TODO

STAT 548 Qualifying Paper

Kenny Chiu

December 20, 2021

Abstract. TODO

1 Introduction

2 Place in the literature

The work by Forastiere et al. (2021) fits into the niche literature that looks at performing causal inference on observational network data in the presence of interference. Forastiere et al. (2021) formulate the network interference problem under the potential outcome framework (TODOcite), and propose a procedure to estimate causal treatment and spillover effects (as defined under their formulation) based on a joint propensity score that they define for network data. Their proposed propensity-adjusted estimators are unbiased under three assumptions, two of which form the Stable Unit Treatment on Neighbourhood Value Assumption (SUTNVA, a generalization of SUTVA that relaxes the no interference assumption to allow interference from directly connected nodes) and the third being an unconfoundedness assumption that says the treatment assignment mechanism is conditionally independent of the outcomes for some set of covariates. In addition, Forastiere et al. (2021) also derive the bias of SUTVA-assuming estimators when SUTVA does not hold or when the unconfoundedness assumption does not hold for the given set of covariates. The problem formulation, the proposed estimation procedure (based on a defined joint propensity score), and the derived bias of the naive estimator are the main contributions of Forastiere et al.

TODO: limitations here or conclusion?

There appears to be a limited number of works in the literature that examine the similar problem of causal inference in observational data with general forms of interference. As noted by Forastiere et al., Liu et al. (2016) proposed inverse probability-weighted (IPW) and Hájek-type estimators for the causal treatment effect, and van der Laan (2014) proposed a novel targeted maximum likelihood estimator for the effect that was later further developed by Sofrygin and van der Laan (2017) and Ogburn et al. (2017). Jackson et al. (2020) use propensity-based estimators similar to Forastiere et al. (2021) but take homophily into account by modeling correlated treatment assignments as an incomplete information game. The recent work by Sánchez-Becerra (2021) questioned the justification of the unconfoundedness assumption with respect to a constructed statistic (e.g., the joint propensity score) and proposed a model-based estimator that is obtained by optimizing a loss function. TODO

More commonly, related works in the literature examine the inference problem under the assumption of partial interference. Under this assumption, individuals can be partitioned into groups where it is assumed that there are no spillover effects between groups. The focus on partial interference settings seems to be primarily due to momentum of earlier works (e.g., Sobel, 2006; Hudgens & Halloran, 2008) that looked at causal inference in randomized studies with interference, in which group-randomization tends to be more practical. Examples of recent work that assume partial interference include the work by Liu et al. (2019), Barkley et al. (2020), and Qu et al. (2021). It is notable that these works all propose IPW estimators despite exploring slightly difference contexts of the interference problem. Estimators based on IPW appear to be another idea in the literature that persisted since its original introduction by Tchetgen and VanderWeele (2012) for grouped observational data.

Several other works in the literature consider inference in observational studies with interference under specific contexts. For example, Toulis et al. (2018) explore the problem of treatment entanglement where treatment assignments are assumed to satisfy certain restrictions. Zigler and Papadogeorgou (2021) focus on the problem of bipartite causal inference with interference where treatments are applied to one unit and the outcome is measured on another. Outside of the observational setting, there are many works that study causal inference in randomized studies with potential interference. These works generally examine study designs that allow for treatment effect estimation in the presence of interference under varying contexts (e.g., Saveski et al., 2017; Doudchenko et al., 2020; Jagadeesan et al., 2020; Imai et al., 2021)

3 Questions about the theory

4 Empirical study

5 Concluding remarks

References

Barkley, B. G., Hudgens, M. G., Clemens, J. D., Ali, M., & Emch, M. E. (2020). Causal inference from observational studies with clustered interference, with application to a cholera vaccine study. *The Annals of Applied Statistics*, *14*(3), 1432–1448.

Doudchenko, N., Zhang, M., Drynkin, E., Airoldi, E. M., Mirrokni, V., & Pouget-Abadie, J. (2020). Causal inference with bipartite designs. *Available at SSRN 3757188*.

Forastiere, L., Airoldi, E. M., & Mealli, F. (2021). Identification and estimation of treatment and interference effects in observational studies on networks. *Journal of the American Statistical Association*, *116*(534), 901–918. https://doi.org/10.1080/01621459.2020.1768100

Hudgens, M. G., & Halloran, M. E. (2008). Toward causal inference with interference. *Journal of the American Statistical Association*, *103*(482), 832–842.

Imai, K., Jiang, Z., & Malani, A. (2021). Causal inference with interference and noncompliance in two-stage randomized experiments. *Journal of the American Statistical Association*, *116*(534), 632–644. https://doi.org/10.1080/01621459.2020.1775612

Jackson, M. O., Lin, Z., & Yu, N. N. (2020). Adjusting for peer-influence in propensity scoring when estimating treatment effects. *Available at SSRN 3522256*.

Jagadeesan, R., Pillai, N. S., & Volfovsky, A. (2020). Designs for estimating the treatment effect in networks with interference. *The Annals of Statistics*, *48*(2), 679–712.

Liu, L., Hudgens, M. G., & Becker-Dreps, S. (2016). On inverse probability-weighted estimators in the presence of interference. *Biometrika*, *103*(4), 829–842. https://doi.org/10.1093/biomet/asw047

Liu, L., Hudgens, M. G., Saul, B., Clemens, J. D., Ali, M., & Emch, M. E. (2019). Doubly robust estimation in observational studies with partial interference. *Stat*, *8*(1), e214.

Ogburn, E. L., Sofrygin, O., Diaz, I., & Van der Laan, M. J. (2017). Causal inference for social network data. *arXiv preprint arXiv:1705.08527*.

Qu, Z., Xiong, R., Liu, J., & Imbens, G. (2021). Efficient treatment effect estimation in observational studies under heterogeneous partial interference. *arXiv preprint arXiv:2107.12420*.

Sánchez-Becerra, A. (2021). Spillovers, homophily, and selection into treatment: The network propensity score. https://economics.sas.upenn.edu/system/files/2021-03/AlejandroSanchez_JMP_March2021_0.pdf

Saveski, M., Pouget-Abadie, J., Saint-Jacques, G., Duan, W., Ghosh, S., Xu, Y., & Airoldi, E. M. (2017). Detecting network effects: Randomizing over randomized experiments. In *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*.

Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate? causal inference in the face of interference. *Journal of the American Statistical Association*, *101*(476), 1398–1407.

Sofrygin, O., & van der Laan, M. J. (2017). Semi-parametric estimation and inference for the mean outcome of the single time-point intervention in a causally connected population. *Journal of Causal Inference*, *5*(1).

Tchetgen, E. J. T., & VanderWeele, T. J. (2012). On causal inference in the presence of interference. *Statistical Methods in Medical Research*, *21*(1), 55–75.

Toulis, P., Volfovsky, A., & Airoldi, E. M. (2018). Propensity score methodology in the presence of network entanglement between treatments. *arXiv preprint arXiv:1801.07310*.

van der Laan, M. J. (2014). Causal inference for a population of causally connected units. *Journal of Causal Inference*, *2*(1), 13–74.

Zigler, C. M., & Papadogeorgou, G. (2021). Bipartite causal inference with interference. *Statistical science: a review journal of the Institute of Mathematical Statistics*, *36*(1), 109.

A