# STA 640 — Causal Inference

# Chapter 4.1 Treatment Effect Heterogeneity: Causal Subgroup Analysis

Fan Li

Department of Statistical Science

Duke University

# Motivation

- Acknowledgement: Large part of this lecture is written by Joey Antonelli of University of Florida
- There is huge interest in understanding whether a treatment or policy affects certain individuals more than others
  - Referred to as treatment effect heterogeneity or heterogeneous treatment effects
- Personalized medicine is a huge area of interest
  - What treatment should an individual get
  - Physicians are implicitly considering how treatment effects vary when determining what treatment to assign a patient
    - Given their characteristics, treatment history, etc.

# Motivation

- There are countless other applications for which heterogeneity of the treatment effect is of scientific interest
- Many cancer treatments only work on a subset of the population
  - Why? What subsets of the population?
- Limited resource settings where not everyone can be assigned treatment
  - Give it to those individuals most likely to benefit
- Helps to transport causal effects from one population to another
  - Two populations might have different characteristics and therefore different ATEs

# Motivation

- An additional issue is that sometimes average or marginal treatment effects can mask the effect of a policy
- What if a policy has a positive impact on some individuals and a negative impact on others?
  - ATE will likely be very close to zero
  - Hypothesis tests indicate no treatment effect
  - In truth the treatment is very important
- Looking at heterogeneous treatment effects provides more scientific information than marginal effects alone
  - Immediately recover marginal effects from heterogeneous ones

# Motivation

- There are many questions one can answer in a study of heterogenous treatment effects
  - Which covariates modify the treatment effect?
  - Is there any heterogeneity whatsoever?
  - For a given  - , what is the expected treatment effect (CATE)
  - For a given individual, what is their treatment effect (ITE)
- Choice of statistical approach will depend on the goals of the study

# Estimands of interest

- The most common target estimand is the conditional average treatment effect

$$\text{CATE} = g[G^0] = E[Y^1 - Y^0 \mid G]$$

- Note the ATE is simply the average CATE

$$\text{ATE} = E[Y^1 - Y^0] = \int_G g[G^0] \, dF_G$$

- This shows how the CATE provides additional information over the ATE
  - Once we know the CATE, we immediately know the ATE

# Estimands of interest

- Another relevant estimand refers to subgroup analysis
- Assume we have a subset of the covariate space defined by $C$, e.g. specific age or gender or medical history
- A subgroup specific estimand is given by

$$E[Y^1 - Y^0 \mid C]$$

- Commonly we will have non-overlapping regions given by $C_1, \ldots, C$ , and we estimate

$$E[Y^1 - Y^0 \mid C_6] \text{ for } 6 = 1, \ldots,$$

- And again we can easily recover the ATE by marginalizing over these

# Estimands of interest

- Sometimes the CATE is not of interest, but focus is on a subset of predictors given by + - :

$$E\left[\tau(X) \mid X_j = x_j\right] = E[\tau(X) \mid X_j = x_j]$$

- Maybe we simply care whether a particular covariate modifies the treatment effect

- This construction is really useful in high-dimensional settings where - is high-dimensional, but we care more about heterogeneity by certain covariates
  - Still need to account for - when adjusting for confounding, but not when estimating heterogeneous treatment effects

# Estimands of interest

- Individual treatment effects (ITE) are also of concern

$$g_i = .^1 1^o \quad .^1 0^o$$

- For example, this is the question that personalized medicine looks to address
  - How will the treatment affect this particular individual
- Generally speaking, these are much harder to estimate
  - More uncertainty
  - Prediction intervals are wider than intervals for a mean
  - Stronger assumptions

# Estimands of interest

- The literature often conflates the ITE and the CATE
- Clearly, we have that

$$\tau(x) = \mathbb{E}[Y(1) - Y(0) \mid X = x] = \mathbb{E}[Y(1) \mid X = x] - \mathbb{E}[Y(0) \mid X = x]$$

- Related concepts, and certainly the CATE evaluated at $x$ is a good point estimate for the ITE of individual $i$
- Under outcome modeling approach, all estimands are estimated in the same fashion

# Identifying assumptions

- Estimation of heterogeneous treatment effects (HTE) differs from that of marginal treatment effects, but identification is effectively the same
- Easy to see that under SUTVA and unconfoundedness we have

$$
\begin{aligned}
g^{1}G^{o} &= E\text{»}.\ ^{1}1^{o} \quad .\ ^{1}0^{o}j\ -\ = G¼ \\
&= E\text{»}.\ ^{1}1^{o}j\ -\ = G¼ \quad E\text{»}.\ ^{1}0^{o}j\ -\ = G¼ \\
&= E\text{»}.\ ^{1}1^{o}j\ /\ = 1-\ -\ = G¼ \quad E\text{»}.\ ^{1}0^{o}j\ /\ = 0-\ -\ = G¼ \\
&= E\text{»}.\ j\ /\ = 1-\ -\ = G¼ \quad E\text{»}.\ j\ /\ = 0-\ -\ = G¼
\end{aligned}
$$

- Unconfoundedness allows us to use data with $/_{8} = 0$ to estimate $E\text{»}.\ ^{1}0^{o}j\ -\ = G¼$ in the whole population
  - Same for $.\ ^{1}1^{o}$

# Identifying assumptions

- Overlap is still a fundamental assumption for heterogeneous treatment effects as well
  - With little overlap, causal inference is problematic conceptually and has large uncertainity operationally
- Suppose we have certain regions of the covariate space that are always treated
- We have to then extrapolate our estimates of E».j/ = 0- - = G¼ to these individuals with different covariate values
  - Heavily reliant on model specification
  - Difficult to understand the degree of extrapolation
  - Unclear impacts on uncertainty quantification
- We will discuss overlap a bit more in subgroup analysis

# Identifying assumptions

- In this section, we will mostly cover estimation issues
  - There are a lot!
- A lot of other issues inherent to a causal analysis apply here as well
  - Considering plausibility of causal assumptions
  - Sensitivity analysis (to be covered in a couple of weeks)
  - Overlap and balance checks
- When these issues differ in ways unique to heterogeneous treatment effect estimation, we will cover them as they come up

# Subgroup analysis

- This sub-chapter focuses on the simplest form of HTE: subgroup analysis
- The simplest form of heterogeneity is subgroup analysis (SGA)
- Again suppose we have non-overlapping subsets of the covariate space given by $C_1 - \cdots - C$
- Our goal is estimation of

$$E\left[\tau(X) \mid X \in C_6\right] \text{ for } 6 = 1 - \cdots -$$

- We will see that many of the same estimation strategies we've already learned about can be utilized here analogously

# Subgroup analysis

- Important that these groups are chosen beforehand, often in a one-variable-at-a-time fashion
  - Might look old-fashioned, by still informative and widely used in practice, e.g. medical research
- There are data-driven approaches for finding the subsets of the population that benefit from treatment
- Generally speaking using the data to find subgroups complicates analyses
  - Valid inference becomes challenging
  - Post selection inference issues
  - Can use data splitting to alleviate these issues
- We will focus for now on situations where these groups are known beforehand

# Subgroup analysis (SGA): weighting

- All balancing weights can be directly applied to SGA. Below we will focus on IPW for simplicity
- Recall the original IPW estimator of the ATE

$$\frac{1}{\#}\left(\sum_{i=1} \frac{Y_i A_i}{\hat{e}(X_i)} - \sum_{i=1} \frac{Y_i (1-A_i)}{1-\hat{e}(X_i)}\right)$$

  which is used as a sample estimate of

$$E\left[\frac{YA}{e(X)} - \frac{Y(1-A)}{1-e(X)}\right]$$

- We are using the empirical distribution from the sample to approximate this expectation that is with respect to the overall target population

# Subgroup analysis: weighting

- Now our target population is the subset of individuals within $C_6$:

$$E(Y(1) - Y(0) \mid X \in C_6) = E\left[\frac{Y \cdot W}{P(W=1 \mid X \in C_6)} - \frac{Y \cdot (1-W)}{1 - P(W=1 \mid X \in C_6)} \,\middle|\, X \in C_6\right]$$

so a natural estimator of this is simply

$$\frac{1}{\#_6}\left(\sum_{i:X_i \in C_6} \frac{Y_i W_i}{P(W=1 \mid X_i \in C_6)} - \sum_{i:X_i \in C_6} \frac{Y_i (1-W_i)}{1 - P(W=1 \mid X_i \in C_6)}\right)$$

where $\#_6 = \sum_{i=1}^{N} I(X_i \in C_6)$

# Subgroup analysis: weighting

- We simply use the IPW estimator but instead average over just the individuals in the desired subgroup
- The same procedure applies to other balancing weights, e.g. overlap weights, ATT weights
- Can apply this procedure separately within each subgroup to estimate subgroup specific effects
- Remember that balancing weights are intended to construct a weighted population for which the covariates are balanced across treatment groups
  - Does that happen here?

# Subgroup analysis: variance-bias tradeoff

- An important question is how the PS is estimated
  - Using the entire sample
  - Using just the individuals in $C_6$
- Using the full sample aims to ensure balance in the entire target population, not the subgroup specific one
- Nonetheless, if the PS is correctly specified, using the full sample should work well
- Using just the individuals in $C_6$ will improve balance within the subgroup
  - Less efficient. Bias/variance trade-off
- A simple logistic propensity score model with only main effects of all covariates is not usually adequate in SGA

# Subgroup analysis: outcome modeling

- Similar issues occur for outcome modeling or doubly robust estimators
- Recall the outcome modeling estimator

$$\frac{1}{\#} \sum_{i=1} \hat{\mu}_1(x_i) - \hat{\mu}_0(x_i)$$

- Can similarly replace this with

$$\frac{1}{\#_A} \sum_{i: x_i \in A} \hat{\mu}_1(x_i) - \hat{\mu}_0(x_i)$$

in order to estimate the subgroup effect

# Subgroup analysis

- Similar decisions need to be made here
- Fitting the outcome model only on individuals with $-_8\ 2\ C_6$ is more flexible, but also less efficient
- One alternative is to fit a model on the full sample, but include interactions between covariates and indicators of subgroup index
  - Similar to fitting separate regression models
  - Can use penalization on interaction terms to shrink/regularize towards the standard model fit on the full data
  - Balance bias and variance concerns
- One solution is the subgroup balancing propensity score (Dong et al. 2020): estimating PS that reaches a comprise between global and subgroup balance
- In general, cumbersome to implement

# Post-LASSO Algorithm to Balance Bias-Variance Tradeoff

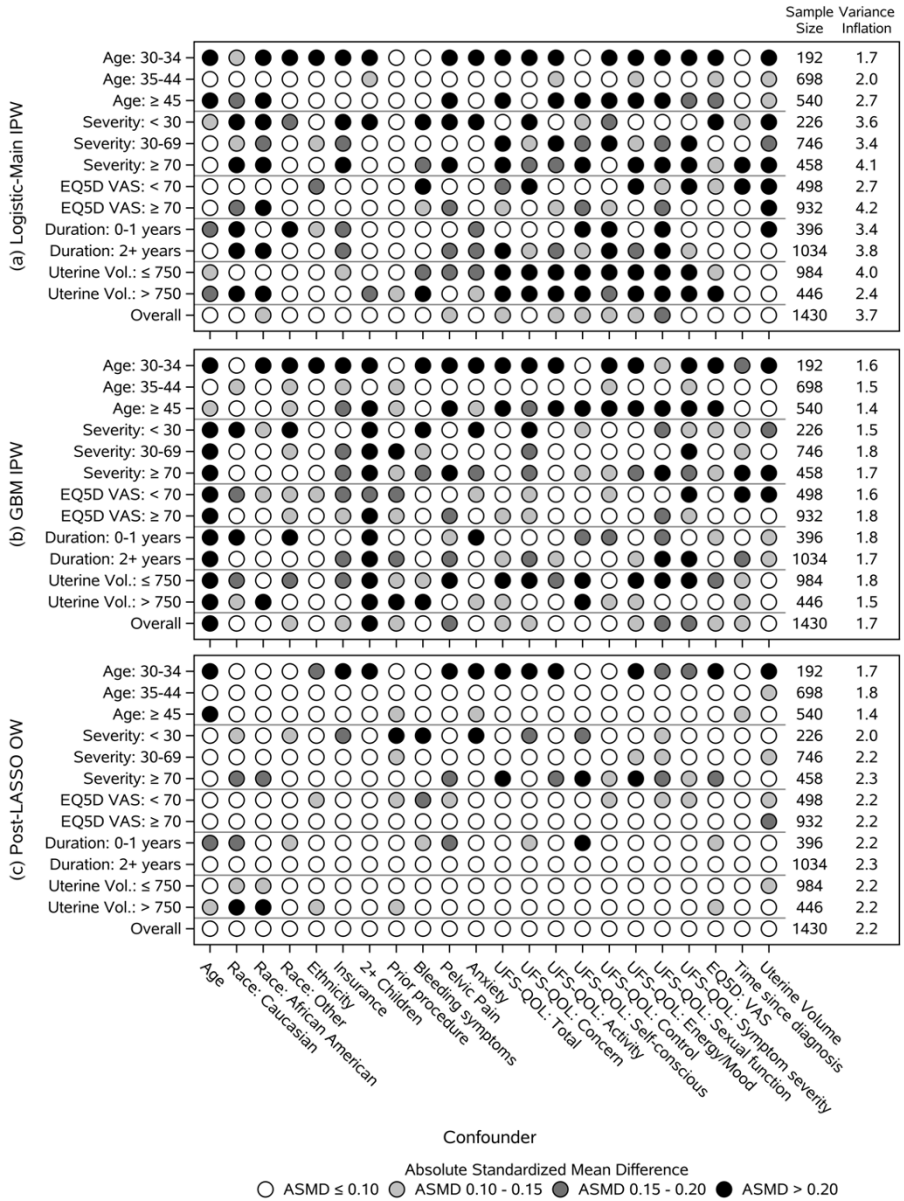
- Yang et al. 2021 (SIM) proposed to use post-LASSO estimate PS. The procedure is:
  - S1. Fit a logistic PS model with all pre-specified covariates and subgroup variables along with pairwise covariate-subgroup interactions, and perform LASSO to select covariate-subgroup interactions (without penalizing the main effects in the model).
  - S2. Estimate PS by refitting the logistic regression with all main effects and selected covariate-subgroup interactions from S1.
  - S3. Calculate a chosen type of weights (e.g. IPW or OW) based on the PS estimated from S2, and check subgroup balance before and after weighting.
  - S4. Estimate the causal effects for all prespecified subgroups using the Hajek estimator within subgroup with the weights from S3.

# Visualizing Subgroup Balance: Connect-S plot

- Difficult to visualize subgroup balance. For subgroups and ? covariate, there are ? standardized differences
- One can draw love plots, each for ? covariates, but still cumbersome
- Connect-S plot (Yang et al. 2021) visualizes ? balance statistics all at once
  - each row represents a subgroup variable, (e.g. a race group)
  - each column represents a confounder/covariate that we want to balance (e.g. age).
  - Each dot corresponds to a specific subgroup and confounder, and the shade of the dot is coded based on the corresponding balance statistics, with darker color meaning more severe imbalance.

# Connect S plot: example of COMPARE UF

Yang et al. 2021



# Case study of subgroup analysis - COMPARE UF

Yang et al. 2021

- Goal: determine whether certain patient subgroups should receive myomectomy versus hysterectomy (two treatments)
- Pre-specified 35 subgroups: categories of 16 variables including race, age, and baseline symptom severity
- 20 covariates/confounders: including demographics, disease history, quality of life and symptoms
- Total sample size: 1430, 567 in the myomectomy group and 863 patients in the hysterectomy group
- Outcome: quality of life score after 1 year
- Connect-S plot shows imbalance in many subgroup-confounder combinations