Estimating Intervals for CATE in Target Sample

Moderation Meeting: 3/6/23

Overview

Goal: Estimate conditional average treatment effect (CATE) for target group of individuals after fitting a multi-study CATE model.

1. Background
2. Notation
3. Methods
4. Open Questions

Background

- Have been using methods to estimate the conditional average treatment effect by combining data from multiple randomized controlled trials
- Several methods yield study-specific functions where study matters at varying levels
- Goal is either a universal treatment effect function or a way to apply these functions to individuals who are not coming from the original sample

Background continued

- Can we use prediction intervals from meta-analysis and apply them to a linear combination of parameter estimates?
- Other option: bootstrapping procedure where we repeat the individual observation N times, and for each iteration we randomly assign study and then predict the CATE

Notation

- $A \in \{0, 1\}$ indicates treatment status
- $\boldsymbol{X}$ are covariates (continuous)
- Y is a continuous outcome
  - $Y(1)$ is the potential outcome under treatment
  - $Y(0)$ is the potential outcome under control
- $S \in \{1, ..., K\}$ is a study indicator

Estimand

The estimand is the conditional average treatment effect (CATE):

$$\tau(\boldsymbol{X}) = E(Y(1)|\boldsymbol{X}) - E(Y(0)|\boldsymbol{X})$$

Prediction Intervals

After a random effects meta-analysis, a prediction interval can be calculated to give a range for the predicted parameter value in a new study. Assuming the random effects (that is, the individual study parameter values) are normally distributed with between-study standard deviation (τ), then the prediction interval is approximately[6]

$$\hat{\mu} - t_{k-2}\sqrt{\hat{\tau}^2 + \mathrm{SE}(\hat{\mu})^2}\ , \quad \hat{\mu} + t_{k-2}\sqrt{\hat{\tau}^2 + \mathrm{SE}(\hat{\mu})^2}$$

where $\hat{\mu}$ is the estimate of the average parameter value across studies; $\mathrm{SE}(\hat{\mu})$ is the standard error of $\hat{\mu}$; $\hat{\tau}$ is the estimate of between study standard deviation; t_{k-2} is the $100(1-\alpha/2)$ percentile of the t distribution with $k-2$ degrees of freedom, where k is the number of studies in the meta-analysis and α is usually chosen as 0.05, to give a 5% significance level and thus 95% prediction interval. A t distribution, rather than a normal distribution, is used to help account for the uncertainty of $\hat{\tau}$. The correct number of degrees of freedom for this t distribution is complex, and we use a value of $k-2$ largely for pragmatic reasons.

(Riley et al., 2011)

Prediction Intervals continued

- A typical formula for IPD MA could be

$$E(Y) = \alpha_s + \beta_{1s}X + \beta_{2s}A$$

  either in one study at a time and then aggregated (two-stage) or in one model combining all study information (one-stage). Either way, we can ultimately assume $\beta_{2s} \sim N(\beta_2, \sigma^2_{\beta_2})$. Then a prediction interval can be calculated for β_2.

- What if we add in heterogeneity, so the model looks more like

$$E(Y) = \alpha_s + \beta_{1s}X + \beta_{2s}A + \delta_s AX$$

  Then the CATE is $\theta_s(X) = \beta_{2s} + \delta_s X$. Can we create a prediction interval for this CATE?

Prediction Intervals: Two Stage MA

- It seems that prediction intervals are geared towards two-stage meta-analysis, where study-specific models are fit first and then the coefficients are combined in a second model
- To do this with an interaction term, we can fit the following model *within each study*:

$$E(Y) = \alpha_s + \beta_{1s}X + \beta_{2s}A + \delta_s XA$$

where δ_s is the average change in the treatment effect for one unit increase in X (assume X is one covariate for now). Then we would pool using a random effects model:

$$\hat{\delta}_s \sim N(\delta_s, var(\hat{\delta}_s)), \delta_s \sim N(\delta, \tau^2)$$

We could then get a prediction interval for δ using an estimate of $\hat{\delta}$, $SE(\hat{\delta})$, and $\hat{\tau}^2$.

Two Stage MA continued

- The previous slide gives us a prediction interval for the interaction coefficient, but what if we want a prediction interval of the CATE: $\theta(X) = \beta_2 + \delta X$, so that we can see what the CATE might be in our target sample and/or a different study.
- Option 1: Estimate $\theta_s(X)$ for a fixed X within each study; then fit random effects model like:

$$\hat{\theta}_s \sim N(\theta_s, var(\hat{\theta}_s)), \theta_s \sim N(\theta, \tau^2)$$

  to then apply the prediction interval formula to θ.
- Option 2: Estimate β_{2s} and δ_s within each study; fit random effects for both parameters; apply prediction interval to linear combination

Variance Calculation

- If we assume x is one covariate for now and s is some fixed value:

$$\begin{aligned} Var(\hat{\theta}_s(x)) &= Var(\hat{\beta}_{2s} + \hat{\delta}_s x) \\ &= Var(\hat{\beta}_{2s}) + Var(\hat{\delta}_s x) + 2Cov(\hat{\beta}_{2s}, \hat{\delta}_s x) \\ &= Var(\hat{\beta}_{2s}) + x^2 Var(\hat{\delta}_s) + 2xCov(\hat{\beta}_{2s}, \hat{\delta}_s) \end{aligned}$$

- All of these components should be easy to get from model results.
- For $\hat{\tau}^2$, it depends on our approach. If we fit a second stage random effects model to the $\hat{\theta}_s$'s, then we can pull $\hat{\tau}^2$ directly. If we have random effects separately associated with each coefficient, then we can take $\hat{\tau}^2 = \hat{\tau}^2_{\beta_2} + \hat{\tau}^2_{\delta}$, assuming independent random effects. (*Thanks TJ for helping with these calculations!)

One Stage IPD MA

- Can we do this in one stage? Fit a mixed effects model like

$$E(Y) = \alpha_s + \beta_{1s}X + \beta_{2s}A + \delta_s XA$$

  but this time in all studies combined, where $\alpha_s \sim N(\alpha, \sigma^2_\alpha)$, $\beta_{1s} \sim N(\beta_1, \sigma^2_{\beta_1})$, $\beta_{2s} \sim N(\beta_2, \sigma^2_{\beta_2})$, and $\delta_s \sim N(\delta, \sigma^2_\delta)$
- Then directly estimate $\theta(X) = \beta_2 + \delta X$ and use the standard errors and random effect variances to get an interval...?

Questions

- Has this been done much? What is done instead to discuss treatment effect heterogeneity?
- It looks like there are different types of prediction intervals and some debate around how to calculate the standard error - any thoughts?
- Are these prediction intervals meant for aggregate data MA only, and can it still be used with IPD MA?
- Does it make sense computationally to be calculating intervals for CATEs when we have so many possible covariate values?
- Extending to multiple covariates

Supplementary Info

Bootstrapping Methods

1. **Completely random:** Repeat each observation N times and for each repetition, randomly sample S with equal probability for each value. Predict and then calculate a confidence interval using the mean and standard deviation of the N predictions.
2. **Study membership model:** Same as completely random but when assigning study, do it according to a distribution defined by the probabilities of a study membership multinomial logistic regression model.
3. **Within-forest default:** Follow the default process of the causal forest, meaning split such that all missing values go in the same direction. This won't be a bootstrap anymore.
4. **Within-forest random sampling:** assign study based on the probabilities in each split of the tree.

Simulation Steps: Training Data

Training Data: Mimic MDD dataset: RCTs comparing Duloxetine versus Vortioxetine for reduction in depressive symptoms, measured by MADRS score

- 6 studies with 200 or 500 people in each
- Probability of treatment is 1/2
- 5 covariates per person: sex, smoking status, weight, baseline MADRS (depression scale), age such that:
  1. Same covariate distribution
  2. MADRS varies
  3. MADRS and age vary
  4. Age completely distinguishable across studies

Simulation Steps: Target Data

- Simulate target data using one of three options
  1. Random sample of 100 individuals from the training data
  2. Random sample of 100 individuals from the training data, where individuals from studies 3 and 5 have three times the chance of getting selected
  3. Different sample of 100 individuals who are younger and have less severe depression according to their MADRS score
- In all three options, study is assigned to be missing

Simulation Steps: Outcome and Treatment Effect

- Define m and τ function:
  - **Simple:**
    $m = -0.02 * age - 0.7 * madrs - 0.15 * sex + \epsilon_{study-main}$ and $\tau = -8.5 + 0.07 * age + 0.20 * madrs + \epsilon_{study-tau}$, where $\epsilon_{study} \sim N(0, \epsilon_{sd}^2)$ in the training data and $\epsilon_{study} = 0$ in the target data
- Define $Y = m + A * \tau + \epsilon$ where $\epsilon \sim N(0, 0.05^2)$
- Fit model to training data using causal forest with pooling with trial indicator to estimate CATE

Simulation Steps: Method Comparison

- Predict on target data and create confidence intervals for each individual using each imputation method
- Calculate measures of accuracy
  1. MSE in training data
  2. MSE in target data
  3. Confidence interval coverage in target data
  4. Confidence interval length in target data

Simulation Parameters

Overall, the simulation will vary depending on the following parameters:

1. $N = 100$, where N is the number of repetitions in the bootstrap
2. $K = 6$, where K is the number of RCTs
3. $n_k \in \{200, 500\}$ where n_k is the sample size for study k
4. m and τ scenario: Simple with $\epsilon_{sd} \in \{0.01, 0.05, 1, 3\}$
5. Training data distributional shift: same, varying MADRS, varying MADRS and age, distinguishable age
6. Target data setup: random sample, random with upweighted studies, different distribution