Causal Analysis

Impact Evaluation and Causal Machine Learning with Applications in R

Chapter 5: Causal Machine Learning (2)

Table of Contents

Effect Heterogeneity with CML

Effect Heterogeneity

Effect heterogeneity refers to differences in causal effects across subgroups, e.g. based on observed characteristics (covariates X), such as gender, age, and income.

Exploring effect heterogeneity in multiple subgroups is not straightforward due to multiple hypothesis testing:

- Conventional t-statistics and related p-values are only valid for testing a single hypothesis, and would need to be adjusted for multiple hypothesis testing.
- CML approaches can be applied for investigating effect heterogeneity across X in a way that avoids inferential issues of multiple hypothesis testing.

Effect Heterogeneity with Causal Trees

Causal trees (Athey and Imbens, 2016) permit studying effect heterogeneity. With random treatment D, decision trees for prediction must be modified in two ways:

1. Instead of Y, the difference in Y across treatment groups serves as the outcome to be predicted, when recursively splitting the covariate space into subgroups.
   $\rightarrow$ Maximize effect homogeneity within subgroups and effect heterogeneity across subgroups.
2. Apply sample splitting, such that the tree's structure and treatment effect estimation is based on distinct parts (or folds) of the data.
   $\rightarrow$ Avoid spuriously large effect heterogeneities due to overfitting.

Effect Heterogeneity with Causal Random Forests (1)

Causal random forests (Wagner and Athey, 2018; Athey, Tibshirani, and Wager, 2019) combine sample splitting and partialling out to control for confounders with causal tree approach for finding effect heterogeneity:

1. Predict both Y and D as a function of X using random forests and leave-one-out cross-validation.
2. Compute residuals of Y and D using the predictions, which corresponds to the partialling out strategy.
3. Predict heterogeneous effects of the residuals of D on the residuals of Y as a function of X using a causal random forest, obtained by averaging across many causal trees.

Effect Heterogeneity with Causal Random Forests (2)

Causal trees vs. causal forests:

- Causal trees provide easy-to-interpret splitting rules for defining subgroups, but tend to have higher variance.
- Causal forests reduce variance by averaging across many trees, but lack straightforward guidance for defining subgroups.

Assessing effect heterogeneity:

- Causal forest yields an estimate of the conditional average treatment effect (CATE), $\Delta_X = E[Y(1) - Y(0)|X = x]$, such that we can investigate the CATE as a function of X.

Conditional Average Treatment Effects

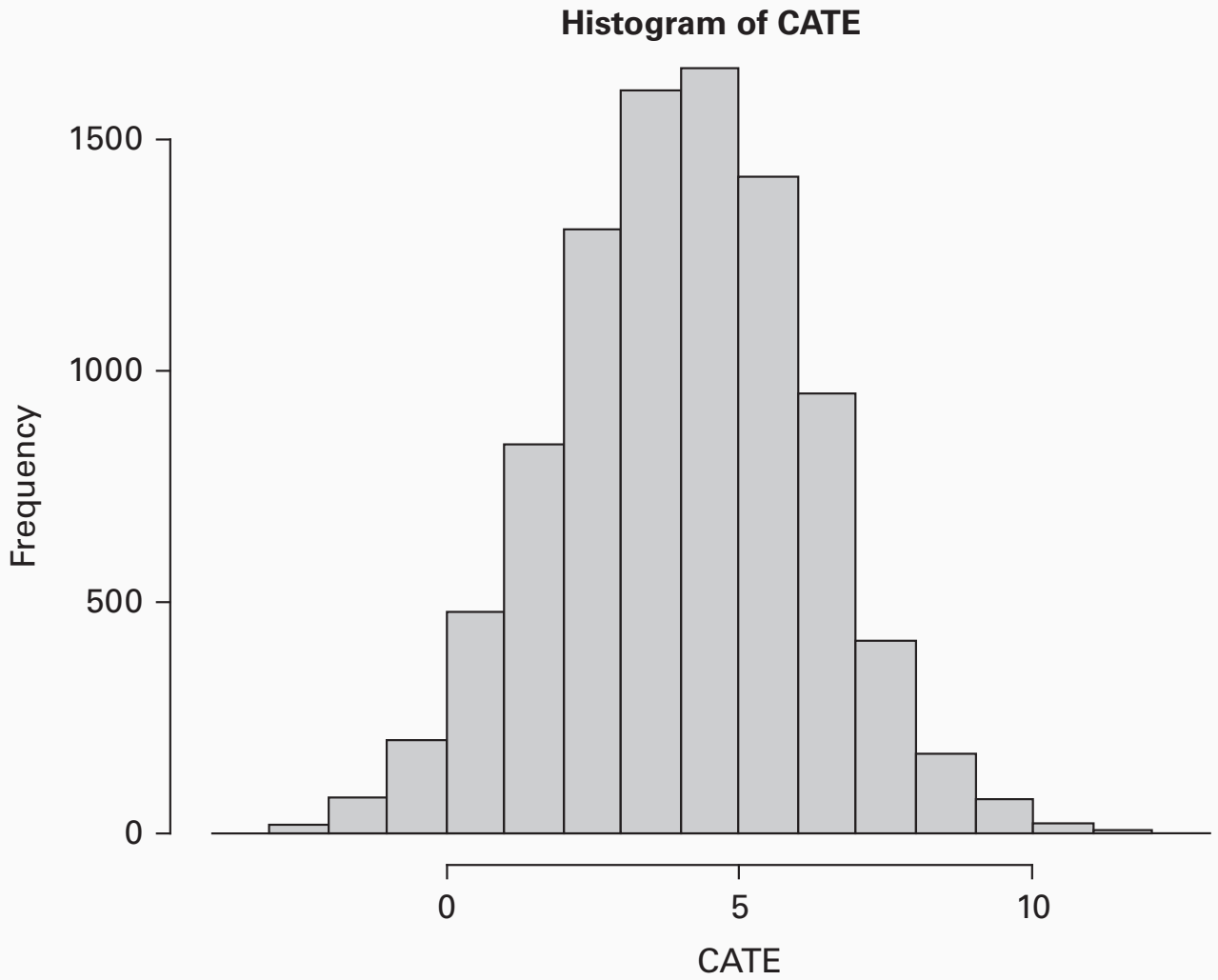

Figure 1: CATE distribution.

Causal Forests with Continuous Treatments

- Causal forests can be applied to continuous treatments.
- This approach yields the conditional average partial effect (CAPE) of marginally increasing the current treatment intensity given x, instead of the CATE.
- Formally, the CAPE is the average derivative of the conditional mean of Y with respect to D, taken over the distribution of D given $X = x$:

$$\Delta_x = E\left[\frac{\partial E[Y(D)|X=x]}{\partial D}\middle|X=x\right] = E\left[\frac{\partial E[Y \mid D, X=x]}{\partial D}\middle|X=x\right]. \quad (5.8)$$

Effect Heterogeneity with Double Machine Learning (1)

Double machine learning approach to effect heterogeneity:

- Use machine learning-based estimate of the influence function $\phi(X)$ in equation (4.45).
- (Unconditional) average treatment effect (ATE) is given by $E[\phi(X)]$.
- CATE is given by $E[\phi(X)|X = x]$.
- Regress $\hat{\phi}(X)$ on covariates X or a subset of the covariates to explore effect heterogeneity.

Effect Heterogeneity with Double Machine Learning (2)

Semenova and Chernozhukov (2021):

- OLS regression of $\hat{\phi}(X)$ on a limited number of preselected covariates can, under certain conditions, yield asymptotically correct coefficient estimates and standard errors for testing effect heterogeneity.
- Two conditions must be met:
  1. $\hat{\phi}(X)$ is estimated by cross-fitting (i.e., in a different part (or fold) of the data than the plug-in parameters).
  2. The convergence rate of the plug-in parameters is faster than $n^{-1/4}$.

When assessing effect heterogeneity across continuously distributed covariates (such as income), non-linear methods such as nonparametric kernel or series regression can be considered.

Effect Heterogeneity with Double Machine Learning (3)

Detect covariates that most importantly predict effect heterogeneity in a data-driven way:

- Predict $\hat{\phi}(X)$ as a function of X and assess which covariates X are most predictive (e.g., based on lasso regression).
- Most predictive covariates could be strongly correlated with other covariates, which rank lower in terms of predictive power but are not necessarily unimportant in terms of influencing effect heterogeneity.
- Under data-driven covariate selection for effect herogeneity analysis, statistical inference generally requires further sample splitting steps, in order to avoid overfitting.
- Use different data folds (1) for selecting predictive covariates and (2) for statistical inference (assessing whether the selected covariates statistically significantly drive effect heterogeneity).

Effect Heterogeneity with Double Machine Learning (4)

For optimal DML-based estimation of the CATE with the smallest possible bound on the estimation error, cross-fitting approach can be further refined, see Kennedy (2020):

1. Estimate propensity scores in the first fold.
2. Estimate conditional mean outcomes in the second fold.
3. Estimate the efficient influence function and CATE in the third fold.
4. Swap the roles of the folds to obtain the final CATE estimates by averaging over the CATE estimates in the various folds.

   Repeat random data-splitting into three folds multiple times and take the median CATE estimate to further reduce variance.

Table of Contents

Optimal Policy Learning

Optimal policy learning, as discussed e.g. in Manski (2004), Hirano and Porter (2009), and Kitagawa and Tetenov (2018), is related to effect heterogeneity:

- Optimal allocation of a treatment in a population as a function of covariates X when also considering costs of treatment.

Example

- Optimal price discount targeting: Only offer price discount to customers if benefits (e.g., additional sales) outweigh costs (e.g., reduced profit margin) given customers' characteristics X.

Policy Tree Illustration

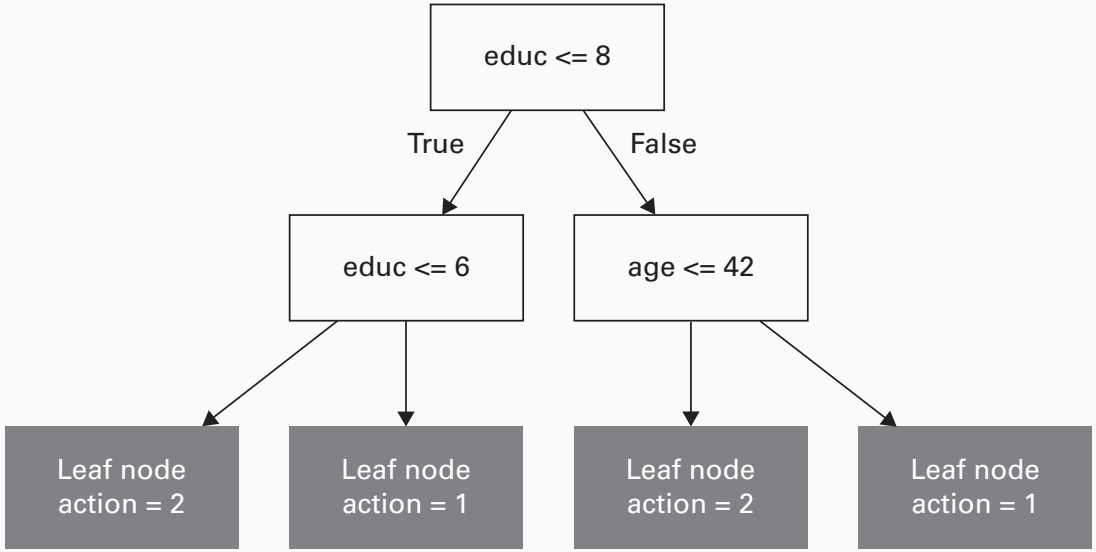

Figure 2: Policy tree.

Notation

- Consider maximizing average outcomes under a binary treatment based on covariate-dependent treatment assignment.
- $\pi(X)$ is a treatment policy defined as a function of X.
- For example, $\pi(X) = I\{age \geq 65\}$ could indicate that a medical treatment is assigned to individuals who are 65 years old and older, such that $\pi(age = 30) = 0$ and $\pi(age = 70) = 1$.

Optimal Policy

- $\Delta(\pi(X))$ is the average effect of policy $\pi(X)$, which is defined as the difference in mean potential outcomes under $\pi(X)$ versus nontreatment of everyone:

$$\begin{aligned}\Delta(\pi(X)) &= E[Y(\pi(X)) - Y(0)] = E[\pi(X) \cdot (Y(1) - Y(0))] \\ &= E[\pi(X) \cdot E[Y(1) - Y(0)|X]] = E[\pi(X) \cdot \Delta_X]. \end{aligned} \tag{5.9}$$

- The optimal policy $\pi^*(X)$ maximizes the average effect among the set of feasible policies:

$$\pi^*(X) = \max_{\pi \in \Pi} \Delta(\pi(X)), \tag{5.10}$$

where Π denotes a finite (e.g., countable) set of feasible policies.

Regret Function

- Based on equation (5.09) and (5.10), we can define a regret function:

$$R(\pi(X)) = \Delta(\pi^*(X)) - \Delta(\pi(X)), \tag{5.11}$$

  corresponding to the reduction in average policy effect $\Delta(\pi(X))$ due to implementing a suboptimal policy $\pi(X)$ instead of the optimal policy $\pi^*(X)$.
- Therefore, finding the optimal policy among the set of feasible policies implies that the average policy effect is maximized and regret R is equal to zero.

Optimal Policy with DML (1)

- Solving the optimization problem in equation (5.10) to learn the optimal policy is equivalent to solving the following maximization problem:

$$\pi^*(X) = \max_{\pi \in \Pi} E[(2\pi(X) - 1) \cdot \phi(X)], \tag{5.12}$$

  where $\phi(X)$ is the efficient influence function from equation (4.45).
- Term $(2\pi(X) - 1)$ implies that the CATEs of treated and nontreated individuals enter the expectation positively and negatively, respectively.
- Maximizing the expectation thus requires a trade-off between treated and nontreated individuals according to their CATEs, when choosing the optimal policy among the feasible policies.

Optimal Policy with DML (2)

- When applying optimal policy learning based on DML, the upper bound (or worst case value) of the regret function,

$$R(\hat{\pi}^*(X)) = \Delta(\pi^*(X)) - \Delta(\hat{\pi}^*(X)),$$

  can go to zero at $\sqrt{n}$-rate under certain conditions.
- As shown in Athey and Wager (2021), this is the case if:
  1. Plug-in parameters are estimated at a convergence rate faster than $n^{-1/4}$.
  2. Set of possible policies Π is not too complex (satisfied in a policy tree with a moderate number of subsets).

Incorporating Treatment Costs

- Until now, we did not explicitly consider treatment costs.
- But costs are relevant because:
  1. A policy should only be implemented if it outweighs the costs.
  2. Costs might vary across policy rules as a function of covariates X.
- Denote the treatment costs of policy $\pi(X)$ as $C(\pi(X))$, which might vary with the values of X.
- We modify the average effect of policy $\Delta(\pi(X))$ in equation (5.9) such that it accounts for treatment costs:

$$\Delta(\pi(X)) = E[\pi(X) \cdot \Delta_X - C(\pi(X))], \tag{5.13}$$

  which corresponds to the policy's net effect rather than its gross effect.

Table of Contents

Reinforcement Learning

Reinforcement learning:

- Machine-learning based method for learning optimal treatment policies.
- Treatment assignment is dynamic across time periods in this approach.
- Aims at learning most effective treatment by repeated assignment of multiple treatments and their evaluation across time periods (Sutton and Barto, 1998).

Notation

- Treatment D_t and outcome Y_t in period $T = t$, with $t \in \{1, 2, \ldots, \mathcal{T}\}$ and $\mathcal{T}$ denoting the total number of periods.
- D_t may take values $d_t \in \{0, 1, \ldots, J\}$, where 0 indicates no treatmant and $1, \ldots, J$ index different nonzero treatments.
- Mean potential outcome of treatment d in period t: $\mu_t(d_t) = E[Y_t(d_t)|T = t]$
- Average of the mean potential outcomes for a fixed treatment $d = d_1 = \cdots = d_{\mathcal{T}}$ across all treatment periods: $\mu(d) = E[Y_T(d)]$.

Homogeneity Assumptions

Two assumptions rule out that treatment effects interact with the time periods:

- Assume that treatments only affect outcomes in the same period t, which implies that there are no dynamic treatment effects.
- Assume that ATEs are stationary, i.e., treatment effects do not change over time.

These homogeneity assumptions imply that $\mu_t(d_t) = \mu_{t'}(d_{t'})$, for two time periods $t \neq t'$, and that $\mu_t(d_t) = \mu(d)$.

Optimal Treatment and Regret Function

- Assuming homogeneous ATEs across time, we aim to find the treatment that maximizes $\mu(d)$, denoted by $d^* = \max_{d \in \{0,1,\ldots,J\}} \mu(d)$.
- Related to optimal policy learning in chapter 5.4, we define a regret function $R_{\mathcal{T}}(d)$.
- The regret function is the difference in mean potential outcomes under the optimal treatment d^* versus another treatment d, summed over all periods $\mathcal{T}$:

$$R_{\mathcal{T}}(d) = \sum_{t=1}^{\mathcal{T}} \mu(d^*) - \mu_t(d_t). \tag{5.14}$$

Exploration-Exploitation Trade-off

Reinforcement learning faces a trade-off:

- On the one hand, continuing learning (exploration) over more time periods (and thus, data) increases the chance of finding the optimal treatment.
- On the other hand, this decreases the number of individuals to which the currently optimal treatment can be assigned (exploitation).
- If too few periods are used for exploration, we may end up with a suboptimal treatment choice.
- If too many periods are used, we increase regret by not assigning everyone to the optimal treatment.

  $\rightarrow$ Reinforcement learning must balance exploration and exploitation to minimize regret.

Adaptive Randomization

- Let us consider random treatment assignment under assumptions stated in expression (3.53) in chapter 3.
- Random assignment with equal probability in all periods is suboptimal for balancing exploration and exploitation.
- Use adaptive randomization (e.g., Thompson sampling) instead to gradually assign more individuals to better-performing treatments:
  - Start with a prior distribution for the potential outcomes of each treatment.
  - Randomly assign treatments based on the likelihood that they are optimal.
  - Update the posterior distribution based on the effects found in each newly run experiment.

Thompson Sampling

Thompson sampling (Thompson, 1933):

- Based on Bayesian updating to modify treatment assignment as more information on treatment effectiveness becomes available.
- Starts with a prior distribution $\tilde{F}_{Y(d)_0}$ for potential outcomes under each treatment d in the nonobserved period $t = 0$.
  - Neutral or uninformative prior: assuming a constant mean across potential outcome distributions of any treatment, in line with a null hypothesis of zero ATE for any treatment.
- Uses probability matching to base treatment assignment on the likelihood that a specific treatment is optimal, according to effects in past data and the prior distribution.
- Continues to explore possibly optimal treatments while gradually shifting away from underperforming treatments.
- Can attain a near-optimal regret bound.

Thompson Sampling Steps

Thompson sampling steps:

- Compute probabilities: Calculate $\tilde{p}_{d,t-1}$, the probability that treatment d is optimal, based on its performance in previous periods. In the first period $t = 0$, this is based on the prior distribution.
- Random treatment assignment: Draw treatments D_t from a multinomial distribution in which treatment assignment probabilities correspond to their probabilities to be optimal, $\tilde{p}_{d,t-1}$.
- Bayesian updating: Assess outcomes Y_t to update the posterior distribution $\tilde{F}_{Y(\cdot)_t}$, which will serve as the prior for the next period $t + 1$.

Time-Varying Effects

Possible approaches to adapt Thompson sampling to time-varying effects (nonstationary $\mu_t(d_t)$):

- Discard observations beyond a specific number of periods in the past to focus on more recent data.
- Introduce a discount factor to reduce the weight of older periods.

Inference

Due to the adaptive randomization and the time-dependence it introduces:

- The assumption of independent and identically distributed data is violated.
- Conventional t-statistics do not yield valid p-values and confidence intervals.

Apply adaptively weighted estimators for inference such as IPW-based estimation (Hadad et al., 2021):

$$\hat{\mu}(d) = \sum_{t=1}^{T} \frac{I\{D_t = d\} \cdot Y_t}{\sqrt{\tilde{p}_{d,t}}} \Big/ \sum_{t=1}^{T} \frac{I\{D_t = d\}}{\sqrt{\tilde{p}_{d,t}}}, \tag{5.15}$$

which satisfies asymptotic normality.

Heterogeneous Treatment Effects

To account for/explore heterogeneity in treatment effects, apply reinforcement learning within subsets of data defined by covariates X (Caria et al., 2020):

- More subsets enable better-tailored treatments for specific subpopulations.
- However, more subsets also imply a longer exploration period due to smaller sample sizes within each subset.
- This creates a trade-off between optimal treatment specificity and exploration duration.