Causal Inference: prediction, explanation, and intervention

Lecture 3: Probabilistic causality

Samantha Kleinberg
samantha.kleinberg@stevens.edu

Why probabilistic causality?

- Non-deterministic relationships
  - Radioactive decay
- Incomplete knowledge
  - Even if food poisoning were deterministic, we still would not know all the conditions needed to bring it about
- Capture *how much* of a difference cause makes to effect
  - Difference between oxygen and lit match for house fire

Quick probability review

- Sample space: set of all possible outcomes
- Event: subset of the sample space

- Flipping a coin twice in row
  - $\Omega = \{HH, HT, TH, TT\}$
  - Event of getting exactly one tail {HT, TH}

Probability function

- For A being subsets of Ω
  - $P(A) \in [0,1]$, for A in set of events
  - $P(\Omega)=1$
  - For disjoint events A,B: $P(A \vee B) = P(A) + P(B)$

- $\Omega = \{HH,HT,TH,TT\}$
- P any combination of heads/tails = 1
- $P(HH \vee TT) = 1/2$

Probabilities from data

How to determine probability of heads/tails from coin flip?

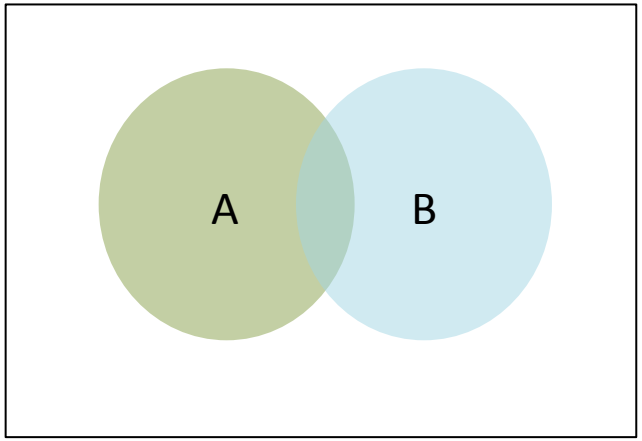
- Frequency of event x
  - #(x) = number of times it is observed
- Relative frequency
  - #(x)/N, where N is number of observations
- As N→∞, relative frequency →P(x)

Probabilities from data - challenges

- Sampling bias
  - E.g. survey FOX news or MSNBC viewers about president
- Confirmation bias
  - Ignoring some events
  - Coding events differently

Disjunction

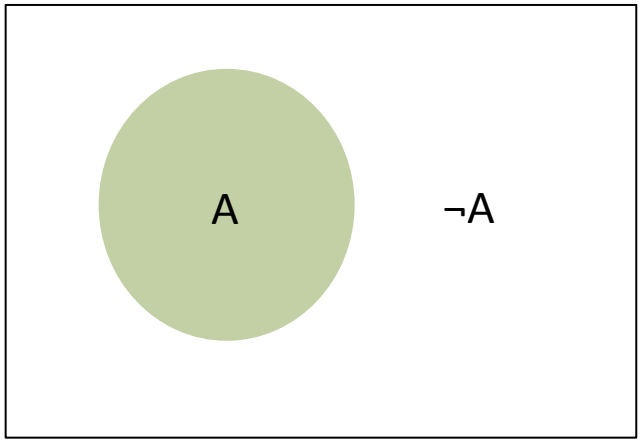
- Addition rule
  - $P(A \vee B)=P(A)+P(B)-P(A \wedge B)$
- Mutually exclusive events
  - $P(A \vee B)=P(A)+P(B)$
  - Flipping $H \vee T$ versus increasing unemployment or decreasing interest rates



Complement rule

$1=P(A)+P(\neg A)$

$P(\neg A)=1-P(A)$



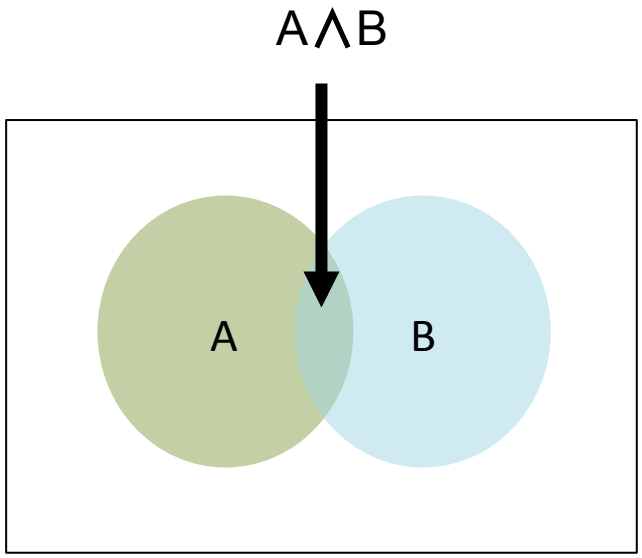
Dependence/independence

- Events that are independent
  - Flipping heads and then tails
  - Day of week and whether a patient had a heart attack (probably?)
- Events that are dependent
  - Vice presidential candidate and presidential nominee
  - Diagnostic test being positive and whether patient has a disease

Conjunction

- For independent events

  $P(A \wedge B)=P(A)P(B)$



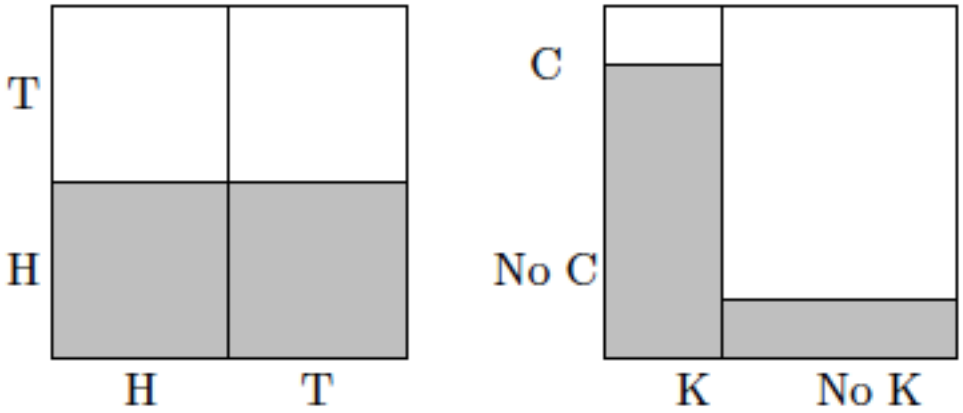
Conditional probability

P(A⋀B)=P(A|B)P(B), where P(B)>0

Note, if A and B independent, P(A⋀B)=P(A)P(B)

Then:

$$P(A|B) = \frac{P(A)P(B)}{P(B)} = P(A)$$



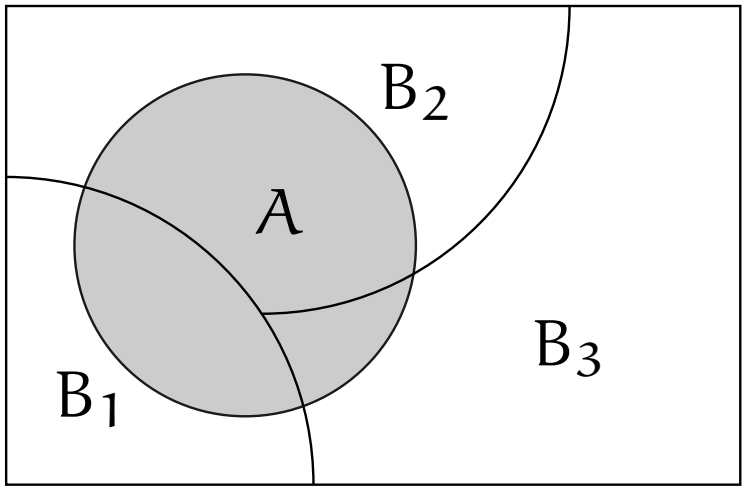
(a) (b)

Law of total probability

$$P(A) = P(A \wedge B_1) + P(A \wedge B_2) + \ldots + P(A \wedge B_3)$$

$$= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots + P(A|B_n)P(B_n)$$

$$= \sum_{i=1}^{n} P(A|B_i)P(B_i)$$



Also called marginal probability,
and this process is marginalization

$$P(A|B) = \frac{P(A \wedge B)}{P(B)}$$

$$P(A \wedge B) = P(A|B)P(B) = P(B|A)P(A)$$

Bayes rule $$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Example of Bayes rule

- Probability of disease (d) is 0.001
- But diagnostic test not foolproof
  - $P(+|d) = 0.98$
  - $P(+|\neg d) = 0.1$

Given a positive test, what's probability patient has disease?

Example continued

$$P(d \mid +) = \frac{P(+ \mid d)P(d)}{p(+)}$$

$$P(\neg d) = 1 - P(d)$$

$$= \frac{P(+ \mid d)P(d)}{P(+ \mid d)P(d) + P(+ \mid \neg d)P(\neg d)}$$

$$= \frac{0.98 \times 0.001}{0.98 \times 0.001 + 0.1 \times (1 - 0.001)}$$

$$= \frac{0.00098}{0.10088} \approx 0.01$$

Why is understanding conditional probability important?

Sally Clark case

- Mother whose first son died suddenly in 1996
- Second son died similarly in 1998
- SIDs or murder?

Case focused on statistic

- Probability of death from SIDs = 1/8543
- Probability of two deaths from SIDs = 1/8543*1/8543 ≈ 1/73 million!

See: www.statslab.cam.ac.uk/~apd/SallyClark_report.doc

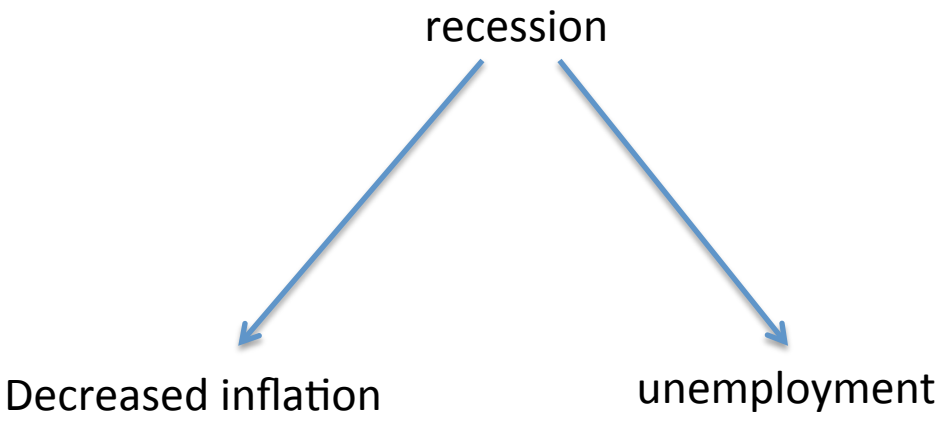
Probabilistic causality: basic idea

C causes E if

$P(E|C)>P(E)$

Or $P(E|C)>P(E|\neg C)$

Probability raising and common causes



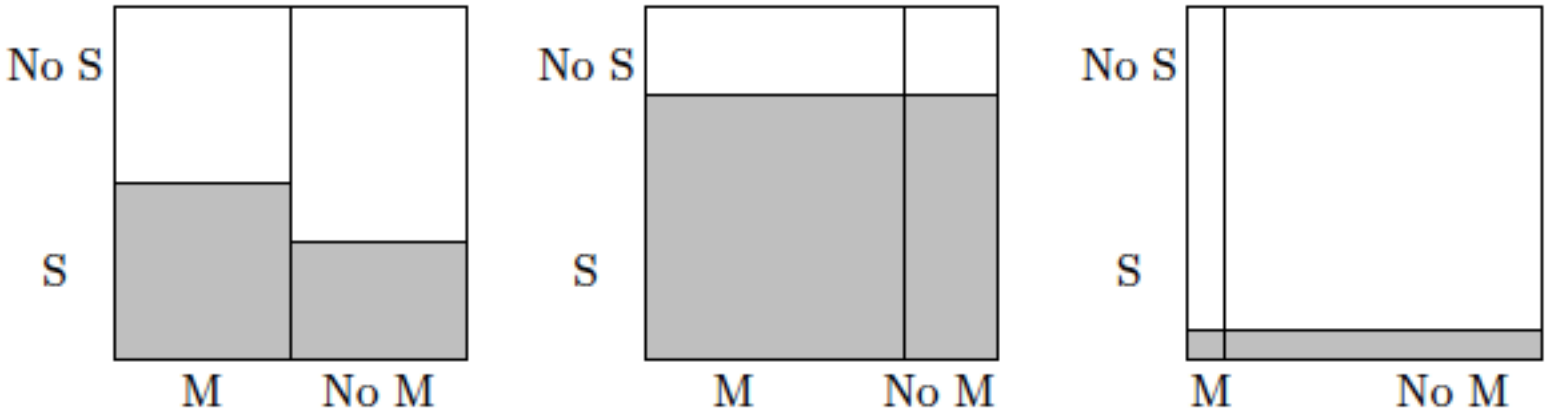
Screening off

How to weed out non-causes that raise probability

- Yellowed fingers and lung cancer
- Falling barometer and rain

C *screens off* A from B if

P(B|A^C)=P(B|C)



(a) (b) Disease = true (c) Disease = false

Screening off and causal chains

Disease → medication → side effect

$P(S|M,D)=P(S|M)$

Common cause principle

C is common cause of A and B if C is earlier than both and:

1. $P(A \wedge B|C)=P(A|C)P(B|C)$
2. $P(A \wedge B|\neg C)=P(A|\neg C)P(B|\neg C)$
3. $P(A|C)>P(A|\neg C)$
4. $P(B|C)>P(B|\neg C)$

Reichenbach

If C and E are dependent either C causes E, E causes C or they have a common cause (that is earlier than both) that screens them off from one another

Problem: no single common cause

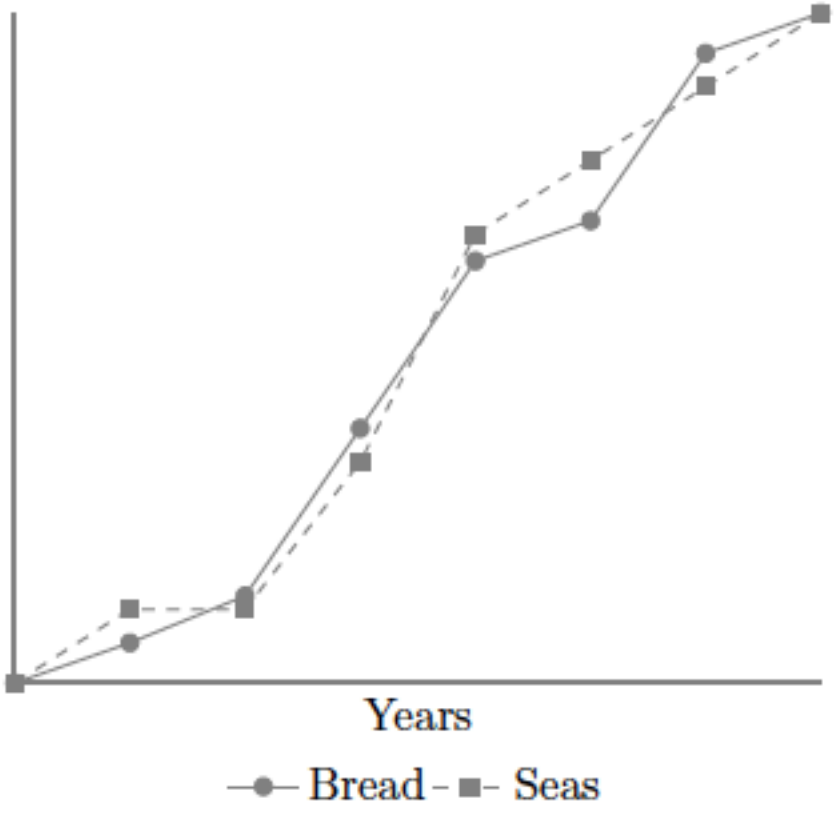
Alice and Bob both like to take AI (i) classes that meet in the afternoon (t), so they're likely to be in the same classes

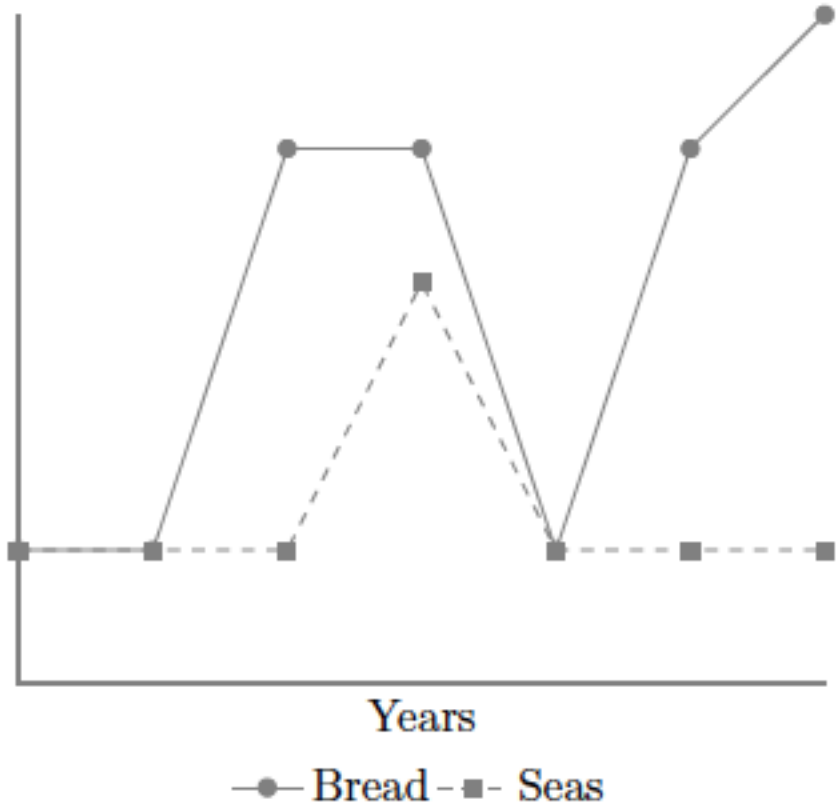
$P(A|B)>P(A)$

$P(A|B \wedge i)>P(A|i)$

$P(A|B \wedge t)>P(A|t)$

Problem: Nonstationarity

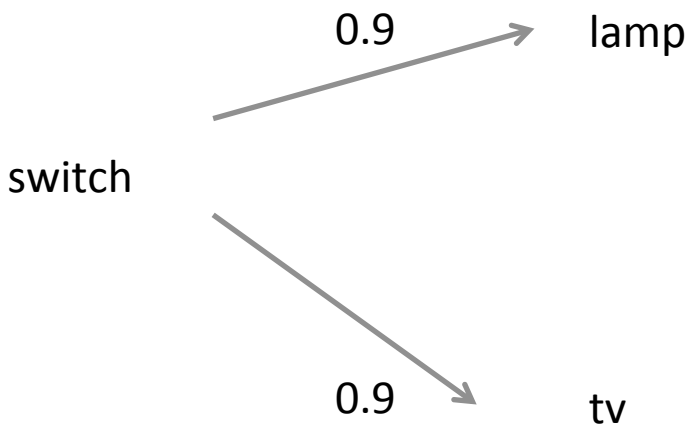




Sober, E. (2001). Venetian sea levels, British bread prices, and the principle of the common cause. *The British Journal for the Philosophy of Science*, *52*(2), 331

Problem: Indeterminism

If lamp on, tv also on, so P(L|T)=1



Does switch screen off L and T?

Bias in graduate admissions?

	Admit	Deny
Male	3738	4704
Female	1494	2827

P(admitted)= 5232/12763≈0.41

P(admitted|female) =1494/4321≈0.35
P(admitted|male)=3738/8442≈0.44

Bickel, P. J., Hammel, E. A., & O'connell, J. W. (1975). Sex bias in graduate admissions: Data from berkeley. *Science*, *187*(4175), 398

Problem: Simpson's paradox

	Men		Women	
Dept	Applied	% Admitted	Applied	Admitted
A	825	62	108	82
B	560	63	25	68
C	325	37	393	34
D	417	33	375	35
E	191	28	393	24
F	272	6	341	7

Suppes

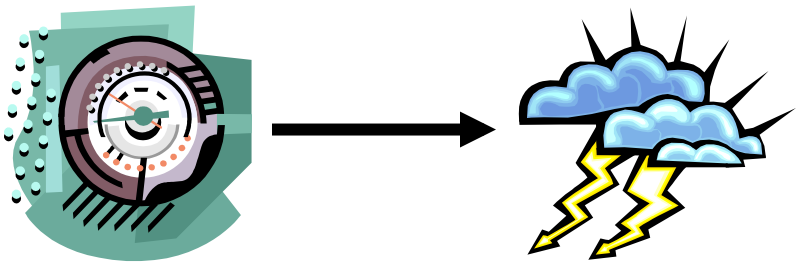
- Causes raise probability of their effects
- Causes are temporally prior to their effects

How can we identify spurious causes?

Prima facie causes

C is a **prima facie cause** of E iff:

- C is earlier than E
- $P(C) > 0$
- $P(E|C) > P(E)$



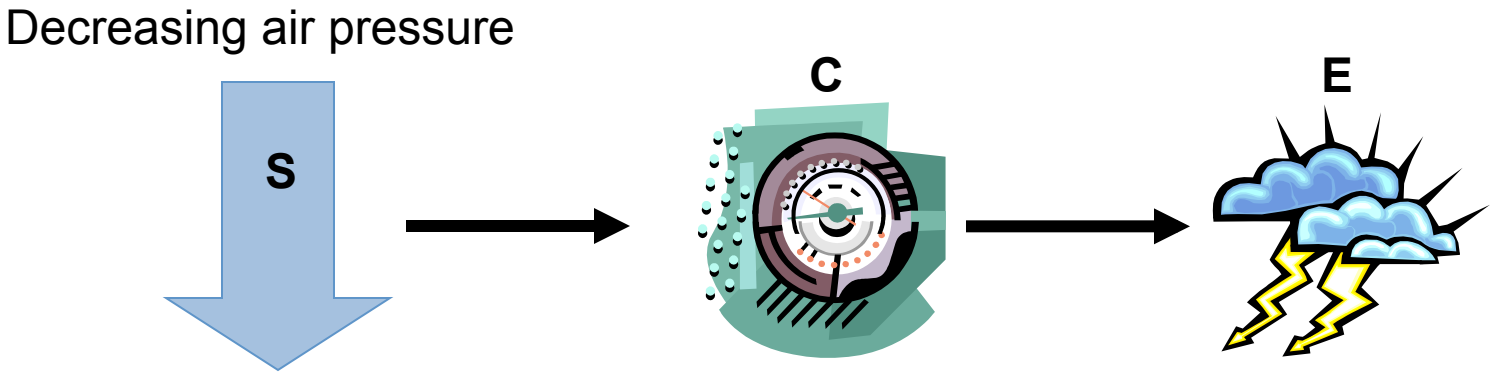
Spurious causes

- Many non-causal correlations
- Main idea: finding things earlier than prima facie cause that account equally well for effect

Spurious

C is **a spurious cause** of E iff:

- C is prima facie cause of E
- Exists S earlier than C such that:
  - P(C,S) > 0
  - P(E|C,S)=P(E|S)
  - P(E|C,S) ≥ P(E|C)



Why is this insufficient?

Cause is spurious if

1. We have an event that eliminates cause's effectiveness for predicting the effect
2. S is at least as good a predictor of effect as cause alone

Genuine causes

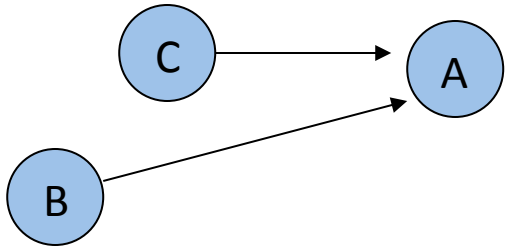
- A prima facie cause that is not spurious is genuine

ε-spurious

- Before, $P(A|B,C)=P(A|C)$
- What if not exactly =, but still small effect?
- B is ε-spurious cause of A iff:
  - Same conditions as before, but:
  - $|P(A|B,C)-P(A|C)|< \varepsilon$

Direct causes

- B is a direct cause of A iff:
  - B is a prima facie cause of A
  - There is no partition such that for each C in the partition:
    - C is temporally in between B and A
    - P(B,C)>0
    - P(A|C,B)=P(A|C)



- ε-direct
  - partition such that $P(A|C,B)-P(A|C) \leq \varepsilon$

Supplementary causes

- Two prima facie causes may aid each other in producing an effect
- In this case, one may be earlier or they may be co-temporary
- B and C are supplementary causes of A iff:
  - $P(B,C)>0$
  - $P(A|B,C)>\max(P(A|B),P(A|C))$
- ε-supplementary

Negative causes

- Prevent effects
  - Ex: vaccine may prevent getting a disease
- B is **prima facie negative** cause of A iff:
  - B earlier than A
  - $P(B)>0$
  - $P(A|B)<P(A)$

Complement of negative cause

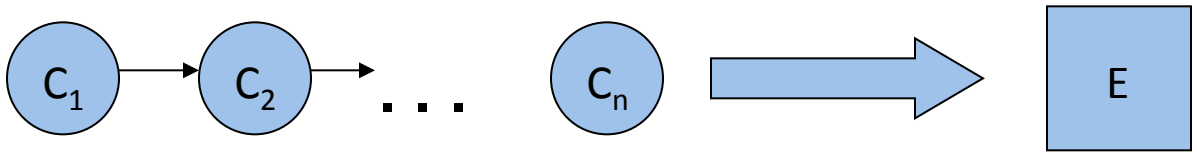
- If $P(A|B)<P(A)$ and
  $P(A)=P(A|B)P(B)+P(A|\neg B)P(\neg B)$
- Then $P(\neg B)>0$ (otherwise $P(A)=P(A|B)$
- And $P(A) < P(A|\neg B)$

- $P(A)=P(A|B)(1-P(\neg B))+P(A|\neg B)P(\neg B)$
  $=P(A|B)+P(\neg B)[P(A|\neg B)-P(A|B)]$
  $P(A)-P(A|B)=P(\neg B)[P(A|\neg B)-P(A|B)]$

Difficulties with probabilistic causality?

Problem: Causal chain

- Each element in chain produces next with P=1
- What happens?
- Each element spurious aside from first – always an earlier cause to account for effect

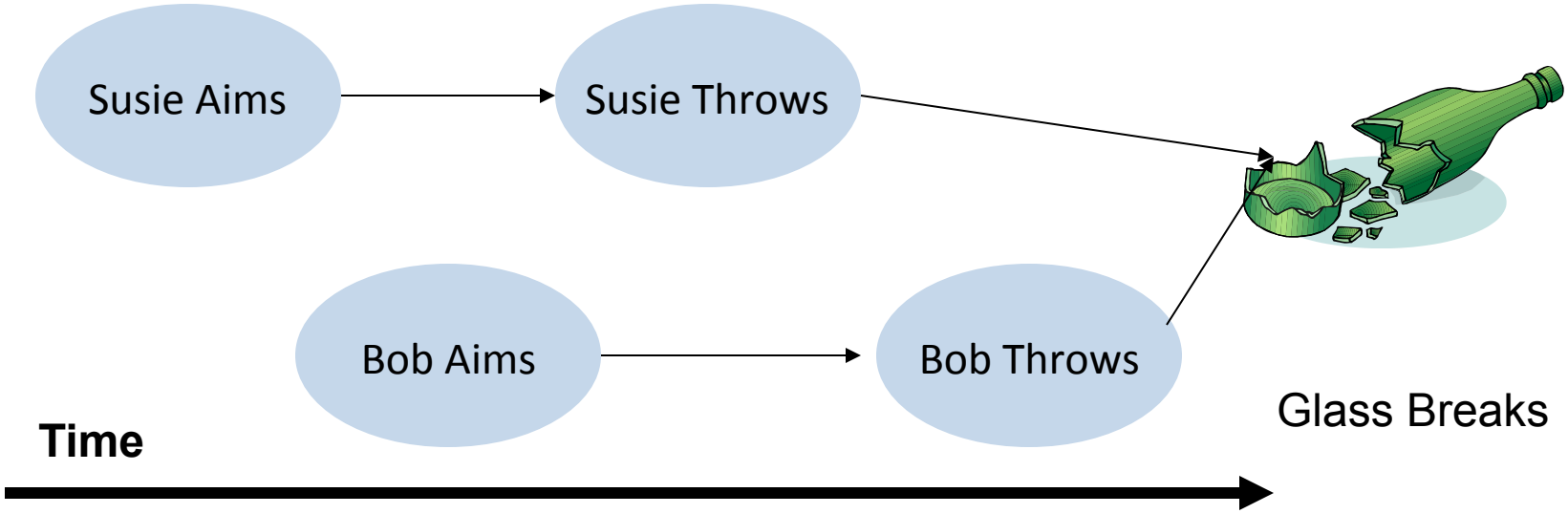


Problem: Simpson's paradox

Same problem as with common cause principle

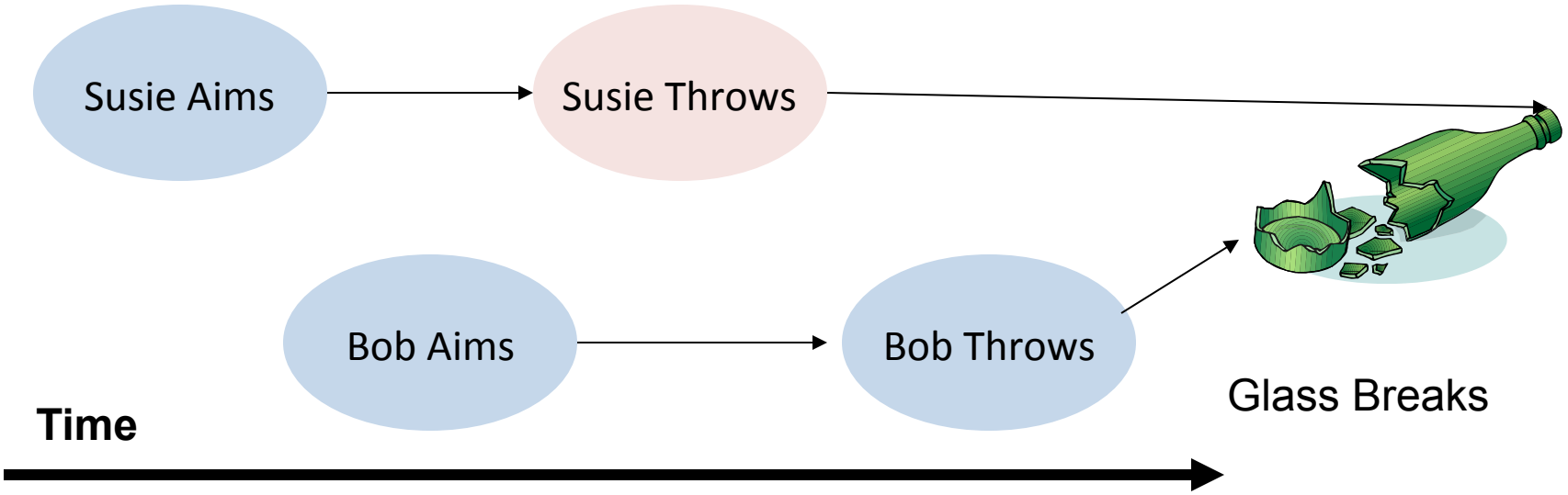
Problem: Overdetermination

- Bob and Susie throwing rocks at glass bottle
- Rocks hit simultaneously, but Susie threw first

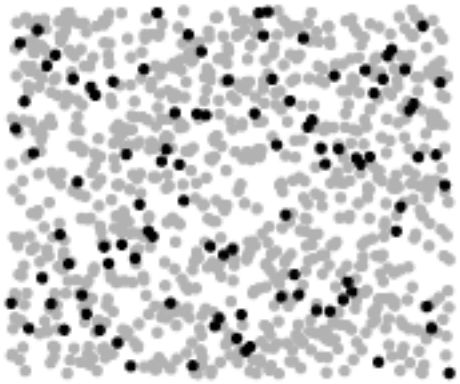


Problem: Preemption

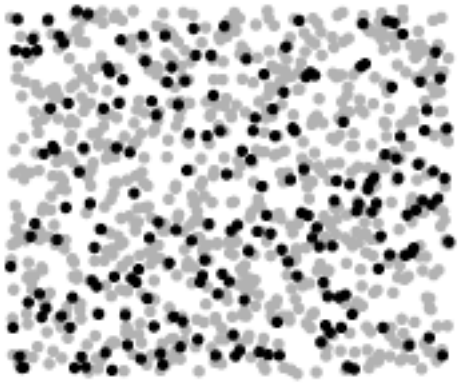
- Bob's rock arrives first and breaks glass
- But Susie still threw first and has perfect aim (breaks glass with P=1)



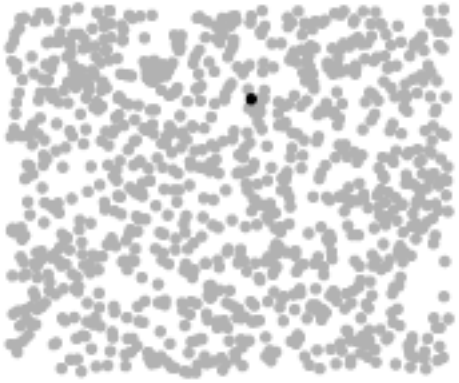
Problem: effect size



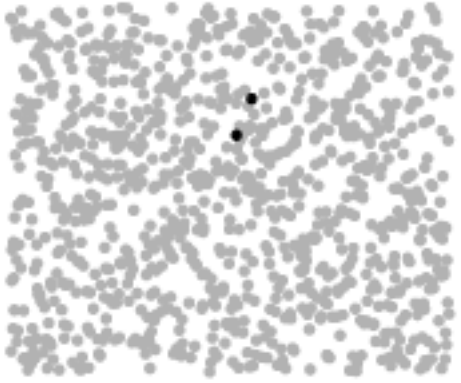
(a) Probability .1



(b) Probability .2



(c) Probability .0000001



(d) Probability .0000002

Eells

- Instead of single event making something spurious, average over contexts
- Separated type/token
  - Probability in general != probability in single case

Background contexts

- Holding fixed other factors, does cause still raise probability of effect?
- Idea is to hold fixed everything relevant to the effect that's not caused by C

Causal background contexts

- n factors relevant to E independently of C
- $K=2^n$ ways of holding these factors fixed
- Background contexts are K_i's such that:
  - $P(K_i \wedge C)>0$
  - $P(K_i \wedge \neg C)>0$

E.g. factors A,B,C, one K_i = {A, ¬B,C}

Comparison to Suppes

- Suppes factors: earlier than prima facie cause
- Eells: factors earlier than effect, not caused by cause

Context unanimity

- For C to be positive cause, must raise probability of E with respect to **every** background context
  - $P(E|C\wedge K_i)>P(E|\neg C\wedge K_i)$ for all K_i
- Can't lower probability or remain same
- What about medications with rare side effects?

Four types of causal factors

Note: all relative to population

1. Positive
   $P(E|C \wedge K_i) > P(E|\neg C \wedge K_i)$ for all K_i
2. Negative
   $P(E|C \wedge K_i) < P(E|\neg C \wedge K_i)$ for all K_i
3. Neutral
   $P(E|C \wedge K_i) = P(E|\neg C \wedge K_i)$ for all K_i
4. Mixed
   Neither positive, negative, or neutral

Strength of relationship

- Not just relevant/not relevant, but how relevant
- Weed out factors that are always positively relevant, but make small difference to probability
- Average degree of causal significance

$$\sum_i [P(E|C \wedge K_i) - P(E|\neg C \wedge K_i)]P(K_i)$$

Example

- To assess whether smoking causes lung cancer, would hold fixed perhaps genetic factors (G), and asbestos exposure (A)
  - Both relevant to LC but independent of S
- Test whether P(LC) raised by S in each context {A ⋀ G, A ⋀ ¬G, ¬A ⋀ G, ¬A ⋀ ¬G}

Difficulties

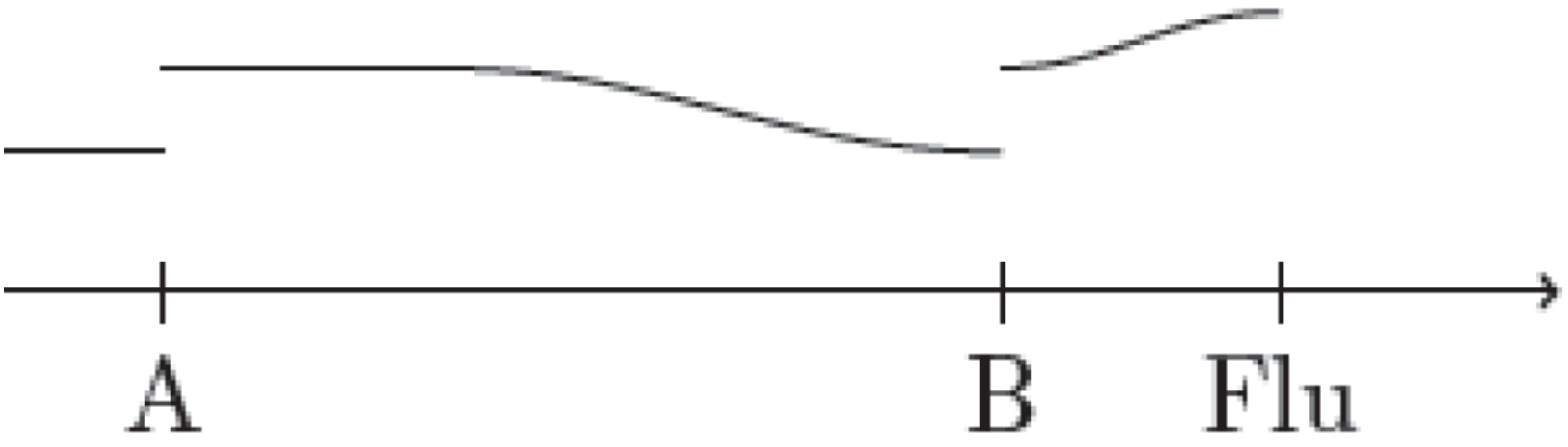
- Circular
  - Need to hold fixed causally relevant factors… to figure out if something is causally relevant
- Indirect
  - What happens in a causal chain? (not holding fixed other effects of cause)

Token causality

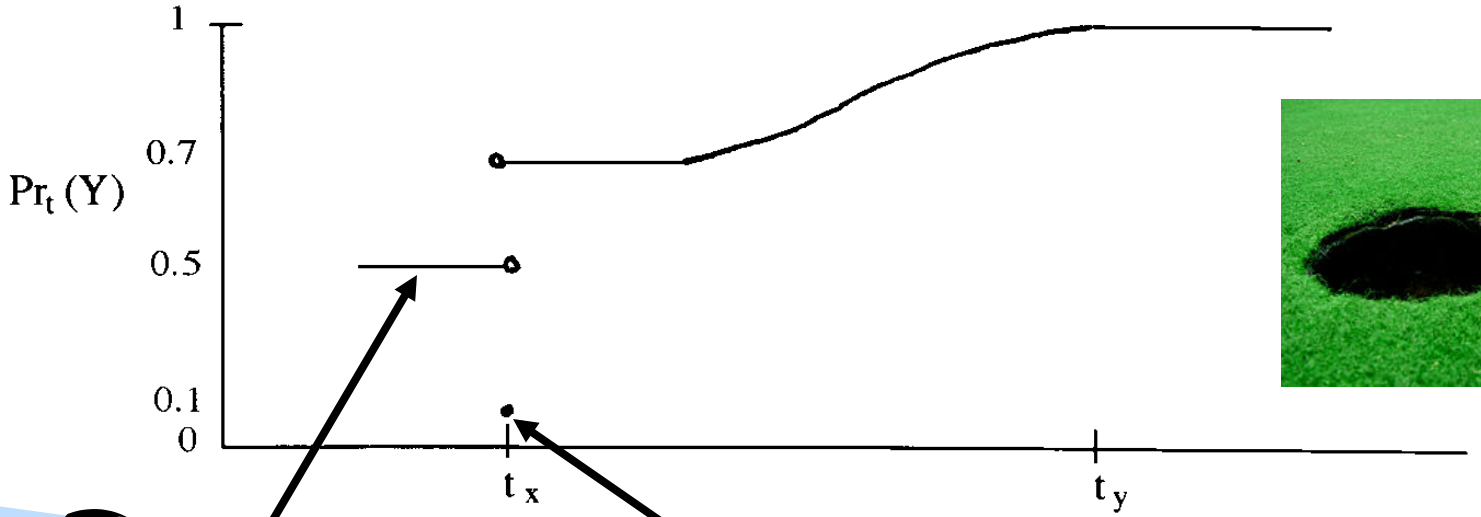
- Difference between P(LC|S) for individual and population
- Timing: lit match and house fire today versus match lit weeks ago
- Differing roles: Positive causes at type level may be negative at token level

Eells on token causality

- Type and token are distinct
- Look at how probability changes over time after cause
- Can distinguish between probability in general and probability on specific occasion
  - Note: this lets us update explanation to be consistent with observation


A
B
Flu

Probability trajectory



Y





x

- P(Y) changes at x
- After x P(Y) is high
- P is higher than before x
- P remains high until y

Was Because Of!

Ways of answering "why"?

Not just caused/didn't cause, but what was role of cause in effect's occurrence?

Anne drives drunk and crashes car. What's significance of her drinking for the crash?

Could be

- Because of
- Despite
- Independently of

y is because of x if

- P(y) changes after x
- P(y) is high after x
- P(y) is higher than it was before x
- P(y) remains high until it occurs

Four relations

Despite

- y is Y despite x if the probability of Y is lowered after $x_{t'}$

Because

- y is Y because *of* x if the probability of Y is increased after $x_{t'}$ and remains increased until y_t

Autonomously

- y is Y autonomously *of* x if the probability of Y changes at $x_{t'}$, this probability is high, but then decreases before y_t

Independently

- y is Y independently *of* x if the probability of Y is the same just after $x_{t'}$ as it is just before $x_{t'}$

The probability of death (D) within 30 minutes of poisoning (O) is 0.8

But if an antidote (A) is given within 10 minutes of poisoning, the probability of death is reduced to 0.001

However the antidote has occasional side effects, so the probability of death when an antidote is given in the absence of the poison is still 0.001

Is O a cause of death? Is A?

Poison and antidote

P(D|O)=0.8 in 30 minutes

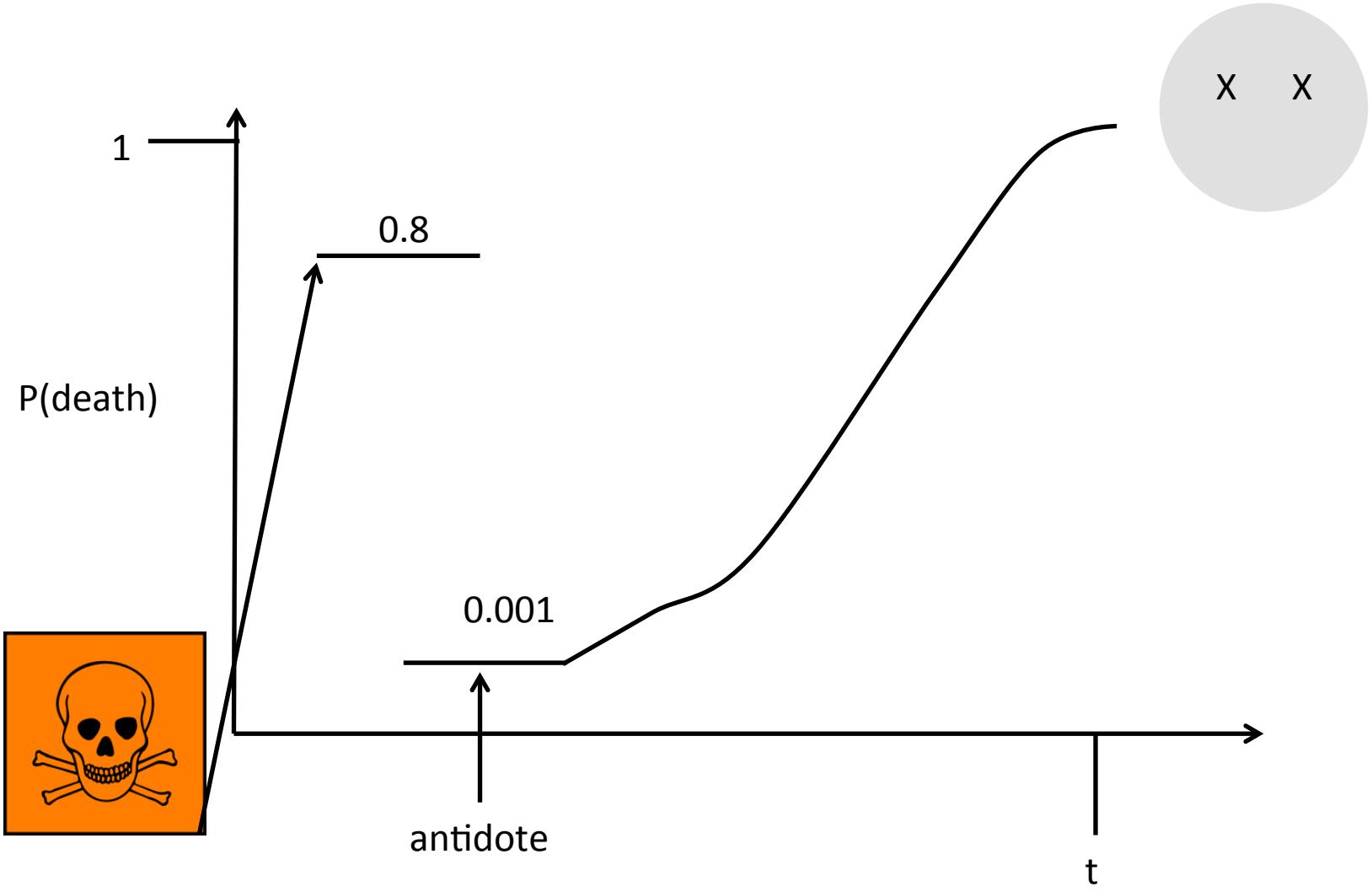
P(D|A^O)=0.001 if within 10 minutes of O

P(D|A^¬O)=0.001

Are A and O positive, negative, neutral or mixed? Better handled by Suppes or Eells?

Can impact on a specific instance be different than that at type level? When?

Poison and antidote



Why are background contexts needed?

- What if poison and antidote are administered at same time?
- What about exposure that provides resistance to later bacteria?

More generally

- Need to account for temporally intermediate factors that affect outcome – independently of the cause
- Want to focus on impact of cause on probability, separating it from impact of other factor

Background factors

- Not relevant
  - Things caused by cause
  - Golf ball's motion near cup
- Relevant
  - Factors interacting with cause, or causes of effect (not caused by cause)
  - Radon exposure for lung cancer

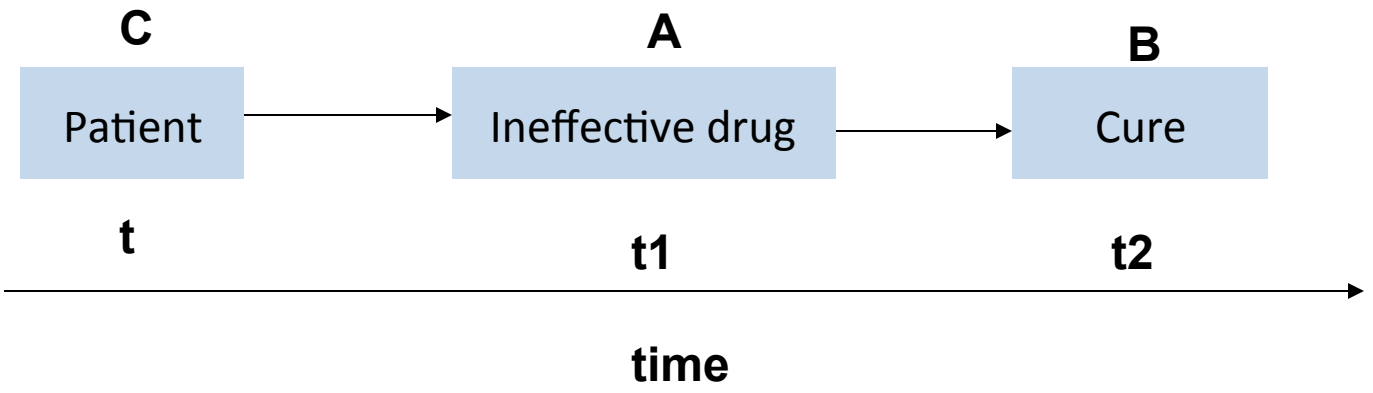
Token causal relevance

x is relevant to y if it happened either because of or despite x, after holding fixed two sets of factors

1. Factors that occur, are not token caused by x being X and are type-level causally relevant to y's being Y holding fixed how things are before these factors actually occur.
2. Factors that token occur in the particular case, are not token caused by x being X and that interact with X with respect to Y holding fixed what is actually true before x_t.

Counterexample (sort of)

- C likely to survive until t2 but not a later time
- P(death|A)=P(death|disease)
- P(death|cure)=0
- BUT residual chance of death from A



Caveats

- How can we know all the important factors?
- How can we find the probability trajectory?

Another approach

- Strength of relationship at token level proportional to that at type level
  - Radiation exposure and death / Marie Curie's radiation exposure causing her death
- Something is a type-level cause because it's often a token cause
  - If poison is fatal with P=0.8, have good reason to believe that when someone is poisoned and dies, this is another instance of death by poison

Connecting principle

If C is a causal factor for producing E in population P of magnitude m, then the support for the hypothesis that C at time t1 caused E at time t2 given that they both token occur in P at these times is m.

Sober, E., & Papineau, D. (1986). Causal factors, causal inference, causal explanation. *Proceedings of the Aristotelian Society, Supplementary Volumes*, *60*, 97-136

Connecting principle

- C and E are types of events, but occur at specific times/places
- S(c caused e|c, e occur in P)=m

$$m = \sum_i [P(E \mid C \wedge K_i) - P(E \mid \neg C \wedge K_i)]P(K_i)$$

But what population?

- Sober: most specific
  - In explaining illness, if we know age/weight, the population is people with those characteristics
  - If only know gender, population is all men or women
- In practice..
  - Never have arbitrarily specific data
  - Would uniquely specify an individual if taken to extremes

Also…

- What if we do not know cause happened for certain?
- What about timing?

Recap of first section

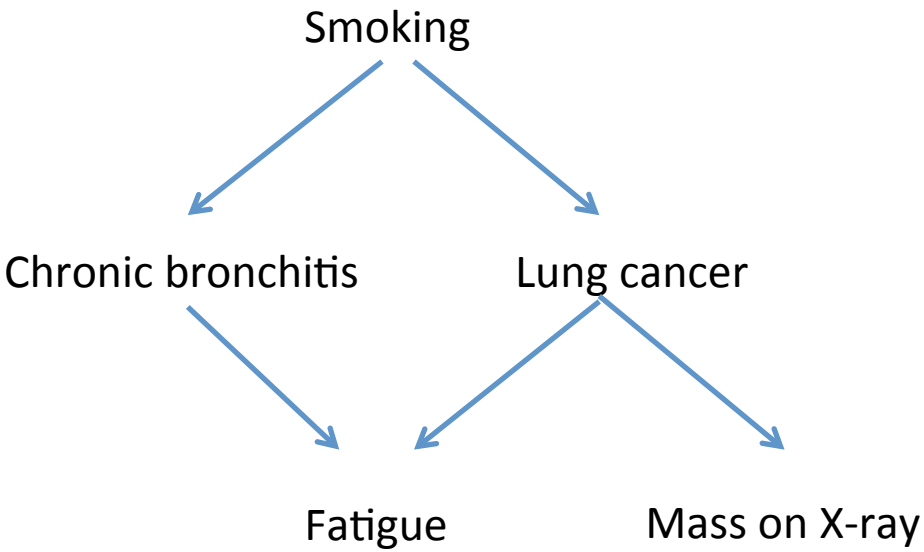
- Main approaches
  - Regularities
  - Counterfactuals
  - Probabilistic causality
  - Omitted: interventionist
- Two problems
  - Type level
  - Token level

Looking forward

- Next few weeks
  - Graphical models
  - Bayes nets, DBNs
  - Granger causality
  - Other computational methods
- Then
  - Experiment design, applications

For next time…

What's probability a smoker with fatigue has cancer?



Read D. Heckerman. A tutorial on learning with Bayesian networks. http://research.microsoft.com/en-us/um/people/heckerman/tutorial.pdf
Sections 1-5