Causal Inference

By: Miguel A. Hernán and James M. Robins

Part I: Causal inference without models

Chapter 1: A definition of causal effect

27th November, 2013

Chapter 1: A definition of causal effects

Contents

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 $a = i$, $i \in \{0, 1\}$.

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.
- 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

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

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

More generally (nondichotomous outcomes):

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

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

- 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**

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

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

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

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.
- **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			

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

Chapter 1.5: Causation versus association

Association measures

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

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

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

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

Example: $\text{ARD} = 0.54 - 0.43 = 0.11$, $\text{ARR} = 0.54/0.43 = 1.26$.

Chapter 1.5: Causation versus association

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

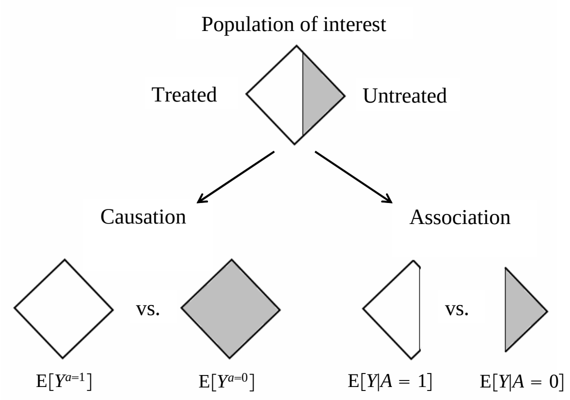

Figure : Association-causation difference

Concluding question:

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

CONTINUARÁ...