CAUSAL INFERENCE IN STATISTICS

Judea Pearl
University of California
Los Angeles
(www.cs.ucla.edu/~judea/jsm09)

• Inference: Statistical vs. Causal, distinctions, and mental barriers
• Unified conceptualization of counterfactuals, structural-equations, and graphs
• Inference to three types of claims:
  1. Effect of potential interventions
  2. Attribution (Causes of Effects)
  3. Direct and indirect effects
• Frills

OUTLINE

TRADITIONAL STATISTICAL INFERENCE PARADIGM

Data
\[ P \]
Joint Distribution
\[ Q(P) \]
(Aspects of \( P \))
Inference

e.g.,
Infer whether customers who bought product \( A \) would also buy product \( B \).
\[ Q = P(B \mid A) \]

FROM STATISTICAL TO CAUSAL ANALYSIS: 1. THE DIFFERENCES

<table>
<thead>
<tr>
<th>Statistical</th>
<th>Causal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability and statistics deal with static relations</td>
<td>Probability and statistics deal with static relations</td>
</tr>
<tr>
<td>Data</td>
<td>Data</td>
</tr>
<tr>
<td>[ P ] Joint Distribution</td>
<td>[ P ] Joint Distribution</td>
</tr>
<tr>
<td>Inference</td>
<td>Inference</td>
</tr>
<tr>
<td>[ Q(P) ] (Aspects of ( P ))</td>
<td>[ Q(P') ] (Aspects of ( P' ))</td>
</tr>
<tr>
<td>What happens when ( P ) changes?</td>
<td>What happens when ( P ) changes?</td>
</tr>
<tr>
<td>e.g., Infer whether customers who bought product ( A ) would still buy ( A ) if we were to double the price.</td>
<td>e.g., Infer whether customers who bought product ( A ) would still buy ( A ) if we were to double the price.</td>
</tr>
</tbody>
</table>

FROM STATISTICAL TO CAUSAL ANALYSIS: 1. THE DIFFERENCES (CONT)

What remains invariant when \( P \) changes say, to satisfy \( P'(\text{price}=2)=1 \)

Data
\[ P \]
Joint Distribution
\[ P \]
Joint Distribution
\[ Q(P') \]
(Aspects of \( P' \))
Inference

Note: \( P'(v) \neq P(v \mid \text{price} = 2) \)
\( P \) does not tell us how it ought to change
e.g. Curing symptoms vs. curing diseases
e.g. Analogy: mechanical deformation
**FROM STATISTICAL TO CAUSAL ANALYSIS: 2. MENTAL BARRIERS**

1. Causal and statistical concepts do not mix.
   - CAUSAL
     - Spurious correlation
     - Randomization / Intervention
     - Confounding / Effect
     - Instrumental variable
     - Strong Exogeneity
     - Explanatory variables
   - STATISTICAL
     - Regression
     - Association / Independence
     - "Controlling for" / Conditioning
     - Odd and risk ratios
     - Collapsibility / Granger causality
     - Propensity score

2. No causes in – no causes out (Cartwright, 1989)
   - statistical assumptions + data
   - causal assumptions
   - causal conclusions


4. Non-standard mathematics:
   - a) Structural equation models (Wright, 1920; Simon, 1960)
   - b) Counterfactuals (Neyman-Rubin, Lewis)

---

**WHY CAUSALITY NEEDS SPECIAL MATHEMATICS**

Scientific Equations (e.g., Hooke’s Law) are non-algebraic

- e.g., Length ($Y$) equals a constant (2) times the weight ($X$)

\[
Y = 2X \quad \quad \quad X = 1 \\
X = 1 \quad \quad \quad Y = 2 \\
\text{Process information} \quad \quad \quad \text{The solution}
\]

Had $X$ been 3, $Y$ would be 6.
If we raise $X$ to 3, $Y$ would be 6.
Must “wipe out” $X = 1$.

---

**WHY CAUSALITY NEEDS SPECIAL MATHEMATICS**

Scientific Equations (e.g., Hooke’s Law) are non-algebraic

- e.g., Length ($Y$) equals a constant (2) times the weight ($X$)

\[
Y \leftarrow 2X \quad \quad \quad X = 1 \\
X = 1 \quad \quad \quad Y = 2 \\
\text{Process information} \quad \quad \quad \text{The solution}
\]

Had $X$ been 3, $Y$ would be 6.
If we raise $X$ to 3, $Y$ would be 6.
Must “wipe out” $X = 1$. 

---

**WHY CAUSALITY NEEDS SPECIAL MATHEMATICS**

Scientific Equations (e.g., Hooke’s Law) are non-algebraic

- e.g., Length ($Y$) equals a constant (2) times the weight ($X$)

Correct notation:

\[
Y := 2X \quad \quad \quad X = 1 \\
X = 1 \quad \quad \quad Y = 2 \\
\text{Process information} \quad \quad \quad \text{The solution}
\]

Had $X$ been 3, $Y$ would be 6.
If we raise $X$ to 3, $Y$ would be 6.
Must “wipe out” $X = 1$. 

---

**WHY CAUSALITY NEEDS SPECIAL MATHEMATICS**

Scientific Equations (e.g., Hooke’s Law) are non-algebraic

- e.g., Length ($Y$) equals a constant (2) times the weight ($X$)

Correct notation:

\[
Y \leftarrow 2X \quad \quad \quad X = 1 \\
X = 1 \quad \quad \quad Y = 2 \\
\text{Process information} \quad \quad \quad \text{The solution}
\]

Had $X$ been 3, $Y$ would be 6.
If we raise $X$ to 3, $Y$ would be 6.
Must “wipe out” $X = 1$. 

---
THE STRUCTURAL MODEL PARADIGM

Joint Distribution → Data Generating Model → Q(M)

Inference

M – Invariant strategy (mechanism, recipe, law, protocol) by which Nature assigns values to variables in the analysis.

“Think Nature, not experiment!”

FAMILIAR CAUSAL MODEL ORACLE FOR MANIPULATION

INPUT... OUTPUT

STRUCTURAL CAUSAL MODELS

Definition: A structural causal model is a 4-tuple \( \langle V, U, F, P(u) \rangle \), where
- \( V = \{V_1, \ldots, V_n\} \) are endogenous variables
- \( U = \{U_1, \ldots, U_m\} \) are background variables
- \( F = \{f_1, \ldots, f_n\} \) are functions determining \( V \), \( v_i = f_i(v, u) \)
- \( P(u) \) is a distribution over \( U \)

\( P(u) \) and \( F \) induce a distribution \( P(v) \) over observable variables

Example: Price – Quantity equations in economics

\[ q = b_1 p + d_1 q + u_1 \]
\[ p = b_2 q + d_2 w + u_2 \]

STRUCTURAL MODELS AND INTERVENTION

Let \( X \) be a set of variables in \( V \).
The action \( \text{do}(x) \) sets \( X \) to constants \( x \) regardless of the factors which previously determined \( X \).

\( \text{do}(x) \) replaces all functions \( f_i \) determining \( X \) with the constant functions \( X=x \), to create a mutilated model \( M_x \).

\[ q = b_1 p + d_1 q + u_1 \]
\[ p = b_2 q + d_2 w + u_2 \]

STRUCTURAL MODELS AND INTERVENTION

Let \( X \) be a set of variables in \( V \).
The action \( \text{do}(x) \) sets \( X \) to constants \( x \) regardless of the factors which previously determined \( X \).

\( \text{do}(x) \) replaces all functions \( f_i \) determining \( X \) with the constant functions \( X=x \), to create a mutilated model \( M_x \).

\[ q = b_1 p + d_1 q + u_1 \]
\[ p = b_2 q + d_2 w + u_2 \]
CAUSAL MODELS AND COUNTERFACTUALS

Definition:
The sentence: “Y would be y (in situation u), had X been x,” denoted $Y(u) = y$, means:
The solution for Y in a mutilated model $M_x$, (i.e., the equations for X replaced by $X = x$) with input $U = u$, is equal to y.

The Fundamental Equation of Counterfactuals:

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

REGRESSION VS. STRUCTURAL EQUATIONS
(The Confusion of the Century)

Regression (claimless, nonfalsifiable):

$$Y = ax + \epsilon$$

Structural (empirical, falsifiable):

$$Y = bx + u_Y$$

Claim: (regardless of distributions):

$$E(Y \mid do(x)) = E(Y \mid do(x), do(z)) = bx$$

The mothers of all questions:

Q. When would $b$ equal $a$?
A. When all back-door paths are blocked, $(u_Y \perp X)$

Q. When is $b$ estimable by regression methods?
A. Graphical criteria available

THE FOUR NECESSARY STEPS FOR EFFECT ESTIMATION

Define:
Express the target quantity $Q$ as a function $Q(M)$ that can be computed from any model $M$.

Assume:
Formulate causal assumptions using ordinary scientific language and represent their structural part in graphical form.

Identify:
Determine if $Q$ is identifiable.

Estimate:
Estimate $Q$ if it is identifiable; approximate it, if it is not.

THE FOUR NECESSARY STEPS FOR CAUSAL ANALYSIS

Define:
Express the target quantity $Q$ as a function $Q(M)$ that can be computed from any model $M$.

Assume:
Formulate causal assumptions using ordinary scientific language and represent their structural part in graphical form.

Identify:
Determine if $Q$ is identifiable.

Estimate:
Estimate $Q$ if it is identifiable; approximate it, if it is not.

THE FOUR NECESSARY STEPS FOR EFFECT ESTIMATION

Define:
Express the target quantity $Q$ as a function $Q(M)$ that can be computed from any model $M$.

Assume:
Formulate causal assumptions using ordinary scientific language and represent their structural part in graphical form.

Identify:
Determine if $Q$ is identifiable.

Estimate:
Estimate $Q$ if it is identifiable; approximate it, if it is not.
**THE FOUR NECESSARY STEPS FOR POLICY ANALYSIS**

**Define:** Express the target quantity $Q$ as a function $Q(M)$ that can be computed from any model $M$.

$P(Y_{X=g(Z)} = y)$ or $P(Y | do(x = g(z)))$

**Assume:** Formulate causal assumptions using ordinary scientific language and represent their structural part in graphical form.

**Identify:** Determine if $Q$ is identifiable.

**Estimate:** Estimate $Q$ if it is identifiable; approximate it, if it is not.

---

**THE FOUR NECESSARY STEPS FROM DEFINITION TO ASSUMPTIONS**

**Define:** Express the target quantity $Q$ as a function $Q(M)$ that can be computed from any model $M$.

$P(Y | do(x))$

**Assume:** Formulate causal assumptions using ordinary scientific language and represent their structural part in graphical form.

**Identify:** Determine if $Q$ is identifiable.

**Estimate:** Estimate $Q$ if it is identifiable; approximate it, if it is not.

---

**INFERRING THE EFFECT OF INTERVENTIONS**

The problem:
To predict the impact of a proposed intervention using data obtained prior to the intervention.

The solution (conditional):
Causal Assumptions + Data → Policy Claims

1. Mathematical tools for communicating causal assumptions formally and transparently.
2. Deciding (mathematically) whether the assumptions communicated are sufficient for obtaining consistent estimates of the prediction required.
3. Deriving (if (2) is affirmative) a closed-form expression for the predicted impact

---

**FORMULATING ASSUMPTIONS THREE LANGUAGES**

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

2. Counterfactuals: $Z_X(u) = Z_Y(u)$,$X(u) = X_Y(u) = X_Z(u) = X(u)$,$Y_X(u) = Y_Z(u)$,$Z_X \perp \{Y_Z, X\}$

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

4. Structural:

---

**INFERRING THE EFFECT OF INTERVENTIONS**

The problem:
To predict the impact of a proposed intervention using data obtained prior to the intervention.

The solution (conditional):
Causal Assumptions + Data → Policy Claims

1. Mathematical tools for communicating causal assumptions formally and transparently.
2. Deciding (mathematically) whether the assumptions communicated are sufficient for obtaining consistent estimates of the prediction required.
3. Suggesting (if (2) is negative) a set of measurements and experiments that, if performed, would render a consistent estimate feasible.
IDENTIFIABILITY

Definition:
Let $Q(M)$ be any quantity defined on a causal model $M$, and let $A$ be a set of assumption.

$Q$ is identifiable relative to $A$ iff

$$P(M_1) = P(M_2) \Rightarrow Q(M_1) = Q(M_2)$$

for all $M_1, M_2$ that satisfy $A$.

* 

THE PROBLEM OF CONFOUNDING

Find the effect of $X$ on $Y$, $P(y \mid do(x))$, given the causal assumptions shown in $G$, where $Z_1, \ldots, Z_k$ are auxiliary variables.

Can $P(y \mid do(x))$ be estimated if only a subset, $Z$, can be measured?

ELIMINATING CONFOUNDING BIAS

THE BACK-DOOR CRITERION

$P(y \mid do(x))$ is estimable if there is a set $Z$ of variables such that $Z \not\alpha$-separates $X$ from $Y$ in $G_x$.

Moreover, $P(y \mid do(x)) = \sum_z P(y \mid x, z) P(z)$ ("adjusting" for $Z$).

EFFECT OF INTERVENTION BEYOND ADJUSTMENT

Theorem (Tian-Pearl 2002)
We can identify $P(y \mid do(x))$ if there is no child $Z$ of $X$ connected to $X$ by a confounding path.
EFFECT OF WARM-UP ON INJURY (After Shrier & Platt, 2008)

EFFECT OF WARM-UP ON INJURY (After Shrier & Platt, 2008)

No, no!

• Complete calculus for reducing $P(y|do(x), z)$ to expressions void of $do$-operators.

• Complete graphical criterion for identifying causal effects (Shpitser and Pearl, 2006).

• Complete graphical criterion for empirical testability of counterfactuals (Shpitser and Pearl, 2007).

EFFECT OF INTERVENTION COMPLETE IDENTIFICATION

COUNTERFACTUALS AT WORK ETT – EFFECT OF TREATMENT ON THE TREATED

1. Regret:
   I took a pill to fall asleep.
   Perhaps I should not have?

2. Program evaluation:
   What would terminating a program do to those enrolled?

   $P(Y_x = y | x')$

THE FOUR NECESSARY STEPS EFFECT OF TREATMENT ON THE TREATED

Define: Express the target quantity $Q$ as a function $Q(M)$ that can be computed from any model $M$.

$\text{ETT} \triangleq P(Y_x = y | X = x')$

Assume: Formulate causal assumptions using ordinary scientific language and represent their structural part in graphical form.

Identify: Determine if $Q$ is identifiable.

Estimate: Estimate $Q$ if it is identifiable; approximate it, if it is not.
### ETT - IDENTIFICATION

Theorem (Shpitser-Pearl, 2009)

ETT is identifiable in $G$ iff $P(y \mid do(x), w)$ is identifiable in $G'$

Moreover, $ETT = P(Y_x = y \mid x') = P(y \mid do(x), w)$

Complete graphical criterion

### ETT - THE BACK-DOOR CRITERION

$P(Y_x = y \mid x')$ is identifiable in $G$ if there is a set $Z$ of variables such that $Z$ $d$-separates $X$ from $Y$ in $G'$

Moreover, $ETT = \sum_Z P(y \mid z) P(z \mid x')$

"Standardized morbidity"

### FROM IDENTIFICATION TO ESTIMATION

Define: Express the target quantity $Q$ as a function $Q(M)$ that can be computed from any model $M$.

Assume: Formulate causal assumptions using ordinary scientific language and represent their structural part in graphical form.

Identify: Determine if $Q$ is identifiable.

Estimate: Estimate $Q$ if it is identifiable; approximate it, if it is not.

### PROPENSITY SCORE ESTIMATOR

(Rosenbaum & Rubin, 1983)

$P(y \mid do(x)) = ?$

$\sum_L P(z \mid z_L) P(X = 1 \mid z_1, z_2, z_3, z_4, z_5)$

Theorem: $\sum_L P(y \mid z, x) P(z) = \sum_L P(y \mid L = l, x) P(L = l)$

Adjustment for $L$ replaces Adjustment for $Z$

### WHAT PROPENSITY SCORE (PS) PRACTITIONERS NEED TO KNOW

$L(z) = P(X = 1 \mid Z = z)$

$\sum_z P(y \mid z, x) P(z) = \sum_l P(y \mid l, x) P(l)$

1. The asymptotic bias of PS is EQUAL to that of ordinary adjustment (for same $Z$).
2. Including an additional covariate in the analysis CAN SPOIL the bias-reduction potential of others.
3. Choosing sufficient set for PS, requires knowledge about the model.

### WHICH COVARIATES MAY / SHOULD BE ADJUSTED FOR?

Question: Which of these eight covariates may be included in the propensity score function (for matching) and which should be excluded.

Answer:

Must include: Age

Must exclude: $B_1, M, B_2, \text{Follow-up, Assignment without Age}$

May include: Cost, Hygiene, {Assignment + Age}, {Hygiene + Age + $B_1$}, more . . .
WHAT PROPENSITY SCORE (PS) PRACTITIONERS NEED TO KNOW

1. The asymptotic bias of PS is EQUAL to that of ordinary adjustment (for same $Z$).
2. Including an additional covariate in the analysis CAN SPOIL the bias-reduction potential of others.
3. Choosing sufficient set for PS, requires knowledge about the model.
4. That any empirical test of the bias-reduction potential of PS, can only be generalized to cases where the causal relationships among covariates, observed and unobserved is the same.

\[ L(z) = P(X = 1 \mid Z = z) \]
\[ \sum P(y \mid z, x) P(z) = \sum P(y \mid l, x) P(l) \]

TWO PARADIGMS FOR CAUSAL INFERENCE

Observed: \( P(X, Y, Z,...) \)
Conclusions needed: \( P(Y,=y), P(X,=x \mid Z=Z) \)

How do we connect observables, \( X,Y,Z,... \) to counterfactuals \( Yx, Xz, Zy, ... \)?

N-R model
Counterfactuals are primitives, new variables
Super-distribution
\( P^*(X, Y, ..., X_x, X_z, ...) \)
\( X, Y, Z \) constrain \( Y, Z_y, ... \)

Structural model
Counterfactuals are derived quantities
Subscripts modify the model and distribution
\( P(Y, =y) = P_{M_x}(Y = y) \)

“SUPER” DISTRIBUTION IN N-R MODEL

<table>
<thead>
<tr>
<th>( X )</th>
<th>( Y )</th>
<th>( Z )</th>
<th>( Y_{e=0} )</th>
<th>( Y_{e=1} )</th>
<th>( X_{e=0} )</th>
<th>( X_{e=1} )</th>
<th>( U )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>( u_1 )</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>( u_2 )</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>( u_3 )</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>( u_4 )</td>
</tr>
</tbody>
</table>

inconsistency: \( x = 0 \Rightarrow Y_{e=0} = Y = x_Y + (1-x) Y_0 \)

Defines:
\( P^*(X, Y, Z, ... Y_x, Z_y, ... Y_x, Z_x, ... \) \)
\( P^*(Y_x = y \mid Z, X_x) \)
\( Y_x \perp \perp X \mid Z \)

THE FOUR NECESSARY STEPS IN POTENTIAL-OUTCOME FRAMEWORK

Define: Express the target quantity \( Q \) as a counterfactual formula
Assume: Formulate causal assumptions using the distribution:
\( P(X \mid Y, Z, Y(1), Y(0)) \)
Identify: Determine if \( Q \) is identifiable.
Estimate: Estimate \( Q \) if it is identifiable; approximate it, if it is not.

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

Find \( P^*(Y = y) \) given covariate \( Z \),
\( P^*(Y_x = y) = \sum P^*(Y_x = y \mid z) P(z) \)
Assume ignorability:
\( Y_x \perp \perp X \mid Z \)
Assume consistency:
\( X \Rightarrow Y_x = Y \)
\( P(Y_x = y) = \sum P(y \mid x, z) P(z) \)

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

TYPICAL INFERENCE IN N-R MODEL

Find \( P^*(Y = y) \) given covariate \( Z \),
**GRAPHICAL – COUNTERFACTUALS SYMBIOSIS**

Every causal graph expresses counterfactuals assumptions, e.g., \( X \rightarrow Y \rightarrow Z \)

1. Missing arrows \( Y \leftarrow Z \)
2. Missing arcs \( Y \perp Z \)

consistent, and readable from the graph.

Every theorem in SCM is a theorem in Potential-Outcome Model, and conversely.

**DEMYSTIFYING STRONG IGNORABILITY**

\( \{Y(0), Y(1)\} \perp X \mid Z \) (Ignorability)

\( P(y \mid do(x)) = \sum_{z} P(y \mid z, x) P(z) \) (Z-admissibility)

\( (X \perp \!\!\!\!\!\!\perp Y \mid Z)_{G} \) (Back-door)

Is there a \( W \) in \( G \) such that \( (W \perp \!\!\!\!\!\!\perp X \mid Z)_{G} \Rightarrow \) Ignorability?

**DETERMINING THE CAUSES OF EFFECTS**

* Your Honor! My client (Mr. A) died because he used that drug.

* Court to decide if it is more probable than not that \( A \) would be alive but for the drug!

\( PN = P(\text{?} \mid A \text{ is dead, took the drug}) > 0.50 \)

**THE ATTRIBUTION PROBLEM**

Definition:

1. What is the meaning of \( PN(x, y) \):
   “Probability that event \( y \) would not have occurred if it were not for event \( x \), given that \( x \) and \( y \) did in fact occur.”

Answer:

\[ PN(x, y) = P(Y_{x} = y' \mid x, y) \]

Computable from \( M \)
**TYPICAL THEOREMS**
(Tian and Pearl, 2000)

- Bounds given combined non-experimental and experimental data
  \[ \max \left\{ \frac{P(y) - P(y|\bar{x}_{y})}{P(x,y)} \right\} \leq PN \leq \min \left\{ \frac{1}{P(x,y)} \right\} \]
- Identifiability under monotonicity (Combined data)
  \[ PN = \frac{P(y|x) - P(y|x')}{P(y|x)} + \frac{P(y|x') - P(y|x')}{P(x,y)} \]
  corrected Excess-Risk-Ratio

**CAN FREQUENCY DATA DECIDE LEGAL RESPONSIBILITY?**

<table>
<thead>
<tr>
<th>Experimental</th>
<th>Nonexperimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deaths (y)</td>
<td>16 14 2 28</td>
</tr>
<tr>
<td>Survivals (y')</td>
<td>984 986 988 972</td>
</tr>
<tr>
<td>Total</td>
<td>1,000 1,000 1,000 1,000</td>
</tr>
</tbody>
</table>

- Non-experimental data: drug usage predicts longer life
- Experimental data: drug has negligible effect on survival
- Plaintiff: Mr. A is special.
  1. He actually died
  2. He used the drug by choice
- Court to decide (given both data):
  Is it more probable than not that $A$ would be alive but for the drug?

**SOLUTION TO THE ATTRIBUTION PROBLEM**

- WITH PROBABILITY ONE $1 \leq P(y^*|x,y) \leq 1$
- Combined data tell more that each study alone

**EFFECT DECOMPOSITION**
(direct vs. indirect effects)

1. Why decompose effects?
2. What is the definition of direct and indirect effects?
3. What are the policy implications of direct and indirect effects?
4. When can direct and indirect effect be estimated consistently from experimental and non-experimental data?

**WHY DECOMPOSE EFFECTS?**

1. To understand how Nature works
2. To comply with legal requirements
3. To predict the effects of new type of interventions:
   Signal routing, rather than variable fixing

**LEGAL IMPLICATIONS OF DIRECT EFFECT**

Can data prove an employer guilty of hiring discrimination?

\[ E(Y \mid do(x), do(z)) - E(Y \mid do(x), do(z)) \]
(averaged over $z$) Adjust for $Z$? No! No!

\[ (\text{Gender}) \quad X \xrightarrow{\text{Z}} \quad Z \ (\text{Qualifications}) \]
\[ (\text{Hiring}) \quad Y \]

What is the direct effect of $X$ on $Y$?
NATURAL INTERPRETATION OF AVERAGE DIRECT EFFECTS

Robins and Greenland (1992) – “Pure”

$X \rightarrow Z \leftarrow Y$

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 $Z$ constant at whatever value it attained before the change.

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

In linear models, $DE = \text{Controlled Direct Effect} = \beta(x_1 - x_0)$

DEFINITION AND IDENTIFICATION OF NESTED COUNTERFACTUALS

Consider the quantity $Q = E_0[Y_{Z(x_1)\{u\}}(u)]$

Given $(M, P(u))$, $Q$ is well defined

Given $u$, $Z_{x_1}(u)$ is the solution for $Z$ in $M_{x_1}$, call it $z = f(x, u)$

$Y_{Z(x_1)\{u\}}(u)$ is the solution for $Y$ in $M_{x_1}$

Can $Q$ be estimated from experimental and nonexperimental data?

Experimental: nest-free expression

Nonexperimental: subscript-free expression

DEFINITION OF INDIRECT EFFECTS

$X \rightarrow Z \leftarrow Y$

Indirect Effect of $X$ on $Y$: $IE(x_0, x_1; Y)$

The expected change in $Y$ when we keep $X$ constant, say at $x_0$, and let $Z$ change to whatever value it would have attained had $X$ changed to $x_1$.

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

In linear models, $IE = TE - DE$

POLICY IMPLICATIONS OF INDIRECT EFFECTS

What is the indirect effect of $X$ on $Y$?

The effect of Gender on Hiring if sex discrimination is eliminated.

GENDER $X$ QUALIFICATION

IGNORE $Z$ QUALIFICATION

HIRING $Y$

Blocking a link – a new type of intervention

EXPERIMENTAL IDENTIFICATION OF NATURAL DIRECT EFFECTS

Theorem: If there exists a set $W$ such that $Y_{Zx} \perp Z_{x^*} \mid W$ for all $z$ and $x$

Then the average direct effect

$DE(x, x^*; Y) = E[Y_{x}Z_{x^*}] - E[Y_{x^*}]$

Is identifiable from experimental data and is given by

$DE(x, x^*; Y) = \sum_{W,z} E[Y_{Zx} \mid w] - E[Y_{Zx^*} \mid w]P(Z_{x^*} = z \mid w)P(w)$

Example:

GRAPHICAL CONDITION FOR EXPERIMENTAL IDENTIFICATION OF DIRECT EFFECTS

Theorem: If there exists a set $W$ such that $(Y \perp Z \mid W)_{G_{xZ}}$ and $W \subseteq ND(X \cup Z)$

then,

$DE(x, x^*; Y) = \sum_{w,z} E[Y_{Zx} \mid w] - E[Y_{Zx^*} \mid w]P(Z_{x^*} = z \mid w)P(w)$

Example:
CONCLUSIONS

I TOLD YOU CAUSALITY IS SIMPLE

CONCLUSIONS

• Formal basis for causal and counterfactual inference (complete)
• Unification of the graphical, potential-outcome and structural equation approaches
• Friendly and formal solutions to century-old problems and confusions.

CONCLUSIONS

He is wise who bases causal inference on an explicit causal structure that is defensible on scientific grounds.

(Aristotle 384-322 B.C.)

From Charlie Poole

QUESTIONS???

They will be answered