Standardization and Generalization

INFO/STSCI/ILRST 3900: Causal Inference

9 Sep 2025

Learning goals for today

At the end of class, you will be able to:

1. Reason about treatment effect heterogeneity
2. Estimate the average causal effect using data from a conditionally randomized experiment

Logistics

- Ch 2.3 and 4.1-4.3 in Hernan and Robins
- Problem Set 1 due today at 11pm (turn in pdf on canvas)
- Peer reviews will be assigned, due Sep 16

Conditional randomization review

Exchangeability may not hold in every randomized experiment

- Age ≥ 55 receive vaccine with $2/3$; more likely to get COVID if treated
- Age < 55 get vaccine with probability $1/2$; less likely to get COVID if treated

Conditional randomization review

Exchangeability may not hold in every randomized experiment

- Age ≥ 55 receive vaccine with $2/3$; more likely to get COVID if treated
- Age < 55 get vaccine with probability $1/2$; less likely to get COVID if treated
- Exchangeability does not hold in entire population
- Exchangeability holds within each sub-population
- Two separate experiments; both are exchangeable

Conditional randomization review

- **Marginal exchangeability:** $Y^a \perp\!\!\!\perp A$ for all a
- **Conditional exchangeability:** $Y^a \perp\!\!\!\perp A \mid L$ for all a
  The potential outcomes are independent of treatment **conditional on** L

Conditional randomization

- Can be useful in designing experiments
- Most useful as an idealized experiment to target with observational analysis
- Marginal exchangeability is very unlikely in observational data
- Conditional exchangeability may be more reasonable

Conditional randomization

- **Stratification**: We can directly estimate causal effect within each sub-population (or stratum)

$$
\begin{aligned}
E(Y \mid A = a, L = \ell) &\overset{\text{consis}}{=} E(Y^a \mid A = a, L = \ell) \\
&\overset{\text{exchange}}{=} E(Y^a \mid L = \ell)
\end{aligned}
$$

- The causal effect within a stratum is the Conditional ACE

$$
E(Y^{a=1} \mid L = \ell) - E(Y^{a=0} \mid L = \ell)
$$

Conditional randomization

- **Stratification**: We can directly estimate causal effect within each sub-population (or stratum)

$$\begin{aligned} E(Y \mid A = a, L = \ell) &\overset{\text{consis}}{=} E(Y^a \mid A = a, L = \ell) \\ &\overset{\text{exchange}}{=} E(Y^a \mid L = \ell) \end{aligned}$$

- The causal effect within a stratum is the Conditional ACE

$$E(Y^{a=1} \mid L = \ell) - E(Y^{a=0} \mid L = \ell)$$

- If the treatment effect varies across sub-population, we say there is **treatment effect heterogeneity**

$$\begin{gathered} E(Y^{a=1} \mid L = 55+) - E(Y^{a=0} \mid L = 55+) \\ \neq \\ E(Y^{a=1} \mid L =< 55) - E(Y^{a=0} \mid L =< 55) \end{gathered}$$

Measuring a causal effect

- Average Causal Effect $\mathrm{E}(Y^{a=1}) - \mathrm{E}(Y^{a=0})$

Measuring a causal effect

- Average Causal Effect $\mathrm{E}(Y^{a=1}) - \mathrm{E}(Y^{a=0})$
- Some people could be harmed, some could be helped
- Suppose new policy results in .99 of individuals losing 1 dollar; .01 gains 100 dollars
- Positive ACE, but maybe still not a good idea
- Conditional ACE helps give more information

Measuring a causal effect

- Average Causal Effect $\mathrm{E}(Y^{a=1}) - \mathrm{E}(Y^{a=0})$
- Some people could be harmed, some could be helped
- Suppose new policy results in .99 of individuals losing 1 dollar; .01 gains 100 dollars
- Positive ACE, but maybe still not a good idea
- Conditional ACE helps give more information

- **Sharp null hypothesis**: $Y_i^{a=1} = Y_i^{a=0}$ for all i
- Sharp null hypothesis also means ACE $= 0$, but not the other way around!

Standardization

- Under conditional exchangeability, we can directly estimate the conditional average causal effect for each sub-population
- Conditional ACE (or ATE)

$$\mathrm{E}(Y^{a=1} \mid L = \ell) - \mathrm{E}(Y^{a=0} \mid L = \ell)$$

- When different treatments can be applied to different individuals
- Ex: medicine, job training program

Standardization

- Under conditional exchangeability, we can directly estimate the conditional average causal effect for each sub-population
- Conditional ACE (or ATE)

$$E(Y^{a=1} \mid L = \ell) - E(Y^{a=0} \mid L = \ell)$$

- When different treatments can be applied to different individuals
- Ex: medicine, job training program
- ACE (or ATE)

$$E(Y^{a=1}) - E(Y^{a=0})$$

- When everyone gets the same treatment
- Ex: Tax policy, new product feature
- **Standardization** allows us to estimate the ACE by combining estimates from each sub-population

Standardization

- **Standardization** allows us to estimate the ACE by combining estimates from each sub-population
- Use law of total expectation

Standardization

- **Standardization** allows us to estimate the ACE by combining estimates from each sub-population
- Use law of total expectation

Standardization

- **Standardization** allows us to estimate the ACE by combining estimates from each sub-population
- For $L = 0, 1$

$$
\begin{aligned}
\mathrm{E}(Y^a) &= Pr(L=1)\mathrm{E}(Y^a \mid L=1) + Pr(L=0)\mathrm{E}(Y^a \mid L=0) \\
&= Pr(L=1)\mathrm{E}(Y \mid L=1, A=a) \\
&\quad + Pr(L=0)\mathrm{E}(Y \mid L=0, A=a)
\end{aligned}
$$

Standardization

- **Standardization** allows us to estimate the ACE by combining estimates from each sub-population
- For $L = 0, 1$

$$\begin{aligned} E(Y^a) &= Pr(L = 1)E(Y^a \mid L = 1) + Pr(L = 0)E(Y^a \mid L = 0) \\ &= Pr(L = 1)E(Y \mid L = 1, A = a) \\ &\quad + Pr(L = 0)E(Y \mid L = 0, A = a) \end{aligned}$$

- More generally, for each a

$$E(Y^a) = \sum_{\ell} Pr(L = \ell)E(Y \mid L = \ell, A = a)$$

Exercise

		L	A	Y
1	Rheia	0	0	0
2	Kronos	0	0	1
3	Demeter	0	0	0
4	Hades	0	0	0
5	Hestia	0	1	0
6	Poseidon	0	1	0
7	Hera	0	1	0
8	Zeus	0	1	1

		L	A	Y
9	Artemis	1	0	1
10	Apollo	1	0	1
11	Leto	1	0	0
12	Ares	1	1	1
13	Athena	1	1	1
14	Hephaestus	1	1	1
15	Aphrodite	1	1	1
16	Polyphemus	1	1	1
17	Persephone	1	1	1
18	Hermes	1	1	0
19	Hebe	1	1	0
20	Dionysus	1	1	0

$$E(Y^a) = Pr(L=1)E(Y \mid L=1, A=a)$$

$$+ Pr(L=0)E(Y \mid L=0, A=a)$$

Excercise

		L	A	Y
1	Rheia	0	0	0
2	Kronos	0	0	1
3	Demeter	0	0	0
4	Hades	0	0	0
5	Hestia	0	1	0
6	Poseidon	0	1	0
7	Hera	0	1	0
8	Zeus	0	1	1

		L	A	Y
9	Artemis	1	0	1
10	Apollo	1	0	1
11	Leto	1	0	0
12	Ares	1	1	1
13	Athena	1	1	1
14	Hephaestus	1	1	1
15	Aphrodite	1	1	1
16	Polyphemus	1	1	1
17	Persephone	1	1	1
18	Hermes	1	1	0
19	Hebe	1	1	0
20	Dionysus	1	1	0

$$\mathrm{E}(Y^{a=1}) = Pr(L=1)\mathrm{E}(Y \mid L=1, A=1)$$

$$+ Pr(L=0)\mathrm{E}(Y \mid L=0, A=1)$$

Excercise

		L	A	Y
1	Rheia	0	0	0
2	Kronos	0	0	1
3	Demeter	0	0	0
4	Hades	0	0	0
5	Hestia	0	1	0
6	Poseidon	0	1	0
7	Hera	0	1	0
8	Zeus	0	1	1

		L	A	Y
9	Artemis	1	0	1
10	Apollo	1	0	1
11	Leto	1	0	0
12	Ares	1	1	1
13	Athena	1	1	1
14	Hephaestus	1	1	1
15	Aphrodite	1	1	1
16	Polyphemus	1	1	1
17	Persephone	1	1	1
18	Hermes	1	1	0
19	Hebe	1	1	0
20	Dionysus	1	1	0

$$\mathrm{E}(Y^{a=1}) = \underbrace{Pr(L=1)}_{12/20}\underbrace{\mathrm{E}(Y \mid L=1, A=1)}_{6/9} + Pr(L=0)\mathrm{E}(Y \mid L=0, A=1)$$

Excercise

		L	A	Y
1	Rheia	0	0	0
2	Kronos	0	0	1
3	Demeter	0	0	0
4	Hades	0	0	0
5	Hestia	0	1	0
6	Poseidon	0	1	0
7	Hera	0	1	0
8	Zeus	0	1	1

		L	A	Y
9	Artemis	1	0	1
10	Apollo	1	0	1
11	Leto	1	0	0
12	Ares	1	1	1
13	Athena	1	1	1
14	Hephaestus	1	1	1
15	Aphrodite	1	1	1
16	Polyphemus	1	1	1
17	Persephone	1	1	1
18	Hermes	1	1	0
19	Hebe	1	1	0
20	Dionysus	1	1	0

$$\begin{aligned} \mathrm{E}(Y^{a=1}) &= \underbrace{Pr(L=1)}_{12/20}\underbrace{\mathrm{E}(Y \mid L=1, A=1)}_{6/9} \\ &\quad + \underbrace{Pr(L=0)}_{8/20}\underbrace{\mathrm{E}(Y \mid L=0, A=1)}_{1/4} = 1/2 \end{aligned}$$

Exercise

		L	A	Y
1	Rheia	0	0	0
2	Kronos	0	0	1
3	Demeter	0	0	0
4	Hades	0	0	0
5	Hestia	0	1	0
6	Poseidon	0	1	0
7	Hera	0	1	0
8	Zeus	0	1	1

		L	A	Y
9	Artemis	1	0	1
10	Apollo	1	0	1
11	Leto	1	0	0
12	Ares	1	1	1
13	Athena	1	1	1
14	Hephaestus	1	1	1
15	Aphrodite	1	1	1
16	Polyphemus	1	1	1
17	Persephone	1	1	1
18	Hermes	1	1	0
19	Hebe	1	1	0
20	Dionysus	1	1	0

$$\mathrm{E}(Y^{a=0}) = \underbrace{Pr(L=1)}_{12/20}\underbrace{\mathrm{E}(Y \mid L=1, A=0)}_{2/3} + \underbrace{Pr(L=0)}_{8/20}\underbrace{\mathrm{E}(Y \mid L=0, A=0)}_{1/4}$$

Exercise

		L	A	Y
1	Rheia	0	0	0
2	Kronos	0	0	1
3	Demeter	0	0	0
4	Hades	0	0	0
5	Hestia	0	1	0
6	Poseidon	0	1	0
7	Hera	0	1	0
8	Zeus	0	1	1

		L	A	Y
9	Artemis	1	0	1
10	Apollo	1	0	1
11	Leto	1	0	0
12	Ares	1	1	1
13	Athena	1	1	1
14	Hephaestus	1	1	1
15	Aphrodite	1	1	1
16	Polyphemus	1	1	1
17	Persephone	1	1	1
18	Hermes	1	1	0
19	Hebe	1	1	0
20	Dionysus	1	1	0

$$\mathrm{E}(Y^{a=0}) = \underbrace{Pr(L=1)}_{12/20}\underbrace{\mathrm{E}(Y \mid L=1, A=0)}_{2/3} + \underbrace{Pr(L=0)}_{8/20}\underbrace{\mathrm{E}(Y \mid L=0, A=0)}_{1/4} = 1/2$$

Generalization (transportability)

- Randomized control trials provide convincing evidence for a causal effect

Generalization (transportability)

- Randomized control trials provide convincing evidence for a causal effect
- Unless the sharp null is true, no such thing as "The" average causal effect
- If treatment effect is heterogeneous, ACE is defined with respect to a population of interest
- Ex: most psychology experiments consist primarily of undergraduate students. Results may be accurate for undergraduate students, but do not say much about broader population

Generalization (transportability)

- Drawing representative sample for RCTs yields more reliable conclusions
- Transporting effect to different population
  - Assume conditional ACE is same across populations
  - Standardization should use population weights in target population

Generalization (transportability)

- Consider experiment conducted on students in a psychology course at Cornell
- Let $L = 0$ denote undergraduate and $L = 1$ denote graduate student
- Suppose in the psychology course $P(L = 0) = .95$ and $P(L = 1) = .05$
- In an anthropology course, $P(L = 0) = .6$ and $P(L = 1) = .4$
- Suppose

$$\mathrm{E}(Y^1 - Y^0 \mid L = 0) = .2$$
$$\mathrm{E}(Y^1 - Y^0 \mid L = 1) = -.4$$

- ACE for the experiment would be

$$.2 \times .95 + (-.4) \times .05 = 0.17$$

- ACE in the anthropology course would be

$$.2 \times .6 + (-.4) \times .4 = -0.04$$

Learning goals for today

At the end of class, you will be able to:

1. Reason about treatment effect heterogeneity
2. Estimate the average causal effect using data from a conditionally randomized experiment