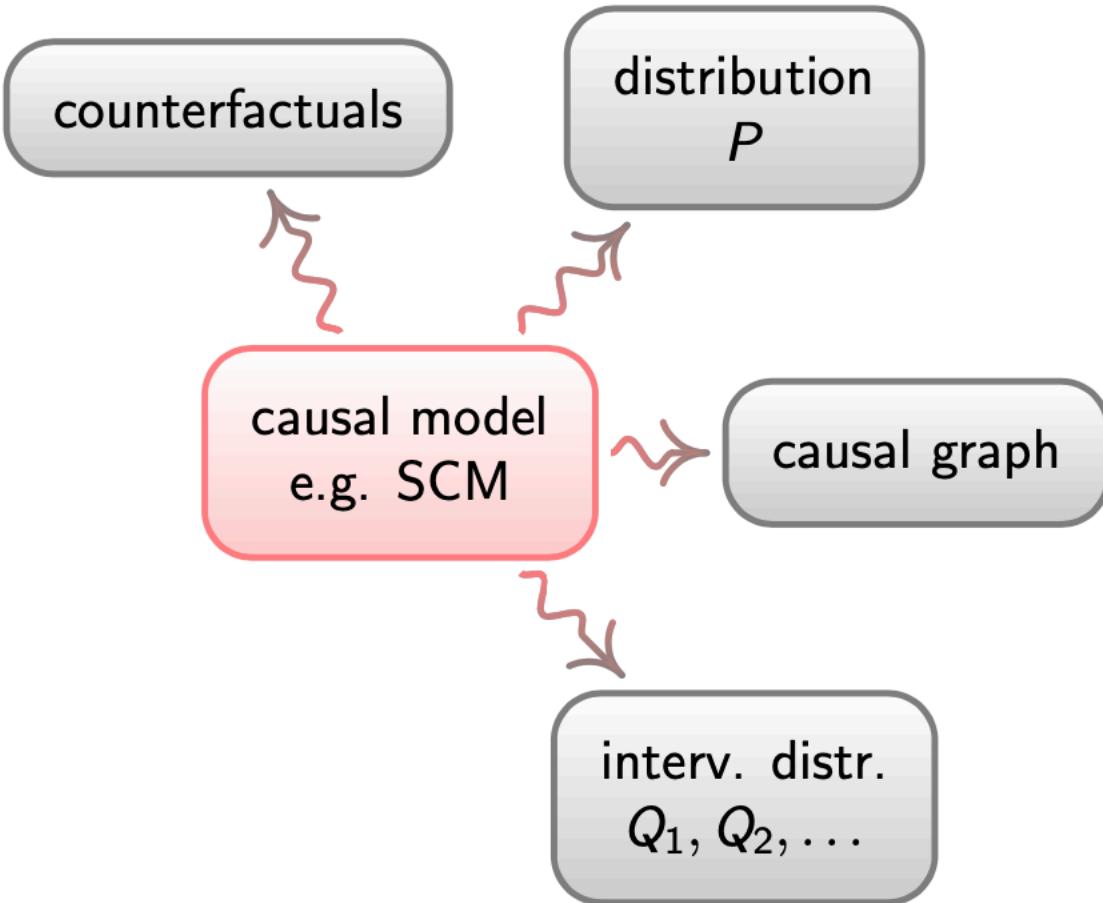


Step 3: What is a causal model?



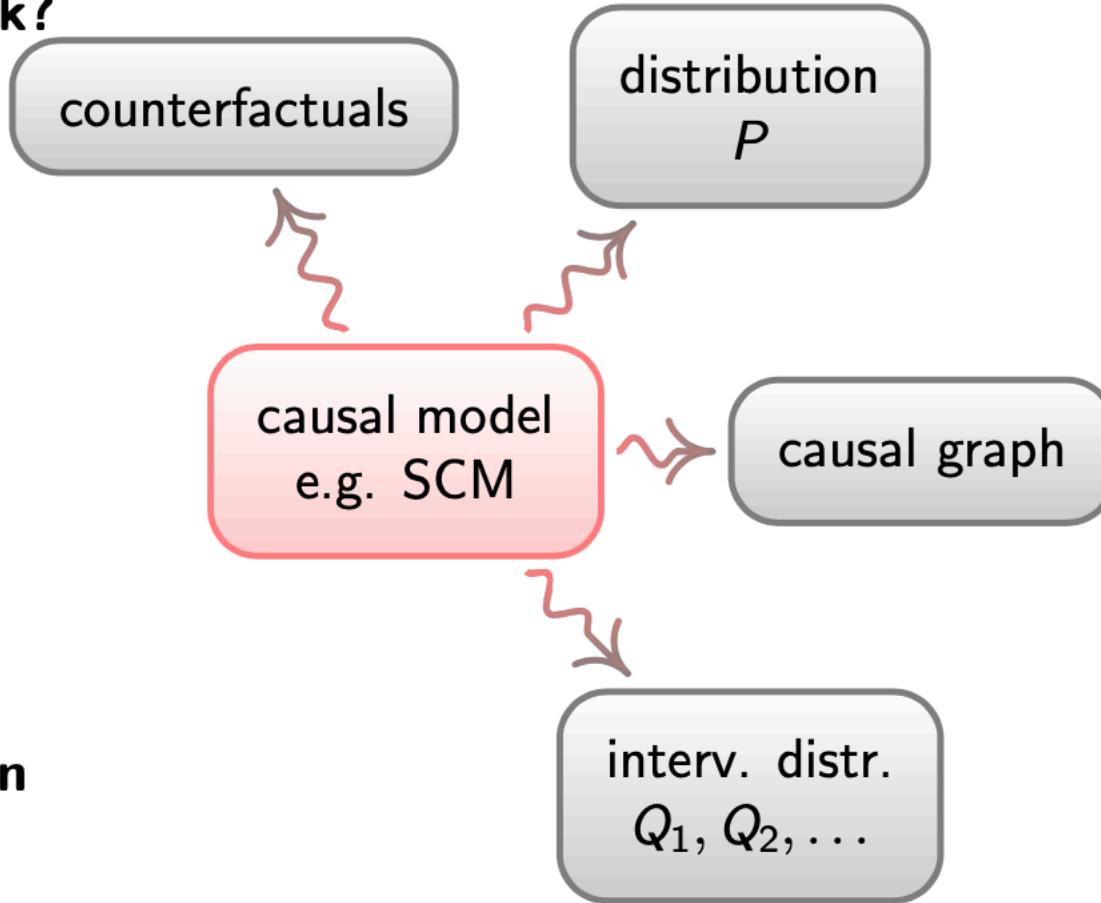
Step 4: What questions are being asked?

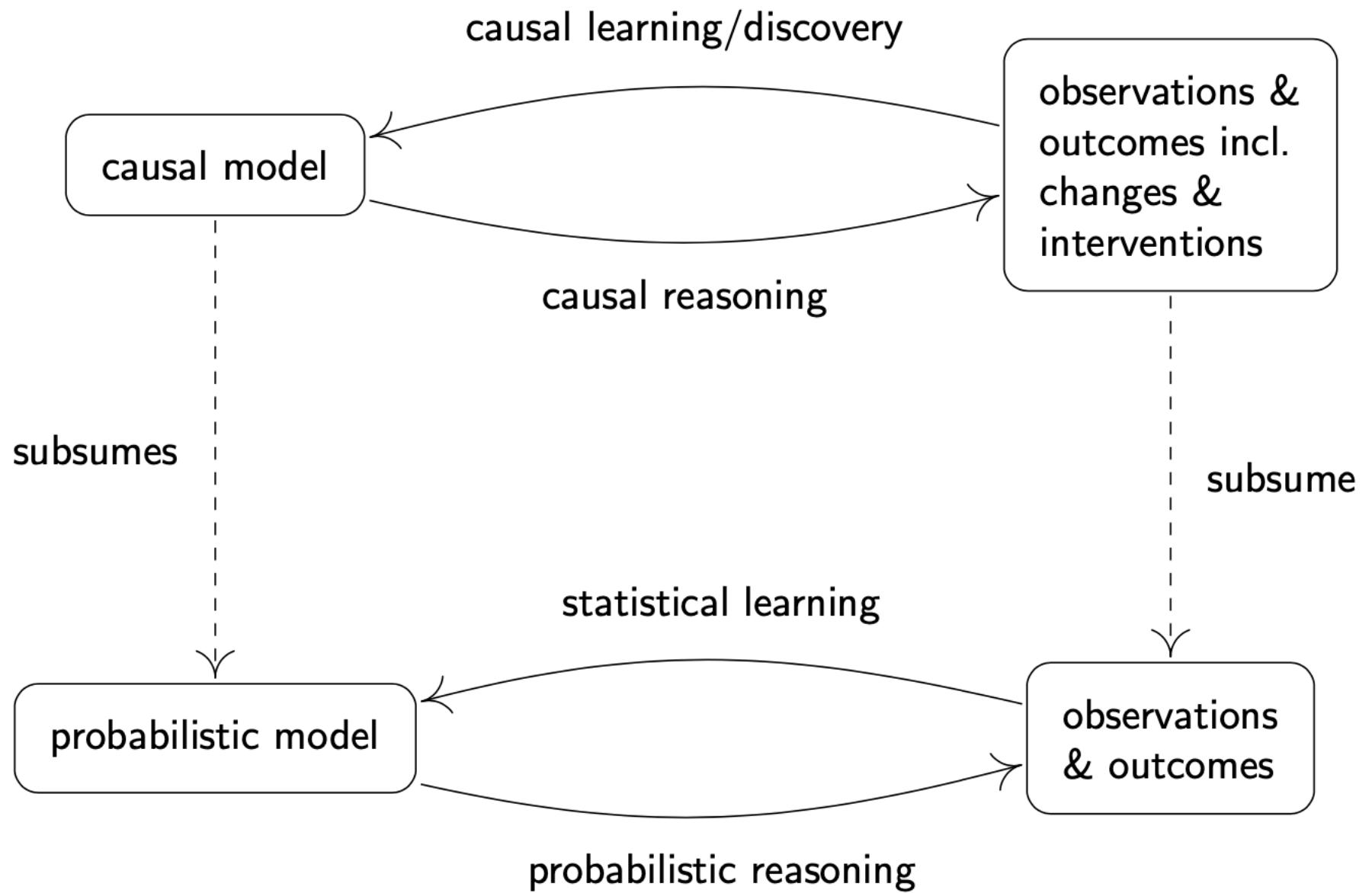
- How does the model work?
- What if there are hidden variables or feedback?
- What are nice graphical representations?
- Can we test counterfactual statements?
- Can we infer the graph structure from data?
- Is causality useful, even in classical ML/statistics settings?



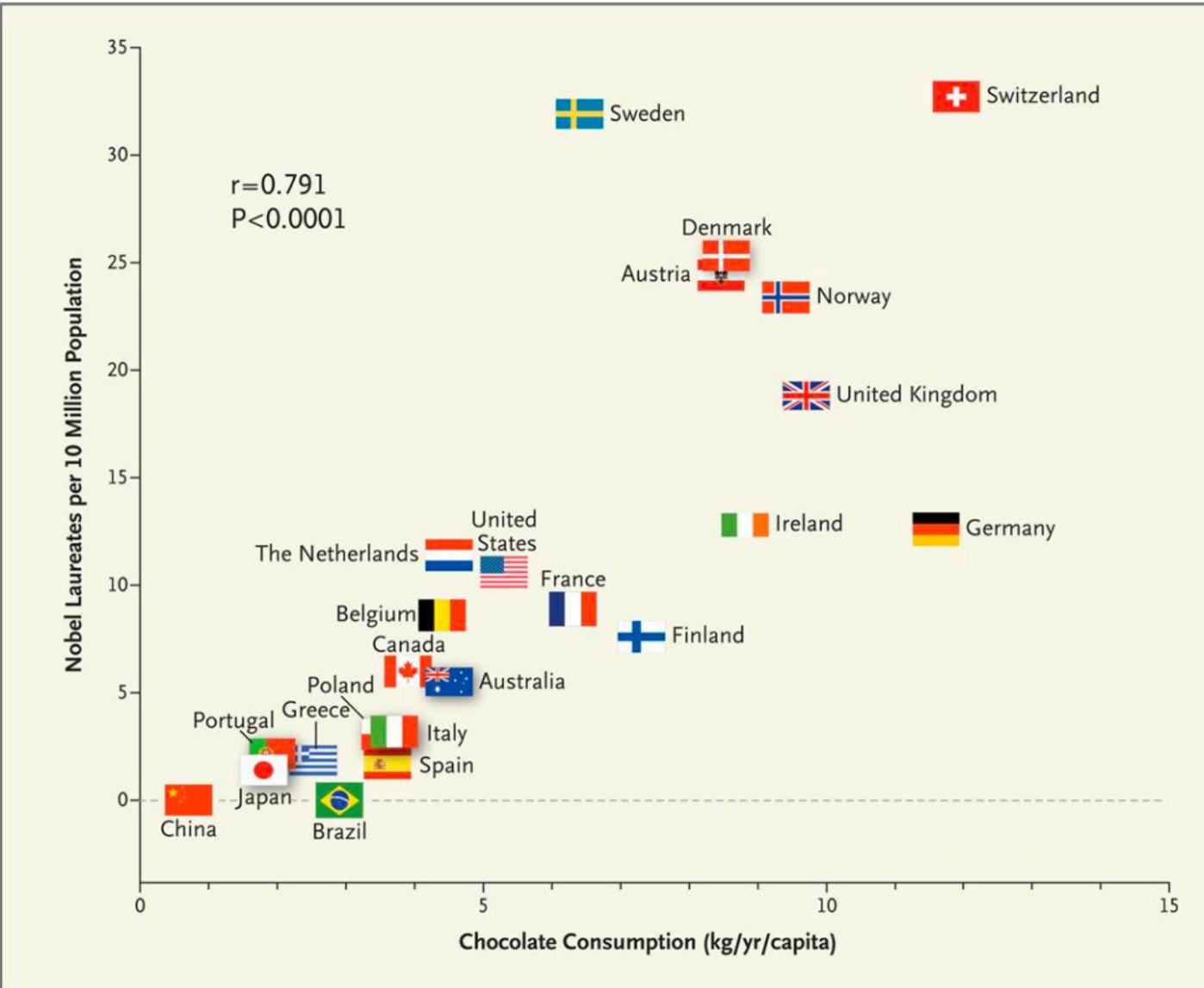
Step 4: What questions are being asked?

- **How does the model work?**
- What if there are hidden variables or feedback?
- What are nice graphical representations?
- Can we test counterfactual statements?
- **Can we infer the graph structure from data?**
- **Is causality useful, even in classical ML/statistics settings?**



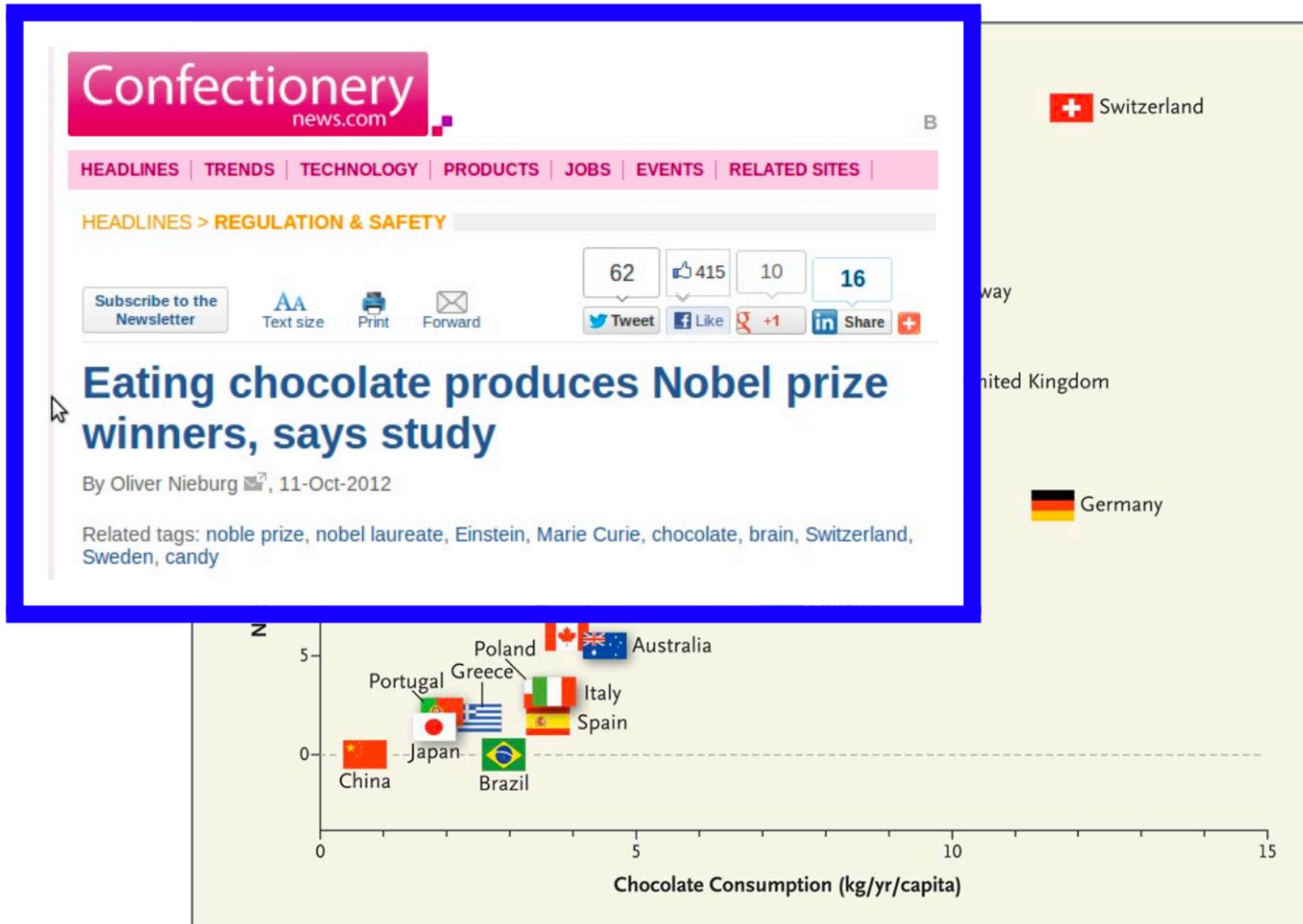


Example: chocolate



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

Example: chocolate



Example: chocolate

Confectionery

HEADLINES | T
HEADLINES >

Subscribe to the Newsletter

Eating winners

By Oliver Niebuhr

Related tags: nutrition, Sweden, candy

Forbes

New Posts +10 posts this hour

Most Popular Google's Driverless Car

List Business

I'm a medical journalist covering cardiology news.

+ Follow (73)

PHARMA & HEALTHCARE | 10/10/2012 @ 5:02PM | 14,700 views

Chocolate And Nobel Prizes In Study

4 comments, 2 called-out

+ Comment Now + Follow Comments

You don't have to be a genius to like chocolate, but geniuses are more likely to eat lots of chocolate, at least according to a new paper published in the August *New England Journal of Medicine*. Franz Messerli reports a highly



F. H. 12

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

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:

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%)	

UNG

ouncil

y Director of the Statistical

In no one would deny that it isutory. As a corollary, it is for other causes.

s of the Increase

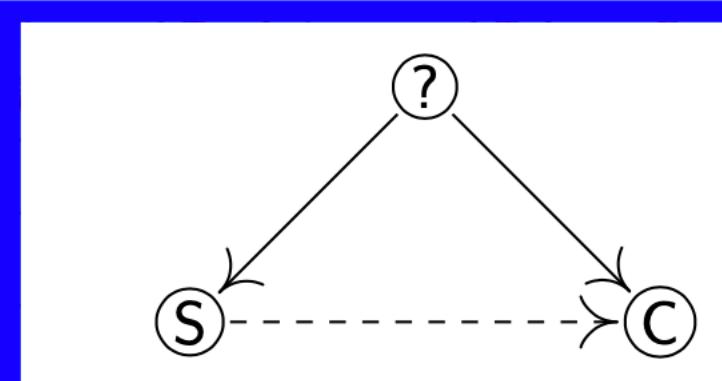
om time to time been put for-

Example: smoking

BRITISH MEDICAL JOURNAL

TABLE VII.—*E... by Smoker Diseases O...*

Disease Group	Consumed patients with				
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 cigarettes a day grouped together)

JUNG

council

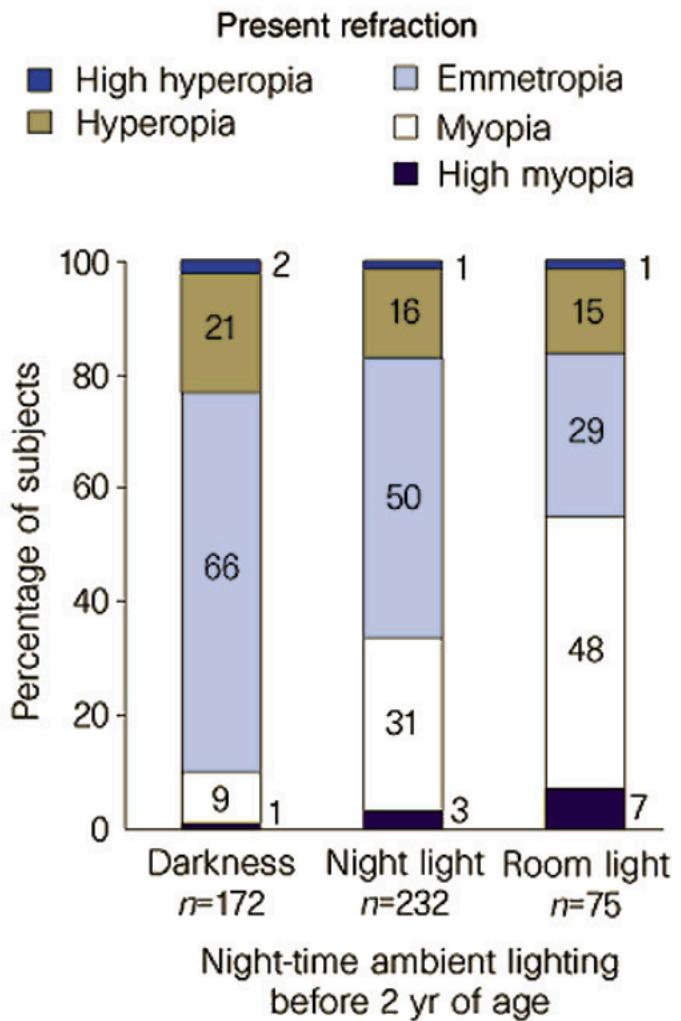
by Director of the Statistical

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

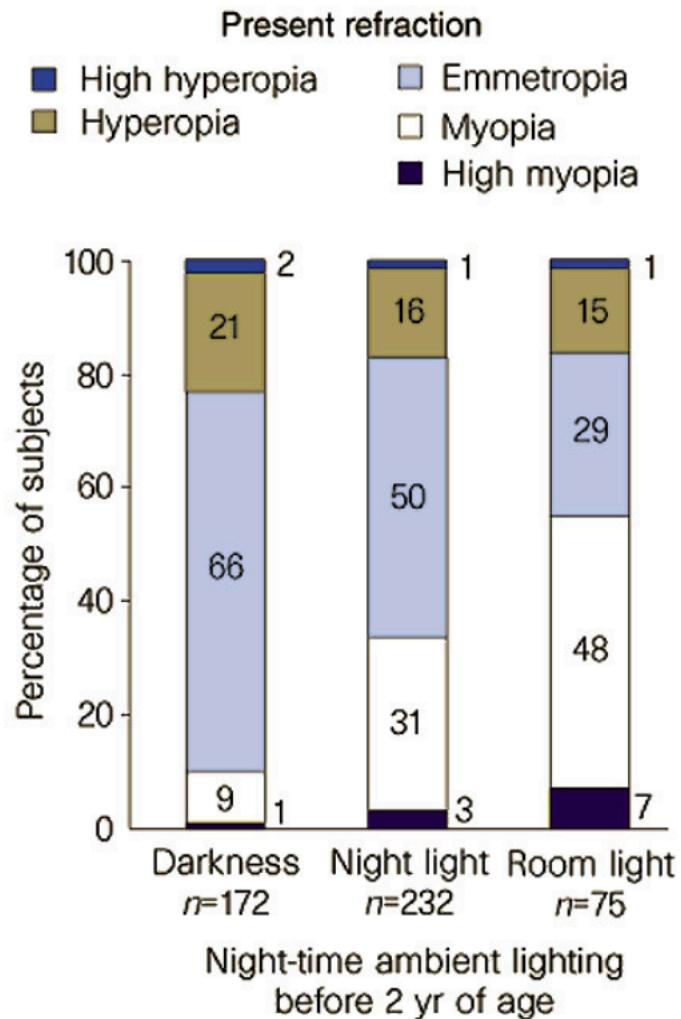
of the Increase

from 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”

Example: myopia

Patente

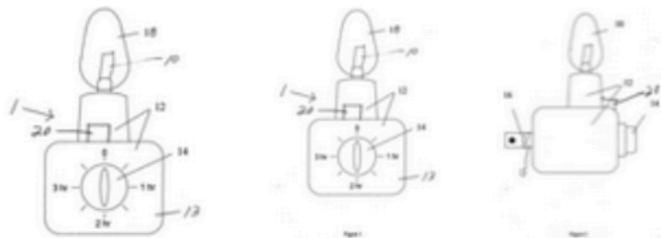
Night light with sleep timer

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

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

ANSPRÜCHE (18)

Example: myopia

Patente

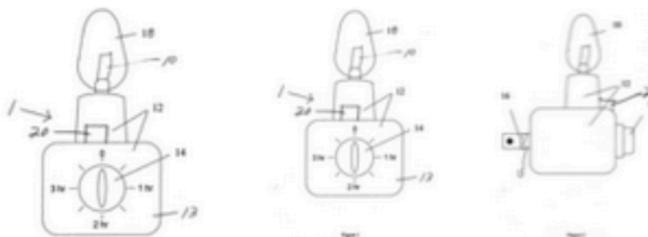
Night light with sleep timer

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

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

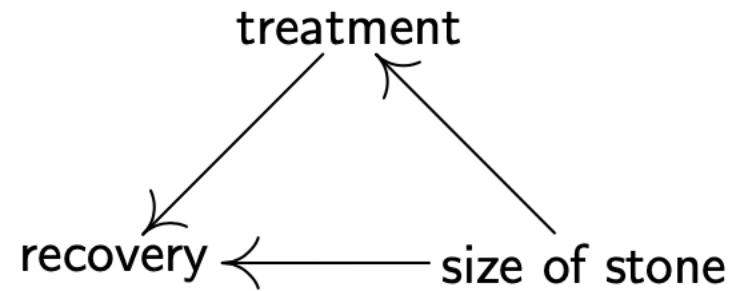
Example: kidney stones

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

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

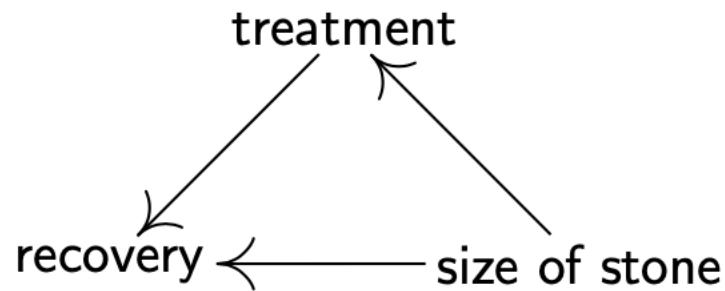
Example: kidney stones

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

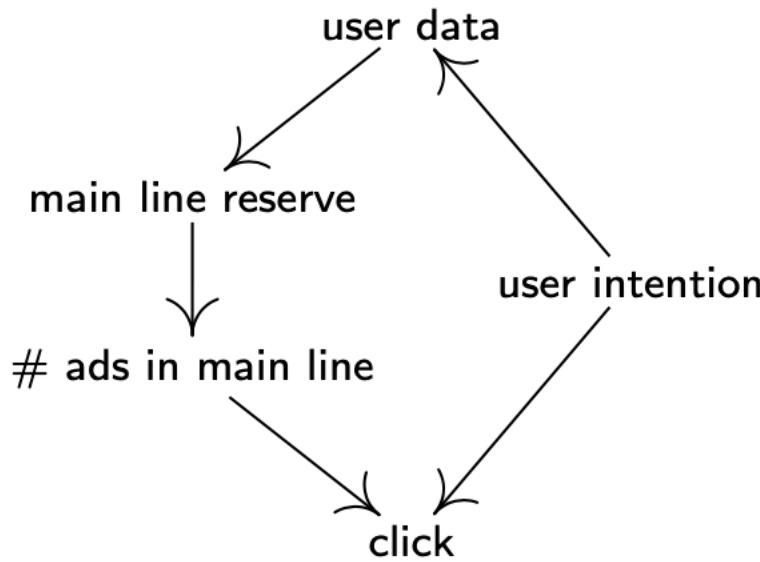
A screenshot of a Google search results page for the query "buy coffee beans". The search was performed in Chromium on a Mac OS X system. The results are displayed in a standard Google search interface with a dark header.

The search bar at the top shows the query "buy coffee beans". Below the search bar, the results are categorized by type: All, Images, Maps, Videos, More, and Search tools. A "Sign in" button is also visible in the top right corner.

The search results include:

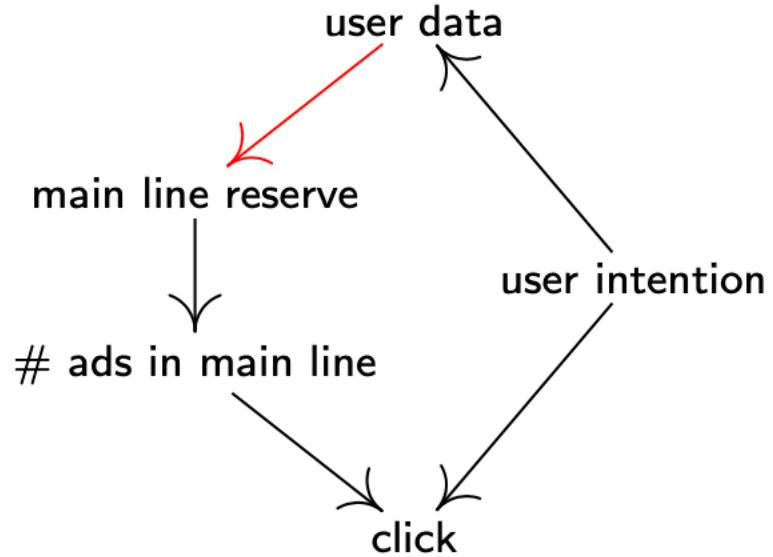
- Buy Coffee Beans Online - NextDayCoffee.co.uk**
Ad www.nextdaycoffee.co.uk/CoffeeBeans +44 1698 842528
Big Savings On Coffee Beans, Buy Now - Next Day Delivery
Coffee Beans Single Bags 100% Arabica Coffee
Coffee Pods & Capsules Caffe Roma Coffee
- Trade Commodities Online - Buy and Sell Oil,Gold,Silver,Wheat**
Ad www.plus500.dk/Commodities
Kr 185 Trading Bonus! Plus500 CFDs.
Listed on the AIM · CFD Provider · Tight Spreads · 25 € Trading Bonus · Free demo account
Fastest growing UK CFD platform – LeapRate
Gold CFDs · Oil CFDs · Silver CFDs
- Kicking Horse, 454 Horse Power, Dark, Whole Bean Coffee, 12.3 oz**
Ad www.iherb.com/
\$5 Off Your First iHerb Order! Affordable Shipping to Denmark.
100k + Product Reviews · Referral Rewards · 24/7 Customer Service
Trial Pricing Products · International Shipping · \$5 Off for New Customers
- Fair Trade Beans - Purchase Fair Trade Certified - FairTradeUSA.org**
Ad www.fairtradeusa.org/
Empower Farmers Around the World
- Buy Coffee Beans Online from Coffee Bean Shop**
<https://www.coffeebeanshop.co.uk/>
You can now buy some of the finest coffee beans from around the world. Order superb coffee blends and tea infusions from the UK coffee bean shop.
Coffees · Single Origin Coffees · Promotions · Login / Register

Example: advertisement



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

Example: advertisement

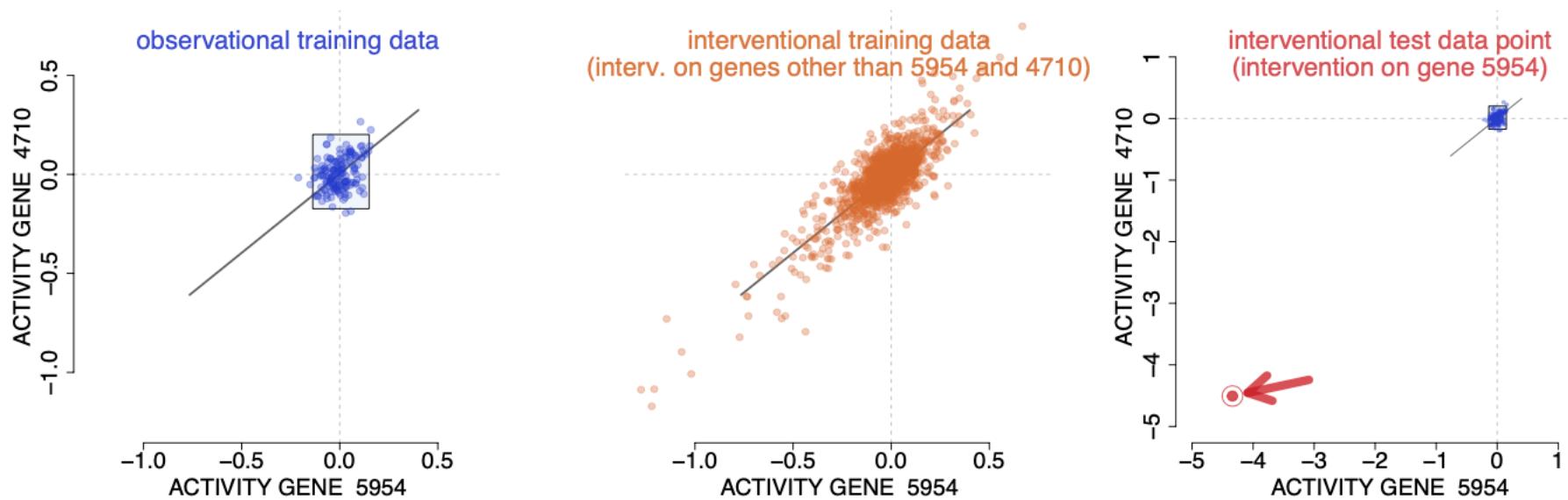


Question: How do we choose an optimal main line reserve?

Example: gene interactions

genetic perturbation experiments for yeast

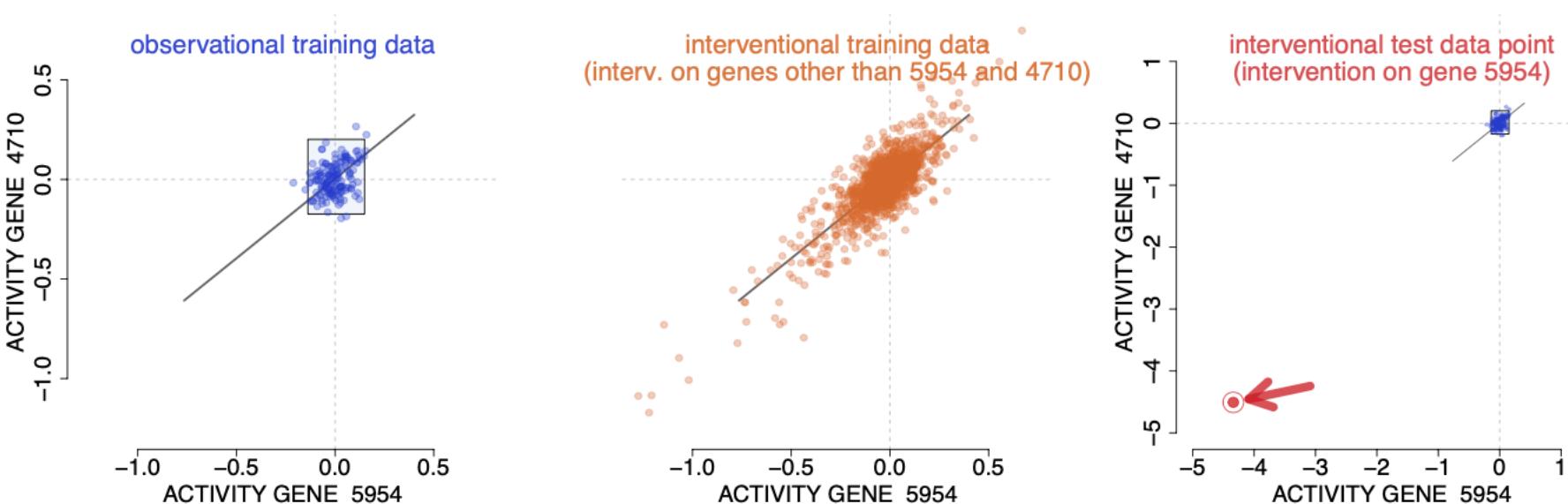
- $p = 6170$ genes
- $n_{obs} = 160$ wild-types
- $n_{int} = 1479$ gene deletions (targets known)



Example: gene interactions

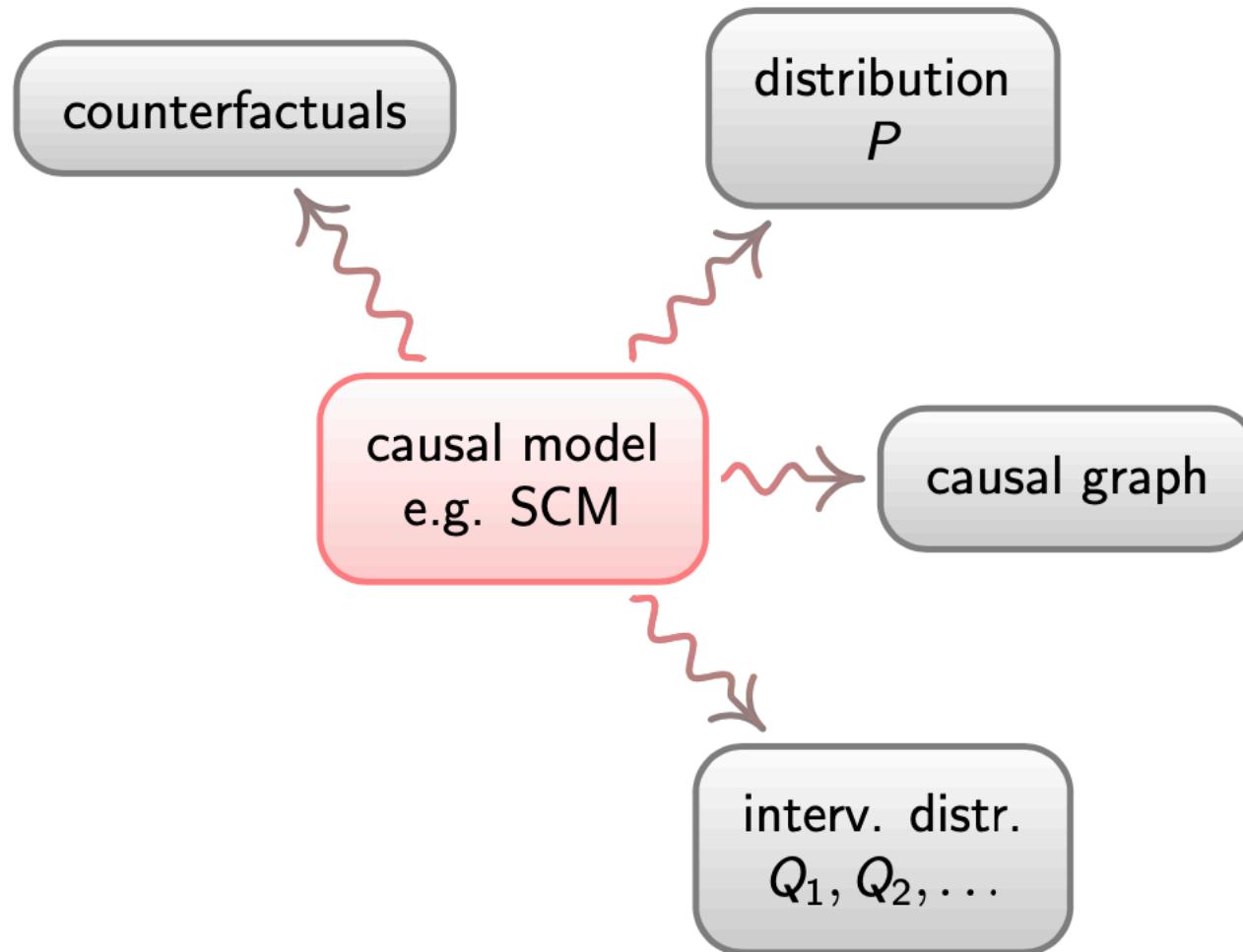
genetic perturbation experiments for yeast

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



- Causal relationships are often stable!

Causal Language and Causal Reasoning

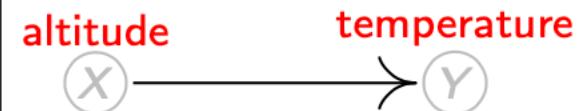


Example: Two variables

SCMs ($\mathbf{S}, P^{\mathbf{N}}$) model observational distributions.

$$X := N_X$$
$$Y := -6X + N_Y$$

$$N_X, N_Y \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$$



Example: Two variables

SCMs (S, P^N) model observational distributions.

$$Y := -6X + N_Y$$

$$N_X, N_Y \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$$



$$(X, Y) \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & -6 \\ -6 & 37 \end{pmatrix}\right)$$

Example: Two variables

SCMs ($\mathbf{S}, P^{\mathbf{N}}$) model interventions, too.

$$X := N_X \quad X := 3$$

$$Y := -6X + N_Y$$

$$N_X, N_Y \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$$



Example: Two variables

SCMs ($\mathbf{S}, P^{\mathbf{N}}$) model interventions, too.

$$X := N_X \quad X := 3$$

$$Y := -6X + N_Y$$

$$N_X, N_Y \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$$



$$P(X = 3) = 1 \quad \text{and} \quad Y \sim \mathcal{N}(-18, 1)$$

Example: Two variables

SCMs (\mathbf{S}, P^N) model interventions, too.

$$X := N_X$$

$$Y := -6X + N_Y \quad Y := \mathcal{N}(2, 2)$$

$$N_X, N_Y \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$$

altitude
X

temperature
Y

Example: Two variables

SCMs ($\mathbf{S}, P^{\mathbf{N}}$) model interventions, too.

$$X := N_X$$

$$Y := -6X + N_Y \quad Y := \mathcal{N}(2, 2)$$

$$N_X, N_Y \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$$

altitude
 X

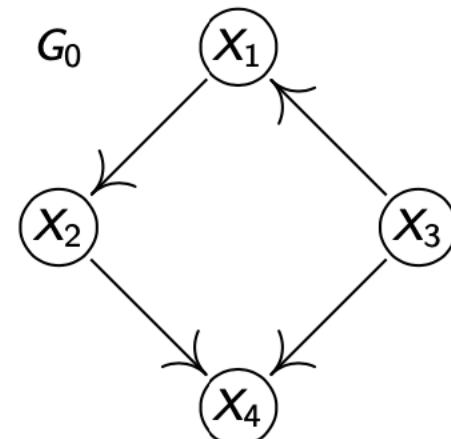
temperature
 Y

$$(X, Y) \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\right)$$

SCMs $(\mathbf{S}, P^{\mathbf{N}})$: structural equations with noise distribution.

$$\begin{aligned}X_1 &:= f_1(X_3, N_1) \\X_2 &:= f_2(X_1, N_2) \\X_3 &:= f_3(N_3) \\X_4 &:= f_4(X_2, X_3, N_4)\end{aligned}$$

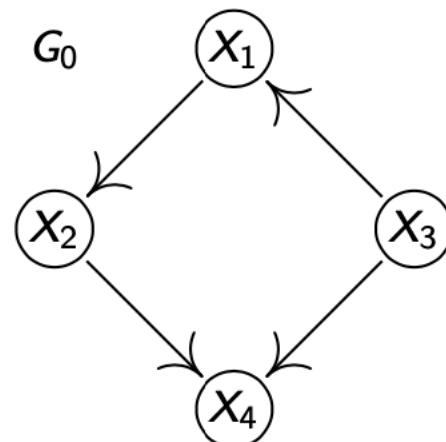
- N_i jointly independent
- G_0 has no cycles



SCMs $(\mathbf{S}, P^{\mathbf{N}})$ model **observational distributions** over X_1, \dots, X_d . Call it P .

$$\begin{aligned}X_1 &:= f_1(X_3, N_1) \\X_2 &:= f_2(X_1, N_2) \\X_3 &:= f_3(N_3) \\X_4 &:= f_4(X_2, X_3, N_4)\end{aligned}$$

- N_i jointly independent
- G_0 has no cycles



SCMs $(\mathbf{S}, P^{\mathbf{N}})$ model **interventions**, too. Call it: $P_{do(X_1:=0)}$.

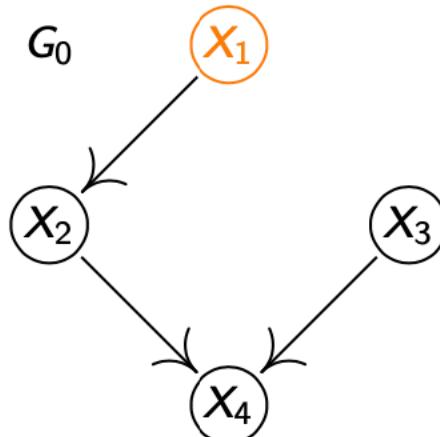
$$X_1 := 0$$

$$X_2 := f_2(X_1, N_2)$$

$$X_3 := f_3(N_3)$$

$$X_4 := f_4(X_2, X_3, N_4)$$

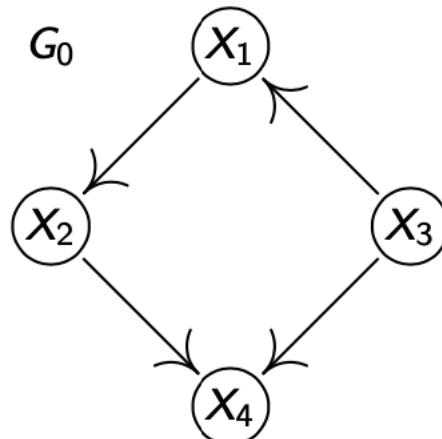
- N_i jointly independent
- G_0 has no cycles



SCMs model **observational distributions** over X_1, \dots, X_d . Call it: P .

$$\begin{aligned}X_1 &:= f_1(X_3, N_1) \\X_2 &:= f_2(X_1, N_2) \\X_3 &:= f_3(N_3) \\X_4 &:= f_4(X_2, X_3, N_4)\end{aligned}$$

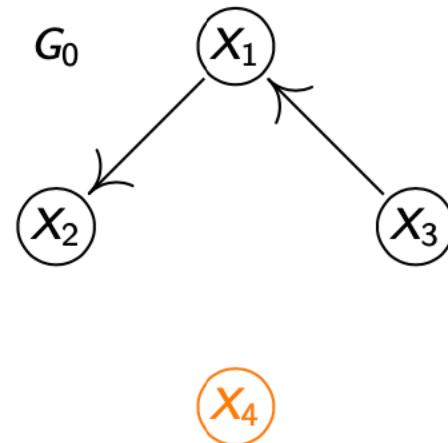
- N_i jointly independent
- G_0 has no cycles



SCMs model **interventions**, too. Call it $P_{do(X_4:=13)} \neq P(. | X_4 = 13)$

$$\begin{aligned}X_1 &:= f_1(X_3, N_1) \\X_2 &:= f_2(X_1, N_2) \\X_3 &:= f_3(N_3) \\X_4 &:= 13\end{aligned}$$

- N_i jointly independent
- G_0 has no cycles

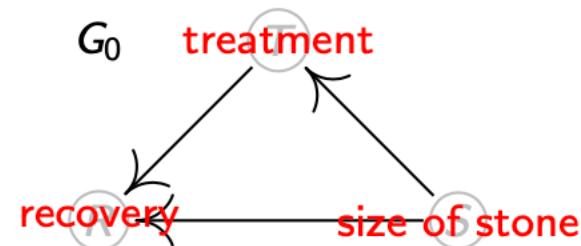


Example: kidney stones

Given: graph and P , i.e., only the structure, not the functions.

$$\begin{aligned}T &:= f_1(S, N_1) \\R &:= f_2(T, S, N_2) \\S &:= f_3(N_3)\end{aligned}$$

- N_i jointly independent
- G_0 has no cycles



Example: kidney stones

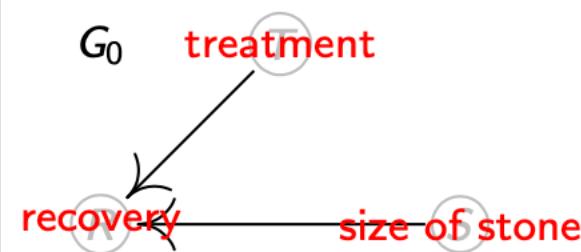
Given: graph and P . We want to compute $P_{\text{do}(T:=A)}$.

$$T := f_1(S, N_1) \quad T := A$$

$$R := f_2(T, S, N_2)$$

$$S := f_3(N_3)$$

- N_i jointly independent
- G_0 has no cycles



IMPORTANT: modularity, autonomy: Aldrich 1989, Pearl 2009, Schölkopf et al. 2012, ...

Most Useful Tautology Ever

If you intervene only on X_j , you intervene only on X_j (MUTE).

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 ($\frac{343}{700} = 0.49$)	$\frac{192}{263} = 0.73$	$\frac{55}{80} = 0.69$
	$\frac{273}{350} = 0.78$	$\frac{289}{350} = 0.83$
		$\frac{562}{700} = 0.80$

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



wanted:

$$P_{do(T:=A)}(R = 1)$$

use: $P(R | S, T) =$

$$P_{do(T:=A)}(R | S, T)$$

Example: kidney stones

$$\begin{aligned} E_{do(T:=A)} R &= P_{do(T:=A)}(R = 1) \\ &= \sum_s P_{do(T:=A)}(R = 1, S = s, T = A) \\ &= \sum_s P_{do(T:=A)}(R = 1 \mid S = s, T = A)P_{do(T:=A)}(S = s, T = A) \\ &= \sum_s P_{do(T:=A)}(R = 1 \mid S = s, T = A)P_{do(T:=A)}(S = s) \\ &= \sum_s P(R = 1 \mid S = s, T = A)P(S = s) \\ &= 0.832 \\ &> 0.782 \\ &= \dots \\ &= P_{do(T:=B)}(R = 1) = E_{do(T:=B)} R \end{aligned}$$

Example: kidney stones

$$\begin{aligned} E_{do(T:=A)} R &= P_{do(T:=A)}(R = 1) \\ &= \sum_s P_{do(T:=A)}(R = 1, S = s, T = A) \\ &= \sum_s P_{do(T:=A)}(R = 1 \mid S = s, T = A)P_{do(T:=A)}(S = s, T = A) \\ &= \sum_s P_{do(T:=A)}(R = 1 \mid S = s, T = A)P_{do(T:=A)}(S = s) \\ &= \sum_s P(R = 1 \mid S = s, T = A)P(S = s) \\ &= 0.832 \quad \neq P(R = 1 \mid T = A) \\ &> 0.782 \\ &= \dots \\ &= P_{do(T:=B)}(R = 1) = E_{do(T:=B)} R \end{aligned}$$

This idea holds more generally.

Definition

Given an SCM over (X, Y, \mathbf{W}) . We call $\mathbf{Z} \subseteq \mathbf{W}$ a valid adjustment set for (X, Y) if

$$p_{do(X:=x)}(y) = \sum_{\mathbf{z}} p(y|x, \mathbf{z})p(\mathbf{z}) \neq p(y|x)$$

This idea holds more generally.

Definition

Given an SCM over (X, Y, \mathbf{W}) . We call $\mathbf{Z} \subseteq \mathbf{W}$ a valid adjustment set for (X, Y) if

$$p_{do(X:=x)}(y) = \sum_{\mathbf{z}} p(y|x, \mathbf{z})p(\mathbf{z}) \neq p(y|x)$$

Proposition (Parent Adjustment)

Assume $Y \notin PA(X)$. Then

$PA(X)$ is a valid adjustment set for (X, Y) .

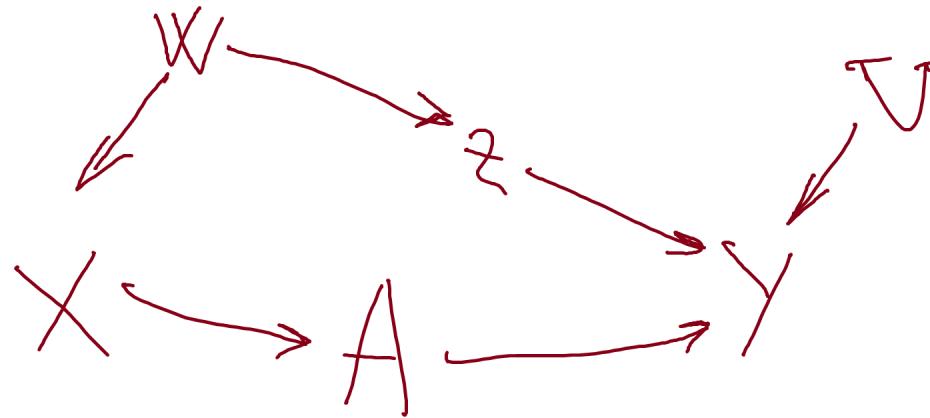
Adjusting in Linear Gaussian Models

Measure of variable influence

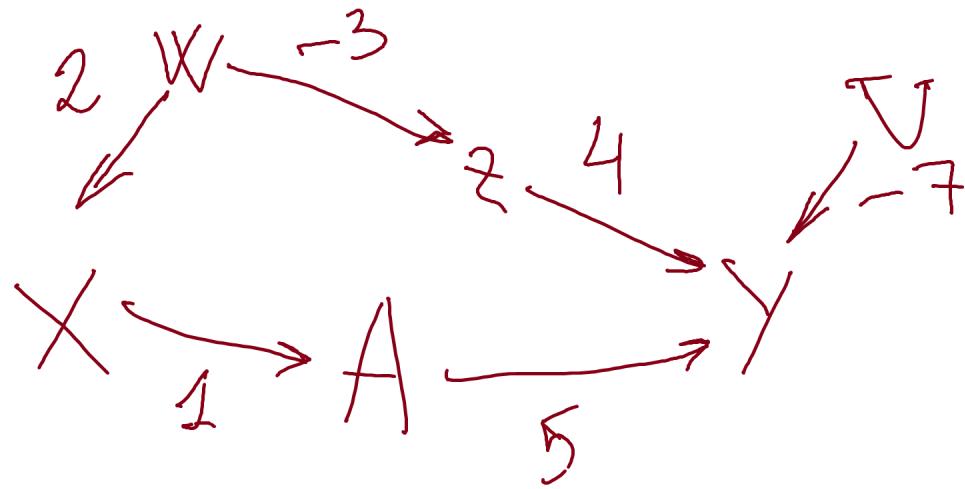
Compare $p(y)$ and $P_{do}(X=x)(y)$

$$C_{X \rightarrow Y} \approx \frac{\partial}{\partial X} \underset{do X=x}{\mathbb{E}} [Y]$$

Linear models

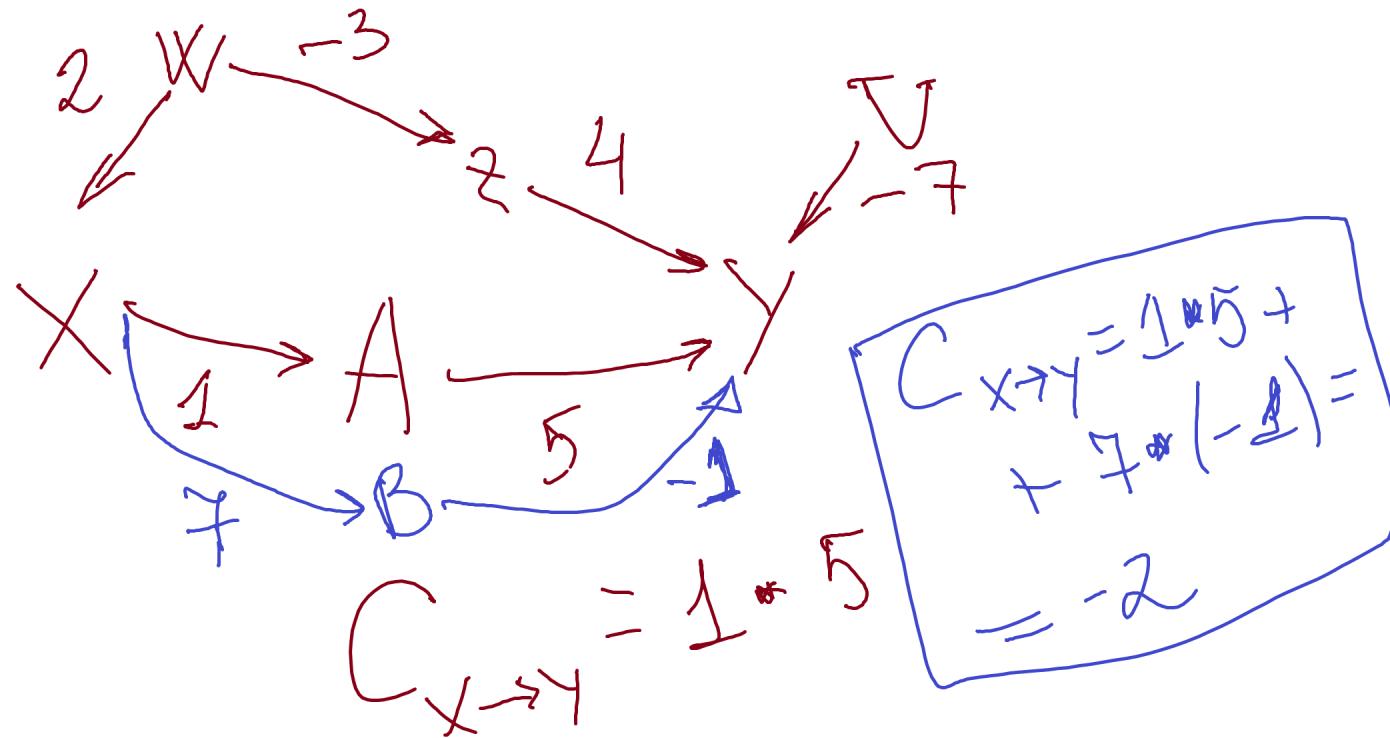


Linear models



$$C_{X \rightarrow Y} = 1 * 5$$

Linear models



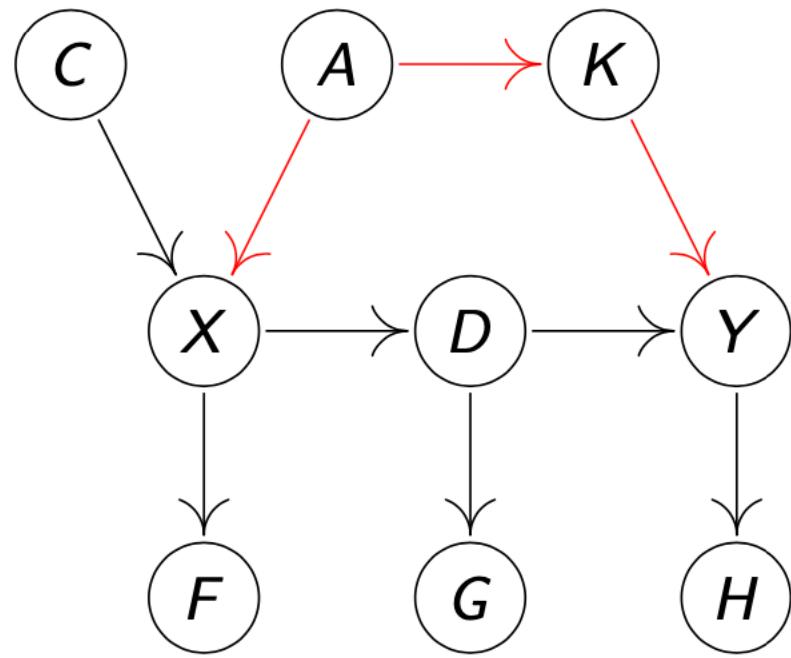
Adjustment and Linear models

If Z is a valid adjustment set

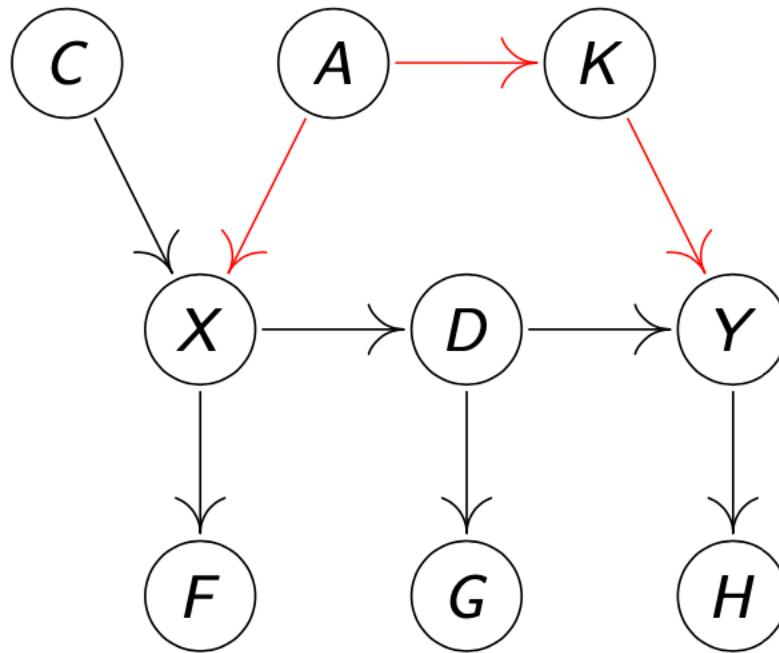
$C_{X \rightarrow Y}$ = regn. coefficient for X
in a linear model.

$Y \sim aX + bZ$, i.e.

$$C_{X \rightarrow Y} = a$$



$X \leftarrow A \rightarrow K \rightarrow Y$ is a “backdoor path” from X to Y .



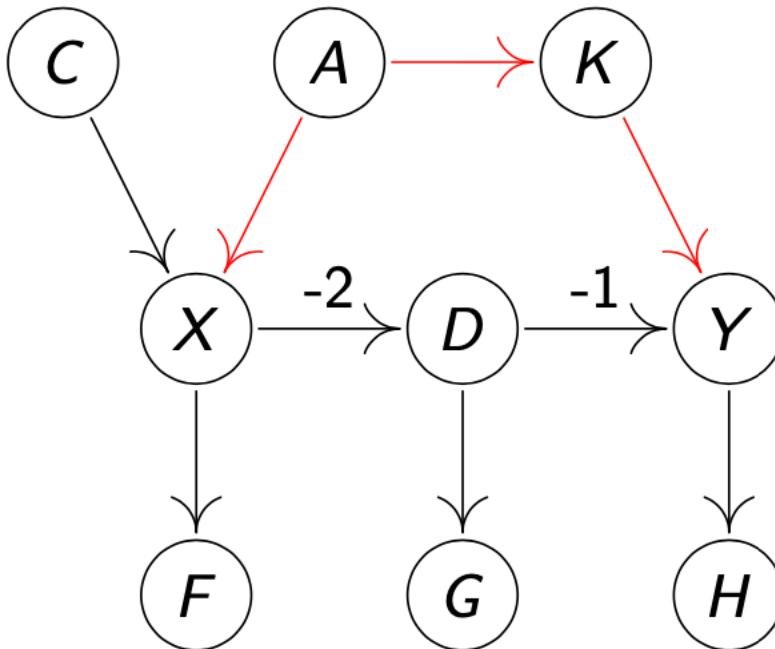
$X \leftarrow A \rightarrow K \rightarrow Y$ is a “backdoor path” from X to Y .

$$\mathbf{Z} = \{C, A\},$$

$$\mathbf{Z} = \{K\},$$

$$\mathbf{Z} = \{F, C, K\}$$

are valid adjustment sets for (X, Y) (no proof).



$X \leftarrow A \rightarrow K \rightarrow Y$ is a “backdoor path” from X to Y .

$$\mathbf{Z} = \{C, A\},$$

$$\mathbf{Z} = \{K\},$$

$$\mathbf{Z} = \{F, C, K\}$$

are valid adjustment sets for (X, Y) (no proof).

```
n <- 500

# generate a sample from the distr. ent. by the SCM
set.seed(1)
C <- rnorm(n)
A <- 0.8*rnorm(n)
K <- A + 0.1*rnorm(n)
X <- C - 2*A + 0.2*rnorm(n)
F <- 3*X + 0.8*rnorm(n)
D <- -2*X + 0.5*rnorm(n)
G <- D + 0.5*rnorm(n)
Y <- 2*K - D + 0.2*rnorm(n)
H <- 0.5*Y + 0.1*rnorm(n)

lm(Y~X)$coefficients
lm(Y~X+K)$coefficients
lm(Y~X+F+C+K)$coefficients
lm(Y~X+F+C+K+H)$coefficients
```

```

n <- 500

# generate a sample from the distr. ent. by the SCM
set.seed(1)
C <- rnorm(n)
A <- 0.8*rnorm(n)
K <- A + 0.1*rnorm(n)
X <- C - 2*A + 0.2*rnorm(n)
F <- 3*X + 0.8*rnorm(n)
D <- -2*X + 0.5*rnorm(n)
G <- D + 0.5*rnorm(n)
Y <- 2*K - D + 0.2*rnorm(n)
H <- 0.5*Y + 0.1*rnorm(n)

lm(Y~X)$coefficients
lm(Y~X+K)$coefficients
lm(Y~X+F+C+K)$coefficients
lm(Y~X+F+C+K+H)$coefficients

```

$$Y \sim X \Rightarrow C_{X \rightarrow Y} = 1.30$$

$$Y \sim X + K \Rightarrow C_{X \rightarrow Y} \approx 2.01$$

$$Y \sim X + F + C + K \Rightarrow C_{X \rightarrow Y} \approx 1.98$$

$$Y \sim X + F + C + K + H \Rightarrow C_{X \rightarrow Y} \approx 0.11$$

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

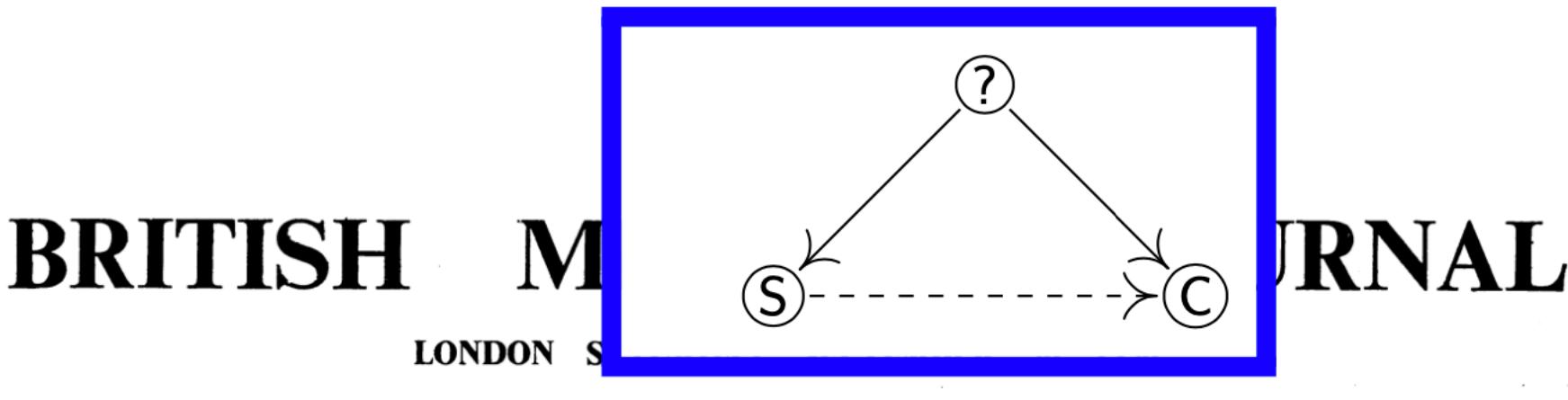
In England and Wales the phenomenal increase in the number of deaths attributed to cancer of the lung provides one of the most striking changes in the pattern of mortality recorded by the Registrar-General. For example, in the quarter of a century between 1922 and 1947 the annual number of deaths recorded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

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:

Example: smoking



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

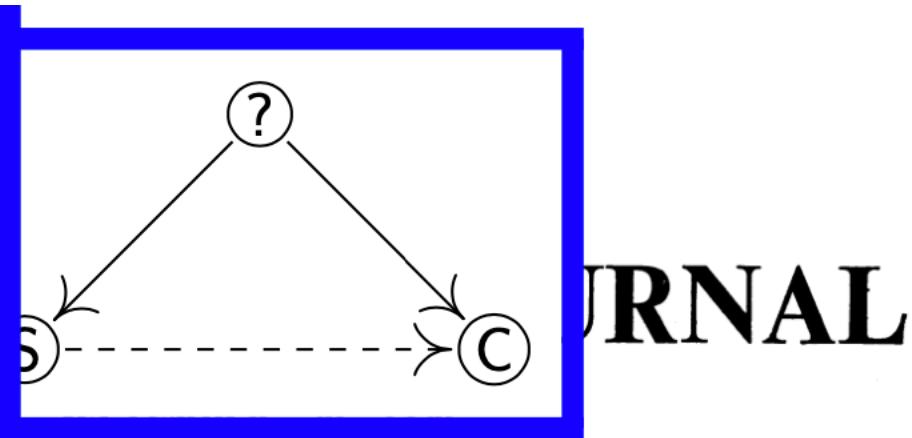
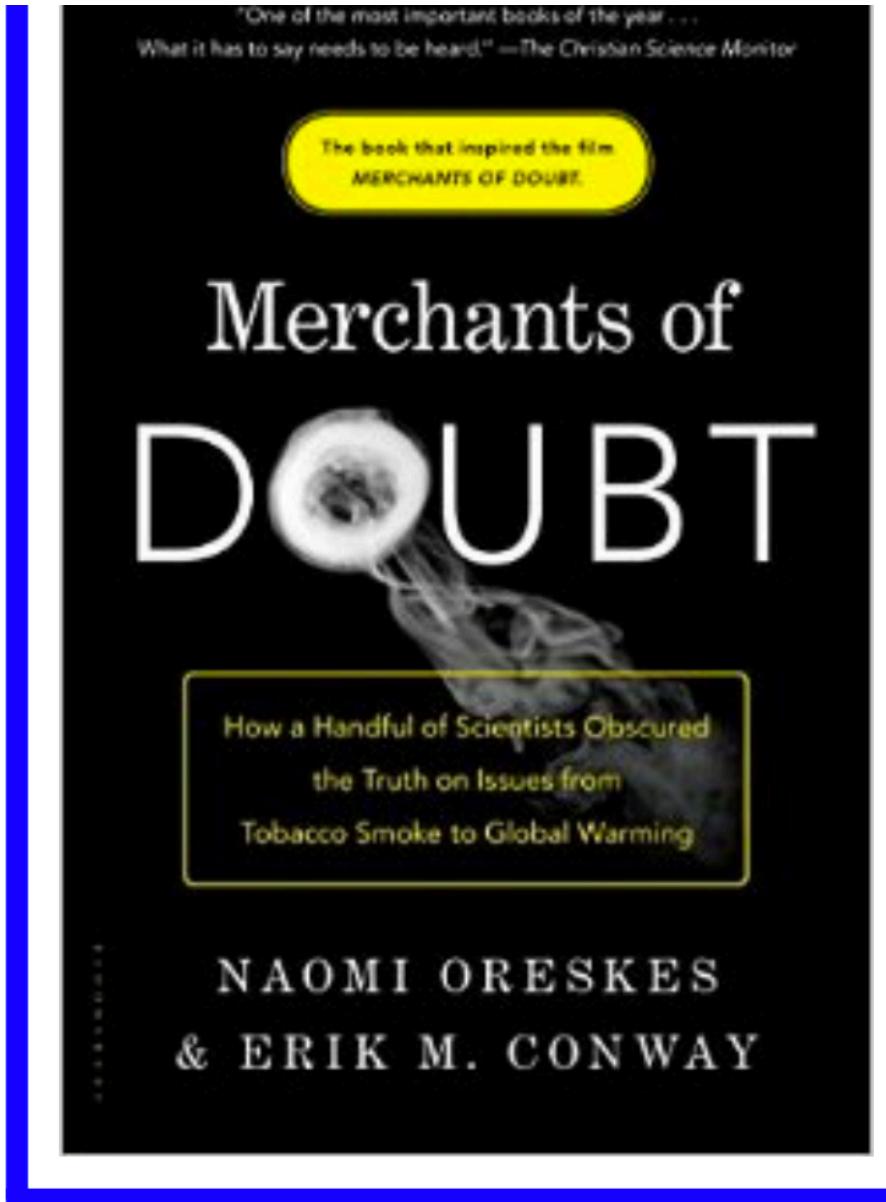
In England and Wales the phenomenal increase in the number of deaths attributed to cancer of the lung provides one of the most striking changes in the pattern of mortality recorded by the Registrar-General. For example, in the quarter of a century between 1922 and 1947 the annual number of deaths recorded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

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:

Example: smoking



JRNAL

INOMA OF THE LUNG ARY REPORT

BY

L, M.D., M.R.C.P.

Unit of the Medical Research Council

ND

HILL, Ph.D., D.Sc.

*nd Tropical Medicine; Honorary Director of the Statistical
Medical Research Council*

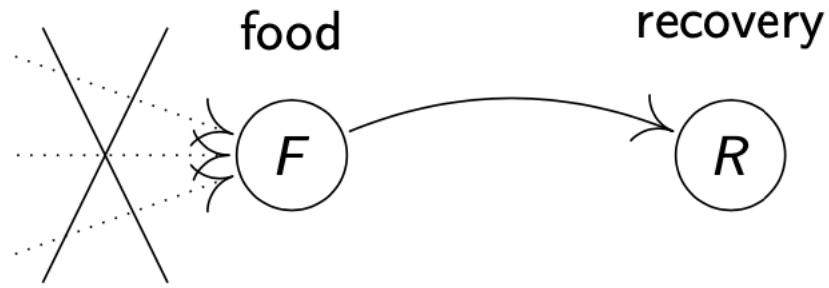
whole explanation, although no one would deny that it may well have been contributory. As a corollary, it is right and proper to seek for other causes.

Possible Causes of the Increase

Two main causes have from time to time been put forward:

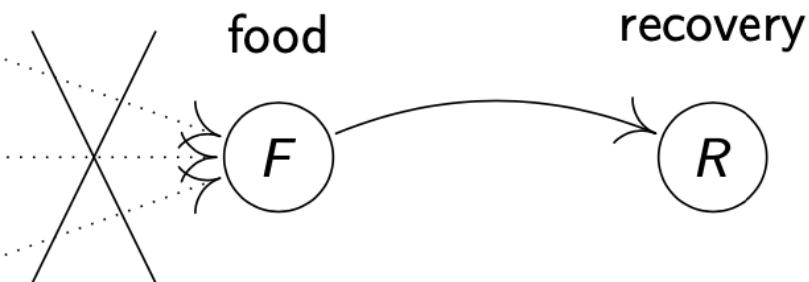
James Lind (1716–94):

James Lind (1716–94): Randomized Experiment



Then: $P_{do(F:=f)}(R) = P(R|F = f)$

James Lind (1716–94): Randomized Experiment



$$\text{Then: } P_{do(F:=f)}(R) = P(R|F=f)$$

“On the 20th of May 1747, I selected twelve patients in the scurvy, on board the Salisbury at sea. [...] Two were ordered each a quart of cyder a day. Two others took twenty-five drops of elixir vitriol three times a day [...] Two others took two spoonfuls of vinegar three times a day [...] Two of the worst patients were put on a course of sea-water [...] Two others had each two oranges and one lemon given them every day [...] The two remaining patients, took [...] an electuary recommended by a hospital surgeon [...] The consequence was, that the most sudden and visible good effects were perceived from the use of oranges and lemons; one of those who had taken them, being at the end of six days fit for duty.”

Definition (Equivalence of causal models)

Two models are called

{probabilistically / interventionally} equivalent

if they entail the same

{observational / observational & interventional}

distributions. Here, it suffices to consider interventions that set a variable X_j to a fully supported \tilde{N}_j (“randomized experiments”).