THE SCIENCE OF CAUSAL INFERENCE
With A Glimpse at Personalized Decision Making

Judea Pearl
UCLA

GENENTECH
March 26, 2025

1

OUTLINE

- Why we need "New Science" to answer causal questions
- Why we need a new inference engine to operationalize the new science
- The fundamental laws ("double-helix") of causal inference
- The Seven+1 Pillars of Causal Wisdom
  - The tools of causal inference,
  - how they impact several application areas
  - Personalized Decision Making

2

TYPICAL CAUSAL QUESTIONS

1. How effective is a given treatment in preventing a disease?
2. Was it the new tax break that caused our sales to go up? Or our marketing campaign?
3. What is the annual health-care costs attributed to obesity?
4. Can hiring records prove an employer guilty of sex discrimination?
5. I am about to quit my job, will I regret it?

* Unarticulatable in the standard grammar of science.

$$Y = aX \qquad \text{vs.} \qquad Y \longleftarrow aX$$

3

3

THE TWO FUNDAMENTAL LAWS OF CAUSAL INFERENCE

1. The Law of Counterfactuals (and Interventions)

$$\boxed{Y_x(u) = Y_{M_x}(u)}$$

(Y_x is equal to Y in a mutilated model M_x.)

2. The Law of Conditional Independence (d-separation)

$$\boxed{(X\ sep\ Y|Z)_{G(M)} \Rightarrow (X \perp\!\!\!\perp Y|Z)_{P(v)}}$$

(Separation in the model ⇒ independence in the distribution.)

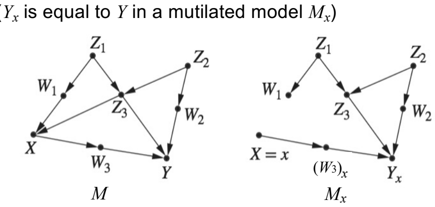
4

SCM: AN ORACLE FOR COUNTERFACTUALS

1. The Law of Counterfactuals (and Interventions)

$$\boxed{Y_x(u) = Y_{M_x}(u)}$$

(Y_x is equal to Y in a mutilated model M_x)



- Counterfactuals are embarrassingly **simple**

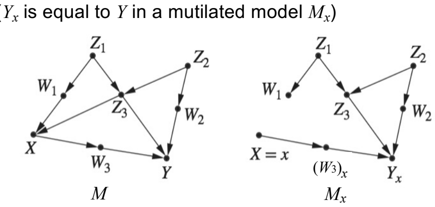
5

SCM: AN ORACLE FOR COUNTERFACTUALS

1. The Law of Counterfactuals (and Interventions)

$$\boxed{Y_x(u) = Y_{M_x}(u)}$$

(Y_x is equal to Y in a mutilated model M_x)

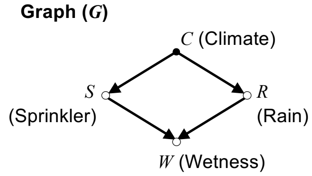


- So are interventions: $P(Y=y \mid do(X=x)) = P(Y_x = y)$

6

READING INDEPENDENCIES

Graph (G)



Model (M)

$$C = f_C(U_C)$$
$$S = f_S(C, U_S)$$
$$R = f_R(C, U_R)$$
$$W = f_W(S, R, U_W)$$

Miracles do happen

If the U's are independent, the observed distribution $P(C,R,S,W)$ satisfies constraints that are:

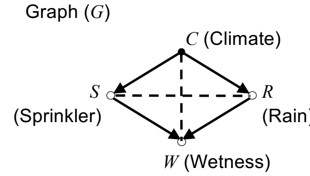
(1) independent of the f's and of $P(U)$,

(2) readable from the graph.

7

READING INDEPENDENCIES (Cont)

Graph (G)



Model (M)

$$C = f_C(U_C)$$
$$S = f_S(C, U_S)$$
$$R = f_R(C, U_R)$$
$$W = f_W(S, R, U_W)$$

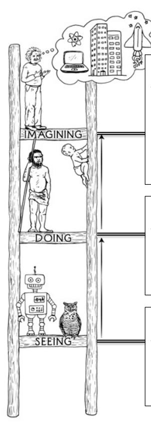
$\text{e.g.}, C \perp\!\!\!\perp W | (S, R) \qquad S \perp\!\!\!\perp R | C$

Applications:

1. Model testing
2. Structure learning
3. Reducing **interventional** questions to adjustments
4. Reducing **interventional** questions to symbolic calculus

8

3-LEVEL HIERARCHY



3. COUNTERFACTUALS
ACTIVITY: Imagining, Retrospection, Understanding
QUESTIONS: *What if I had done . . . ? Why?*
(Was it X that caused Y? What if X had not occurred? What if I had acted differently?)
EXAMPLES: Was it the aspirin that stopped my headache? Would Kennedy be alive if Oswald had not killed him? What if I had not smoked the last 2 years?

2. INTERVENTION
ACTIVITY: Doing, Intervening
QUESTIONS: *What if I do . . . ? How?*
(What would Y be if I do X?)
EXAMPLES: If I take aspirin, will my headache be cured? What if we ban cigarettes?

1. ASSOCIATION
ACTIVITY: Seeing, Observing
QUESTIONS: *What if I see . . . ?*
(How would seeing X change my belief in Y?)
EXAMPLES: What does a symptom tell me about a disease? What does a survey tell us about the election results?

9

THE SEVEN+1 PILLARS (TOOLS) OF CAUSAL INFERENCE

Tool 1: Encoding causal information in transparent and testable way
Tool 2: Predicting the effects of actions and policies
Tool 3: Computing counterfactuals and finding causes of effects (**attribution, explanation, susceptibility**)
Tool 4: Computing direct and indirect effects (Mediation) (discrimination, inequities, fairness)
Tool 5: Integrating data from diverse sources (fusion, transportability, transfer-learning)
Tool 6: Recovering from missing Data
Tool 7: Causal Discovery
Tool 8: Personalized Decision Making

10

PILLAR 1: ENCODING CAUSAL INFORMATION IN A TRANSPARENT AND TESTABLE WAY

- In the 1980's graphs and probabilities come together to create a machine for revising beliefs in light of new evidence.
- Bayesian Networks captured probabilistic knowledge in graphical form and permitted updating by message passing.
- Gift of the Gods:
  The axioms of conditional independence and the axioms of graph separation share a common core.
- d-separation: If we find all paths from X to Y intercepted by vertices of set Z, we can conclude that X is independent of Y conditioned on Z.
  (The alpha-bet of causal modeling)

11

11

PILLAR 2: THE CONTROL OF CONFOUNDING

Problem: Determine if effects of a given policy/action can be estimated from data and how.

Solution: Demystified and reduced to a game

•"back-door" – adjustment for covariates
•"front door" – extends it beyond adjustment
•*do-calculus* – predicts the effect of policy interventions whenever feasible

12

PILLAR 3: THE ALGORITHMIZATION OF COUNTERFACTUALS

Task: Given {Model + Data}, determine what Joe's salary would be, had he had one more year of education.

Solution: The probability of every counterfactual can be computed or bounded using the "surgery" procedure.

Corollary: "Causes of effects" and "Attribution" formalized.

13

ATTRIBUTION

- Your Honor! My client (Mr. A) died BECAUSE he used that drug.

-

14

ATTRIBUTION

- Your Honor! My client (Mr. A) died BECAUSE he used this drug.

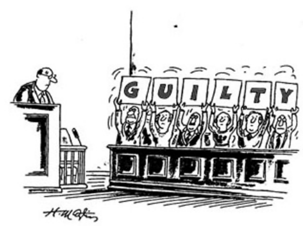
- Court to decide if it is MORE PROBABLE THAN NOT that A would be alive BUT FOR the drug!
- $PN = P(alive_{\{no\ drugs\}} \mid dead, drug) \geq 0.50$

15

CAN FREQUENCY DATA DETERMINE LIABILITY?

Sometimes:

When PN is bounded above 0.50.



- WITH PROBABILITY ONE $1 \leq PN \leq 1$
- Combined data tell more that each study alone

16

PILLAR 4: MEDIATION ANALYSIS – DIRECT AND INDIRECT EFFECTS

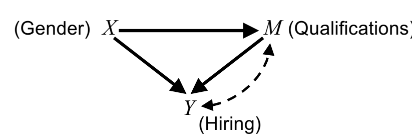
Task: Given {Data + Model}, unveil and quantify the mechanisms that transmit changes from a cause to its effects.

Result: The graphical representation of counterfactuals tells us when direct and indirect effects are estimable from data, and, if so, how necessary (or sufficient) mediation is for the effect.

17

LEGAL IMPLICATIONS OF DIRECT EFFECT

Can data prove an employer guilty of hiring discrimination?



What is the direct effect of X on Y?

$$CDE = E(Y|do(x_1), do(m)) - E(Y|do(x_0), do(m))$$

(m-dependent) Adjust for M? No! No!

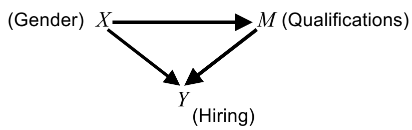
CDE identification is completely solved

18

18

LEGAL DEFINITION OF DISCRIMINATION

Can data prove an employer guilty of hiring discrimination?

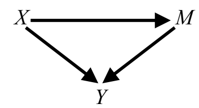


The Legal Definition:
Find the probability that “the employer would have acted differently had the employee been of different sex and qualification had been the same.”

19

NATURAL INTERPRETATION OF AVERAGE DIRECT EFFECTS

Robins and Greenland (1992) – Pearl (2001)



$m = f(x, u)$
$y = g(x, m, u)$

Natural Direct Effect of X on Y: $DE(x_0, x_1; Y)$
The expected change in Y, when we change X from x_0 to x_1 and, for each u, we keep M constant at whatever value it attained before the change.

$$E[Y_{x_1 M_{x_0}} - Y_{x_0}]$$

Note the nested counterfactuals 20

20

PILLAR 5:
TRANSFER LEARNING, EXTERNAL VALIDITY, AND SAMPLE SELECTION BIAS

Task: A machine trained in one environment finds that environmental conditions changed. When/how can it amortize past learning to the new environment?

Solution: Complete formal solution obtained through the *do*-calculus and “selection diagrams” (Bareinboim et al., 2016)

21

APPLICATION: GENERALIZABILITY AND DATA FUSION

The problem

- How to combine results of several experimental and observational studies, each conducted on a different population and under a different set of conditions,
- so as to construct a valid estimate of effect size in yet a new population, unmatched by any of those studied.

22

THE PROBLEM IN REAL LIFE

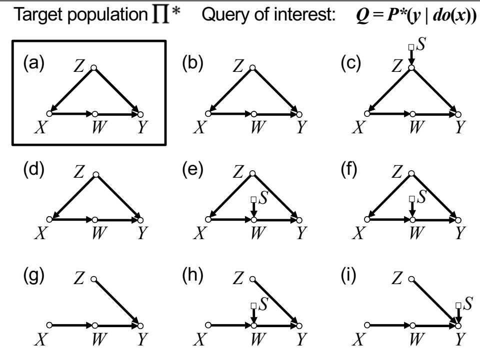
Target population $\prod^*$ Query of interest: $Q = P^*(y \mid do(x))$

(a) Arkansas Survey data available	(b) New York Survey data Resembling target	(c) Los Angeles Survey data Younger population
(d) Boston Age not recorded Mostly successful lawyers	(e) San Francisco High post-treatment blood pressure	(f) Texas Mostly Spanish subjects High attrition
(g) Toronto Randomized trial College students	(h) Utah RCT, paid volunteers, unemployed	(i) Wyoming RCT, young athletes

23

THE PROBLEM IN MATHEMATICS

Target population $\prod^*$ Query of interest: $Q = P^*(y \mid do(x))$



24

PILLAR 6: MISSING DATA (Mohan, 2017)

Problem: Given data corrupted by missing values and a model of what causes missingness. Determine when relations of interest can be estimated consistently "as if no data were missing."

Results: Graphical criteria unveil when estimability is possible, when it is not, and how.

25

25

MISSING DATA: A STATISTICAL PROBLEM TURNED CAUSAL

Sam-ple #	X	Y	Z
1	1	0	0
2	1	0	1
3	1	m	m
4	0	1	m
5	m	1	m
6	m	0	1
7	m	m	0
8	0	1	m
9	0	0	m
10	1	0	m
11	1	0	1
-			

Question:
Is there a consistent estimator of $P(X,Y,Z)$?
That is, is $P(X,Y,Z)$ estimable (asymptotically) as if no data were missing.

26

MISSING DATA: A STATISTICAL PROBLEM TURNED CAUSAL

Sam-ple #	X	Y	Z
1	1	0	0
2	1	0	1
3	1	m	m
4	0	1	m
5	m	1	m
6	m	0	1
7	m	m	0
8	0	1	m
9	0	0	m
10	1	0	m
11	1	0	1
-			

Question:
Is there a consistent estimator of $P(X,Y,Z)$?
That is, is $P(X,Y,Z)$ estimable (asymptotically) as if no data were missing.

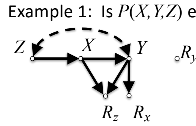
Answers:
1. There is no Model-blind estimator, but,
2. Given a missingness model, we can tell you yes/no, and how.
3. Given a missingness model, we can tell you whether or not it has testable implications.

27

SMART ESTIMATION OF $P(X,Y,Z)$

Sam-ple #	X	Y	Z
1	1	0	0
2	1	0	1
3	1	m	m
4	0	1	m
5	m	1	m
6	m	0	1
7	m	m	0
8	0	1	m
9	0	0	m
10	1	0	m
11	1	0	1
-			

Example 1: Is $P(X,Y,Z)$ estimable?



$R_x = 0 \Rightarrow X$ observed
$R_x = 1 \Rightarrow X$ missing

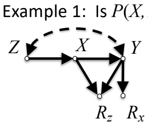
$$P(X,Y,Z) \neq P(X,Y,Z \mid R_x = 0, R_y = 0, R_z = 0)$$

28

SMART ESTIMATION OF $P(X,Y,Z)$

Sam-ple #	X	Y	Z
1	1	0	0
2	1	0	1
3	1	m	m
4	0	1	m
5	m	1	m
6	m	0	1
7	m	m	0
8	0	1	m
9	0	0	m
10	1	0	m
11	1	0	1
-			

Example 1: Is $P(X,Y,Z)$ estimable?



$R_x = 0 \Rightarrow X$ observed
$R_x = 1 \Rightarrow X$ missing

$$\begin{aligned} P(X,Y,Z) &= P(Z \mid X,Y) P(X \mid Y) P(Y) \\ &= P(Z \mid X,Y,R_x = 0, R_y = 0, R_z = 0) \\ &\times P(X \mid Y, R_x = 0, R_y = 0) \\ &\times P(Y \mid R_y = 0) \end{aligned}$$

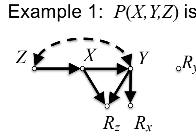
Testable implications:
$Z \perp\!\!\!\perp R_y \mid R_z = 0$
$R_z \perp\!\!\!\perp R_x \mid Y, R_y = 0$

29

SMART ESTIMATION OF $P(X,Y,Z)$

Sam-ple #	X	Y	Z
1	1	0	0
2	1	0	1
3	1	m	m
4	0	1	m
5	m	1	m
6	m	0	1
7	m	m	0
8	0	1	m
9	0	0	m
10	1	0	m
11	1	0	1
-			

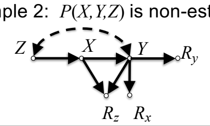
Example 1: $P(X,Y,Z)$ is estimable



$R_x = 0 \Rightarrow X$ observed
$R_x = 1 \Rightarrow X$ missing

$$\begin{aligned} P(X,Y,Z) &= P(Z \mid X,Y) P(X \mid Y) P(Y) \\ &= P(Z \mid X,Y,R_x = 0, R_y = 0, R_z = 0) \\ &\times P(X \mid Y, R_x = 0, R_y = 0) \\ &\times P(Y \mid R_y = 0) \end{aligned}$$

Example 2: $P(X,Y,Z)$ is non-estimable



30

SUMMARY: MISSING DATA

Results:
Causal models unveil (1) when estimability is possible, (2) how, (3) when it is not, (4) when model-blind estimators can do it, and (5) when they cannot.

Corollary:
Only by taking models seriously we can learn when they are not needed.

31

PILLAR 7: CAUSAL DISCOVERY

Task: Search for a set of models (graphs) that are compatible with the data, and represent them compactly.

Results: In certain circumstances, and under weak assumptions, causal queries can be estimated directly from this compatibility set.

(Spirtes, Glymour and Scheines (2000); Jonas Peters etal (2018))

32

INTRODUCTION TO PERSONALIZED DECISION MAKING

Counterfactual analysis permits us to take population data and estimate the probability that a **given individual** $\boldsymbol{u}$ would **benefit** (or be harmed) by a given treatment X, as opposed to the **average** recovery rate in the subpopulation **resembling** the individual.

$$\boldsymbol{P(Benefit)} = P[Y(1) = 1 \;\&\; Y(0) = 0 | C(u)]$$

vs.

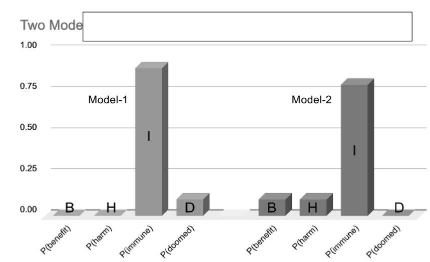
$$\boldsymbol{ATE} = E[Y(1) - Y(0) | C(u)]$$

Example: no effect vs. cure and kill

33

TWO MODELS, INDISTINGUISHABLE IN RCT, YET DIFFERENT FOR DECISIONS

Model-1: Drug has no effect whatsoever on any individual
Model-2: Drug saves 10% of the population and kills another 10%



34

FOR WHOM THE DIFFERENCE MAKES

Model-1 No effect on any individual
Model-2 Saves 10% and Kills 10%

1. **Patient**: M-1 useless but safe, M-2 Scary or Life-saver
2. **Policy maker**: Zero population efficacy; M-2 may cause outrage.
3. **Researcher**: From differentiating markers to understanding underlying mechanism

35

HOW POPULATION DATA DIFFERENTIATE INDIVIDUAL BEHAVIORS

- In general, $P(Benefit)$ and $P(Harm)$ cannot be point-estimated from population data, but they can be **bounded**.
- The **bounds** shrink when both experimental and observational data are combined.
- Why?
- **Intuition**: *M-1(no effect)* can be ruled out if we find no deaths in Observational Studies.
- **Mathematics**: Tight **bounds** on $P(Benefit)$ for any RCT + OS data.
- Further shrinkage with structural information.

36

IDENTIFYING "PATIENTS IN NEED"

Counterfactual: Patients susceptible to treatment.

$P(Benefit)$ = Probability that a patient with characteristics c will improve IF AND ONLY IF treated.

$$P(Benefit) = P(Y(1) = 1, Y(0) = 0 \mid C = c)$$

Experimental and observational studies provide informative bounds on $P(Benefit)$.

In general, going from group data to individual behavior requires counterfactual logic.

37

IDENTIFYING "PATIENTS IN NEED"

Counterfactual: Patients susceptible to treatment.

$P(Benefit)$ = Probability that a patient with characteristics c will improve IF AND ONLY IF treated.

$$P(Benefit) = P(Y(1) = 1, Y(0) = 0 \mid C = c)$$

Experimental and observational studies provide informative bounds on $P(Benefit)$.

In general, going from group data to individual behavior requires counterfactual logic.

- Situation-specific decisions

38

IDENTIFYING "PATIENTS IN NEED"

Counterfactual: Patients susceptible to treatment.

$P(Benefit)$ = Probability that a patient with characteristics c will improve IF AND ONLY IF treated.

$$P(Benefit) = P(Y(1) = 1, Y(0) = 0 \mid C = c)$$

Experimental and observational studies provide informative bounds on $P(Benefit)$.

Going from group data to individual behavior requires counterfactual logic.

- Personalized medicine

39

IDENTIFYING "PATIENTS IN NEED"

Counterfactual: Patients susceptible to treatment.

$P(Benefit)$ = Probability that a patient with characteristics
c will improve IF AND ONLY IF treated.

$$P(Benefit) = P(Y(1) = 1, Y(0) = 0 \mid C = c)$$

Experimental and observational studies provide informative bounds on $P(Benefit)$.

Going from group data to individual behavior requires counterfactual logic.

- Identify customers worthy of offer/recommendation

40

IDENTIFYING "PATIENTS IN NEED"

Counterfactual: Patients susceptible to treatment.

$P(Benefit)$ = Probability that a patient with characteristics c will improve IF AND ONLY IF treated.

$$P(Benefit) = P(Y(1) = 1, Y(0) = 0 \mid C = c)$$

Experimental and observational studies provide informative bounds on $P(Benefit)$.

Going from group data to individual behavior requires counterfactual logic.

- Characterize voters swayable by a slogan

41

IDENTIFYING "PATIENTS IN NEED"

Counterfactual: Patients susceptible to treatment.

$P(Benefit)$ = Probability that a patient with characteristics c will improve IF AND ONLY IF treated.

$$P(Benefit) = P(Y(1) = 1, Y(0) = 0 \mid C = c)$$

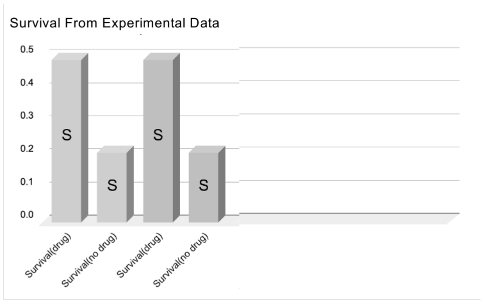
Experimental and observational studies provide informative bounds on $P(Benefit)$.

Going from group data to individual behavior requires counterfactual logic.

- Unit Selection: Li, Mueller and Pearl (2021)

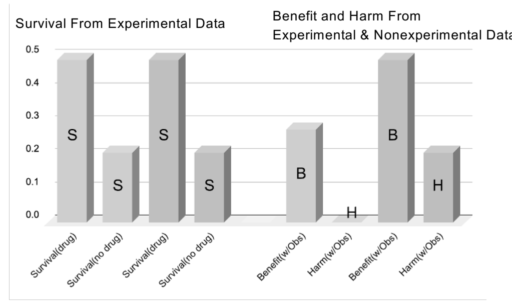
42

HOW COUNTERFACTUAL ANALYSIS DISTINGUISHES SURVIVAL, BENEFIT, AND HARM



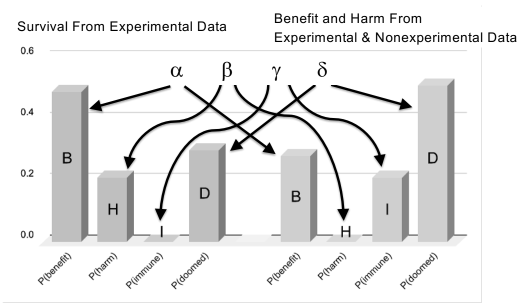
43

HOW COUNTERFACTUAL ANALYSIS DISTINGUISHES SURVIVAL, BENEFIT, AND HARM



44

HOW COUNTERFACTUAL ANALYSIS DISTINGUISHES SURVIVAL, BENEFIT, AND HARM



45

PROBABILITIES AND UTILITIES FOR 4 TYPES OF PATIENTS, IN MEN AND WOMEN

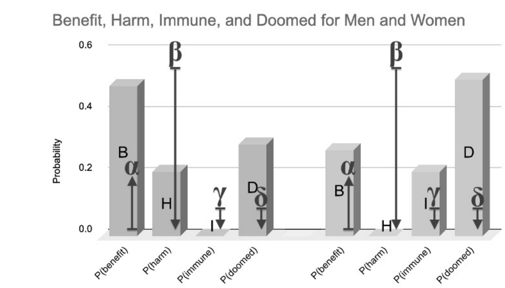


Figure 1: Probabilities of benefit, harm, immunity, and doomed among men (blue bars) and women (pink bars). Arrows represent positive (up) and negative (down) utilities of benefit, harm, immunity, and doom.

46

WEIGHTED UTILITIES FOR 4 TYPES OF PATIENTS, IN MEN AND WOMEN

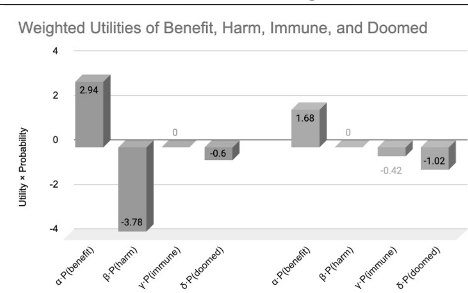


Figure 2: Expected utilities of benefit, harm, immunity, and doom among men (blue bars) and women (pink bars).

47

OVERALL UTILITY FOR MEN VS. WOMEN

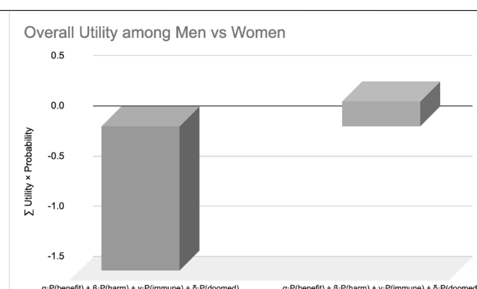


Figure 3: Expected utility among men (blue bar) and women (pink bar).

48

PROBABILITY BOUNDS UNDER NON-IDENTIFIABILITY

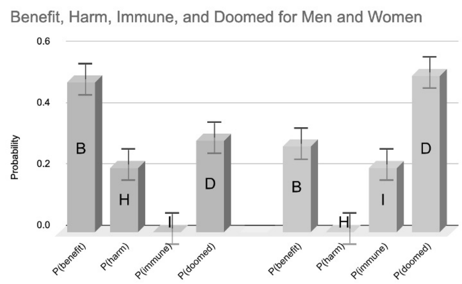


Figure 4: Probabilities of benefit, harm, immunity, and doom among men (blue bars) and women (pink bars) with bounds.

49

WEIGHTED UTILITIES FOR 4 TYPES OF PATIENTS, IN MEN AND WOMEN

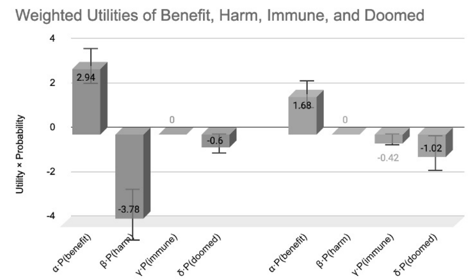


Figure 5: Expected utilities of benefit, harm, immunity, and doom among men (blue bars) and women (pink bars) with bounds.

50

OVERALL UTILITY BOUNDS FOR MEN VS. WOMEN

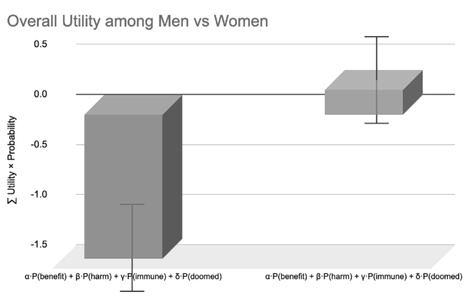


Figure 6: Expected utility among men (blue bar) and women (pink bar) with non-overlapping bounds.

51

OVERLAPPING UTILITY BOUNDS FOR MEN VS. WOMEN

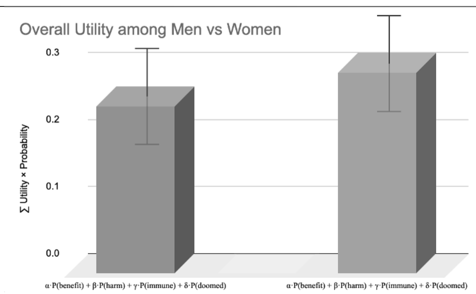


Figure 7: Expected utility among men (blue bar) and women (pink bar) with overlapping bounds.

52

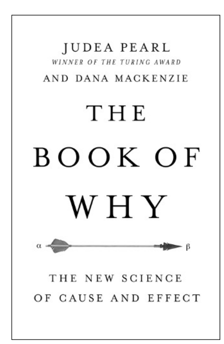
CONCLUSIONS

"More has been learned about causal inference in the last few decades than the sum total of everything that had been learned about it in all prior recorded history."
(Gary King, Harvard, 2014)

"Medicine is on the brink of a major revolution, as we can now answer situation-specific questions that scientists have always longed to ask but couldn't."
(The Author, UCLA, 2024)

53

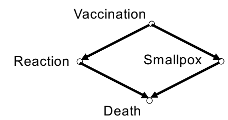
For a trailer, click WHY on my home page.
For discussions: @yudapearl



54

55

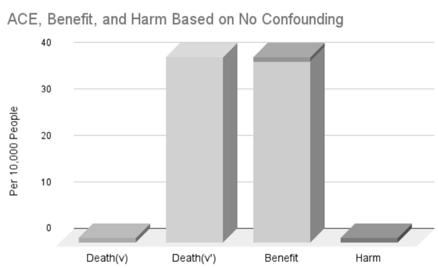
MORE PEOPLE DIED FROM SMALLPOX VACCINE THAN FROM SMALLPOX ITSELF (France, 1840s)



- Is vaccine (V) harmful or beneficial for an individual
- P(benefit) = P(Live if Vaxed & Die if not Vaxed)
  = $P(L_v \,\&\, D_{v'})$
  ≠ $P(L_v) - P(L_{v'})$ = ACE (from RCT)
- Counterfactual logic, Algorithmitized (1994) and analyzed (1999), applied (2022).

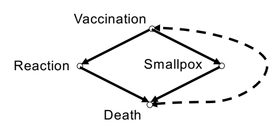
56

AVERAGE CAUSAL EFFECTS (ACE), BENEFIT AND HARM, ASSUMING NO-CONFOUNDING



57

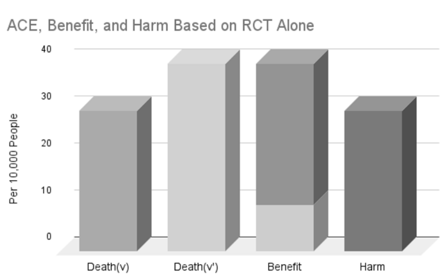
MORE PEOPLE DIED FROM SMALLPOX VACCINE THAN FROM SMALLPOX ITSELF (France, 1840s)



- What if we have confounding?
- What if the average causal effect (ACE) is not as pronounced?

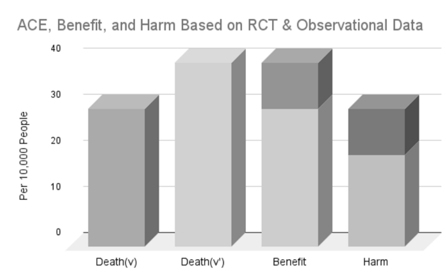
58

ACE AND INDIVIDUALIZED BENEFIT AND HARM BASED ON RCT ALONE



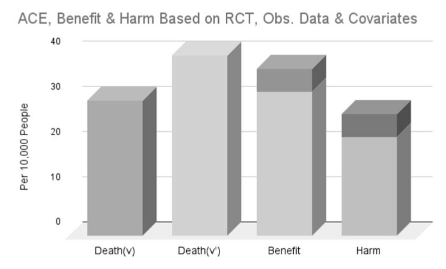
59

COMBINING RCT WITH OBSERVATIONAL STUDIES NARROWS BENEFIT-HARM BOUNDS



60

ADDING COVARIATE DATA FURTHER SEPARATES BENEFIT FROM HARM



61

CONCLUSIONS

"More has been learned about causal inference in the last few decades than the sum total of everything that had been learned about it in
all prior recorded history."

(Gary King, Harvard, 2014)

"The next revolution will be even more impactful upon realizing that data science is the science of interpreting reality, not of summarizing data."

(The Author, UCLA, 2022)

62

CONCLUSIONS (cont.)

Q. So What is Causal Inference?

A. It's the leverage that elevates Data Science from Rung-1 of the Ladder to Rungs 2 and 3,

i.e., from data-fitting to deep understanding.

63

Paper available: http://ftp.cs.ucla.edu/pub/stat_ser/r475.pdf
Refs: http://bayes.cs.ucla.edu/jp_home.html

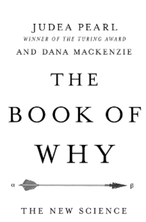
Every science that has thriven has thriven upon its own symbols
~Augustus de Morgan (1864)

THANK YOU

Joint work with:
Elias Bareinboim
Karthika Mohan
Ilya Shpitser
Jin Tian
Many more . . .

64

For a trailer, click WHY on my home page.
For discussions: @yudapearl



65

66

67

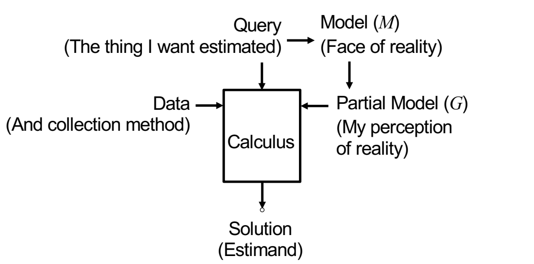
DATA SCIENCE − A CLASH OF TWO PARADIGMS

1. **The data-centric paradigm**
   - How best to fit the data so as to maximize some "**goodness of fit**" measure over the same data.
2. **The scientific paradigm**
   - How best to answer questions about the **world**, given the available data and what we know about the **world**.

Alternatively: Extract understanding from Data

68

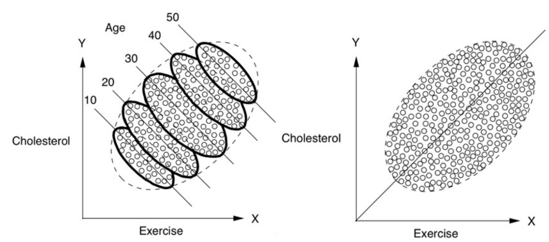
THE SCIENTIFIC PARADIGM



A problem is solvable if my model permits the calculus to express the Query in terms of the available data.

69

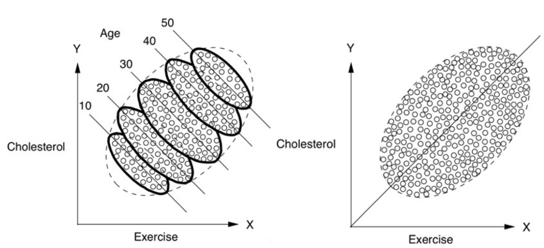
WHY THE LOGIC OF ASSOCIATION IS INSUFFICIENT



Exercise is helpful in every age group but harmful for a typical person. Why not?

70

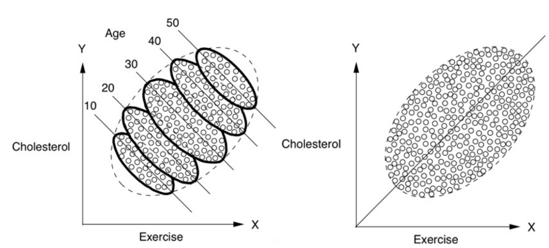
WHY THE LOGIC OF ASSOCIATION IS INSUFFICIENT



Exercise is helpful in every age group but harmful for a typical person. Is exercise helpful or not?

71

WHY THE LOGIC OF ASSOCIATION IS INSUFFICIENT



Exercise is helpful in every age group but harmful for a typical person. Is exercise helpful or not? More specific? What about seatbelt usage?

72

FORMULATING ASSUMPTIONS THE MERITS OF THE FIRST LAW

1. English: Smoking (X), Cancer (Y), Tar (Z), Genotypes (U)

2. Counterfactuals:
$$Z_x(u) = Z_{yx}(u)$$
$$X_y(u) = X_{zy}(u) = X_z = X(u),$$
$$Y_z(u) = Y_{zx}(u),$$
$$Z_x(u) \perp\!\!\!\perp \{Y_z, X\}$$

Not too friendly:
Consistent?, complete?, redundant?, arguable?

4. Structural:

$$X \longrightarrow Z \longrightarrow Y$$

73

EFFECT OF WARM-UP ON INJURY IN POTENTIAL-OUTCOME FRAMEWORK

X Warm-up → $P(X|Y, Z, Y(1), Y(0))$ → Y Injury

Z:
Team motivation
Coach
Pre-game proprioception
Fitness level
Genetics
Neuromuscular fatigue
Connective tissue disorder
Tissue weakness

74

TYPICAL INFERENCE IN *N-R* MODEL

Find $P^*(Y_x{=}y)$ given covariate Z,

$$P^*\,(Y_x = y) = \sum_z P^*\big(Y_x = y|z)P(z)\big)$$

Assume ignorability: $Y_x \perp\!\!\!\perp X|Z$
$$= \sum_z P^*\big(Y_x = y|x,z)P(z)\big)$$

Assume consistency: $X = x \Rightarrow Y_x = Y$
$$= \sum_z P^*\big(Y = y|x,z)P(z)\big)$$
$$= \sum_z P\big(y|x,z)P(z)\big)$$

Problems:
1) $Y_x \perp\!\!\!\perp X|Z$ judgmental & opaque
2) Is consistency the only connection between X, Y and Y_x?

75

TYPICAL INFERENCE IN *N-R* MODEL

Find $P^*(Y_x{=}y)$ given covariate Z,

$$P^*\,(Y_x = y) = \sum_z P^*\big(Y_x = y|z)P(z)\big)$$

Assume ignorability: $Y_x \perp\!\!\!\perp X|Z$
$$= \sum_z P^*\big(Y_x = y|x,z)P(z)\big)$$

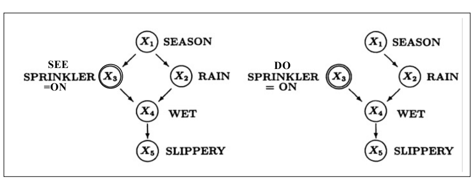
Assume consistency: $X = x \Rightarrow Y_x = Y$
$$= \sum_z P^*\big(Y = y|x,z)P(z)\big)$$
$$= \sum_z P\big(y|x,z)P(z)\big)$$

? Try it: $X \to Y \to Z$

Problems:
1) $Y_x \perp\!\!\!\perp X|Z$ judgmental & **opaque**
2) Is consistency the **only** connection between X, Y and Y_x?

76

THE SECRET TO CAUSAL REASONING DISTINGUISH SEEING FROM DOING

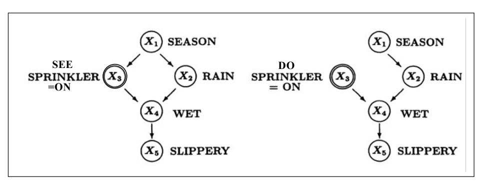


What if we see the Sprinkler ON?

What if we turn the Sprinkler ON?

77

THE SECRET TO CAUSAL REASONING DISTINGUISH SEEING FROM DOING



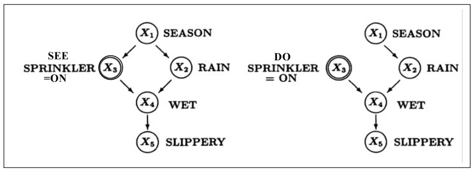
What if we see the Sprinkler ON?

What if we turn the Sprinkler ON?

- Actions can be simulated analytically by “blocking backdoor paths” = adjustmentz

78

THE SECRET TO CAUSAL REASONING
DISTINGUISH SEEING FROM DOING

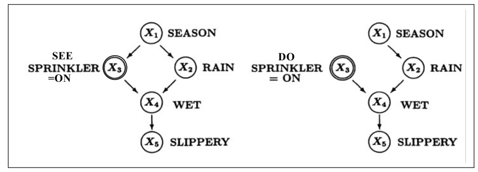


What if we see the Sprinkler ON?

What if we turn the Sprinkler ON?
What if the Sprinkler were ON?

79

THE SECRET TO CAUSAL REASONING
DISTINGUISH SEEING FROM DOING



What if we see the Sprinkler ON?

3 steps to counterfactuals

What if we turn the Sprinkler ON?
What if the Sprinkler were ON?

80

THE SEVEN PILLARS
OF CAUSAL WISDOM

Pillar 1: Graphical models for prediction and diagnosis
Pillar 2: Deconfounding policy analysis
Pillar 3: Algorithmitization of counterfactuals
Pillar 4: Mediation analysis and the assessment of direct and indirect effects
Pillar 5: External validity and sample selection bias
Pillar 6: Missing data
Pillar 7: Causal discovery

81

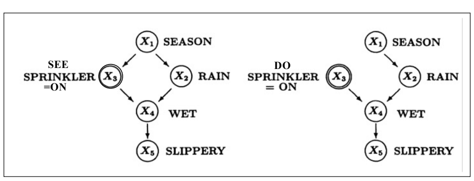
WHAT CAPABILITIES DOES
DEEP UNDERSTANDING ENTAIL?

A state of knowledge evoking a sensation of "**being in control**."

1. Predict future events from past/present **observations**
2. Predict consequence of contemplated **actions**
3. Provide **explanations** of unanticipated events
4. **Imagine** alternative worlds or "Roads not Taken"
5. Design new experiments, seek new observations (attention, **curiosity**, and conjectures)

82

THE SECRET TO CAUSAL REASONING
DISTINGUISH SEEING FROM DOING

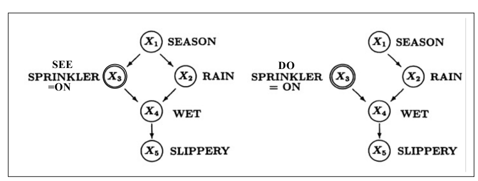


What if we **see** the Sprinkler ON?

What if we **turn** the Sprinkler ON?

83

THE SECRET TO CAUSAL REASONING
DISTINGUISH SEEING FROM DOING

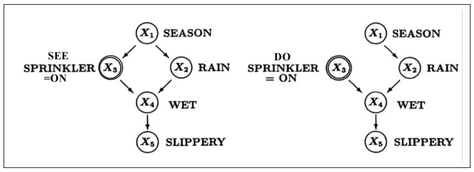


What if we **see** the Sprinkler ON?

What if we **turn** the Sprinkler ON?
What if the Sprinkler **were** ON?

84

THE SECRET TO CAUSAL REASONING
DISTINGUISH SEEING FROM DOING



What if we **see** the Sprinkler ON?

What if we **turn** the Sprinkler ON?

What if the Sprinkler **were** ON?

3 steps to **counterfactuals**

85