Causal and statistical inference with social network data: Massive challenges and meager progress

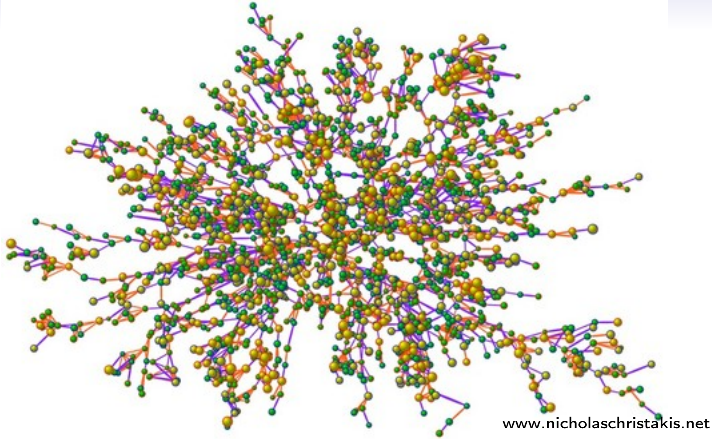
Elizabeth L. Ogburn

Department of Biostatistics,
Johns Hopkins University

outline

- Brief history of causal and statistical inference using network data.
- What is network dependence, and why is it a problem?
    - Standard errors and effective sample size.
- Some solutions:
    - test for dependence,
    - harness conditional independences,
    - adapt results for spatial-temporal dependence to the network setting.

Disclaimer: this talk is about the problem more than the solutions.



- Two challenges for causal inference using data sampled from a single social network:
  - nonparametric identification of causal effects with interference,
  - valid statistical inference under complex forms of dependence among observations.

brief history of inference using network data

- Computer scientists, physicists and mathematicians have been researching networks for decades:
  - topology, diffusion properties, generation,...
  - very little statistics for outcomes on network nodes.
- Statisticians have been researching dependent data for decades:
  - time series
  - spatial data
  - interference
  - surprisingly little of this work is directly applicable to networks.

brief history of inference using network data

- Christakis and Fowler (2007, 2008, 2009, 2010, 2011, 2012) initiated a wave of interest in estimating peer effects from social network data.
  - To examine peer effects, they fit models

$$Y^t_{ego} \sim Y^{t-1}_{alter}, Y^{t-2}_{alter}, Y^{t-2}_{ego}, \boldsymbol{C}_{ego}$$

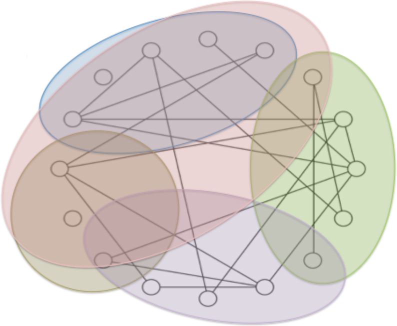
  - Widely publicized results include significant peer effects for obesity, smoking, alcohol consumption, sleep habits, etc.
  - Researchers began using similar models to assess peer effects across a wide range of disciplines and problems (e.g. Ali and Dwyer, 2009; Cacioppo et al., 2009; 2008; Lazer et al., 2010; Rosenquist et al., 2010, Wasserman 2012).

brief history of inference using network data

- Randomization based inference facilitates hypothesis testing (e.g. Toulis & Kao, 2013; Bowers et al., 2013; Aronow & Samii, 2013; Eckles et al., 2014).
- Work on interference also relies on randomization but may provide a solution to the problem of network dependence in cluster randomized trials (e.g. Sobel, 2006; Hong & Raudenbush, 2006; Rosenbaum, 2007; Hudgens & Halloran, 2008; Tchetgen Tchetgen & VanderWeele, 2012).
- Mathematical modeling of contagious processes avoids these problems but is highly dependent on parametric assumptions about agent-based processes (e.g. Steglich, Snijders & Pearson, 2007; Railsback & Grimm, 2011).

sources of network dependence

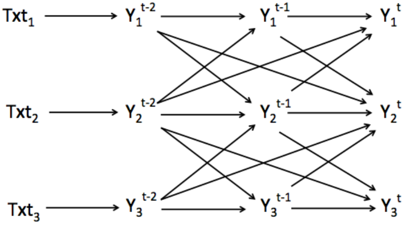
- **Latent variables** cause outcomes among close social contacts to be more correlated than among distant contacts. (E.g. homophily, geography, shared culture, shared genetics.)



- Similar to spatial dependence.

sources of network dependence

- **Contagion** implies information barrier structures, e.g. $\left[Y_1^t \perp Y_2^t \mid Y_1^{t-2}, Y_2^{t-2}, Y_1^{t-1}, \text{ and } Y_2^{t-1}\right]$ and $\left[Y_1^{t-2} \perp Y_3^{t-1}\right]$.



- When a network is observed at a single time point, this will resemble latent variable dependence.
- If the network is observed frequently, so that the outcome can't diffuse very far between observations, we can harness conditional independence restrictions to facilitate inference.

why is dependence a problem?

- Statistical analysis that incorrectly assumes independence will be invalid.
- Two problems for traditional frequentist inference:
  - CLT may not hold,
  - Standard error estimates and resulting inference will be anticonservative.

- If $E[\bar{Y}] \to \mu$, the rate of convergence is determined by

$$var(\bar{Y}) = \frac{1}{n^2}\left\{\sum_{i=1}^{n}\sigma^2 + \sum_{i\neq j} cov(Y_i, Y_j)\right\}$$

- Define

$$b_n = \frac{1}{n}\sum_{i\neq j} cov(Y_i, Y_j)$$

- Now

$$var(\bar{Y}) = \frac{\sigma^2}{n/\left(1+\frac{b_n}{\sigma^2}\right)}$$

- If a CLT holds, then

$$\sqrt{\frac{n}{1+\frac{b_n}{\sigma^2}}}\left\{\bar{Y}-\mu\right\} \xrightarrow{d} N(0,\sigma^2)$$

solutions

- Test for dependence.
- Estimate b_n.
- Perform inference on conditionally independent units.
- Use conditional independences to derive new IF asymptotics.
- Extend results for spatial / temporal dependence to this new topology.
    - weakly dependent clusters
    - subsampling methods
    - k-dependence

test for network dependence

- Sample K non-overlapping pairs of observations from the network, where the k^{th} pair is (Y_{1k}, Y_{2k}).
- The test statistic

$$\Psi \equiv \sum_{k=1}^{K} \left(Y_{1k} - \bar{Y}\right)\left(Y_{2k} - \bar{Y}\right)$$

  will equal 0 under the null hypothesis of independence.
  - But it will also be close to 0 under the alternative hypothesis!
- If we select K pairs of nodes who are friends with one another, then we can derive tests that are powered for the alternative hypothesis of network dependence.

test for network dependence

- But it's difficult to calculate the power, and to be conservative in this test we want to minimize type II rather than type I error. Use informal diagnostics instead:
- Define

$$\Psi_\Delta \equiv \sum_{k \in \mathscr{K}_\Delta} \left(Y_{1k} - \bar{Y}\right)\left(Y_{2k} - \bar{Y}\right)$$

$$\mathscr{K}_\Delta = \{(i,j) : |i-j| = \Delta\}$$

  - Does Ψ_Δ decrease with Δ?
- Let

$$\Phi_\Delta = \sum_{k \in \mathscr{K}_\Delta}^{K} I\left\{\left|\left(Y_{1k} - \bar{Y}\right)\left(Y_{2k} - \bar{Y}\right)\right| > 0\right\}$$

  - Does Φ_Δ decrease with Δ?

estimate b_n

What if tests and diagnostics lead to the conclusion that network dependence is present?

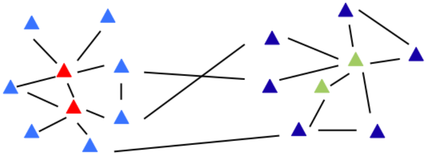
- If we observe M independent networks, we can estimate $var(\bar{Y})$ directly.
- If M is small we may be able to estimate b_n via $Cov_n(\underset{\sim}{Y})$:

$$\left[\hat{Cov}_n(\underset{\sim}{Y})\right]_{i,j} = \frac{1}{M}\sum_{m=1}^{M}\left(Y_i^m - \bar{Y}_i\right)\left(Y_j^m - \bar{Y}_j\right)$$

$$\begin{aligned}\hat{b}_n &= \sum_{i=1}^{n}\sum_{j\neq i}^{n}\left[\hat{Cov}_n(\underset{\sim}{Y})\right]_{i,j} \\ &= \frac{1}{n}\frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{n}\sum_{j\neq i}^{n} Y_i^m Y_j^m - \frac{1}{n}\frac{1}{M^2}\sum_{m=1}^{M}\sum_{k=1}^{M}\sum_{i=1}^{n}\sum_{j\neq i}^{n} Y_i^m Y_j^k\end{aligned}$$

conditionally independent units

- Create conditionally independent units; analyze with standard, i.i.d. models, but **conditional** on "information barriers."
- Randomly sample non-overlapping groups from the network.



- This will allow us to condition on an "information barrier."
- Now we can estimate conditional estimands using standard statistical machinery like GLMs.
    - The residuals will be uncorrelated across subjects despite the dependence structure.

For details see Ogburn & VanderWeele (2014). Vaccines, contagion, and social networks. (available on arXiv)

conditional independences + IFs

ongoing joint work with Ivan Diaz, Mark van der Laan, Oleg Sofrygin

- Extension of semiparametric, influence-function-based inference from the iid setting.
- We define a model $\mathscr{M}$, which restricts the observed data distribution in some way(s).
- We are interested in estimating a parameter ψ under model $\mathscr{M}$, i.e. a functional of the observed data.
- Under $\mathscr{M}$, there is a class of **influence functions** for ψ.
  - Each (RAL) estimator $\hat{\psi}$ is paired with an IF φ, and in the iid setting

$$\sqrt{n}(\hat{\psi} - \psi) \overset{p}{\approx} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \varphi(O_i)$$

  - Because the IF has mean 0 at the true parameter value, we can use it to create unbiased estimating functions for ψ.

- We have a model in which each individual's covariates, treatment, and outcome can depend on his alters' covariates, treatments, and outcomes.
    - All dependence must be due to observed contagion!
- Under the usual assumptions of positivity, no unmeasured confounding, and consistency, we can identify $E\left[\bar{Y}^*\right]$: the expected mean counterfactual under any well-defined intervention.
- The observed data quantity that identifies $E\left[\bar{Y}^*\right]$ is ψ; we have derived the efficient influence function $\varphi(\mathbf{O})$ for ψ (in a particular semiparametric model).
    - Unlike in the iid setting, $\varphi(\mathbf{O}) \neq \frac{1}{n}\sum_{i=1}^{n}\varphi(\mathbf{O}_i)$. Instead, $\varphi(\mathbf{O}) = \sum_{i=1}^{n}\varphi_i(\mathbf{O})$.

- Turning the efficient IF into an estimating equation and solving it gives us an estimate $\tilde{\psi}$ of ψ.
- $\tilde{\psi}$ is asymptotically efficient and doubly robust.
- If each subject interferes with $\leq K$ other subjects, then (van der Laan, 2014)

$$\sqrt{n}(\tilde{\psi} - \psi) \rightarrow N(0, var(IF))$$

- If $K \rightarrow \infty$ as $n \rightarrow \infty$ s.t. $\frac{K}{n} \rightarrow 0$, then

$$\sqrt{\frac{n}{K}}(\tilde{\psi} - \psi) \rightarrow N(0, var(IF))$$

extend results from spatial/temporal dependence

- Some definitions:
  - Stationarity: features of the distribution of observations does not depend on location in the network.
  - M-dependence: $Y_i \perp Y_j$ if $\|i,j\| > m$.
  - Mixing conditions: $Cov\left(Y_i, Y_j\right) \rightarrow 0$ as $\|i,j\| \rightarrow \infty$.

Why can't we use spatial dependence results?

Network topology doesn't naturally correspond to Euclidean space.

- In order to embed a network in $\mathbb{R}^d$, we would have to let d grow with sample size n.
  - Spatial results require d to be fixed or to grow slowly with n.
- Population growth is usually assumed to occur at the boundaries of the d-dimensional space.
  - It's not clear how to define boundaries in networks.
- Mixing assumptions and m-dependence don't imply bounded correlation structure.
  - In spatial data most observations are distant from one another.
  - The maximum network-based distance between two observations may be very small.
  - The distance distribution may not be right-skewed enough.

subsampling

- Subsampling has been used in many spatial dependence contexts (cf. Lahiri, 2003; Politis, Romano & Wolf, 1999), but neither the implementation nor the conditions under which it is appropriate are immediately applicable to networks.
- Under mild stationarity and dependence conditions, we can subsample to estimate $Var(\bar{Y})$:

  1. Select B subsamples of "consecutive" observations.
  2. In each subsample, calculate the subsample variance estimator $\widehat{\sigma}_b^2$.
  3. Estimate $Var(\bar{Y})$ with the average of the subsample estimators: $\widehat{\sigma}_{\bar{Y}}^2 = \frac{1}{B}\sum_{b=1}^{B}\widehat{\sigma}_b^2$.

subsampling

This is reasonable if, as $n_b \to \infty$ and $n \to \infty$,

1. $\widehat{\sigma}^2_{\bar{Y}}$ is asymptotically unbiased for $Avar(\bar{Y})$.
   - This will hold if key features of the network and the mean and variance of Y are stationary over groups smaller than the subsample sizes.
2. $Var\left(\widehat{\sigma}^2_{\bar{Y}}\right) \to 0$, where

$$
\begin{aligned}
Var\left(\widehat{\sigma}^2_{\bar{Y}}\right) =& \frac{1}{B^2}\sum_{b=1}^{B} Var\left(\widehat{\sigma}^2_{b}\right) && (1)\\
&+ \frac{1}{B^2}\sum_{\|I_b, I_d\| \leq m} 2Cov\left(\widehat{\sigma}^2_{b}, \widehat{\sigma}^2_{d}\right) && (2)\\
&+ \frac{1}{B^2}\sum_{\|I_b, I_d\| > m} 2Cov\left(\widehat{\sigma}^2_{b}, \widehat{\sigma}^2_{d}\right) && (3)
\end{aligned}
$$

weakly dependent clusters

If K clusters are asymptotically mean independent from one another, there are two approaches we might consider:

1. $T-$distribution based confidence intervals (Ibragimov & Muller, 2010; Bester, Conley & Hansen, 2011).
    - Requires asymptotic normality and mean stationarity at the cluster level.
2. Bootstrap the weakly dependent communities.
    - Stationarity is required only at the cluster level.

In the spatial dependence literature mean independence is justified with conditions on the relative size of the boundaries and interiors of the clusters; growth in d dimensions uniformly.

- These conditions don't translate into the network setting...

k-dependence

In some settings it may be expedient to estimate $Cov\left(\underset{\sim}{Y}\right)$ directly.

- K-dependence: $Cov(Y_i, Y_j) = \sigma_k$, where $k = \|i,j\|$.
- Under k-dependence, m-dependence, and mean stationarity, we can get an unbiased and consistent estimate of $Cov\left(\underset{\sim}{Y}\right)$ by this procedure:
  1. For each $k < m$, select pairs of nodes that are k units apart, such that the pairs themselves are at least m units apart from on another.
  2. Estimate $\widehat{\sigma}_k$ with the average covariance across the selected pairs.
  3. Estimate $Cov\left(\underset{\sim}{Y}\right)$ with the plug-in estimator.
- This doesn't demand as much from m-dependence as other procedures do...

summary and other directions

- Although it is accepted practice in many areas, it can be very dangerous to assume that observations are independent when they may not be!
- Different types of asymptotics:
    - combine infill and increasing domain asymptotics,
    - “fractal” asymptotics.
- Learn a new, latent distance metric. (E.g. work by Adrian Raftery & colleagues)

Thank you

references

Rubin, D. B. (1990). On the application of probability theory to agricultural experiments. essay on principles. section 9. comment: Neyman (1923) and causal inference in experiments and observational studies. Statistical Science 5, 472–480.

Halloran, M. E. & Struchiner, C. J. (1995). Causal inference in infectious diseases. Epidemiology , 142–151.

Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate? Journal of the American Statistical Association 101, 1398–1407.

Hong, G. & Raudenbush, S. W. (2008). Causal inference for time-varying instruc- tional treatments. Journal of Educational and Behavioral Statistics 33, 333–362.

Vansteelandt, S. (2007). On confounding, prediction and efficiency in the analysis of longitudinal and cross-sectional clustered data. Scandinavian Journal of Statistics 34, 478–498.

Rosenbaum, P. R. (2007). Interference between units in randomized experiments. Journal of the American Statistical Association 102, 191–200.

Hudgens, M. G. & Halloran, M. E. (2008). Toward causal inference with inter- ference. Journal of the American Statistical Association 103, 832–842.

references

Graham, B. S., Imbens, G. W. & Ridder, G. (2010). Measuring the effects of segregation in the presence of social spillovers: A nonparametric approach. Tech. rep., National Bureau of Economic Research.

Manski, C. F. (In press). Identification of treatment response with social interactions. The Econometrics Journal.

Tchetgen Tchetgen, E. J. & VanderWeele, T. J. (2012). On causal inference in the presence of interference. Statistical Methods in Medical Research 21, 55–75.

Aronow, P. & Samii, C. (2012). Estimating Causal EffectsUnder General Interference. Working paper (http://pantheon.yale.edu/~pma5/ate-interference.pdf).

van der Laan, M. J. (2012). Causal Inference for Networks. U.C. Berkeley Division of Biostatistics Working Paper Series Paper 300.

Lyons, R. (2011). The spread of evidence-poor medicine via áawed social- network analyses. Statistics, Politics and Policy 2(1), Article 2, 1-26.

VanderWeele, T.J., Ogburn, E.L. & Tchetgen Tchetgen, E.J. (2012). Why and When "Flawed" Social Network Analyses Still Yield Valid Tests of no Contagion. Statistics, Politics, and Policy 3(1).

references

Christakis, N.A. and Fowler, J.H. (2011). Social contagion theory: examining dynamic social networks and human behavior. Statistics in Medicine.

Shalizi, C.R., Thomas, A.C. (2011). Homophily and contagion are generically confounded in observational social network studies. Sociological Methods and Research, 40:211-239.

Shalizi C.R. (2012). Comment on Why and When 'Flawed' Social Network Analyses Still Yield Valid Tests of no Contagion. Statistics, Politics, and Policy 3(1).

Ogburn, E.L. & VanderWeele, T.J. (2012). Causal diagrams for interference and contagion. Under revision.

Chen, L.H.Y. (1978). Two central limit problems for dependent random variables. Probability Theory and Related Fields 43, no. 3: 223-243.

Barbour, A. D., Karoński, M., & Ruciński, A. (1989). A central limit theorem for decomposable random variables with applications to random graphs. Journal of Combinatorial Theory, Series B, 47(2), 125-145.

Rinott, Y., & Rotar, V. (1996). A Multivariate CLT for Local Dependence with $n^{-1/2} \log n$ Rate and Applications to Multivariate Graph Related Statistics. Journal of multivariate analysis, 56, 333-350.

references

Raic, M. (2003). Normal approximation by Stein's method. In Proceedings of the seventh young statisticians meeting.

Chen, L. H., & Shao, Q. M. (2005). Stein's method for normal approximation. An introduction to Stein's method, 4, 1-59.

Chen, L. H., & Shao, Q. M. (2004). Normal approximation under local dependence. The Annals of Probability, 32(3), 1985-2028.

Ibragimov, R., & Müller, U. K. (2010). T-statistic based correlation and heterogeneity robust inference. Journal of Business & Economic Statistics, 28(4), 453-468.

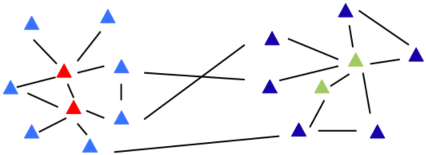
Bester, C. A., Conley, T. G., & Hansen, C. B. (2011). Inference with dependent data using cluster covariance estimators. Journal of Econometrics, 165(2), 137-151.

Lahiri, S. N. (2003). Resampling methods for dependent data. Springer.

Politis, D., Romano, J. P., & Wolf, M. (1999). Weak convergence of dependent empirical measures with application to subsampling in function spaces. Journal of statistical planning and inference, 79(2), 179-190.

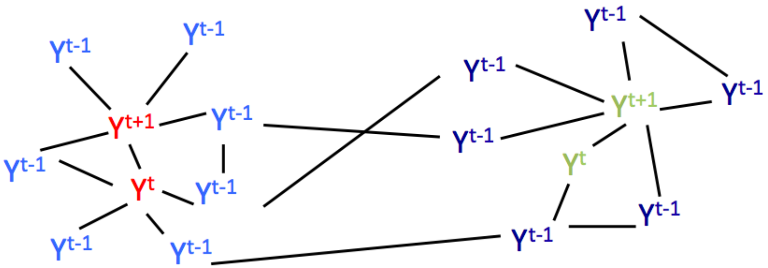
create conditionally independent units

- Randomly sample non-overlapping groups from the network.



- This will allow us to condition on an "information barrier."
- Now can estimate conditional estimands using standard statistical machinery like GLMs.
  - The residuals will be uncorrelated across subjects despite the dependence structure.

create conditionally independent units



- Regress Y^{t+1} on Y^{t} conditional on $\{Y^{t-1}\}$
  - This is more appropriate for causal effects than for sample means.

more principled solution

ongoing joint work with Ivan Diaz, Mark van der Laan, Oleg Sofrygin

- Extension of semiparametric, influence-function-based inference from the iid setting.
- We define a model $\mathscr{M}$, which restricts the observed data distribution in some way(s).
- We are interested in estimating a parameter ψ under model $\mathscr{M}$, i.e. a functional of the observed data.
- Under $\mathscr{M}$, there is a class of **influence functions** for ψ.
    - Each (RAL) estimator $\hat{\psi}$ is paired with an IF φ, and in the iid setting

$$\sqrt{n}(\hat{\psi} - \psi) \overset{p}{\approx} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \varphi(O_i)$$

    - Because the IF has mean 0 at the true parameter value, we can use it to create unbiased estimating functions for ψ.

- van der Laan (2014) extended this approach to settings with interference and/or contagion.
  - Not partial interference, but each subject can only interfere with $\leq K$ other subjects.
- We extend van der Laan (2014) to social network settings:
  - K grows with n
  - highly connected "hubs" may exert undo influence
  - estimation of causal effects of interventions on features of network topology
- This framework can handle longitudinal data, but for simplicity we focus on the single-time-point setting.

- The data are generated by this structural equation model:

$$C_i = f_C\left[\varepsilon_{C_i}\right] \qquad i = 1, \ldots, n$$

$$X_i = f_X\left[s_X\big(\underbrace{\{C_j : T_{ij} = 1\}}_{\equiv W_i}\big), C_i, \varepsilon_{X_i}\right] \qquad i = 1, \ldots, n$$

$$Y_i = f_Y\left[s_Y\big(\underbrace{\{X_j : T_{ij} = 1\}}_{\equiv V_i}, \{C_j : T_{ij} = 1\}\big), X_i, C_i, \varepsilon_{Y_i}\right] \qquad i = 1, \ldots, n,$$

  - $T_{ij} = I\{\text{subject } i \text{ and subject } j \text{ share a tie}\}$
  - f_C, f_X, and f_Y are unknown and unspecified functions
  - $\varepsilon_i = (\varepsilon_{C_i}, \varepsilon_{X_i}, \varepsilon_{Y_i})$ is a vector of exogenous, unobserved errors

- We make independence assumptions for the errors so that

  - there is no unmeasured confounding.
  - $C_i \perp C_j$ if i and j have no friends in common.
  - $Y_i \perp Y_j | parents$ and $X_i \perp X_j | parents$.

- A hypothetical intervention that deterministically sets $\mathbf{X}$ to a user-specified value $\mathbf{x}^*$ is given by

$$
\begin{aligned}
C_i &= f_C\left[\varepsilon_{C_i}\right] && i = 1, \ldots, n \\
X_i &= x_i^* && i = 1, \ldots, n \\
Y_i^* &= f_Y\left[V_i^*, \varepsilon_{Y_i}\right] && i = 1, \ldots, n
\end{aligned}
$$

  - V_i^* is a counterfactual random variable, but its value is determined by the observed realization of $\mathbf{C}$ and by the user-specified value $\mathbf{x}^*$, and it is therefore known.

- Y_i^* is the potential outcome of individual i in a hypothetical world in which $P(\mathbf{X} = \mathbf{x}^*) = 1$.
  - Peer effects: X_i could be a function of alters' outcomes at a previous time point.

- We are interested in $E\left[\bar{Y}^*\right]$, where $\bar{Y}^* = \frac{1}{n}\sum_{i=1}^n Y_i^*$.

- Under the usual assumptions, $E\left[\bar{Y}^*\right]$ is identified by the parameter

$$\psi = \frac{1}{n}\sum_{i=1}^{n} E\left[\sum_y y\, p_Y(y|V_i^*)\right] = \frac{1}{n}\sum_{i=1}^{n}\sum_v \left[\sum_y y\, p_Y(y|v)\right] P\left[V_i^* = v\right].$$

- The efficient influence function for ψ (in a particular semiparametric model) is

$$\begin{aligned}\varphi(\mathrm{O}) = &\sum_{j=1}^{n}\frac{1}{n}\sum_{i=1}^{n} E\left[\sum_y y\, p_Y(y|V_i^*) \mid C_j = c_j\right] - \psi \\ &+ \frac{1}{n}\sum_{i=1}^{n}\frac{\frac{1}{n}\sum_{j=1}^{n} P(V_j^* = v_i)}{\frac{1}{n}\sum_{j=1}^{n} P(V_j = v_i)}\left\{y_i - \sum_y y\, p_Y(y|v_i)\right\}\end{aligned}$$

- Turning the efficient IF into an estimating equation and solving it gives us an estimate $\tilde{\psi}$ of ψ.
- $\tilde{\psi}$ is asymptotically efficient and doubly robust.
- If each subject interferes with $\leq K$ other subjects, as in van der Laan (2014), then

$$\sqrt{n}(\tilde{\psi} - \psi) \to N(0, var(IF))$$

- Instead, we let $K \to \infty$ as $n \to \infty$ s.t. $\frac{K}{n} \to 0$. Then

$$\sqrt{\frac{n}{K}}(\tilde{\psi} - \psi) \to N(0, var(IF))$$

stochastic network interventions

- We can also identify the effects of interventions that replace f_X with a new, user-specified function:

$$\begin{aligned} C_i &= f_C\left[\varepsilon_{C_i}\right] && i = 1, \ldots, n \\ X_i^* &= r_X\left[W_i^*, \varepsilon_{X_i}\right] && i = 1, \ldots, n \\ Y_i^* &= f_Y\left[V_i^*, \varepsilon_{Y_i}\right] && i = 1, \ldots, n, \end{aligned}$$

- This is an example of a stochastic intervention: the intervention changes the distribution of X but does not eliminate the stochasticity introduced by ε_X.

stochastic network interventions

- Examples include
  - interventions on the network, i.e. an intervention that adds, removes, or relocates ties in the network.
  - interventions that change the dependence of a subject's treatment on other subjects' covariates, or of a subject's outcome on other subjects' covariates and treatments.
  - Interventions on summary features of the adjacency matrix
    - An intervention on features of the network topology replaces $\mathbf{T}$ with the members of a class $\mathscr{T}^*$ of $n \times n$ adjacency matrices that share the intervention features, stochastically according to some probability distribution $g_{\mathbf{T}^*}$ over $\mathscr{T}^*$.
    - Whether or not we can define, identify, and estimate interventions involving these features of network topology hinges crucially on the positivity assumption.
    - e.g. degree / centrality

principled approach

- Pros
  - uses all of the available data
  - estimands are unconditional
  - efficient and doubly robust estimation
- Cons
  - hard(er) to understand, hard to implement
  - may not be clear in finite samples what to do with K and with hubs