Causal Inference Under Interference

Sheyan Lalmohammed

Wharton Directed Reading Program (WDRP)

Spring 2025

Four Foundational Assumptions of Causal Inference

- **Consistency**
  For every unit i, if the treatment received is $A_i = a$, then the observed outcome equals the potential outcome under a:
  $$A_i = a \implies Y_i = Y_i(a).$$
- **Ignorability**
  Conditional on covariates X_i, the vector of potential outcomes is independent of treatment assignment:
  $$\{Y_i(a) : a \in \mathbf{A}\} \perp\!\!\!\perp A_i \bigm| X_i.$$
- **Positivity**
  Every treatment level must be feasible at each point of the covariates which means no value of X rules out any $A = a$:
  $$0 < P\big(A_i = a \mid X_i = x\big) < 1 \quad \forall\, a, x \text{ with } P(X_i = x) > 0.$$
- **SUTVA (Stable Unit Treatment Value Assumption)**
  Treatments are well-defined and any unit's outcome depends only on its own treatment, not others treatments:
  $$Y_i(a_1, \ldots, a_n) = Y_i(a_i).$$

What is Interference?

A violation of SUTVA where one unit's treatment affects another unit's outcome:

$$Y_i(\mathbf{a}) \neq Y_i(a_i) \quad \text{if } a_j \text{ influences } Y_i.$$

Examples

- A_i: Vaccination status of individual i.
- Y_i: Whether i becomes infected.
- *Spillover Effect (Herd Immunity)*: Even unvaccinated i benefit if many neighbors are vaccinated.

Modeling Strategies

• *Partial interference:* Assume units only spill over within defined groups, and model each individual outcome as a function of both their own treatment and the proportion of those treated within their group.

• *Exposure mapping:* Define a summary variable to capture exposure and estimate both how outcomes respond to both personal treatment and the network exposure. Common examples in the literature include the fraction of treated neighbors or distance-weighted sum of neighbors treatments

Bias in Average Treatment Effect

Standard estimators assume no interference. Thus, bias in estimates of the direct treatment effect can arise if spillovers from interference are ignored.

Average Treatment Effect (ATE)

$$\mathbb{E}\big[Y(1) - Y(0)\big] \rightarrow \mathbb{E}_X\big[\mathbb{E}[Y \mid A = 1, X] \; - \; \mathbb{E}[Y \mid A = 0, X]\big]$$

Examples

$$\textbf{Model:} \;\; Y_i = \alpha_0 + \alpha_1 A_i + \underbrace{\alpha_2 \frac{\sum_{j \in N_i} A_j}{|N_i|}}_{\text{Interference}} + \beta X_i + \underbrace{\gamma \frac{\sum_{j \in N_i} X_j}{|N_i|}}_{\text{Homophily}} + \epsilon_i$$

Bias in ATE:

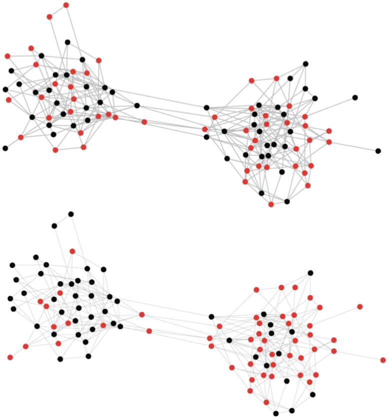
$$\sum_{x \in X^*} (\mathbb{E}[Y_i \mid A_i = 1, X_i = x] - \mathbb{E}[Y_i \mid A_i = 0, X_i = x]) P(X_i = x)$$

$$= \frac{\alpha_2}{|N_i|} \sum_{j \in N_i} (\mathbb{E}[A_j \mid A_i = 1] - \mathbb{E}[A_j \mid A_i = 0]) + \frac{\gamma}{|N_i|} \sum_{j \in N_i} (\mathbb{E}[X_j \mid A_i = 1] - \mathbb{E}[X_j \mid A_i = 0])$$

$$= \alpha_2 \mathbb{E}_{N_i}\{[A_j \mid A_i = 1] - [A_j \mid A_i = 0]\} + \gamma \mathbb{E}_{N_i}\{[X_j \mid A_i = 1] - [X_j \mid A_i = 0]\}$$

Homophily and Contagion

Homophily and Contagion Are Generically Confounded in Observational Social Network Studies(*Cosma Rohilla Shalizi & Andrew C. Thomas*)



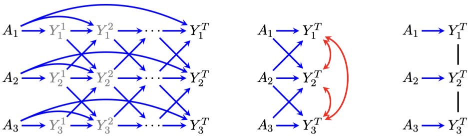
Interpretation in terms of homophily vs contagion:

- **Homophily** creates the structure: two separated communities of like-trait individuals.
- **Contagion** spreads whatever choice happens to dominate in each community.

Together (or even seperately), they produce a strong correlation between trait and choice, just as if X caused Y, demonstrating that homophily + contagion can mimic a direct causal effect.

DAGs and Chain Graphs

Causal Inference, Social Networks, and Chain Graphs (*Elizabeth L. Ogburn, Ilya Shpitser & Youjin Lee*)



(a) (b) (c)

(a) **Temporal DAG with Spillover Effects**: Directed arrows $Y_i^t \to Y_j^{t+1}$ encode one-step interference (contagion) from unit i to unit j.

(b) **Latent-Projection Graph**: When the full time-indexed DAG is collapsed onto only the observed final outcomes Y_i^T, all hidden contagion paths between i and j turn into a single bidirected edge which are unmeasured common causes induced by interference.

(c) **Chain-graph approximation**:
- Directed edges $A_i \to Y_i$: individual treatment effects.
- Undirected edges $Y_i - Y_j$: concise summary of spillover structure among neighbors

Causal Inference under Interference with Unknown Networks

Causal Inference Under Interference And Network Uncertainty
(*Bhattacharya, Malinsky & Shpitser*)

The challenge: We observe only each unit's treatment A_i, covariates X_i, and outcome Y_i, yet interference travels along social ties we haven't measured which leads to biased direct-effect estimates and hidden spillovers.

Step 1: Chain Graph Structure Learning

- Model the joint distribution of Y_i given (A_i, X_i) as a chain graph. There is directed edges $A_i \rightarrow Y_i$,$X_i \rightarrow Y_i$ for direct effects, and undirected edges $Y_i - Y_j$ for potential spillovers.
- Estimate the undirected component by fitting a penalized pseudo-likelihood to the chain-graph, shrinking insignificant $Y_i - Y_j$ links to zero.
- The remaining undirected edges indicate likely interference pathways in the hidden network.

Step 2: Auto-G-Computation

- Simulate counterfactual outcomes $Y_i(\mathbf{a})$ on the learned chain graph.
- Estimate direct and spillover effects by averaging over many $\mathbf{a}$ draws.

Additional Key Takeaways from Causal Inference Papers

- **Average Treatment in the Presence of Unknown Interference**
  Fredrik Sävje, Peter M. Aronow & Michael G. Hudgens — Demonstrates consistency of average treatment estimators under unknown limited interference.
- **Causal Inference for Social Network Data**
  Elizabeth L. Ogburn, Oleg Sofrygin, Ivan Diaz & Mark J. van der Laan — Introduces semiparametric estimators for causal effects allowing contagion and similarity dependence.
- **Identification and Estimation of Treatment and Interference Effects in Observational Studies on Networks**
  Laura Forastiere, Edoardo M. Airoldi & Fabrizia Mealli — Extends unconfoundedness to neighbors and uses generalized propensity scores for spillovers.
- **Causal Inference Under Approximate Neighborhood Interference**
  Michael P. Leung — Proposes a decaying-spillover neighborhood model and shows IPW estimators remain consistent and estimate useful exposure effects.

Questions?

Thank you!