CI = TREATMENT EFFECTS UNDER INTERFERENCE

Applications in — e-commerce (ad effectiveness)
— vaccine trials
— marketplace subsidies.

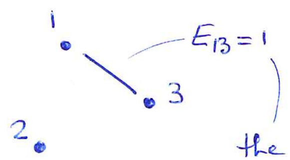
Under no spillover, use the $\{Y_i(0), Y_i(1)\}$ notation
$Y_i = Y_i(W_i)$

Under spillover, we need more potential outcomes:
define $\{Y_i(\underline{w})\}_{i=1}^n$, $\underline{w} \in \{0,1\}^n$
there are 2^n of them.

Q: What structure is there?
Q: What's the causal question here?
⇒ Very rich area & many research papers addressing this topic from ≠ angles.

I_ NETWORK INTERFERENCE

Graph representing interaction between units.



Edge $E_{ij} \in \{0,1\}$
(the graph may or may not be directed)

the P.O. of units 1 and 3 depend on their own treatment and each other's treatment allocation. Not on unit's 2.

$$\Rightarrow Y_i(W) = Y_i(W') \text{ if } W_i = W_i' \;\&\; W_j = W_j' \quad \forall j: E_{ij} = 1$$

[usually a good model in e-commerce applications; less so for vaccine trials who require to incorporate time dynamics]

→ Questions to ask / What are we interested in?

- $\bar{V}(\underline{0}) = \frac{1}{n}\sum_{i=1}^n Y_i(\underline{0})$ all units are in control
  V for Value
- $\bar{V}(\underline{1}) = \frac{1}{n}\sum_{i=1}^n Y_i(\underline{1})$ all units are in treatment

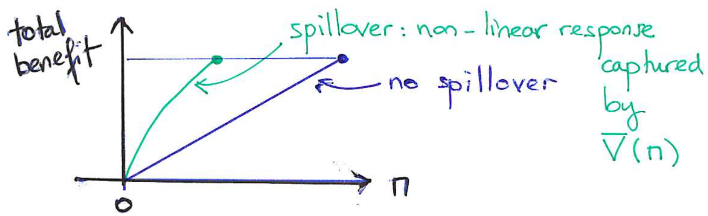
$$\boxed{\bar{\tau}_{TOT} = \bar{V}(\underline{0}) - \bar{V}(\underline{1})}$$

x Remark: we start with a finite population setting, where P.O. are considered fixed or conditioned on.

- $\bar{V}(\pi) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}_{W \sim Bern(\pi)}\left(Y_i \mid \{Y(w)\}\right)$

Under no spillover, this quantity is not interesting since

$$\begin{aligned}
\bar{V}(\pi) &= \frac{1}{n}\sum_{i=1}^n \mathbb{E}_W\left(W_i Y_i(1) + (1-W_i) Y_i(0)\right) \\
&= \frac{1}{n}\sum_{i=1}^n \left(\pi Y_i(1) + (1-\pi) Y_i(0)\right) \\
&= \frac{1}{n}\sum Y_i(0) + \pi \frac{1}{n}\sum Y_i(1) - Y_i(0) \\
&= \bar{V}(\underline{0}) + \pi \bar{\tau} \\
&= \text{linear growth with } \pi
\end{aligned}$$



- $\bar{V}(1;\underline{0}) = \frac{1}{n}\sum_{i=1}^{n} Y_i(W_i=1\ ;\ W_{-i}=0)$

You are treated; but no-one else is

$$\boxed{\tau_{DE-0} = \bar{V}(1;0) - \bar{V}(\underline{0})}$$

Direct Effect

- $\bar{V}(0;\underline{e}_1) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_{W\sim B(\pi)}\Big(Y_i \mid W_i=0,\ \sum_{\substack{j\neq i\\ E_{ij}=1}} W_j = 1,\ P.O.\Big)$

You are not treated, but only one of your neighbour is.

$$\boxed{\tau_{IE-0} = \bar{V}(0;\underline{e}_1) - \bar{V}(\underline{0})}$$

Indirect Effect

Summary:

Can be estimated by changing your policy π (e.g. $\bar{V}(0)$ = treat no one)

More cautious here

	policy rewards	non-policy rewards
mech indpt	$\bar{V}(0)$, $\bar{V}(1)$	$\bar{V}(1;0)$
mech dpt	$\bar{V}(\pi)$	$\bar{V}(0;\underline{e}_1)$

treatment assignment mechanism (i.e. dependence on the probabilistic assignment)

You can easily do $W_i=1$ and $W_i=0$ for a specific unit i; but how to compute the averaging?

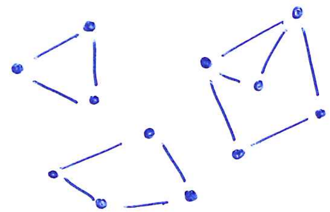
× Example 1: Cluster Interference

Observe (Y_i, W_i, C_i) $\quad C_i \in \{1,\dots,K\}$

K = number of clusters

"cluster interference" means $Y_i = Y_i(\underline{W})$ with

$E_{ij}=1 \Leftrightarrow C_i = C_j$



"cluster experiment" $\begin{cases} Z_k \sim \mathrm{Bern}(\pi) & k=1,\dots,K \\ W_i = Z_{C_i} \end{cases}$

- What do we observe?

$$Y_i = Y_i(\underline{W}) = Y_i(W_i, W[j: E_{ij}=1])$$
$$= \begin{cases} Y_i(\underline{0}) & \text{if control cluster } (Z_{C_i}=0) \\ Y_i(\underline{1}) & \text{if treatment cluster } (Z_{C_i}=1) \end{cases}$$

⇒ Only two types of exposure exist. If our estimand depends on these quantities, then we can estimate (e.g. $\bar{V}(\underline{0})$, $\bar{V}(\underline{1})$) ⇒ $\bar{V}(1;\underline{0})$ and $\bar{V}(0;\underline{e}_1)$ are not identifiable.

- How to estimate τ_{TOT}?

↳ Idea #1 $\quad \hat{\tau}_{DM} = \frac{1}{n_1}\sum_{W_i=1} Y_i - \frac{1}{n_0}\sum_{W_i=0} Y_i$

Fact: Not unbiased for finite n. Under no spillover, the key step to derive unbiasedness of $\hat{\tau}_{DM}$ was the use of $\mathbb{P}(W_i=1 \mid n_0>0, n_1>0) = n_1/n$; which does not hold here

see page 2

↘ Idea #2 $\hat{\tau}_{IPW} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{W_i}{\pi} - \frac{1-W_i}{1-\pi}\right) Y_i$

Fact: Unbiased for τ_{TOT} since

$$\mathbb{E}\left[\left(\frac{W_i}{\pi} - \frac{1-W_i}{1-\pi}\right)Y_i\right] = \mathbb{E}\left(\frac{Z_{c_i} Y_i}{\pi} - \frac{(1-Z_{c_i})Y_i}{1-\pi}\right)$$

all $\mathbb{E}(\dots)$ are conditional on the P.O.

$$= \mathbb{E}\left(\frac{Z_{c_i} Y_i(1)}{\pi} - \frac{(1-Z_{c_i})Y_i(0)}{1-\pi}\right)$$

trt assignmt

$$= Y_i(1) - Y_i(0)$$

& block bootstrap for confidence bounds.

× Example 2 = Interference in Bernoulli experiments.

Observe (Y_i, W_i) $\quad W_i \mid$ P.O. $\sim \text{Bern}(\pi)$

$Y_i = Y_i(w)$ with exposure E_{ij}

Here IPW is very good for estimating τ_{TOT}

let $$\boxed{\hat{V}_{IPW}(\underline{0}) = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{1}(W_i=0, W_j=0 \ \forall E_{ij}=1)}{\mathbb{P}_{\pi}(W_i=0, W_j=0 \ \forall E_{ij}=1)} Y_i}$$

Note that we are throwing away a lot of the data here.

Notation: $N_i = |\{j \neq i \mid E_{ij}=1\}|$

$M_i = |\{j \neq i \mid E_{ij}=1, W_j=1\}|$

Then $$\hat{V}_{IPW}(\underline{0}) = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{1}(W_i=0, M_i=0)}{(1-\pi)^{1+N_i}} Y_i$$

Fact: $\hat{V}_{IPW}(\underline{0})$ is unbiased for $Y_i(\underline{0})$. Indeed,

$$\mathbb{E}(\hat{V}_{IPW}(\underline{0}) \mid \text{P.O.})$$

$$= \mathbb{E}\left(\frac{\mathbb{1}(W_i=0, M_i=0)}{\mathbb{P}_{\pi}(W_i=0, M_i=0)} Y_i \mid \text{P.O.}\right)$$

$$= \mathbb{E}\left(\text{— '' —} \; Y_i(\underline{0}) \mid \text{P.O.}\right)$$

$$= Y_i(\underline{0})$$

likewise, define $$\boxed{\hat{V}_{IPW}(\underline{1}) = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{1}(W_i=1, M_i=N_i)}{\pi^{1+N_i}} Y_i}$$

and $$\boxed{\hat{\tau}_{TOT,IPW} = \hat{V}_{IPW}(\underline{1}) - \hat{V}_{IPW}(\underline{0})}$$

⇒ We may still recover total causal quantities, even in a bernoulli experiment.

[To get confidence bounds, block bootstrap observations. Since we are throwing away a lot of the data, expect wider confidence bounds than with $\hat{\tau}_{DM}$ in a cluster-experiment]

Remarks = We may also consider

- $\hat{V}_{IPW}(1; \underline{0}) = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{1}(W_i=1, M_i=0)}{\pi(1-\pi)^{N_i}} Y_i$
- $\hat{V}_{IPW}(\pi') = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{P}_{\pi'}(W_i, W_j \ \forall j: E_{ij}=1)}{\mathbb{P}_{\pi}(W_i, W_j \ \forall j: E_{ij}=1)} Y_i$

$$= \frac{1}{n} \sum_{i=1}^{n} \frac{\pi'^{\,W_i+M_i} (1-\pi')^{1+N_i-(W_i+M_i)}}{\pi^{W_i+M_i} (1-\pi)^{1+N_i-(W_i+M_i)}} Y_i$$

Fact: $E_\pi(\hat{V}_{IPW}(\pi') \mid P.O.) = \bar{V}(\pi')$

$$E_\pi \left(\frac{P_{\pi'}(\omega)}{P_\pi(\omega)} Y_i \mid P.O. \right)$$ abusing notation here

$$\sum_w \left(\frac{\mathbb{P}_{\pi'}(w)}{\mathbb{P}_{\cancel{\pi}}(w)} Y_i(w) \right) \mathbb{P}_{\pi}(\cancel{W = w}) = \mathbb{E}_{\pi'}(Y_i(\underline{w}))$$

II_ EXPOSURE MAPPINGS

Defined as $A_i : \{0,1\} \longrightarrow \mathcal{A}$

$Y_i(\underline{\omega}) = Y_i(\underline{\omega}')$ if $A_i(\underline{\omega}) = A_i(\underline{\omega}')$

or equivalently $Y_i = Y_i(A_i(\underline{\omega}))$
$(= Y_i(\underline{\omega}))$

Examples =

→ H_0 : No effect exposure $\quad A_i(\underline{\omega}) = \emptyset \quad Y_i(\underline{\omega}) = Y_i(\underline{0})$

→ H_1 : No interference $\quad A_i(\underline{\omega}) = \omega_i \quad Y_i(\underline{\omega}) = Y_i(\omega_i)$

→ H_2 : "At least one" $\quad A_i(\underline{\omega}) = (\omega_i, z_i)$,
where $z_i = \mathbb{1}(\exists j \neq i \text{ s.t. } E_{ij} = 1, \omega_j = 1)$

→ H_3 : General $\quad A_i(\underline{\omega}) = (\omega_i, \omega_j \ \forall j \neq i, E_{ij} = 1)$

Exposure mappings were introduced by Aronow & Samii (2017)
We discuss next an application for testing the presence of interference.

× Permutation test for H_0 [sharp null]

General Procedure :

(i) Pick a test statistic $T(\underline{Y}, \underline{\omega}) = \left| \frac{1}{n_1} \sum_{W_i=1} Y_i - \frac{1}{n_0} \sum_{W_i=0} Y_i \right|$

(ii) Scramble treatment : for $b = 1, \dots, B$, let $\underline{\omega}^b$ be a random permutation of $\underline{\omega}$ and compute $T^b = T(\underline{Y}, \underline{\omega}^b)$ $\quad \underline{Y} = \underline{Y}(\underline{\omega})$

(iii) Compute a "p-value"

$$P_{val} = \frac{1 + \sum_{b=1}^{B} \mathbb{1}(T \leq T^b)}{1 + B}$$

We show below that this is a valid p-value under H_0 ; ie that it satisfies $P_{H_0}(P_{val} \leq u) \leq u \quad \forall u \in [0,1]$.

→ Our experiment is characterized by the P.O $\{Y_i(\omega)\}$ & a randomization distribution $\underline{W} \sim \mathbb{R}$ (considering here a completely randomized experiment)

In the permutation test set-up, all $(\omega, \omega', \dots, \omega^B)$ are iid draws from $\mathbb{R}$ [in particular, it keep the total number of treated customers constant].

Consequence : $(T, T_1^*, \dots, T_B^*)$ are iid conditionally on the P.O, where $T_b^* = T(\mathbf{Y}(\underline{\omega}^b), \underline{\omega}^b)$

Key Step : Under H_0 : $T_b = T_b^*$

$\Rightarrow (T, T_1, \dots, T_B)$ are iid draws cond. on the P.O. under H_0.

$\Rightarrow$ p-value is uniform on $\left\{ \frac{1}{1+B}, \frac{2}{1+B}, \dots, 1 \right\}$. (9)

Remark: Under $R = \mathrm{Bern}(n)$, the $(w, w^1, \dots, w^B)$ are not iid (but cond. iid on the number of treated).

× Permutation test for H_1

Testing for the presence of interference

- Idea #1: permute similarly, generating the w^b
  Issue: can't impute $Y_i(w^b)$ in general from observations (problem if $w_i^b \neq w_i$)

- Idea #2: consider focal units
  Example: $\mathcal{J} \subset \{1, \dots, n\}$ s.t. $W_i = 0 \ \forall i \in \mathcal{J}$
  (select a subset of untreated units)
  & generate w_b s.t. $w_i^b = 0 \ \forall i \in \mathcal{J}$
  (units in the focal group remain untreated)

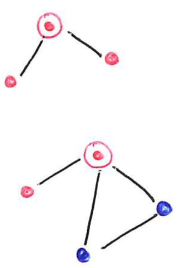
depends only on the outcome of units $i \in \mathcal{J}$

& consider $T(\underline{Y}, \underline{W}) = \dfrac{\sum_{Z_i=1, i\in\mathcal{J}} Y_i}{|\{i\in\mathcal{J} \mid Z_i = 1\}|} - \dfrac{\sum_{Z_i=0, i\in\mathcal{J}} Y_i}{|\{i\in\mathcal{J} \mid Z_i = 0\}|}$

(recall: $Z_i = 1 \Leftrightarrow$ unit i has at least one treated neighbour)

$\Rightarrow T(\underline{Y}, \underline{W})$ captures the mean difference of having one or more treated neighbours with having no one treated

$\Rightarrow$ Need to turn this to a formal test



• = untreated
• = treated
○ = focal unit $\in \mathcal{J}$

---

(10)

- We have a null hypothesis H, P.O. and a randomization scheme $W \sim R$, and observe $\underline{Y}(\underline{W}), \underline{W}$
- Given $\underline{W}$, we are going to choose:

$(*)\begin{cases} \text{– a focal set } \mathcal{J} \subset \{1, \dots, n\} \\ \text{– a test statistic } T(\underline{Y}(\underline{W}), \underline{W}) \text{ that only depends on } Y_i \text{ for } i \in \mathcal{J} \\ \text{– a permutation set } \mathcal{W} \text{ s.t. } W_i = \mathcal{W}_i^b \text{ for } i \in \mathcal{J} \ \forall W^b \in \mathcal{W} \end{cases}$

- Pick $w^1, \dots, w^B \in \mathcal{W}$
- Compute $T_b = T(\underline{Y}(\underline{W}), \underline{w}^b)$

Under H: By $(*)$, $T_b = T_b^*(\underline{Y}(\underline{w}^b), \underline{w}^b) \ \forall b$

Q: What is the right p-value here?

Put $\alpha_{\mathcal{J}}(\underline{w}) = \mathbb{P}_R(\underline{W} = \underline{w} \mid \{\mathcal{J}, T, \mathcal{W}\} \text{ selected})$

Then
$$\text{p-val} = \frac{\alpha_{\mathcal{J}}(\underline{W}) + \sum_{b=1}^{B} \alpha_{\mathcal{J}}(\underline{w}^b)\, \mathbb{1}(T \leq T^b)}{\alpha_{\mathcal{J}}(\underline{W}) + \sum_{b=1}^{B} \alpha_{\mathcal{J}}(\underline{w}^b)}$$

We need these weights to address a Monty-Hall effect here:

Ex: R = treat one person at random
$\mathcal{J}$ = pick two non-adjacent units.
$= \{\{1,3\}, \{1,4\}, \{2,4\}\}$ [one of these]

trt, ctl

	$\{1,3\}$	$\{1,4\}$	$\{2,4\}$
●—○—○—○			× ×
○—●—○—○	×	×	
○—○—●—○		×	×
○—○—○—●	× ×		

We see that conditionally on $\mathcal{F} = \{2, 4\}$, the probability that unit 1 is treated is $2/3$, and not $1/4$.

$\Rightarrow$ the p-value needs to account for this effect (i.e. get the correct conditional distribution, since Pval is defined conditionally on $\mathcal{F}$)

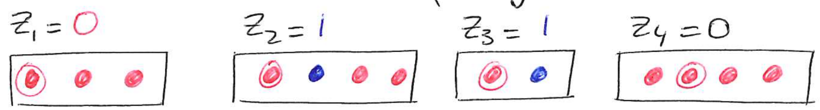
The presence of the $\alpha_{\mathcal{F}}(\underline{w})$ terms make the computation more complicated. However, these sometime cancel out:

× Example: Students in families (Basse, Feller, Toulis (2019))

$C_i = \{1, \dots, K\}$ denotes the family

Two-stage design $Z_k \in \{0, 1\}$ = exposure of the family; completely randomized with n_1 treated families

- If $Z_k = 0$, then $W_i = 0 \; \forall i$ s.t. $C_i = k$
- If $Z_k = 1$, then treat one person at random in the k-th family



• = control
• = treated
○ = focal unit

$\mathcal{F} \equiv$ random untreated person in each family

let S_k = index of the focal unit in the k-th family

$$\cdot\, T(\underline{Y}, \underline{w}) = \left| \frac{1}{n_1} \sum_{Z_k = 1} Y_{S_k} - \frac{1}{n_0} \sum_{Z_k = 0} Y_{S_k} \right|$$

# of treated families $\qquad = K - n_1$

More formally, the exposure mapping for unit i is assumed to be

$$A_i(\underline{W}) = A_i(W_1, \dots, W_n) = A_i(W_i, Z_{[i]})$$

n individuals spread among K families

The P.O. depends on your own treatment, and whether you were assigned to a treatment or control family [Implicit Assumption = it doesn't matter which one of your relatives was treated if you are in a treated family].

where $[i]$ denotes the household wherein i resides, and W_i = treatment indicator of unit i, $i = 1, \dots, n$

We write $Y_i(W_i, Z_{[i]})$ for the value of $Y_i(\underline{w})$.
It can take only three possible values:

$$Y_i(\underline{w}) \in \{\underbrace{Y_i(0,0),\; Y_i(0,1)}, \; Y_i(1,1)\}$$

control unit in a control and treated family, respectively.

$\Rightarrow$ The set $\mathcal{A}$ contains 3 values:

$$\mathcal{A} = \{a = (0,0), \; b = (0,1), \; c = (1,1)\}$$

We consider the null $H_0^S : Y_i(0,0) = Y_i(0,1) \;\; \forall i$

i.e. is there proof of interaction between members of a family?

Under H_0^s, we can swap the P.O. $Y_i(0,1)$ and $Y_i(0,0)$ of all non treated units & thefore of all focal units $\Rightarrow$ we can compute the T^b statistic by scrambling the family assignment (Z^b) at random:

$$T^b(\underline{Y}, \underline{w}^b) = \left| \frac{1}{n_1} \sum_{Z_k^b=1} Y_{S_k} - \frac{1}{K-n_1} \sum_{Z_k^b=0} Y_{S_k} \right|$$

The focal units are the same as before.

Fact: Under this scheme, all $\alpha_{\mathcal{F}}(\underline{w})$ have the same value. This follow immediatly from the two-stage design:

- → First, draw individuals △ that might potentially be treated (1 per family)
- → Second, draw a focal individual □ other than △ within each family (1 per family)
- → Third, draw the family assignment at random.

Since the binary assignment $Z_k \in \{0,1\}$ is completly independent of the selection of △ and □, we have no Monty Hall effect and $\alpha_{\mathcal{F}}(\underline{w}) = \text{constant} \ \forall \underline{w}$.

$\Rightarrow$ The terms $\alpha_{\mathcal{F}}(\underline{w})$ cancel out and

$\left(1 + \sum_{b=1}^{B} \mathbb{1}(T \leq T_b)\right) / (1+B)$ is a valid p-value under H_0^s.

III - AVG DIRECT & INDIRECT EFFECTS

Collect (Y_i, W_i) $i = 1, \dots, n$ $W_i \in \{0,1\}$

$Y_i = Y_i(\underline{W})$; $\underline{W} = (W_1, \dots, W_n)$

Network interference $E_{ij} \in \{0,1\}$ s.t. $Y_i(\underline{w}) = Y_i(\underline{w}')$ if $w_i = w_i'$ and $w_j = w_j'$ $\forall j: E_{ij} = 1$.

$N_i = \sum_{j \neq i} E_{ij}$ = # neighbours of i

$M_i = \sum_{j \neq i} E_{ij} W_j$ = # of treated neighbours of i

In the previous lecture, we had a look at quantities such as

$$V(\underline{0}) = \frac{1}{n} \sum_{i=1}^{n} Y_i(\underline{0})$$

$$V(1; \underline{0}) = \frac{1}{n} \sum_{i=1}^{n} Y_i(w_i = 1; \underline{0})$$

the treatment for each unit other than i was specified & set to a particular value.

In a Bernoulli trial,

$$\hat{V}_{IPW}(1; \underline{0}) = \frac{1}{n} \sum_{i=1}^{n} \frac{\mathbb{1}(W_i = 1;\ W_j = 0 \ \ \forall j \neq i\ E_{ij} = 1)}{\pi (1-\pi)^{N_i}} Y_i$$

is unbiased for $\bar{V}(1;0)$. However, this estimator is very noisy whenever $\pi \approx 1/2$ and N_i is large (large networks) = the denominator is close to 0, inflating the numerator by a large factor

↳ OK for vaccine trials since we can assume sparsity

↳ Not OK for marketplace where many agents interact.

- Other estimands to consider.

$$\bullet\ \tau_{ADE} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_{\mathbb{R}}\left[Y_i(W_i=1; W_{-i}) - Y_i(W_i=0; W_{-i})\right]$$

Average Direct Effects

— the role of all other units is marginalized out —

Depends on $\mathbb{R}$.
If $\mathbb{R} \sim \text{Bern}(\pi)$, the average direct effect will heavily depend on π; i.e. how many people are already affected by the treatment (e.g. vaccine trials)

$$\bullet\ \tau_{AIE} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i} \mathbb{E}_{\mathbb{R}}\left[Y_j(W_i=1; W_{-i}) - Y_j(W_i=0; W_{-i})\right]$$

Average Indirect Effects

same as before: marginalized out.

effect of the i-th unit being treated on j

total effect of treating the i-th unit on all other units

from Hu, Li & Wager (2022)

$$= \frac{1}{n}\sum_{i=1}^{n}\underbrace{\sum_{j\neq i} \mathbb{E}_{\mathbb{R}}\left[Y_i(W_j=1; W_{-j}) - Y_i(W_j=0, W_{-j})\right]}_{\text{average treatment effect of others on you}}$$

x Example: Checking that these two definitions make sense in a simple linear scenario where

$$Y_i(\underline{W}) = \alpha_i + \beta_i W_i + \gamma_i \frac{1}{N_i}\sum_{j\neq i} E_{ij} W_i + \varepsilon$$

β_i response to your own treatment

γ_i response to your neighbours treatment

Then

$$\tau_{ADE} = \frac{1}{n}\sum_{i=1}^{n}\Big[\underbrace{\cancel{\alpha_i} + \beta_i + \cancel{\gamma_i \frac{M_i}{N_i} W_{-i}}}_{= Y_i(W_i=1;\, W_{-i})} - \underbrace{\left(\cancel{\alpha_i} + \cancel{\gamma_i \frac{M_i}{N_i} W_i}\right)}_{Y_i(W_i=0;\, W_{-i})}\Big] = \frac{1}{n}\sum_{i=1}^{n}\beta_i$$

$$\&\ \tau_{AIE} = \frac{1}{n}\sum_{i=1}^{n}\sum_{\substack{j\neq i\\ E_{ij}=1}} E_{ij}\frac{\gamma_i}{N_i} = \frac{1}{n}\sum_{i=1}^{n}\gamma_i$$

Q: What about estimation?

Assume $\mathbb{R} = \text{Bern}(\pi)$

Then

since the W_i are iid $\text{Bern}(\pi)$

$$\mathbb{E}_{\pi}\left[Y_i(W_i=1; W_{-i})\right] = \mathbb{E}_{\pi}\left[Y_i(W_i=1; W_{-i}) \mid W_i=1\right]$$
$$= \mathbb{E}_{\pi}\left[Y_i(W_i; W_{-i}) \mid W_i=1\right]$$
$$= \mathbb{E}_{\pi}\left[Y_i \mid W_i=1\right]$$

only observable quantities remain

$$\&\ \tau_{ADE} = \frac{1}{n}\sum_{i=1}^{n}\left\{\mathbb{E}_{\pi}(Y_i \mid W_i=1) - \mathbb{E}_{\pi}(Y_i \mid W_i=0)\right\}$$

The notation emphasizes that this expression for τ_{ADE} is valid under a Bernoulli scheme

$\Rightarrow$ logically, consider an IPW estimator for τ_{ADE}

$$\hat{\tau}^{IPW}_{ADE} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{W_i}{\pi} - \frac{1-W_i}{1-\pi}\right)Y_i$$

Fact: $\mathbb{E}(\hat{\tau}^{IPW}_{ADE} \mid P.O.) = \bar{\tau}_{ADE}$

Proof: $\mathbb{E}\left[\left(\frac{W_i}{\pi} - \frac{1-W_i}{1-\pi}\right)Y_i\right] = \mathbb{E}\, \mathbb{E}_\pi\left(-''- \mid W_i\right)$

$$= \mathbb{E}\left[\left(\frac{W_i}{\pi} - \frac{1-W_i}{1-\pi}\right) \mathbb{E}(Y_i \mid W_i)\right]$$

$$= \mathbb{E}_{W_i}\left(\frac{W_i}{\pi}\mathbb{E}(Y_i \mid W_i)\right) - \mathbb{E}_{W_i}\left(\left(\frac{1-W_i}{1-\pi}\right)\mathbb{E}(Y_i \mid W_i)\right)$$

$$= \mathbb{E}(Y_i \mid W_i = 1) - \mathbb{E}(Y_i \mid W_i = 0) \quad \blacksquare$$

And apply across all units i

Note that similarly we can prove that $\hat{\tau}^{DM}_{ADE} = \frac{1}{n_1}\sum_{W_i=1} Y_i - \frac{1}{n_0}\sum_{W_i=0} Y_i$ is also unbiased for $\bar{\tau}_{ADE}$.

Take Away: the difference in means estimator is still estimating something meaningful: not an ATE in the presence of spillover, but an ADE

As opposed to an average total effect captured by a clustered randomized experiment

Next, estimating the AIE:

$$\hat{\tau}^{IPW}_{AIE} = \frac{1}{n}\sum_{i=1}^{n} \sum_{\substack{j \neq i \\ E_{ij}=1}} \left(\frac{W_j}{\pi} - \frac{1-W_j}{1-\pi}\right) Y_i$$

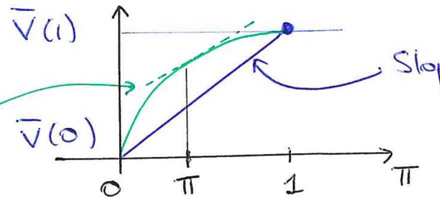
(can prove as before that it is unbiased)

$$\boxed{\hat{\tau}^{IPW}_{AIE} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{M_i}{\pi} - \frac{N_i - M_i}{1-\pi}\right) Y_i}$$

The decomposition theorem

Consider $\mathbb{R} = \text{Bern}(\pi)$.

The total effect of treating no-one ($\pi = 0$) to treating everyone in the population ($\pi = 1$) is $\bar{\tau}_{TOT} = \bar{V}(1) - \bar{V}(0)$.



In the presence of spillovers, the value response is not linear in π

Instead of considering the global slope from 0 to 1, consider a local version of the total effect: the slope at the current policy π.

infinitesimal

$$\boxed{\bar{\tau}_{INF} := \frac{d}{d\pi'}\bar{V}(\pi')\Big|_{\pi'=\pi}}$$

Theorem: $\bar{\tau}_{INF} = \bar{\tau}_{ADE} + \bar{\tau}_{AIE}$

To prove this, we will proceed as follows:

- → $\hat{\tau}^{IPW}_{ADE}$ is unbiased for $\bar{\tau}_{ADE}$
- → $\hat{\tau}^{IPW}_{AIE}$ is unbiased for $\bar{\tau}_{AIE}$
- → Construct $\hat{\tau}^{IPW}_{INF}$ which will be unbiased for $\bar{\tau}_{INF}$ & the notice that $\hat{\tau}^{IPW}_{INF} = \hat{\tau}^{IPW}_{ADE} + \hat{\tau}^{IPW}_{IDE}$
- → Conclude that $\bar{\tau}_{INF} = \bar{\tau}_{ADE} + \bar{\tau}_{IDE}$

- $$\hat{V}^{IPW}(\pi') = \frac{1}{n}\sum_{i=1}^{n} \frac{\mathbb{P}_{\pi'}(W_i,\ W_j \text{ s.t } E_{ij}=1)}{\mathbb{P}_\pi(W_i,\ W_j \text{ s.t. } E_{ij}=1)} Y_i$$

$$= \frac{1}{n}\sum_{i=1}^{n} \frac{\pi'^{(W_i+M_i)}(1-\pi')^{1+N_i-(W_i+M_i)}}{\pi^{(W_i+M_i)}(1-\pi)^{1+N_i-(W_i+M_i)}} Y_i$$

Unbiased for $\bar{V}(\pi')$ under $\mathbb{R} = \text{Bern}(\pi)$

- $$\frac{d}{d\pi'}\left(\hat{V}^{IPW}(\pi')\right) = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{(W_i+M_i)\pi'^{W_i+M_i-1}(1-\pi')^{1+N_i-W_i-M_i}}{\text{denominator}} Y_i\right.$$

$$\left. - \frac{(1+N_i-W_i-M_i)\pi'^{W_i+M_i}(1-\pi')^{1+N_i-W_i-M_i-1}}{\text{denominator}} Y_i\right\}$$

We need to be more rigorous here, but we claim that this derivative is unbiased for $\frac{d}{d\pi'}\bar{V}(\pi')$

$$\left.\frac{d}{d\pi}\hat{V}^{IPW}(\pi')\right|_{\pi=\pi'} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{W_i+M_i}{\pi}Y_i - \frac{1+N_i-W_i-M_i}{1-\pi}Y_i\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{W_i}{\pi} - \frac{1-W_i}{1-\pi}\right)Y_i$$

$$+ \frac{1}{n}\sum_{i=1}^{n}\left(\frac{M_i}{\pi} - \frac{N_i-M_i}{1-\pi}\right)Y_i$$

$$= \hat{\tau}_{ADE} + \hat{\tau}_{AIE}$$ ☒

Remark = $\bar{\tau}_{TOT}$ estimated from a clustered RCT
$\bar{\tau}_{ADE}$ estimated from a unit level Bern RCT

$$\bar{\tau}_{TOT} - \bar{\tau}_{ADE} = \bar{\tau}_{AIE} + (\underbrace{\bar{\tau}_{TOT} - \bar{\tau}_{INF}}_{= V(1)-V(0)-V'(\pi)})$$

& when $\bar{\tau}_{AIE} = 0$ then $V(\pi)$ is linear and the 2nd term vanishes as well