Detecting Latent Heterogeneity

Jeremy Lin

University of California, Irvine

Outline

Heterogeneity Problem

- This presentation is mostly following paper "Detecting Latent Heterogeneity" by Judea Pearl [1]
- Heterogeneity Problem
  - Heterogeneity problem occurs when there exist peculiar groups (in simple case a strata of covariate C) that reacts differently to treatment/ policy (covariate-induced het
- Example of covariate-induced Heterogenity

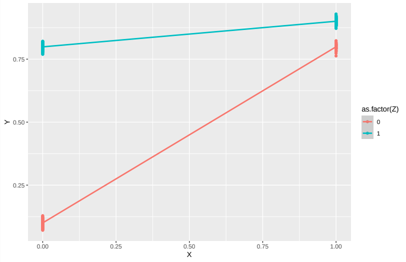


Figure: The effect of treatment is different between two strata of Z

Assessing covariate-induced heterogeneity

- Suppose we are able to measure any characteristic C (baseline covariate that induced heterogeneity) of all individuals
- We can define a measure to calculate the effect difference between two sub-strata of C

$$D(c_i, c_j) = |E(Y_1 - Y_0|C = c_i) - E(Y_1 - Y_0|C = C_j)|$$

- To represent the best measure of heterogeneity in the population, we can find the lower bound LB on the heterogeneity between any two subgroup of C.

$$LB = max_{c_i, c_j} D(c_i, c_j)$$

- Two main problem in this procedure
  - We need to find covariate C for which c-specific effect ($E(Y_1 - Y_0|C)$) is identifiable
  - Perform maximization in over all pairs (i, j) in all vectors of C

Review of Counterfactuals

Consistency Rule

$$E(Y_{X=x}|X = x) = E(Y|X = x)$$

- Consistency rule can be interpreted as the counterfactual Y_x is equal to the observed value of Y whenever X takes the value of x

Theorem 4.3.1 Counterfactual Interpretation of Backdoor

If a set Z of variables satisfies the backdoor condition (Z is an admissible set) relative to (X, Y), then for all x, the counterfactual Y_x is conditionally independent of X given Z

$$E(Y_x|Z) = E(Y_x|X, Z)$$

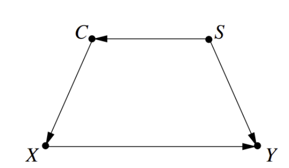
- Conditional independence ($Y \perp\!\!\!\perp X|C$) sometimes referred as conditional ignorability or conditional exchangeabilty.

Special Cases of estimable covariate-based heterogeneity

There are three special cases when covariates C are identifiable.

- C satisfies the back-door criterion (C is admissible)
- C is part of set S that satisfies the backdoor criterion
- C is not part of any admissible sets, but causal effect is identifiable

Case 1 : C satisfies the backdoor criterion



C specific effect is identified through :

$$
\begin{aligned}
E(Y_1 - Y_0|C = c) &= E(Y_1|C = c) - E(Y_0|C = c) \\
&= E(Y_1|X = 1, C = c) - E(Y_0|X = 0, C = c) \quad \text{(Theorem 4.3.1)} \\
&= E(Y|X = 1, C = c) - E(Y|X = 0, C = c) \quad \text{consistency}
\end{aligned}
$$

Case 2 : C is part of an admissible set

- C specific effect is identified through :

$$E(Y_1 - Y_0|C = c) = \sum_s [E(Y_1|C = c, S = s) - E(Y_0|C = c, S = s)] P(s|c)$$
$$= \sum_s [E(Y|X = 1, C = c, S = s) - E(Y|X = 0, C = c, S = s)] P(s|c)$$

Case 3 : C is not part of any admissible sets, but Causal effect is identifiable

- $E(Y_1 - Y_0|C = c)$ is estimable through front door estimator (will be discussed later)

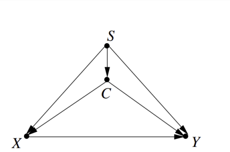


Figure: C is part of admissible set {S,C}

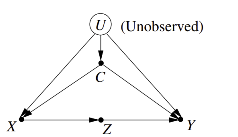


Figure: C is not admissible

Special Cases covariate-based heterogeneity

Example where C-specific effects is not identifiable

- In model 2, S is an admissible set, but conditioning on C will open the collider path
- Model 1 and model 2 is statistically indistinguishable, which imply there is no statistical test can determine whether set S, C is admissible.

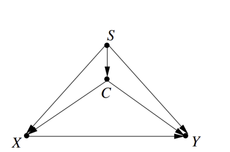


Figure: model 1

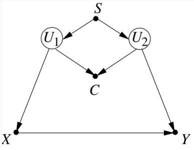


Figure: model 2

Latent Heterogeneity Between The Treated and Untreated

- It is hard to detect heterogeneity through measuring the effect sizes between two subgroups of C
- **Goal** : We would like to find a way to detect heterogeneity in the data without analyzing covariates C
- Latent Heterogeneity : heterogeneity that is not present in any baseline covariates, but manifest itself in effect differences between the Treated and Untreated [1]

Two types of confounding

- Suppose we have a binary treatment, then Average treatment effect (ATE) can be decomposed into several components:

$$
\begin{aligned}
ATE &= E(Y_1 - Y_0) \\
&= E(Y|X=1) - E(Y|X=0) \\
&- (E(Y_0|X=1) - E(Y_0|X=0)) \\
&- (ETT - ETU)/P(X=0)
\end{aligned}
$$

- Where ETT(average effect of treatment on the treated) and ETU(average effect of treatment on the treated) is defined as :

$$
\begin{aligned}
ETT &= E(Y_1 - Y_0|X=1) \\
ETU &= E(Y_1 - Y_0|X=0)
\end{aligned}
$$

- We can define bias as :

$$Bias = E(Y|X=1) - E(Y|X=0) - ATE$$

Detecting Latent Heterogeneity

- From the previous equation We can decompose Bias into two component :

$$Bias = (E(Y_0|X = 1) - E(Y_0|X = 0)) + (ETT - ETU)/P(X = 0)$$

- $E(Y_0|X = 1) - E(Y_0|X = 0)$ is sometimes called *baseline* or *fixed effect bias*.
- ETT - ETU is also called *differential treatment effect bias*, or *variable-effect bias*.
- decomposing bias into Baseline and variable -effect bias can be define counterfactually without conditioning to specific covariates C
- ETT - ETU can be used as an indication of heterogeneity regardless if we know which covariates responsible for heterogeneity.
- We will discus the three classical case where where ETT and ETU are identifiable

Detecting Heterogeneity in Randomized trials

- In binary randomized trial $E(Y_0)$ and $E(Y_1)$ are identifiable

$$E(Y_1) = E(Y|X = 1)p + E(Y_1|X = 0)(1 - p)$$

  Where p = P(X = 1)

- The difference of ETT - ETU is estimable and given by :

$$ETT - ETU = \frac{E(Y|X = 1) - E(Y_1)}{(1 - p)} + \frac{E(Y|X = 0) - E(Y_0)}{p}$$

  .

- based on pre-trial and post trial data we can estimate whether heterogeneity bias exist in the population prior to randomization without measuring any covariates.
- Heterogeneity exist in population whenever experimental findings reveal a non zero ETT - ETU

Detecting Heterogeneity through Adjustments

- Suppose there exist admissible set Z of covariates yielding the the adjustment estimand

$$E(Y_x) = \sum_z E(Y|x,z)P(z)$$

- It can be shown that ;

$$E(Y_x|x') = \sum_z E(Y|x,z)P(z|x')$$

- The difference of ETT - ETU is estimable and given by :

$$ETT - ETU = \sum_z \left[E(Y|X=x',z) - E(Y|X=x,z)\right]\left[P(z|X=x') - P(z|x=x)\right]$$

.

- Note that although we are using set Z to measure ETT and ETU, we don't make any assumption that heterogeneity comes from any of the subset of Z

Detecting Heterogeneity through Mediating Instrument

- Identification by adjustment requires modelling assumption, while
- Instrumental Variable requires milder assumption, but suffers from fundamental limitation (only effective in linear and pseudo linear model)
- Mediating instruments , also known as front door criterion overcomes the limitations.
- Using front door adjustment :

$$E(Y_x|X = x') = \sum_z E(Y|z, x')P(z|x)$$

Where x and x' are any two level of the treatment

- Thus variable effect bias can be estimated by

$$ETT - ETU = \sum_z \left[E(Y|X = x', z) - E(Y|X = x, z)\right] \left[P(z|X = x') - P(z|x = x)\right]$$

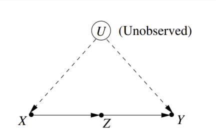


Figure: Z act as a mediating instrument

Example : Heterogeneity in Recruitment

Problem background:

- A government is funding a job training program aimed at getting unemployed people back into the workforce
- A pilot randomized experiment shows that the program is effective, but critics argue that there is no proof that the program works in real life due to heterogeneity in the population
- People who decides to enroll in the program tend to be more informed, more intelligent , and more resourceful and would have found a job regardless of training, while the uniformed people who could benefited from the program is not aggressively recruited

Example : Heterogeneity in Recruitment

We can model this problem as :

- Z is a binary variable for class of individual ($Z = 0$ represent uniformed individual)
- r as the proportion informed individual in the population
- X is a binary variable stands for participation in the program ($X = 1$ represent participation in the program)
- $q_1 = P(X = 1|Z = 0)$ is the propensity for enrollment among the uninformed
- $q_2 = P(X = 1|Z = `)$ is the propensity for enrollment among the informed
- $diff = q2 - q1$ represent the difference in propensity enrollment between two class. Large $diff$ imply informed people are more likely to be enrolled in the program

Results :

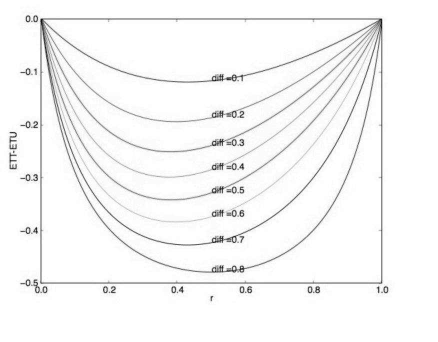


Figure: Z act as a mediating instrument

- ETT- ETU is negative, which indicates loss of opportunity due to recruiting policy
- As *r* approaches to 0 or 1 (population is homogeneous), variable effect bias goes to 0
- as *diff* goes larger, variable effect bias increases

Conclusion

- Heterogeneity problem raises when there exist a group that reacts differently to treatment / policy
- It's often hard to assess heterogeneity through measuring the effect sizes between subgroups.
- Under certain conditions it is possible to measure Latent heterogeneity(i.e measuring heterogeneity without knowing the covariates that causes heterogeneity problem).
- Bias of Average Treatment Effect can be decomposed into baseline bias and variable-effect bias (ETT - ETU).
- $ETT - ETU \neq 0$ implies that Heterogeneity exist in the population.

References I

[1] J. Pearl, "Detecting latent heterogeneity," in *Probabilistic and Causal Inference: The Works of Judea Pearl*, 1st ed. New York, NY, USA: Association for Computing Machinery, 2022, 483–506, ISBN: 9781450395861. [Online]. Available: `https://doi.org/10.1145/3501714.3501742`.

[2] Y. Xie, J. E. Brand, and B. Jann, "Estimating heterogeneous treatment effects with observational data," *Sociological Methodology*, vol. 42, no. 1, 314–347, 2012. DOI: `10.1177/0081175012452652`.

[3] J. Pearl, M. Glymour, and N. P. Jewell, *Causal inference in statistics: A Primer*. Wiley, 2021.

[4] J. Pearl, "Causal diagrams for empirical research," *Biometrika*, vol. 82, no. 4, pp. 669–688, 1995, ISSN: 00063444. [Online]. Available: `http://www.jstor.org/stable/2337329` (visited on 03/08/2023).

[5] I. Shpitser and J. Pearl, "Effects of treatment on the treated: Identification and generalization," in *Conference on Uncertainty in Artificial Intelligence*, 2009.

Thank You!

Question : How do we identify heterogeneity in binary randomized trial ?

data generating model for example

- data generating model

$$E(Y|X,Z) = 0.7X + 0.7Z - 0.6XZ + 0.1$$
$$X|Z \sim Bern(Zq_2 + (1-Z)q_1)$$
$$Z \sim Bern(r)$$

ATE decomposition proof

$$
\begin{aligned}
ATE &= E(Y_1 - Y_0) \\
&= E(Y_1|X=0)P(X=0) + E(Y_1|X=1)P(X=1) \\
&\quad - E(Y_0|X=0)P(X=0) - E(Y_0|X=1)P(X=1) \\
&= E(Y_1|X=1) + E(Y_1|X=1)(P(X=1)-1) \\
&\quad - E(Y_0|X=0) + E(Y_0|X=0)(P(X=0)-1) \\
&\quad + E(Y_1|X=0)P(X=0) - E(Y_0|X=1)P(X=1) \\
&= E(Y_1|X=1) - E(Y_0|X=0) \\
&\quad + E(Y_0|X=1) + E(Y_0|X=1)(1-P(X=1)) \\
&\quad - E(Y_0|X=0) - E(Y_0|X=0)(1-P(X=1)) \\
&\quad + E(Y_1|X=0)P(X=0) - E(Y_1|X=1)P(X=0)
\end{aligned}
$$

ATE decomposition proof

$$
\begin{aligned}
ATE &= E(Y_1|X=1) - E(Y_0|X=0) \\
&\quad + E(Y_0|X=1) - E(Y_0|X=0) \\
&\quad + E(Y_0|X=1)(1-P(X=1)) - E(Y_0|X=0)(1-P(X=1)) \\
&\quad + E(Y_1|X=0)P(X=0) - E(Y_1|X=1)P(X=0) \\
&= E(Y_1|X=1) - E(Y_0|X=0) \\
&\quad + E(Y_0|X=1) - E(Y_0|X=0) \\
&\quad + E(Y_0|X=1)(P(X=0)) - E(Y_0|X=0)(P(X=0)) \\
&\quad + E(Y_1|X=0)P(X=0) - E(Y_1|X=1)P(X=0) \\
&= E(Y|X=1) - E(Y|X=0) \\
&\quad - (E(Y_0|X=1) - E(Y_0|X=0)) \\
&\quad - (ETT - ETU)P(X=0)
\end{aligned}
$$