Overview

- The causal relation
- Conditional Independence
- Example CI
- Blocking and d-separation
- Exercise

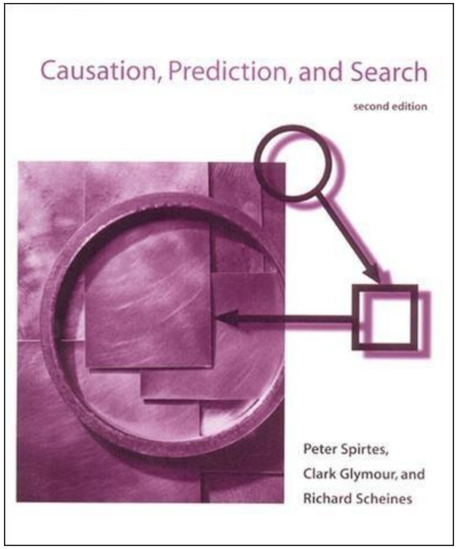
Conditional independence

- 20th century statistics struggled with this issue; at the end of the 20th century many had given up
- However, where there's correlational smoke, there is often a causal fire…
- How to identify that fire?
- Pearl and Glymour et al. then simultaneously developed the insight that not *correlations* or *conditional probabilities* but *conditional independence relations* are key to the identification of causal structure.

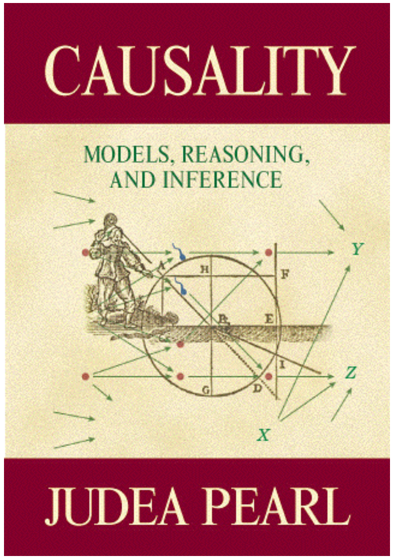
Pittsburgh philosopher Clark Glymour (1942-)

Two major works on this

- Spirtes, Peter, Clark Glymour, and Richard Scheines, (2000) *Causation, Prediction and Search*, Second Edition, Cambridge, MA: MIT Press.



- Pearl, Judea (2009), *Causality: Models, Reasoning, and Inference*, Second Edition, Cambridge: Cambridge University Press.

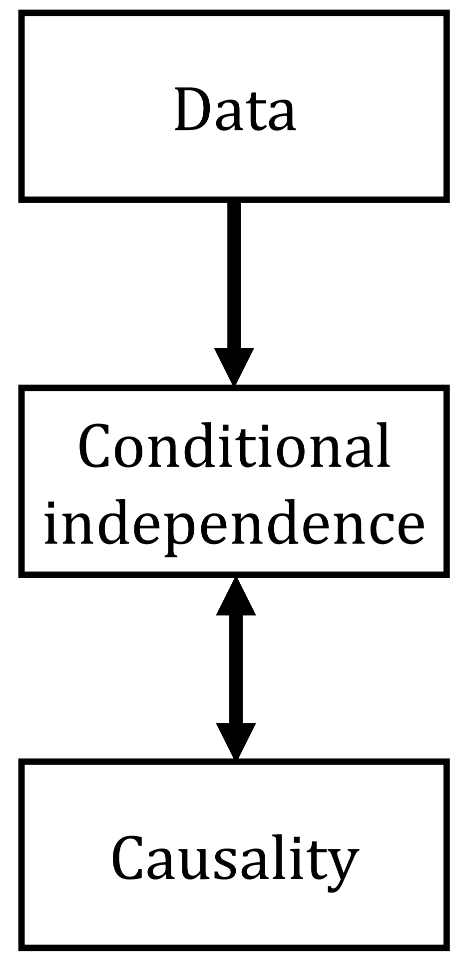


Pearl’s approach (VI)

Trick: shift attention from bivariate to multivariate systems and then ask two new questions:

1) which conditional independence relations are implied by a given causal structure?

2) which causal structures are implied by a given set of conditional independence relations?



Pearl's approach (VI)

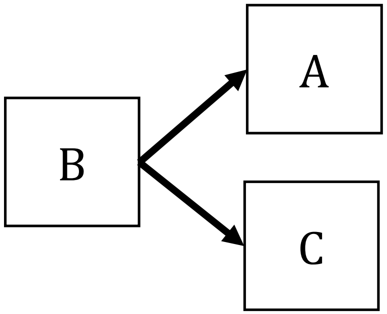
"correlation is not causation"

Judea Pearl: you will never get causal information out without beginning by putting causal hypotheses in

So, correlation is not causation, but it surely helps..

Judea Pearl (1936 -)

Common Cause



Example:

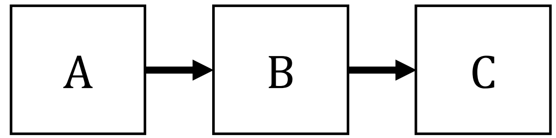
Village Size (B) causes babies (A) and storks (C)

CI:

$A \not\perp\!\!\!\perp C$

$A \perp\!\!\!\perp C | B$

Chain



Example:

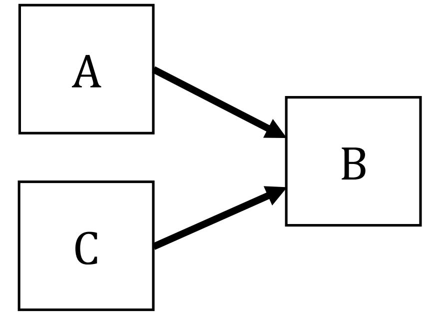
Smoking (A) causes tar (B) causes cancer (C)

CI:

$A \not\perp\!\!\!\perp C$

$A \perp\!\!\!\perp C | B$

Collider



Example:

Firing squad (A & C) shoot prisoner (B)

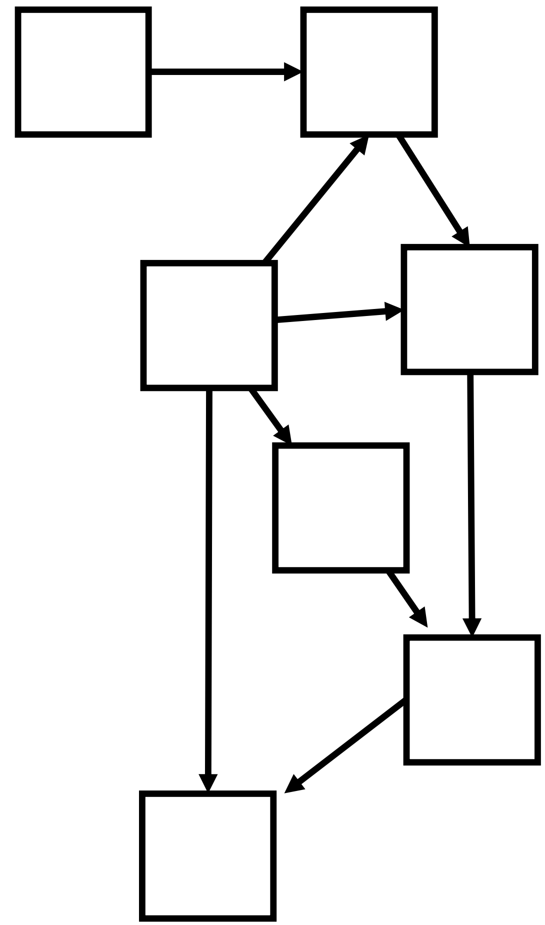
CI:

$A \perp\!\!\!\perp C$

$A \not\perp\!\!\!\perp C | B$

Small conceptual shifts can have significant implications

- If we combine small networks to build larger networks, then we might have a graphical criterion to deduce implied CI relations from a causal graph (i.e., we could look at the graph rather than solve equations)
- If we have a dataset, we can establish which set of possible causal graphs could have generated the CI relations observed
- If certain links cannot be deleted from the graph then it is in principle possible to establish causal relations from non-experimental data…



Overview

- The causal relation
- Conditional Independence
- Example CI
- Blocking and d-separation
- Exercise d-separation

To work!

Example: reasoning with CI

Hypothetical example: the causal relations between smoking, cancer and stained fingers.

Contingency tables for three variables.

Which variables are conditionally independent?

We infer for each pair whether they are *marginally independent* and *conditionally independent*

Example: reasoning with CI

<u>Probabilistic dependence</u>

Given $P(B) > 0$

A and B are *independent* iff
$P(A|B) = P(A)$

$P(A|B = b_1) = P(A|B = b_2)$

A and B are *dependent* iff
$P(A|B) \neq P(A)$

$P(A|B = b_1) \neq P(A|B = b_2)$

Conditional probability:

$P(A|B) = P(A, B)/P(B)$

Conditional independence (CI) *(see handout)*

Table 1a. Contingency table for smoking and cancer with n=100 persons.

	C	¬C	
S	20	40	60
¬S	10	30	40
	30	70	100

Conditional independence (CI) *(see handout)*

Table 1a. Contingency table for smoking and cancer with n=100 persons.

	C	¬C	
S	20	40	60
¬S	10	30	40
	30	70	100

Are smoking and Cancer associated?

Conditional independence (CI) *(see handout)*

Table 1a. Contingency table for smoking and cancer with n=100 persons.

	C	¬C	
S	20	40	60
¬S	10	30	40
	30	70	100

Are smoking and Cancer associated?

We want to know whether $P(C|S) \neq P(C)$

Conditional independence (CI) *(see handout)*

Table 1a. Contingency table for smoking and cancer with n=100 persons.

	C	¬C	
S	20	40	60
¬S	10	30	40
	30	70	100

We want to know whether P(C|S) ≠ P(C)
Formula for conditional probabilities:
P(A|B)=P(A,B)/P(B)

Conditional independence (CI) *(see handout)*

Table 1a. Contingency table for smoking and cancer with n=100 persons.

	C	¬C	
S	20	40	60
¬S	10	30	40
	30	70	100

We want to know whether $P(C|S) \neq P(C)$

Formula for conditional probabilities:

$P(A|B)=P(A,B)/P(B)$

$P(C|S) = P(S,C)/P(S) = 0.2/0.6 = 1/3$

Conditional independence (CI) *(see handout)*

Table 1a. Contingency table for smoking and cancer with n=100 persons.

	C	¬C	
S	20	40	60
¬S	10	30	40
	30	70	100

We want to know whether $P(C|S) \neq P(C)$

Formula for conditional probabilities:

$P(A|B)=P(A,B)/P(B)$

$P(C|S) = 20/60 = 1/3$

$P(C|\neg S) = 10/40 = 1/4$

$P(C) = 30/100 = 3/10$

Conditional independence (CI) *(see handout)*

Table 1a. Contingency table for smoking and cancer with n=100 persons.

	C	¬C	
S	20	40	60
¬S	10	30	40
	30	70	100

We want to know whether P(C|S) ≠ P(C)

Formula for conditional probabilities:

P(A|B)=P(A,B)/P(B)

P(C|S) = 20/60 = 1/3

P(C|¬S) = 10/40 = 1/4

P(C) = 30/100 = 3/10

We conclude: $P(C|S) \neq P(C)$

C and S are dependent!

Conditional independence (CI) *(see handout)*

Table 3a. Contingency table for stained fingers and cancer

	C	¬C	
F	16	74	90
¬F	14	96	110
	30	170	200

In this table:

P(C|F)=16/90=0,18

P(C|¬F)=14/110=0,13

P(C)=30/200=0,15

We conclude: $P(C|F) \neq P(C)$

C and F are dependent!

Conditional independence (CI) *(see handout)*

Now suppose that among the people in 2a there are 100 smokers and 100 non-smokers. Make a separate contingency table for each group.

Table 3b. Contingency table for stained fingers and cancer, given S (only smokers here)

	C	¬C	
F	14	56	70
¬F	6	24	30
	20	80	100

In this table:
P(C|F)=14/70=0,2
P(C|¬F)=6/30=0,2
P(C)=20/100=0,2

Conditional independence (CI) *(see handout)*

Now suppose that among the people in 2a there are 100 smokers and 100 non-smokers. Make a separate contingency table for each group.

Table 3b. Contingency table for stained fingers and cancer, given S (only smokers here)

	C	¬C	
F	14	56	70
¬F	6	24	30
	20	80	100

In this table:
P(C|F)=14/70=0,2
P(C|¬F)=6/30=0,2
P(C)=20/100=0,2

We conclude: $P(C|F) = P(C)$
C and F are independent!

Conditional independence (CI) *(see handout)*

Table 3c. Contingency table for stained fingers and cancer, given ¬S (only non-smokers here)

	C	¬C	
F	2	18	20
¬F	8	72	80
	10	90	100

In this table:
P(C|F)=2/20=0,1
P(C|¬F)=8/80=0,1
P(C)=10/100=0,1

We conclude: $P(C|F) = P(C)$
C and F are independent!

Conditional independence (CI) *(see handout)*

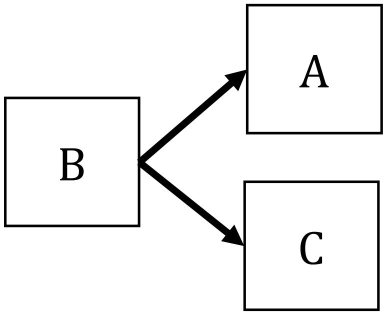
Table 3c. Contingency table for stained fingers and cancer, given ¬S (only non-smokers here)

	C	¬C	
F	2	18	20
¬F	8	72	80
	10	90	100

In this table:
P(C|F)=2/20=0,1
P(C|¬F)=8/80=0,1
P(C)=10/100=0,1

Conclusion: Conditioning on smoking renders F and C probabilistically independent; we say that 'C and F are independent given S'

Common Cause



Example:

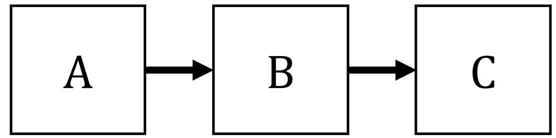
Village Size (B) causes babies (A) and storks (C)

CI:

$A \not\perp\!\!\!\perp C$

$A \perp\!\!\!\perp C | B$

Chain



Example:

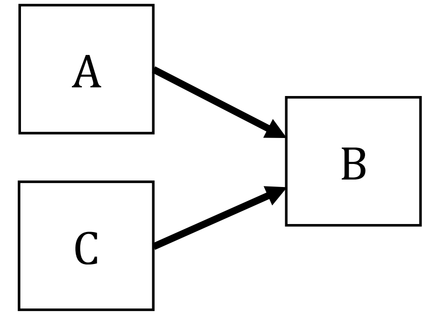
Smoking (A) causes tar (B) causes cancer (C)

CI:

$A \not\perp\!\!\!\perp C$

$A \perp\!\!\!\perp C | B$

Collider



Example:

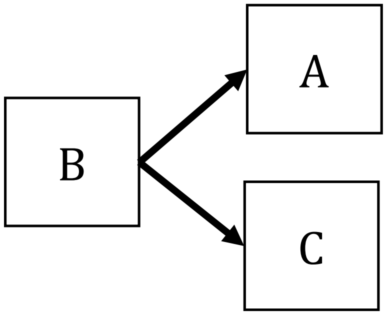
Firing squad (A & C) shoot prisoner (B)

CI:

$A \perp\!\!\!\perp C$

$A \not\perp\!\!\!\perp C | B$

Common Cause



Example:

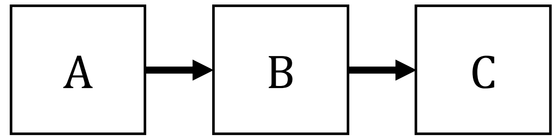
Village Size (B) causes babies (A) and storks (C)

CI:

$A \not\perp\!\!\!\perp C$

$A \perp\!\!\!\perp C \mid B$

Chain



Example:

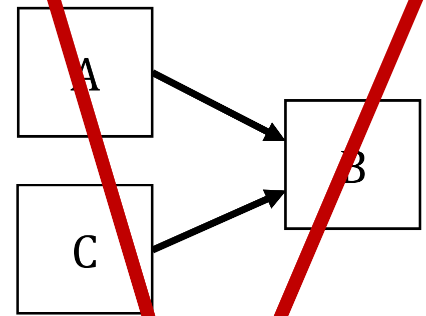
Smoking (A) causes tar (B) causes cancer (C)

CI:

$A \not\perp\!\!\!\perp C$

$A \perp\!\!\!\perp C \mid B$

Collider



Example:

Firing squad (A & C) shoot prisoner (B)

CI:

$A \perp\!\!\!\perp C$

$A \not\perp\!\!\!\perp C \mid B$

Overview

The causal relation

↓

Conditional Independence

↓

Example CI

↓

Blocking and d-separation

↓

Exercise