Causal Inference

By: Miguel A. Hernán and James M. Robins

Part I: Causal inference without models

19th March, 2014

Outline

Chapter 1: A definition of causal effects

Purpose of Chapter 1:

"... is to introduce mathematical notation that formalizes the causal intuition that you already possess."

Chapter 1.1: Individual causal effects

Some notation

- Dichotomous treatment variable: A (1: treated; 0: untreated)
- Dichotomous outcome variable: Y (1: death; 0: survival)
- $Y^{a=i}$: Outcome under treatment $a = i$, $i \in \{0, 1\}$.

Chapter 1.1: Individual causal effects

Some notation

- Dichotomous treatment variable: A (1: treated; 0: untreated)
- Dichotomous outcome variable: Y (1: death; 0: survival)
- $Y^{a=i}$: Outcome under treatment $a = i$, $i \in \{0, 1\}$.

Definition

Causal effect for an individual: Treatment A has a causal effect if

$$Y^{a=1} \neq Y^{a=0}.$$

Chapter 1.1: Individual causal effects

Examples

- Zeus: $Y^{a=1} = 1 \neq 0 = Y^{a=0} \Longrightarrow$ treatment has causal effect.

Chapter 1.1: Individual causal effects

Examples

- Zeus: $Y^{a=1} = 1 \neq 0 = Y^{a=0} \Longrightarrow$ treatment has causal effect.
- Hera: $Y^{a=1} = Y^{a=0} = 0 \Longrightarrow$ treatment has no causal effect.

Chapter 1.1: Individual causal effects

Examples

- Zeus: $Y^{a=1} = 1 \neq 0 = Y^{a=0} \Longrightarrow$ treatment has causal effect.
- Hera: $Y^{a=1} = Y^{a=0} = 0 \Longrightarrow$ treatment has no causal effect.

Definition

Consistency: If $A_i = a$, then $Y_i^a = Y^{A_i} = Y_i$.

Chapter 1.1: Individual causal effects

Examples

- Zeus: $Y^{a=1} = 1 \neq 0 = Y^{a=0} \Longrightarrow$ treatment has causal effect.
- Hera: $Y^{a=1} = Y^{a=0} = 0 \Longrightarrow$ treatment has no causal effect.

Definition

Consistency: If $A_i = a$, then $Y_i^a = Y^{A_i} = Y_i$.

Important:

- $Y^{a=0}$ and $Y^{a=1}$ are **counterfactual** outcomes.
- Only one can be observed, i.e., only one is **factual**.

Chapter 1.1: Individual causal effects

Examples

- Zeus: $Y^{a=1} = 1 \neq 0 = Y^{a=0} \Longrightarrow$ treatment has causal effect.
- Hera: $Y^{a=1} = Y^{a=0} = 0 \Longrightarrow$ treatment has no causal effect.

Definition

Consistency: If $A_i = a$, then $Y_i^a = Y^{A_i} = Y_i$.

Important:

- $Y^{a=0}$ and $Y^{a=1}$ are **counterfactual** outcomes.
- Only one can be observed, i.e., only one is **factual**.
- Hence, in general, individual effects **cannot** be identified.

Chapter 1.2: Average causal effects

Chapter 1.2: Average causal effects

An example: Zeus's extended family

	$Y^{a=0}$	$Y^{a=1}$		$Y^{a=0}$	$Y^{a=1}$
Rheia	0	1	Leto	0	1
Kronos	1	0	Ares	1	1
Demeter	0	0	Athena	1	1
Hades	0	0	Hephaestus	0	1
Hestia	0	0	Aphrodite	0	1
Poseidon	1	0	Cyclope	0	1
Hera	0	0	Persephone	1	1
Zeus	0	1	Hermes	1	0
Artemis	1	1	Hebe	1	0
Apollo	1	0	Dionysus	1	0

Chapter 1.2: Average causal effects

Definition

Average causal effect is present if

$$\Pr(Y^{a=1} = 1) \neq \Pr(Y^{a=0} = 1).$$

Chapter 1.2: Average causal effects

Definition

Average causal effect is present if

$$\Pr(Y^{a=1} = 1) \neq \Pr(Y^{a=0} = 1).$$

More generally (nondichotomous outcomes):

$$\mathrm{E}(Y^{a=1}) \neq \mathrm{E}(Y^{a=0}).$$

Chapter 1.2: Average causal effects

Definition

Average causal effect is present if

$$\Pr(Y^{a=1} = 1) \neq \Pr(Y^{a=0} = 1).$$

More generally (nondichotomous outcomes):

$$\mathrm{E}(Y^{a=1}) \neq \mathrm{E}(Y^{a=0}).$$

Example:

No average causal effect in Zeus's family:

$$\Pr(Y^{a=1} = 1) = \Pr(Y^{a=0} = 1) = 10/20 = 0.5.$$

Chapter 1.2: Average causal effects

Definition

Average causal effect is present if

$$\Pr(Y^{a=1} = 1) \neq \Pr(Y^{a=0} = 1).$$

More generally (nondichotomous outcomes):

$$\mathrm{E}(Y^{a=1}) \neq \mathrm{E}(Y^{a=0}).$$

Example:

No average causal effect in Zeus's family:

$$\Pr(Y^{a=1} = 1) = \Pr(Y^{a=0} = 1) = 10/20 = 0.5.$$

That does **not** imply the absence of individual effects.

Fine Points

Fine point 1.1: Interference between subjects

Fine Points

Fine point 1.1: Interference between subjects

- Present if outcome depends on other subjects' treatment value.

Fine Points

Fine point 1.1: Interference between subjects

- Present if outcome depends on other subjects' treatment value.
- Implies that Y_i^a is not well defined.

Fine Points

Fine point 1.1: Interference between subjects

- Present if outcome depends on other subjects' treatment value.
- Implies that Y_i^a is not well defined.
- Book assumes "stable-unit-treatment-value assumption (SUTVA)" (Rubin 1980)

Fine Points

Fine point 1.1: Interference between subjects

- Present if outcome depends on other subjects' treatment value.
- Implies that Y_i^a is not well defined.
- Book assumes "stable-unit-treatment-value assumption (SUTVA)" (Rubin 1980)

Fine point 1.2: Multiple versions of treatment

Fine Points

Fine point 1.1: Interference between subjects

- Present if outcome depends on other subjects' treatment value.
- Implies that Y_i^a is not well defined.
- Book assumes "stable-unit-treatment-value assumption (SUTVA)" (Rubin 1980)

Fine point 1.2: Multiple versions of treatment

- Different versions of treatment could exist.

Fine Points

Fine point 1.1: Interference between subjects

- Present if outcome depends on other subjects' treatment value.
- Implies that Y_i^a is not well defined.
- Book assumes "stable-unit-treatment-value assumption (SUTVA)" (Rubin 1980)

Fine point 1.2: Multiple versions of treatment

- Different versions of treatment could exist.
- Implies that Y_i^a is not well defined.

Fine Points

Fine point 1.1: Interference between subjects

- Present if outcome depends on other subjects' treatment value.
- Implies that Y_i^a is not well defined.
- Book assumes "stable-unit-treatment-value assumption (SUTVA)" (Rubin 1980)

Fine point 1.2: Multiple versions of treatment

- Different versions of treatment could exist.
- Implies that Y_i^a is not well defined.
- Authors assume "treatment variation irrelevance throughout this book."

Chapter 1.3: Measures of causal effect

Representations of the causal null hypothesis

$$\Pr(Y^{a=1}=1) - \Pr(Y^{a=0}=1) = 0 \qquad \text{(Causal risk difference)}$$

$$\frac{\Pr(Y^{a=1}=1)}{\Pr(Y^{a=0}=1)} = 1 \qquad \text{(Causal risk ratio)}$$

$$\frac{\Pr(Y^{a=1}=1)/\Pr(Y^{a=1}=0)}{\Pr(Y^{a=0}=1)/\Pr(Y^{a=0}=0)} = 1 \qquad \text{(Causal odds ratio)}$$

Chapter 1.3: Measures of causal effect

Representations of the causal null hypothesis

$$\Pr(Y^{a=1}=1) - \Pr(Y^{a=0}=1) = 0 \qquad \text{(Causal risk difference)}$$

$$\frac{\Pr(Y^{a=1}=1)}{\Pr(Y^{a=0}=1)} = 1 \qquad \text{(Causal risk ratio)}$$

$$\frac{\Pr(Y^{a=1}=1)/\Pr(Y^{a=1}=0)}{\Pr(Y^{a=0}=1)/\Pr(Y^{a=0}=0)} = 1 \qquad \text{(Causal odds ratio)}$$

The effect measures quantify the possible causal effect on different scales.

Chapter 1.4: Random variability

Samples: Two sources of random error

- **Sampling variability:**
  We only dispose of $\widehat{\Pr}(Y^{a=1} = 1)$ and $\widehat{\Pr}(Y^{a=0} = 1)$. Statistical procedures are necessary to test the causal null hypothesis.

Chapter 1.4: Random variability

Samples: Two sources of random error

- **Sampling variability:**
  We only dispose of $\widehat{\Pr}(Y^{a=1} = 1)$ and $\widehat{\Pr}(Y^{a=0} = 1)$. Statistical procedures are necessary to test the causal null hypothesis.
- **Nondeterministic counterfactuals:**
  Counterfactual outcomes $Y^{a=1}$ and $Y^{a=0}$ may not be fixed, but rather stochastic.

Chapter 1.4: Random variability

Samples: Two sources of random error

- **Sampling variability:**
  We only dispose of $\widehat{\Pr}(Y^{a=1} = 1)$ and $\widehat{\Pr}(Y^{a=0} = 1)$. Statistical procedures are necessary to test the causal null hypothesis.
- **Nondeterministic counterfactuals:**
  Counterfactual outcomes $Y^{a=1}$ and $Y^{a=0}$ may not be fixed, but rather stochastic.

"Thus statistics is necessary in causal inference to quantify random error from sampling variability, nondeterministic counterfactuals, or both. However, for pedagogic reasons, we will continue to largely ignore statistical issues until Chapter 10."

Chapter 1.5: Causation versus association

A "real world" example

	A	Y		A	Y		A	Y
Rheia	0	0	Zeus	1	1	Aphrodite	1	1
Kronos	0	1	Artemis	0	1	Cyclope	1	1
Demeter	0	0	Apollo	0	1	Persephone	1	1
Hades	0	0	Leto	0	0	Hermes	1	0
Hestia	1	0	Ares	1	1	Hebe	1	0
Poseidon	1	0	Athena	1	1	Dionysus	1	0
Hera	1	0	Hephaestus	1	1			

Chapter 1.5: Causation versus association

A "real world" example

	A	Y		A	Y		A	Y
Rheia	0	0	Zeus	1	1	Aphrodite	1	1
Kronos	0	1	Artemis	0	1	Cyclope	1	1
Demeter	0	0	Apollo	0	1	Persephone	1	1
Hades	0	0	Leto	0	0	Hermes	1	0
Hestia	1	0	Ares	1	1	Hebe	1	0
Poseidon	1	0	Athena	1	1	Dionysus	1	0
Hera	1	0	Hephaestus	1	1			

$$\Pr(Y = 1|A = 1) = 7/13 = 0.54, \qquad \Pr(Y = 1|A = 0) = 3/7 = 0.43.$$

Chapter 1.5: Causation versus association

Association measures

$$\Pr(Y=1|A=1) - \Pr(Y=1|A=0) \quad \text{(Associational risk difference)}$$

$$\frac{\Pr(Y=1|A=1)}{\Pr(Y=1|A=0)} \quad \text{(Associational risk ratio)}$$

$$\frac{\Pr(Y=1|A=1)/\Pr(Y=0|A=1)}{\Pr(Y=1|A=0)/\Pr(Y=0|A=0)} \quad \text{(Associational odds ratio)}$$

Chapter 1.5: Causation versus association

Association measures

$$\Pr(Y=1|A=1) - \Pr(Y=1|A=0) \quad \text{(Associational risk difference)}$$

$$\frac{\Pr(Y=1|A=1)}{\Pr(Y=1|A=0)} \quad \text{(Associational risk ratio)}$$

$$\frac{\Pr(Y=1|A=1)/\Pr(Y=0|A=1)}{\Pr(Y=1|A=0)/\Pr(Y=0|A=0)} \quad \text{(Associational odds ratio)}$$

If $\Pr(Y=1|A=1) = \Pr(Y=1|A=0)$, then $A \coprod Y$ (A, Y independent).

Chapter 1.5: Causation versus association

Association measures

$$\Pr(Y=1|A=1) - \Pr(Y=1|A=0) \quad \text{(Associational risk difference)}$$

$$\frac{\Pr(Y=1|A=1)}{\Pr(Y=1|A=0)} \quad \text{(Associational risk ratio)}$$

$$\frac{\Pr(Y=1|A=1)/\Pr(Y=0|A=1)}{\Pr(Y=1|A=0)/\Pr(Y=0|A=0)} \quad \text{(Associational odds ratio)}$$

If $\Pr(Y=1|A=1) = \Pr(Y=1|A=0)$, then $A \coprod Y$ (A, Y independent).

Example: ARD $= 0.54 - 0.43 = 0.11$, ARR $= 0.54/0.43 = 1.26$.

Chapter 1.5: Causation versus association

$\Pr(Y = 1|A = 1)$ is a conditional, $\Pr(Y^a = 1)$ an unconditional probability.

Chapter 1.5: Causation versus association

$\Pr(Y = 1|A = 1)$ is a conditional, $\Pr(Y^a = 1)$ an unconditional probability.

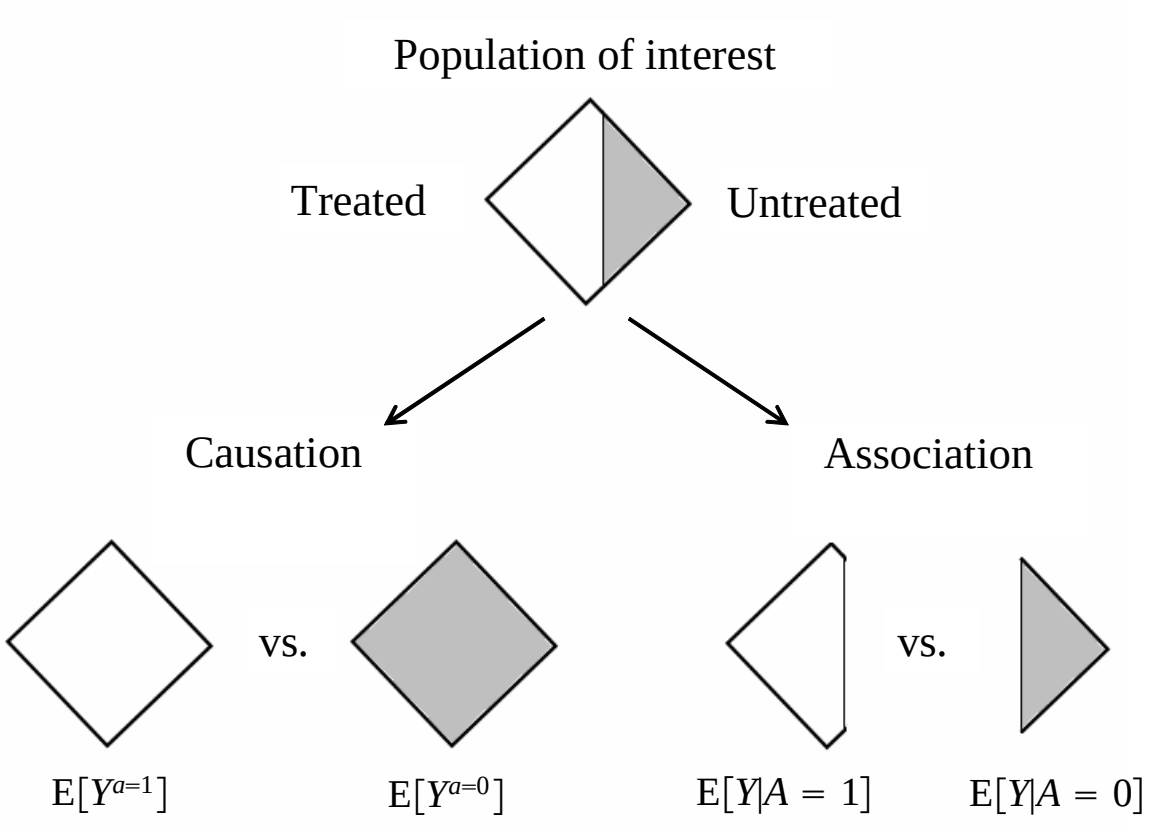


Figure: Association-causation difference

Chapter 1.5: Causation versus association

Concluding question:

"The question is then under which conditions real world data can be used for causal inference."

Chapter 1.5: Causation versus association

Concluding question:

> *"The question is then under which conditions real world data can be used for causal inference."*

CONTINUARÁ...

Chapter 2: Randomized experiments

Does your looking up at the sky make other pedestrians look up too?

Does your looking up at the sky make other pedestrians look up too?

- I can stand on the sidewalk.

Does your looking up at the sky make other pedestrians look up too?

- I can stand on the sidewalk.
- Flip a coin whenever someone approaches.

Does your looking up at the sky make other pedestrians look up too?

- I can stand on the sidewalk.
- Flip a coin whenever someone approaches.
  - If *heads*: I will look up intently

Does your looking up at the sky make other pedestrians look up too?

- I can stand on the sidewalk.
- Flip a coin whenever someone approaches.
  - If *heads*: I will look up intently
  - If *tails*: I will look straight ahead with an absentminded expression

Does your looking up at the sky make other pedestrians look up too?

- I can stand on the sidewalk.
- Flip a coin whenever someone approaches.
  - If *heads*: I will look up intently
  - If *tails*: I will look straight ahead with an absentminded expression
- Repeat the experiment a few thousand times.

Does your looking up at the sky make other pedestrians look up too?

- I can stand on the sidewalk.
- Flip a coin whenever someone approaches.
  - If *heads*: I will look up intently
  - If *tails*: I will look straight ahead with an absentminded expression
- Repeat the experiment a few thousand times.
- Count proportion of pedestrians who looked up within 10 seconds after I did and who did not

Does your looking up at the sky make other pedestrians look up too?

- I can stand on the sidewalk.
- Flip a coin whenever someone approaches.
  - If *heads*: I will look up intently
  - If *tails*: I will look straight ahead with an absentminded expression
- Repeat the experiment a few thousand times.
- Count proportion of pedestrians who looked up within 10 seconds after I did and who did not
  - If $Prop_{lu} > Prop_{nlu}$: My looking up has a causal effect

Was this solution a randomized experiment?

Was this solution a randomized experiment?

Yes

- It was an **experiment** because...
  - The investigator carried out the action of interest
- It was **randomized** because...
  - The decision to act on any study subject was made by a random device

Was this solution a randomized experiment?

Yes

- It was an **experiment** because...
  - The investigator carried out the action of interest
- It was **randomized** because...
  - The decision to act on any study subject was made by a random device

Could not

- You could have looked up when a man approached and looked straight when a woman did
- The assignment of the action would have followed a deterministic rule
  - Up for woman
  - Straight for man

Was this solution a randomized experiment?

Yes

- It was an **experiment** because...
  - The investigator carried out the action of interest
- It was **randomized** because...
  - The decision to act on any study subject was made by a random device

Could not

- You could have looked up when a man approached and looked straight when a woman did
- The assignment of the action would have followed a deterministic rule
  - Up for woman
  - Straight for man

Chapter 2:

"... describes why randomization results in convincing causal inferences."

Refreshing the notation

Observations

- Dichotomous treatment variable A
  - $A = 1$: treated
  - $A = 0$: untreated
- Dichotomous outcome variable: Y
  - $Y = 1$: death
  - $Y = 0$: survival

Refreshing the notation

Observations

- Dichotomous treatment variable A
  - $A = 1$: treated
  - $A = 0$: untreated
- Dichotomous outcome variable: Y
  - $Y = 1$: death
  - $Y = 0$: survival

Counterfactual outcomes

- $Y^{a=i}$: Outcome under treatment $a = i$, $i \in \{0, 1\}$.
  - $Y^{a=0} = 0$: **survival** outcome under **no treatment**
  - $Y^{a=0} = 1$: **death** outcome under **no treatment**
  - $Y^{a=1} = 0$: **survival** outcome under **treatment**
  - $Y^{a=1} = 0$: **death** outcome under **treatment**

2.1 Randomization

Missing values of the counterfactual outcomes

- Randomized experiments generate data with missing values (like any other real world study)
- Randomization ensures missing values occur by chance

	A	Y	Y^0	Y^1		A	Y	Y^0	Y^1
Rheia	0	0	0	?	Leto	0	0	0	?
Kronos	0	1	1	?	Ares	1	1	?	1
Demeter	0	0	0	?	Athena	1	1	?	1
Hades	0	0	0	?	Hephaestus	1	1	?	1
Hestia	1	0	?	0	Aphrodite	1	1	?	1
Poseidon	1	0	?	0	Cyclope	1	1	?	1
Hera	1	0	?	0	Persephone	1	1	?	1
Zeus	1	1	?	1	Hermes	1	0	?	0
Artemis	0	1	1	?	Hebe	1	0	?	0
Apollo	0	1	1	?	Dionysus	1	0	?	0

Ideal Randomized experiment

- No loss of follow-up
- Full adherence to the assigned treatment over the duration of the study
- A single version of treatment
- Double blind assignment

Flip a coin for each subject

- If *heads*: we assigned the subject to the white group
- If *tails*: we assigned the subject to the grey group

- white group= treated group
- grey group = untreated group

Exchangeability

What would have happened if the research assistants had misinterpreted our instructions and had treated the grey group rather the white group?

How does this reversal of treatment status affect our conclusions?

Exchangeability

What would have happened if the research assistants had misinterpreted our instructions and had treated the grey group rather the white group?

How does this reversal of treatment status affect our conclusions?

Not at all.

Exchangeability

Notation: $Y^a \amalg A$ *for all* a

Meaning: The counterfactual outcome and the observed treatment are independent.

i.e.: Treated and untreated would have experienced the same risk of death if they had received the same treatment.

Full and mean exchangeability

Full exchangeability

Randomization makes Y^a independent of $A \Rightarrow$, but $\not\Leftarrow$, exchangeability.

Full and mean exchangeability

Full exchangeability

Randomization makes Y^a independent of $A \Rightarrow$, but $\not\Leftarrow$, exchangeability.

Dichotomous outcome and treatment

- $Y^a \amalg A$, can be written as...
- $Pr[Y^a = 1|A = 1] = Pr[Y^a = 1|A = 0]$, or equivalently, as...
- $E[Y^a|A = 1] = E[Y^a|A = 0]$ for all a.
- The last equality is the **Mean Exchangeability**.

Full and mean exchangeability

Full exchangeability

Randomization makes Y^a independent of $A \Rightarrow$, but $\not\Leftarrow$, exchangeability.

Dichotomous outcome and treatment

- $Y^a \amalg A$, can be written as...
- $Pr[Y^a = 1|A = 1] = Pr[Y^a = 1|A = 0]$, or equivalently, as...
- $E[Y^a|A = 1] = E[Y^a|A = 0]$ for all a.
- The last equality is the **Mean Exchangeability**.

Continuous outcome

- Exchangeability $Y^a \amalg A$ ...
- $\Rightarrow$ **Mean Exchangeability** $E[Y^a|A = a'] = E[Y^a]$
- but **Mean Exchangeability** $\not\Rightarrow$ exchangeability (because distributional parameters other than the mean (e.g., variance) may not be independent of treatment)

Caution

- $Y^a \amalg A \not\Rightarrow Y \amalg A$

- Independence between **counterfactual outcome** and observed treatment $\not\Rightarrow$ independence between **observed outcome** and observed treatment.

Example

Does exchangeability hold in the heart transplant study?

We would need to check if $Y^a \amalg A$ holds for $a = 1$ and for $a = 0$

Table 1.1

	$Y^{a=0}$	$Y^{a=1}$
Rheia	0	1
Kronos	1	0
Demeter	0	0
Hades	0	0
Hestia	0	0
Poseidon	1	0
Hera	0	0
Zeus	0	1
Artemis	1	1
Apollo	1	0
Leto	0	1
Ares	1	1
Athena	1	1
Hephaestus	0	1
Aphrodite	0	1
Cyclope	0	1
Persephone	1	1
Hermes	1	0
Hebe	1	0
Dionysus	1	0

Table 2.1

	A	Y	Y^0	Y^1
Rheia	0	0	0	?
Kronos	0	1	1	?
Demeter	0	0	0	?
Hades	0	0	0	?
Hestia	1	0	?	0
Poseidon	1	0	?	0
Hera	1	0	?	0
Zeus	1	1	?	1
Artemis	0	1	1	?
Apollo	0	1	1	?
Leto	0	0	0	?
Ares	1	1	?	1
Athena	1	1	?	1
Hephaestus	1	1	?	1
Aphrodite	1	1	?	1
Cyclope	1	1	?	1
Persephone	1	1	?	1
Hermes	1	0	?	0
Hebe	1	0	?	0
Dionysus	1	0	?	0

(Suppose the counterfactual data in Table 1.1 are available to us)

Example (cont'd)

Does exchangeability hold in the heart transplant study?

Table 1.1

	$Y^{a=0}$	$Y^{a=1}$
Rheia	0	1
Kronos	1	0
Demeter	0	0
Hades	0	0
Hestia	0	0
Poseidon	1	0
Hera	0	0
Zeus	0	1
Artemis	1	1
Apollo	1	0
Leto	0	1
Ares	1	1
Athena	1	1
Hephaestus	0	1
Aphrodite	0	1
Cyclope	0	1
Persephone	1	1
Hermes	1	0
Hebe	1	0
Dionysus	1	0

Risk of death under no treatment in treated subjects:

$PR[Y^{a=0} = 1|A = 1] = 7/13$

Risk of death under no treatment in no treated subjects:

$PR[Y^{a=0} = 1|A = 0] = 3/7$

$7/13 > 3/7 \Rightarrow$ treated subjects have worse prognosis

$\Rightarrow$ treated and untreated ARE NOT EXCHANGEABLE.

Footnote

Only data in table 2.1 are available in the real world.
It is insufficient to compute counterfactual risks.
We are generally UNABLE to determine whether exchangeability holds in our study.

2.2 Conditional randomization

Example

Besides data on treatment (A) and outcome (Y), we also have data on the prognosis factor (L)

2.2 Conditional randomization

Example

Besides data on treatment (A) and outcome (Y), we also have data on the prognosis factor (L)

- **Design 1**:
  1. Randomly select 65% of individuals
  2. Transplant a new heart in selected individuals

2.2 Conditional randomization

Example

Besides data on treatment (A) and outcome (Y), we also have data on the prognosis factor (L)

- **Design 1**:
  1. Randomly select 65% of individuals
  2. Transplant a new heart in selected individuals
- **Design 2**:
  1. Randomly select 75% of individuals with prognosis $L = 1$
  2. Randomly select 50% of individuals with prognosis $L = 0$
  3. Transplant a new heart in selected individuals

Marginally and Conditionally Randomized Experiments

Observations:

- Both D1 and D2 are randomized experiments

Marginally and Conditionally Randomized Experiments

Observations:

- Both D1 and D2 are randomized experiments
- D1 uses 1 coin:
  - We use a **single unconditional** (marginal) randomization probability common to all subjects
  - D1 are **marginally randomized experiments**

Marginally and Conditionally Randomized Experiments

Observations:

- Both D1 and D2 are randomized experiments
- D1 uses 1 coin:
  - We use a **single unconditional** (marginal) randomization probability common to all subjects
  - D1 are **marginally randomized experiments**
- D2 uses 2 different coins:
  - We use **several** randomization probabilities **that depend** (are conditionally) on the values of the variable L
  - D2 are **conditionally randomized experiments**

Marginally and Conditionally Randomized Experiments

Observations:

- Both D1 and D2 are randomized experiments
- D1 uses 1 coin:
  - We use a **single unconditional** (marginal) randomization probability common to all subjects
  - D1 are **marginally randomized experiments**
- D2 uses 2 different coins:
  - We use **several** randomization probabilities **that depend** (are conditionally) on the values of the variable L
  - D2 are **conditionally randomized experiments**
- D2 is simply the combination of two separate marginally randomized experiments: one over subjects $L = 1$, the other in subjects with $L = 0$.

Marginally and Conditionally Randomized Experiments

Observations:

- Both D1 and D2 are randomized experiments
- D1 uses 1 coin:
  - We use a **single unconditional** (marginal) randomization probability common to all subjects
  - D1 are **marginally randomized experiments**
- D2 uses 2 different coins:
  - We use **several** randomization probabilities **that depend** (are conditionally) on the values of the variable L
  - D2 are **conditionally randomized experiments**
- D2 is simply the combination of two separate marginally randomized experiments: one over subjects $L = 1$, the other in subjects with $L = 0$.

Randomization produces either **marginal exchangeability** (D1) or **conditional exchangeability** (D2).

Missings

MCAR

In Marginally Randomized Experiments, the values of the counterfactual outcomes are missing completely at random (MCAR).

MAR

In Conditionally Randomized Experiments, the values of the counterfactual outcomes are not MCAR, but they are missing at random (MAR) conditional on the covariate L.

2.3 Standardization

Heart transplant study, $L = 0$ noncritical condition, $L = 1$ critical condition.

Table 2.2

	L	A	Y
Rheia	0	0	0
Kronos	0	0	1
Demeter	0	0	0
Hades	0	0	0
Hestia	0	1	0
Poseidon	0	1	0
Hera	0	1	0
Zeus	0	1	1
Artemis	1	0	1
Apollo	1	0	1
Leto	1	0	0
Ares	1	1	1
Athena	1	1	1
Hephaestus	1	1	1
Aphrodite	1	1	1
Cyclope	1	1	1
Persephone	1	1	1
Hermes	1	1	0
Hebe	1	1	0
Dionysus	1	1	0

conditionally randomized experiment $\Rightarrow$ $Y^a \coprod A|L=0$ and $Y^a \coprod A|L=1$

Heart transplant study

8 individuals in noncritical condition

the risk on death among the treated is $\Pr(Y=1|L=0,A=1)=\frac{1}{4}$

the risk on death among the intreated is $\Pr(Y=1|L=0,A=0)=\frac{1}{4}$

12 individuals in critical condition

the risk on death among the treated is $\Pr(Y=1|L=1,A=1)=\frac{2}{3}$

the risk on death among the untreated is $\Pr(Y=1|L=1,A=0)=\frac{2}{3}$

The observed risk = the counterfactual risks

$$\Pr(Y = 1|L = 0, A = 1) = \Pr(Y^{a=1}|L = 0)$$
$$\Pr(Y = 1|L = 0, A = 0) = \Pr(Y^{a=0}|L = 0)$$
$$\Pr(Y = 1|L = 0, A = 1) = \Pr(Y^{a=1}|L = 0)$$
$$\Pr(Y = 1|L = 0, A = 0) = \Pr(Y^{a=0}|L = 0)$$

Goal: compute the causal risk ratio

$$\frac{\Pr(Y^{a=1}=1)}{\Pr(Y^{a=0}=1)}$$

numerator= risk if all the 20 subjects in the population had been treated:

- in the 8 subjects with $L=0$, the risk if all had been treated is $\frac{1}{4}$
- in the 12 subjects with $L=1$, the risk if all had been treated is $\frac{2}{3}$
- the risk if all the 20 subjects had been treated: average of $\frac{1}{4}$ and $\frac{2}{3}$ weighted proportional of its size:

$$\Pr(Y^{a=1}=1) = \frac{1}{4}\times\frac{8}{20} + \frac{2}{3}\times\frac{12}{20} = 0.5$$

denominator= risk if all the 20 subjects in the population had been untreated:

- in the 8 subjects with $L = 0$, the risk if all had been untreated is $\frac{1}{4}$
- in the 12 subjects with $L = 1$, the risk if all had been untreated is $\frac{2}{3}$
- the risk if all the 20 subjects had been untreated: average of $\frac{1}{4}$ and $\frac{2}{3}$ weighted proportional of its size:

$$\Pr(Y^{a=0} = 1) = \frac{1}{4} \times \frac{8}{20} + \frac{2}{3} \times \frac{12}{20} = 0.5$$

Causal risk ratio

$$\frac{\Pr(Y^{a=1} = 1)}{\Pr(Y^{a=0} = 1)} = \frac{0.5}{0.5} = 1$$

Standardized risks in treated and untreated are equal to counterfactual risk under treatment and no treatment, respectively:

$$\Pr(Y^{a=1} = 1) = \sum_l \Pr(Y = 1|L = l, A = 1)\Pr[L = l]$$

$$\Pr(Y^{a=0} = 1) = \sum_l \Pr(Y = 1|L = l, A = 0)\Pr[L = l].$$

Causal risk ratio

By standardization:

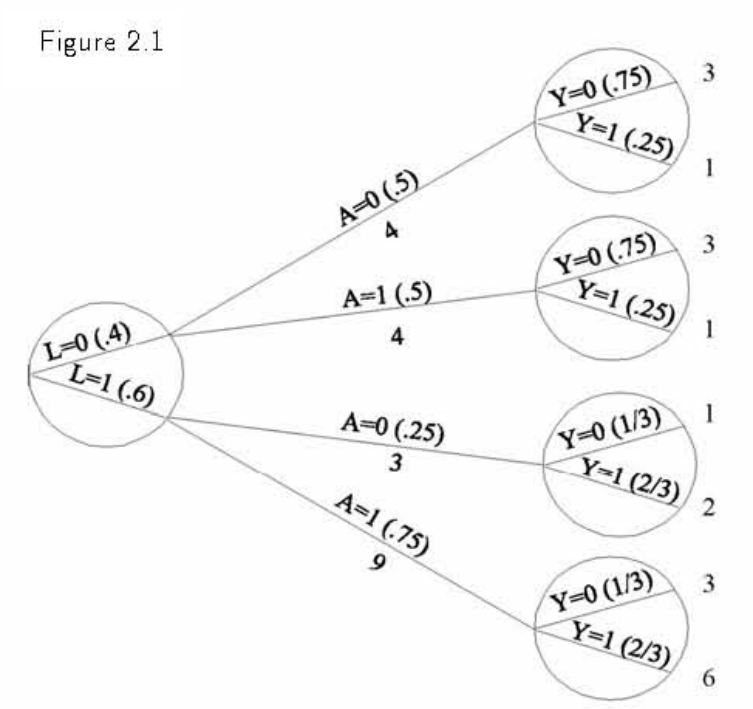
$$\frac{\Pr(Y^{a=1} = 1)}{\Pr(Y^{a=0} = 1)} = \frac{\sum_l \Pr(Y = 1|L = l, A = 1)\Pr[L = l]}{\sum_l \Pr(Y = 1|L = l, A = 0)\Pr[L = l]}$$

2.4 Inverse probability weighting

In this section, we compute causal risk ratio via inverse probability weighting.

2.4 Inverse probability weighting

In this section, we compute causal risk ratio via inverse probability weighting.Data in Figure 2.1 can be displayed as a tree:



Left circle: 8 non critical, 12 critical (probability). Of the 8 individuals in the branch $L = 0$, 4 were untreated and 4 treated (conditional probabilities). Of the 4 individuals in the branch $L = 0$ and $A = 0$, 3 survived and 1 died.

The denominator of casual risk ratio $\Pr(Y^{a=0} = 1)$ is the counterfactual risk of death had everybody in the population remained untreated:

The denominator of casual risk ratio $\Pr(Y^{a=0} = 1)$ is the counterfactual risk of death had everybody in the population remained untreated:

- 4 out of 8 individuals with $L = 0$ were untreated, and 1 of them died. How many deaths would have occurred had the 8 individuals with $L = 0$ remained untreated?:

The denominator of casual risk ratio $\Pr(Y^{a=0} = 1)$ is the counterfactual risk of death had everybody in the population remained untreated:

- 4 out of 8 individuals with $L = 0$ were untreated, and 1 of them died. How many deaths would have occurred had the 8 individuals with $L = 0$ remained untreated?:2 deaths.

The denominator of casual risk ratio $\Pr(Y^{a=0} = 1)$ is the counterfactual risk of death had everybody in the population remained untreated:

- 4 out of 8 individuals with $L = 0$ were untreated, and 1 of them died. How many deaths would have occurred had the 8 individuals with $L = 0$ remained untreated?:2 deaths.
- 3 out of 12 individuals with $L = 1$ were untreated, and 2 of them died. How many deaths would have occurred had the 12 individuals with $L = 1$ remained untreated?:

The denominator of casual risk ratio $\Pr(Y^{a=0} = 1)$ is the counterfactual risk of death had everybody in the population remained untreated:

- 4 out of 8 individuals with $L = 0$ were untreated, and 1 of them died. How many deaths would have occurred had the 8 individuals with $L = 0$ remained untreated?:2 deaths.
- 3 out of 12 individuals with $L = 1$ were untreated, and 2 of them died. How many deaths would have occurred had the 12 individuals with $L = 1$ remained untreated?: 8 deaths.

The denominator of casual risk ratio $\Pr(Y^{a=0} = 1)$ is the counterfactual risk of death had everybody in the population remained untreated:

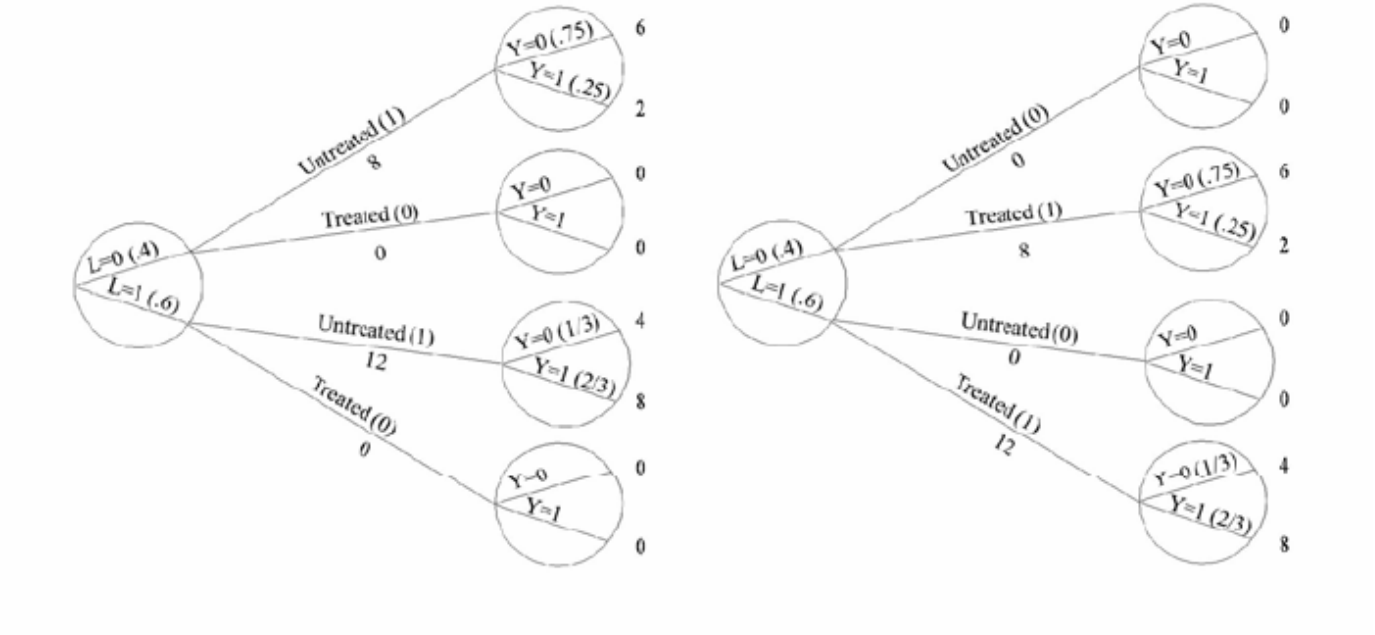
- 4 out of 8 individuals with $L = 0$ were untreated, and 1 of them died. How many deaths would have occurred had the 8 individuals with $L = 0$ remained untreated?:2 deaths.
- 3 out of 12 individuals with $L = 1$ were untreated, and 2 of them died. How many deaths would have occurred had the 12 individuals with $L = 1$ remained untreated?: 8 deaths.
- the risk if all the 20 subjects had been untreated:

The denominator of casual risk ratio $\Pr(Y^{a=0} = 1)$ is the counterfactual risk of death had everybody in the population remained untreated:

- 4 out of 8 individuals with $L = 0$ were untreated, and 1 of them died. How many deaths would have occurred had the 8 individuals with $L = 0$ remained untreated?:2 deaths.
- 3 out of 12 individuals with $L = 1$ were untreated, and 2 of them died. How many deaths would have occurred had the 12 individuals with $L = 1$ remained untreated?: 8 deaths.
- the risk if all the 20 subjects had been untreated:

$$\Pr(Y^{a=0} = 1) = \frac{2 + 8}{20} = 0.5$$

The first tree shows the population had everybody remained untreated. The second tree shows the population had everybody treated.



Of course, these calculations rely on the condition that treated individuals with $L = 0$, had they remained untreated, would have had the same probability of death as those who actually remained untreated. This condition is precisely exchangeability given $L = 0$.

The numerator of casual risk ratio $\Pr(Y^{a=1} = 1)$ is the counterfactual risk of death had everybody in the population been treated:

$$\Pr(Y^{a=1} = 1) = \frac{2+8}{20} = 0.5$$

Casual risk ratio

$$\frac{\Pr(Y^{a=1} = 1)}{\Pr(Y^{a=0} = 1)} = \frac{0.5}{0.5} = 1$$

Technical Point 2.2

Formal definition of IP weights A subject's Inverse Probability (IP) weigh depends on her values of treatment A and covariate L.

Discrete variables A and L

- Treated subject with $L = l$, receives the weight $1/Pr[A = 1|L = l]$
- Untreated subject with $L = l'$, receives the weight $1/Pr[A = 0|L = l']$

In a conditionally randomized experiment, $Pr[A = a|L = l] > 0$ for all l such that $\Pr[L = l] > 0$.

Continuous variable A

We use the probability density function (PDF) $f_{A|L}(a|l) = f(a|l)$. The IP weights are:

$$W^A = 1/f(A|L)$$

IP weighting and standardization

- *IPw* uses the conditional probability of treatment A given the covariate L.
- *STD* uses the probability of the covariate L and the conditional probability of outcome Y given A and L.

IP weighting and standardization

- *IPw* uses the conditional probability of treatment A given the covariate L.
- *STD* uses the probability of the covariate L and the conditional probability of outcome Y given A and L.

Both *IPw* and *STD* simulate what would have been observed if the variable (or variables) L had not been used to decide the probability of treatment. We often say that these methods *'adjust for L'* or *'control for L'*.

Technical Point 2.3

Equivalence of IP weighting and standardization

- The standardization mean for treatment a is defined as

$$\sum_l \mathrm{E}[Y|A=a, L=l]\Pr[L=l]$$

- The IP weighted mean of Y for treatment a is defined as

$$\mathrm{E}\left[\frac{\mathbf{I}(A=a)Y}{f(A|L)}\right]$$

$$\mathrm{E}\left[\frac{\mathbf{I}(A=a)Y}{f(A|L)}\right] = \sum_l \frac{1}{f(a|l)}\{\mathrm{E}[Y|A=a, L=l]f(a|l)\Pr[L=l]\}$$

Technical Point 2.3

If further assume conditional exchangeability then both the IP weighted and the standardized means are equal to the counterfactual mean $\mathrm{E}[Y^a]$.

$$\mathrm{E}\left[\frac{\mathbf{I}(A=a)Y}{f(A|L)}\right] = \mathrm{E}\left[\frac{\mathbf{I}(A=a)Y^a}{f(A|L)}\right]$$

$$= \mathrm{E}\left(\mathrm{E}\left[\frac{\mathbf{I}(A=a)Y^a}{f(A|L)}|L\right]\right)$$

$$= \mathrm{E}\left(\mathrm{E}\left[\frac{\mathbf{I}(A=a)}{f(A|L)}|L\right]\mathrm{E}[Y^a|L]\right) = \mathrm{E}[\mathrm{E}[Y^a|L]] = \mathrm{E}[Y^a]$$

Technical Point 2.3

If further assume conditional exchangeability then both the IP weighted and the standardized means are equal to the counterfactual mean $\mathrm{E}[Y^a]$.

$$\mathrm{E}\left[\frac{\mathbf{I}(A=a)Y}{f(A|L)}\right] = \mathrm{E}\left[\frac{\mathbf{I}(A=a)Y^a}{f(A|L)}\right]$$

$$= \mathrm{E}\left(\mathrm{E}\left[\frac{\mathbf{I}(A=a)Y^a}{f(A|L)}|L\right]\right)$$

$$= \mathrm{E}\left(\mathrm{E}\left[\frac{\mathbf{I}(A=a)}{f(A|L)}|L\right]\mathrm{E}[Y^a|L]\right) = \mathrm{E}[\mathrm{E}[Y^a|L]] = \mathrm{E}[Y^a]$$

When treatment is continuos (unlikely in conditional randomized experiments), effect estimates based on the IP weights $W^A = 1/f(A|L)$ have infinite variance and thus cannot be used.

Why not finish this book here?

We have a study design (an ideal randomized experiment) that, when combined with the appropriate analytic method (standardization or IP weighting), allows us to compute average causal effects.

Why not finish this book here?

We have a study design (an ideal randomized experiment) that, when combined with the appropriate analytic method (standardization or IP weighting), allows us to compute average causal effects.

Unfortunately, randomized experiments are often unethical, impractical, or untimely.

Why not finish this book here?

We have a study design (an ideal randomized experiment) that, when combined with the appropriate analytic method (standardization or IP weighting), allows us to compute average causal effects.

Unfortunately, randomized experiments are often unethical, impractical, or untimely.

Frequently, conducting an observational study is the least bad option.

Why not finish this book here?

We have a study design (an ideal randomized experiment) that, when combined with the appropriate analytic method (standardization or IP weighting), allows us to compute average causal effects.

Unfortunately, randomized experiments are often unethical, impractical, or untimely.

Frequently, conducting an observational study is the least bad option.

CONTINUARA...