Model Based Approach to Network Interference

Cornell STSCI / INFO / ILRST 3900
Fall 2025
causal3900.github.io

Nov 20 2025

Learning goals for today

At the end of class, you will be able to

- Use model based regression to estimate global average treatment effect under interference
- Use inverse probability weighting (IPW) to estimate global average treatment effect with a given exposure mapping
- Explain the implications of the choice of randomized design on the variance of the estimator

Logistics

- Project Check-ins due Nov 25
- PSET 6 due Nov 25; Quiz 6 Dec 2

Review of Network Interference

- Under interference the potential outcome of indiv i can depend on the treatments of others as well
- Requires a change in notation to indicate the additional dependence, e.g. $Y_i^{\mathbf{a}}$ where $\mathbf{a} = (a_1, a_2, \ldots a_n)$
- Assume potential outcome $Y_i^{\mathbf{a}}$ depends only on $\mathbf{a}$ only through treatment a_i and exposure level e_i as given by exposure mapping $e_i = f_i(\mathbf{a})$, e.g. neighborhood interference, anonymous interference

Basic Solutions

- We will focus on estimating the Global Average Treatment Effect from randomized control trials under the neighborhood interference assumption

$$\mathrm{GATE} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i^{(1,\mathbf{1})} - Y_i^{(0,\mathbf{0})}\right)$$

- Two methods for estimation under exchangeability:
  - Standardization & parametric g-formula with outcome model
  - Inverse treatment probability weighted estimator
- Earliest solutions for interference modify these approaches to estimate means under desired treatment and exposure levels

Recap of using outcome modeling

- Learn a parametric model to predict expected outcome Y given treatment and covariates

$$\begin{tikzcd} L \arrow[r] \arrow[rr, bend left] & A \arrow[r] & Y \end{tikzcd}$$

- Estimate Y_i^a using the learned model, $\hat{\mathrm{E}}(Y \mid L = \ell_i, A = a)$
- Average estimates over all units

$$\hat{\mathrm{E}}(Y^a) = \frac{1}{n} \sum_i \hat{\mathrm{E}}(Y \mid L = \ell_i, A = a)$$

- Need L to be a sufficient adjustment set so that we have conditional exchangeability
- Under RCT, don't even need to condition on L

Outcome Modeling under Network Interference

- Key Idea: Learn a parametric model to predict expected outcome $Y_i^{(a,e)}$ given treatment and exposure level
- Fit model to data $\{(A_i, E_i, Y_i)\}_{i \in [n]}$
- Typically requires anonymous interference where exposure level is number or fraction of treated neighbors treated
- For every unit i, use the learned model to predict the outcome under treatment a and exposure level e, denoted $\hat{Y}_i^{(a,e)}$
- Average over all units,

$$\hat{\mathrm{E}}(Y^{a,e}) = \frac{1}{n} \sum_i \hat{Y}_i^{(a,e)}$$

Outcome Modeling under Network Interference

- Linear models are most common, e.g. $Y_i = \alpha A_i + \beta E_i + \gamma$, where E_i is fraction of treated neighbors
- Global Average Treatment Effect

$$\begin{aligned}\widehat{\mathrm{GATE}} &= \tfrac{1}{n}\textstyle\sum_i \hat{Y}_i^{(1,1)} - \tfrac{1}{n}\sum_i \hat{Y}_i^{(0,0)} \\ &= \tfrac{1}{n}\textstyle\sum_i(\hat{\alpha}+\hat{\beta}+\hat{\gamma}) - \tfrac{1}{n}\sum_i \hat{\gamma} = \hat{\alpha}+\hat{\beta}\end{aligned}$$

- Direct Average Treatment Effect

$$\begin{aligned}\widehat{\mathrm{DATE}} &= \tfrac{1}{n}\textstyle\sum_i \hat{Y}_i^{(1,0)} - \tfrac{1}{n}\sum_i \hat{Y}_i^{(0,0)} \\ &= \tfrac{1}{n}\textstyle\sum_i(\hat{\alpha}+\hat{\gamma}) - \tfrac{1}{n}\sum_i \hat{\gamma} = \hat{\alpha}\end{aligned}$$

- Indirect Average Treatment Effect

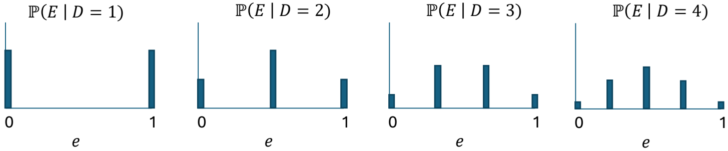
$$\begin{aligned}\widehat{\mathrm{IATE}} &= \tfrac{1}{n}\textstyle\sum_i \hat{Y}_i^{(0,1)} - \tfrac{1}{n}\sum_i \hat{Y}_i^{(0,0)} \\ &= \tfrac{1}{n}\textstyle\sum_i(\hat{\beta}+\hat{\gamma}) - \tfrac{1}{n}\sum_i \hat{\gamma} = \hat{\beta}\end{aligned}$$

Outcome Modeling under Network Interference

- What assumptions are needed?
- What is the relevant causal graph in the network setting?
- When do we need to think about adjusting for confounders?

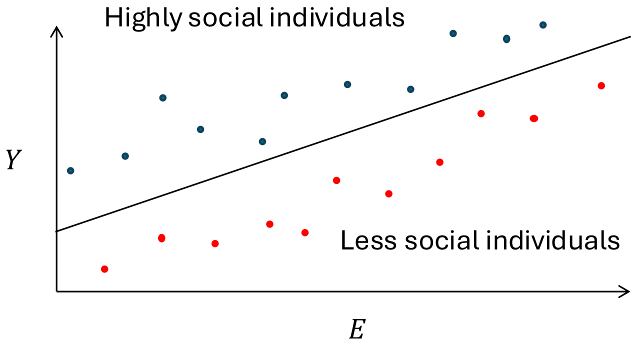
Outcome Modeling under Network Interference

- When could we still need to condition on covariates to get a sufficient adjustment set even if treatments are randomized?
- E.g. let exposure level be fraction of treated neighbors, then distribution of E_i depends on number of neighbors D_i



Outcome Modeling under Network Interference

- When could we still need to condition on covariates to get a sufficient adjustment set even if treatments are randomized?
- E.g. let exposure level be fraction of treated neighbors, then distribution of E_i depends on number of neighbors D_i
- Number of neighbors D_i affects outcome even when conditioned on exposure level



Recap of Inverse probability of treatment weighting

- Estimate means by averaging the outcomes of units with treatment $A_i = a$ multiplied by the inverse of probability of the treatment conditioned on covariates $\mathbb{P}(A_i = a \mid L_i)$

$$\hat{\mathsf{E}}(Y^a) = \frac{1}{n} \sum_{i:A_i=a} \frac{Y_i}{\mathbb{P}(A_i = a \mid L_i)}$$

- π_i denotes probability that i is treated conditioned on its covariates, s.t. $\mathbb{P}(A_i = 1|L_i) = \pi_i$ and $\mathbb{P}(A_i = 0|L_i) = 1 - \pi_i$
- Take difference of estimates for treated and control

$$\hat{\mathsf{E}}(Y^1) - \hat{\mathsf{E}}(Y^0) = \frac{1}{n} \left(\sum_i \frac{A_i Y_i}{\hat{\pi}_i} - \sum_i \frac{(1 - A_i) Y_i}{1 - \hat{\pi}_i} \right)$$

- Requires conditional exchangeability

IPW under network interference

- Modify to use exposure mapping (assume RCT)

$$\hat{\mathrm{E}}(Y^{(a,e)}) = \frac{1}{N} \sum_{i:A_i=a,E_i=e} \frac{Y_i}{\mathbb{P}(A_i = a, E_i = e)}$$

$$\widehat{\mathrm{GATE}} = \hat{\mathrm{E}}(Y^{(1,\mathbf{1})}) - \hat{\mathrm{E}}(Y^{(0,\mathbf{0})})$$

- Does not require anonymous interference, can use any exposure mapping
- Variance will depend on the exposure probabilities $\mathbb{P}(A_i = 1, E_i = \mathbf{1})$ and $\mathbb{P}(A_i = 0, E_i = \mathbf{0})$
- Observational studies are significantly more complex as we now need to care about the joint treatment probability distribution as it relates to the exposure levels

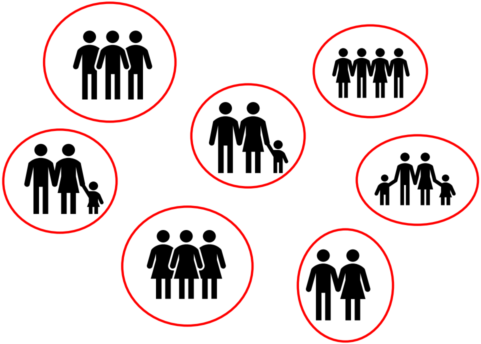
Variance of IPW estimator

$$\hat{\mathrm{E}}(Y^{(a,e)}) = \frac{1}{N} \sum_{i:A_i=a,E_i=e} \frac{Y_i}{\mathbb{P}(A_i = a, E_i = e)}$$

- If exposure probabilities $\mathbb{P}(A_i = 1, E_i = \mathbf{1})$ and $\mathbb{P}(A_i = 0, E_i = \mathbf{0})$ are small, then any measurement noise in the outcomes will be amplified, leading to high variance
- Let D_i denote the number of neighbors (including i itself)
- Under independent treatment w/prob 0.5, exposure probability is exponential in D_i, i.e. $\mathbb{P}(A_i = 1, E_i = \mathbf{1}) = (0.5)^{D_i}$
- For $D_i = 5$, exposure prob is 0.000976, s.t. in a network of 1000 nodes, likely no units observed under full treatment
- Sophisticated clustered treatment assignments reduce variance by increasing probability of full treatment / control

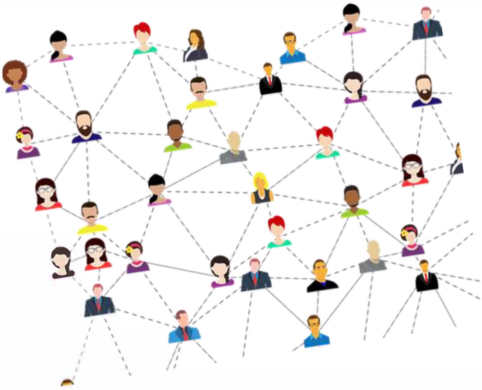
Cluster Randomized Designs

- Initially motivated by networks consisting of many tightly connected households
- No interference edges across households
- Assign treatments to each household jointly
- If household treatment probability is 0.5, then full exposure probability is 0.5



Cluster Randomized Designs

- Can use clustering algorithms on general graphs
- Assign treatments to each cluster jointly
- If cluster treatment probability is 0.5, then full exposure probability is $0.5^{(\#\text{ neighboring clusters})}$



Learning goals for today

At the end of class, you will be able to

- Use model based regression to estimate global average treatment effect under interference
- Use inverse probability weighting (IPW) to estimate global average treatment effect with a given exposure mapping
- Explain the implications of the choice of randomized design on the variance of the estimator