Horvitz-Thompson Estimators of Spillover Effects under Hypothetical Bernoulli Treatment Allocations in Two-Stage Randomized Experiments

Colleen Chan[1], Shinpei Nakamura-Sakai[1], Laura Forastiere[2]

[1]Department of Statistics and Data Science [2]Department of Biostatistics

Yale

Introduction

A fundamental assumption often made in causal inference is the stable unit treatment value assumption (SUTVA), which implies that potential outcomes for each unit are unrelated to the treatment status of other units. This assumption effectively rules out spillover effects, which are often present and of direct interest in many fields where individuals interact, such as epidemiology, education, and labor markets.

Two-stage randomized experiments are incredibly useful designs for estimating causal effects of a given treatment in the presence of interference. Previous two-stage designs have been proposed under complete randomization (CRE) in *both* stages, and simple sample average estimators have been developed under the partial interference assumption. Under CRE in both stages, these estimators are unbiased for average potential outcomes where cluster neighbors' treatments are assigned under a 2nd stage CRE. We propose average potential outcomes, and corresponding direct and spillover effects, where cluster neighbors' treatments are drawn under a hypothetical *Bernoulli assignment*. This allows researchers to evaluate causal effects under hypothetical treatment saturation levels. We develop Horvitz-Thompson estimators for 2nd stage saturations and hypothetical ones under a 2nd stage CRE or Bernoulli.

Two-Stage Randomized Experiments

Closely following Hudgens and Halloran [2], we consider $J > 1$ groups of individuals where for $j = 1, \ldots, J$, n_j denotes the number of individuals in group j and $N = \sum_{j=1}^{J} n_j$ be the total number of individuals. We denote:

- $\mathbf{Z}_j = (Z_{1j}, \ldots, Z_{n_j,j})$ as the treatments that group j receives and $\mathbf{Z}_{-i,j}$ denotes the $n_j - 1$ subvector of $\mathbf{Z}_j$ with the ith entry deleted
- $\mathbf{z}_j$ and z_{ij}as the possible values of $\mathbf{Z}_j$ and Z_{ij}, respectively
- $\Omega(n)$ as the set of vectors of all possible exposure allocations of length n
- $Y_{ij}(\mathbf{z})$ as the potential outcome of individual i in group j under treatment $\mathbf{z}$
- α as the proportion of units assigned to treatment in a group, where we consider M saturations α_m, with $m = 1, \ldots, M$
- S_j as a categorical cluster assignment indicator: $S_j = m$ if cluster j is assigned to α_m, $\mathbf{S} = (S_1, \ldots, S_J)$, where $J_m = \sum_{j=1}^{J} \mathbb{1}(S_j = m)$

Throughout, we assume *partial interference*, i.e., there is no between-group interference such that $\mathbf{Y}_j(\mathbf{Z}) = \mathbf{Y}_j(\mathbf{Z}_j)$. Finally, we let $\mathbb{P}_{\text{2ndStage}(\alpha_m)}(\mathbf{Z}_j = \mathbf{z}_j)$ be the probability of observing a cluster treatment vector equal to $\mathbf{z}_j$ under the second stage assignment with probability/proportion α_m.

Stage 1: **Completely randomized assignment of clusters** to treatment saturations. J_m clusters are randomly assigned to each saturation α_m.

Stage 2: **Randomized assignment of individuals** to treatment and control. For each cluster j assigned to α_m, treatment can be assigned according to either:

- **Bernoulli**: Each unit is assigned independently to treatment with probability α_m
- **CRE**: Exactly $n_j\alpha_m$ units are assigned to treatment

Average Potential Outcomes under Bernoulli Assignment

We define the average individual potential outcome under individual treatment $Z_{ij} = z$ and cluster neighbors' treatments assigned under a Bernoulli assignment with probability α as

$$\overline{Y}_{ij}^{\text{Bern}}(z, \alpha) = \sum_{\mathbf{z}_{j,-i} \in \Omega(n_j - 1)} \left(Y_{ij}(Z_{ij} = z, \mathbf{Z}_{j,-i} = \mathbf{z}_{j,-i}) \prod_{k \neq i} \alpha^{w_{kj}} (1 - \alpha)^{1 - w_{kj}} \right).$$

The cluster or individual-weighted potential outcome under a second-stage Bernoulli assignment is then

$$\overline{Y}^{\text{Bern}}(z, \alpha) = \sum_{j=1}^{J} w_i^* \sum_{i=1}^{n_j} \overline{Y}_{ij}^{\text{Bern}}(z, \alpha_m) \tag{1}$$

where $w_j^* = \frac{1}{J \cdot n_j}$ corresponds to the cluster-weighted estimand and $w_j^* = \frac{1}{N}$ corresponds to the individual-weighted estimand.

Direct and Indirect Effects

The individual direct effect is defined as

$$\text{DE}_{ij}(\alpha) = Y_{ij}(1, \alpha) - Y_{ij}(0, \alpha)$$

and the individual indirect (spillover) effect is defined as

$$\text{IE}_{ij}(\alpha, \alpha'; 0) = Y_{ij}(0, \alpha) - Y_{ij}(0, \alpha').$$

The cluster and sample-weighted direct effect would then be defined as

$$\overline{\text{DE}}(\alpha) = \sum_{j=1}^{J} w_i^* \sum_{i=1}^{n_j} \text{DE}_{ij}(\alpha)$$

and the cluster and sample-weighted indirect effect would then be defined as

$$\overline{\text{IE}}(\alpha, \alpha'; 0) = \sum_{j=1}^{J} w_i^* \sum_{i=1}^{n_j} \text{IE}_{ij}(\alpha, \alpha'; 0).$$

Horvitz-Thompson Estimators

To estimate Eq. 1 under a saturation $\alpha = \alpha_m$, with $m = 1, \ldots, M$, assigned in the first stage, we can use what we call the "**conditional estimator**", similar to the one proposed by Basse and Feller [1]

$$\widehat{\overline{Y}}^{\text{Cond}}(z, \alpha) = \sum_{j=1}^{J} w_j^* \sum_{i=1}^{n_j} Y_{ij}^{\text{obs}} \frac{\mathbb{1}(Z_{ij} = z, S_j = m) \mathbb{P}_{\text{Bern}(\alpha)}(\mathbf{Z}_{j,-i} = \mathbf{z}_{j,-i})}{\mathbb{P}(S_j = m) \mathbb{P}_{\text{2ndStage}(\alpha_m)}(\mathbf{Z}_j = \mathbf{z}_j)}. \tag{2}$$

This estimator is conditional because it only uses data from the clusters assigned to $\alpha = \alpha_m$.

The following estimator is what we call the "**unconditional estimator**" because it extrapolates information from all the clusters:

$$\widehat{\overline{Y}}^{\text{Uncond}}(z, \alpha) = \sum_{j=1}^{J} w_j^* \sum_{i=1}^{n_j} Y_{ij}^{\text{obs}} \frac{\mathbb{1}(Z_{ij} = z) \mathbb{P}_{\text{Bern}(\alpha)}(\mathbf{Z}_{j,-i} = \mathbf{z}_{j,-i})}{\sum_{m=1}^{M} \mathbb{P}(S_j = m) \mathbb{P}_{\text{2ndStage}(\alpha_m)}(\mathbf{Z}_j = \mathbf{z}_j)}. \tag{3}$$

Note that this estimator uses information across *all* clusters, which may make it more efficient than the conditional estimator, depending on the two-stage design.

The unconditional estimator also allows estimation of average potential outcomes under **hypothetical treatment probabilities**, i.e., α that is not equal to any α_m assigned in the first stage.

Advantages

Definition 1 of the average potential outcomes under a Bernoulli assignment has the following advantages:

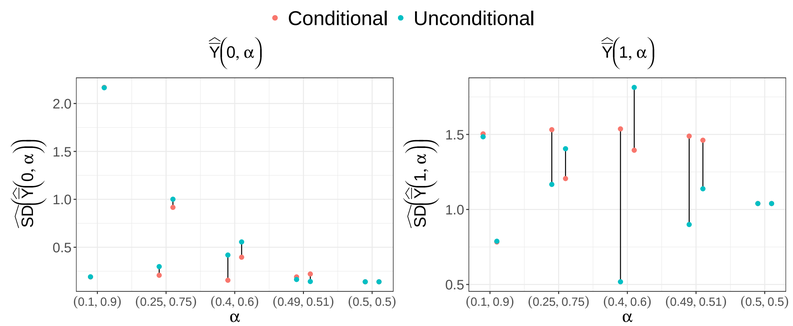
1. Allows researchers to estimate interference effects under **hypothetical treatment probabilities**
2. For treatment saturation levels assigned in the first stage, we can estimate the potential outcomes (and the resulting interference effects) using Eq. 2 or Eq. 3. We compare the finite sample performance of these estimators in the next section.

Simulation Studies

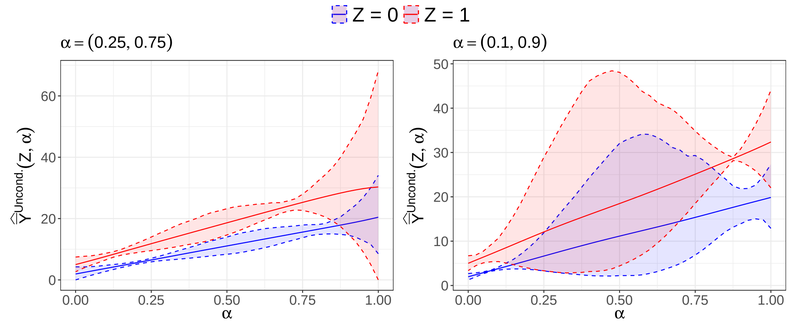
We compare the finite sample performance of the conditional estimator (Eq. 2) and unconditional estimator (Eq. 3) for various treatment saturation levels assigned in the first stage.

Consider $K = 100$ clusters with 10 units in each. We have two treatment saturation levels $\alpha = (\alpha_1, \alpha_2)$, where $\mathbb{P}(\alpha_1) = \mathbb{P}(\alpha_2) = 0.5$. The true potential outcome is generated $Y_i = 2 + 3z_i + 2g_i + z_i g_i$, where g_i is the number of treated cluster neighbors in unit i's cluster (stratified interference). We use first-stage complete randomization of clusters to treatment saturations α and a second-stage Bernoulli assignment. We plot the standard deviation of the individual average potential outcome estimate of each estimator over $B = 500$ simulations for the following treatment saturation pairs: $\alpha = \{(0.1, 0.9), (0.25, 0.75), (0.49, 0.51), (0.5, 0.5)\}$.

For treatment saturation levels that are closer to one another, e.g., $\alpha = (0.49, 0.51)$, we can see that the unconditional estimator is more efficient than the conditional estimator as it is



applying higher weights from the clusters assigned to a different saturation level. In contrast, the conditional estimator only uses information from clusters assigned to α_m. At $\alpha = (0.5, 0.5)$, the standard deviation of the estimators are the same since the estimates are the same. At treatment probabilities that are far apart, e.g., α = (0.1, 0.9), the unconditional estimator borrows almost no information from clusters assigned a different α_m, which explains their similar variances.



Using the same experiment data generating process as before, we allocate clusters with treatment saturations $\alpha = (0.4, 0.6)$ and $(0.25, 0.75)$. For each hypothetical α, we run $B = 300$ simulations and plot the mean estimate and 95% confidence interval of the individual potential outcomes. As expected, at the hypothetical treatment probabilities assigned to α_m, we observe smaller variances.

Discussion and Future Work

We define potential outcomes where the individual and cluster neighbors' treatments are assigned under a Bernoulli and propose an estimator, which we call the "unconditional" estimator, for two-stage randomized experiments where the second stage can be either Bernoulli or CRE. This estimator allows researchers to estimate interference effects of hypothetical treatment probabilities not seen in the actual experiment.

For future work, we aim to develop a closed form expression or bounds for the variance of the conditional estimator. We will also compare our estimator from a 2nd stage Bernoulli or CRE design.

Acknowledgements

We thank Fredrik Sävje for helpful comments and discussion in this work. CC thanks the Adobe Data Science Research Award for funding.

Contact information

colleen.chan@yale.edu s.nakamura.sakai@yale.edu
laura.forastiere@yale.edu

References

[1] Basse, G. and Feller, A. (2018). Analyzing Two-Stage Experiments in the Presence of Interference. *Journal of the American Statistical Association*, 113(521):41–55. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01621459.2017.1323641.

[2] Hudgens, M. G. and Halloran, M. E. (2008). Toward Causal Inference With Interference. *Journal of the American Statistical Association*, 103(482):832–842. Publisher: Taylor & Francis _eprint: https://doi.org/10.1198/016214508000000292.