Conditional Average Treatment Effect

Cholyeon Cho

February 2024

1 Potential Outcome Formulation

1.1 General Formulations

1. $Y(t)$: outcome variable for treatment t.
2. $t_0, t_1 \in \tau$: Treatments
3. X: Matrix of covariates
4. τ: Treatment effect where treatment effect given covariates are defined by $\tau(t_0, t_1, x) = E[Y(t_1) - Y(t_0)|X = x]$.
5. W: Covariates affecting the potential outcome and treatment
6. Z: Variables that affect treatment but not the potential outcome

The marginal CATE(Conditional Average Treatment Effect) in a continuous treatment environment can be calculated by the partial derivative:

$$\partial\tau(t, x) = E[\nabla_t Y(t)|X = x]$$

For the estimation, structural forms are assumed of the outcome.

$$Y = g(T, X, W, Z, \epsilon)$$

$$T = f(X, W, Z, \eta)$$

Therefore, the treatment effects can be expressed as

$$\tau(t_0, t_1, x) = E[g(t_1, X, W, \epsilon) - g(t_0, X, W, \epsilon)|X = x)]$$

1.2 Linear CATE estimator

Many methods depend on a linear relation for potential outcome depending on treatment. The assumptions can be expressed as:

$$Y = H(X, W)T + g(X, W, \epsilon)$$

$$T = f(X, W, Z, \eta)$$

Notice that the outcome doesn't have the variable Z in its formulation. This is because Z purely predict the treatment effect and already included in the T variable.

The resulting ATE of this formulation is:

$$\tau(t_1, t_0, x) = E[H(X, W)|X = x](t_1 - t_0)$$