# Y is always between 0, 1

1. AVERge total effect between 0, 1
2. We looked AIE\_2 means for graphs where no interference for 2nd degree, graphs w/ interference for 2nd degree. The differences were very severe. For ex, when pi = 0.25, the AIE\_2 avg across 3 nonzero betas was [0.1, 0.2, -0.1] AIE\_2 average across 7 zero betas was [400, 500, 600,..]. See full table of stuff in appendix.
3. We saw this visually, wanted to check. Did Welch's ANOVA test because had diff variances, then found that this was significant. Then did posthoc pairwise test.
4. ADE - 0.5 + AIE = 400

# Task List

One - Way ANOVA:

Tukey's honestly significant difference (HSD) post hoc test.

Games Howell post hoc test.

(0.33, 0, 0.67) (0.33, 0.11, 0.56) (0.33, 0.33, 0.33) (0.33, 0.56, 0.11) (0.33, 0.67, 0)

(0.33, 0.11, 0.56) 1 NA NA NA NA

(0.33, 0.33, 0.33) 1 1 NA NA NA

(0.33, 0.56, 0.11) 1 1 1 NA NA

(0.33, 0.67, 0) 0 0 0 0 NA

(0.5, 0, 0.5) 1 1 1 1 0

(0.5, 0.5, 0) 0 0 0 0 1

(0.67, 0, 0.33) 1 1 1 1 0

(0.67, 0.33, 0) 0 0 0 0 1

(1, 0, 0) 0 0 0 0 1

(0.5, 0, 0.5) (0.5, 0.5, 0) (0.67, 0, 0.33) (0.67, 0.33, 0)

(0.33, 0.11, 0.56) NA NA NA NA

(0.33, 0.33, 0.33) NA NA NA NA

(0.33, 0.56, 0.11) NA NA NA NA

(0.33, 0.67, 0) NA NA NA NA

(0.5, 0, 0.5) NA NA NA NA

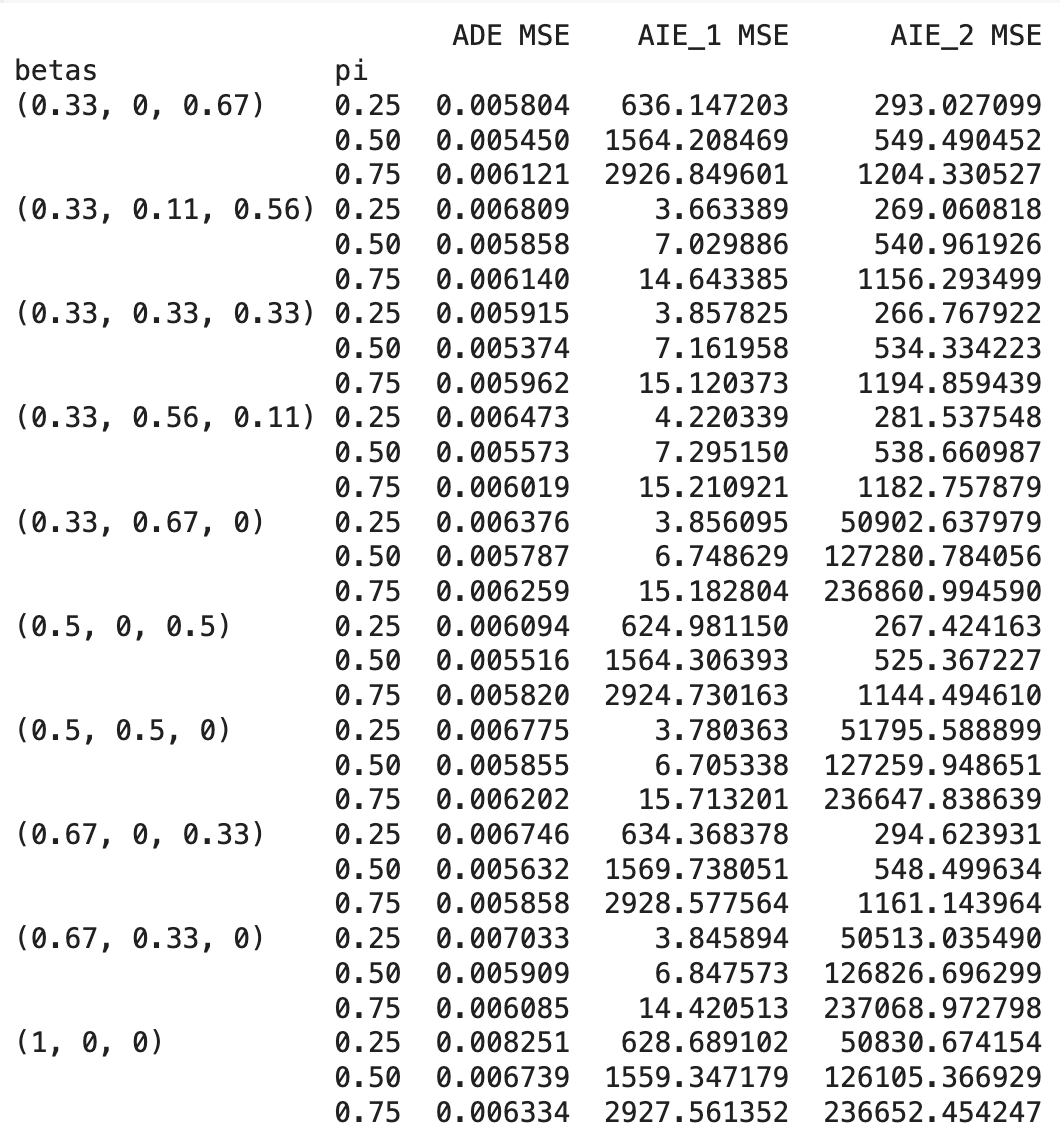
(0.5, 0.5, 0) 0 NA NA NA

(0.67, 0, 0.33) 1 0.00000000 NA NA

(0.67, 0.33, 0) 0 0.05791794 0 NA

(1, 0, 0) 0 0.71225811 0 1

**Graphs and Tables**

* ADE/AIE convergence graph over trials – Poonam
* Fix graphs for MSE over trials for Setting 1 – Amy (only know MSE for setting 1)
  + MSE table
    - 
* Mean of estimators table for Setting 1, 2, 3– Poonam
* Fix replication graphs for each setting (basically just choose a specific beta) - evelyn, done? - check
  + One beta pair across settings; pi vs ADE/AIE1
* Extension to AIE2 fix graphs (basically just choose a specific beta) - evelyn, done
  + One beta pair across settings; pi vs ADE/AIE1/AIE2
  + Just choose one with beta3=0
* Real world data stuff - evelyn??
  + Replication graph - only beta\_1 (2/3) and beta\_2 (1/3)
  + One beta pair across setting 1; pi vs ADE/AIE1/AIE2
  + Convergence graph (screen shot) - Graph structures of dataset
    - bio-grid-plant : <https://networkrepository.com/bio-grid-plant.php>
    - bio-mouse: <https://networkrepository.com/bio-mouse-gene.php>
    - bio-diseasome: <https://networkrepository.com/bio-diseasome.php>

**Things We Need**

1. **Replicate Experiments**
2. **Propose extension via second degree neighbors, redo graphs but with second degree coefficients included**
3. **Compare AIE\_2 for beta\_3 = 0 and beta\_3 != 0 → pretty obvious when there's a linear interference structure, can display the matrix of means**
4. **EM graph, MSE graph**

**Writeup**

* Related works
* Data
* Methods
* Experiments
* Conclusions

# Real World Datasets

* Bio-grid-mouse
* Bio-grid-plant
* biodisease

# Roshni Note

Policy maker can observe that the total effect is (for ex) 2, but is unsure how this can affect their results. They first make an estimate of their direct effect and correspondingly find that AIE is 2 - ADE. Policy maker can test diff interference graphs in order to determine what the interference structure in the population is.

Is the total effect easy to calculate?

* Should be observable ?
* If they know what population looks like when no one is treated, then they treat everyone, then they can observe the difference b/w the two?????
* AOE = ADE + AIE

If E is the same as what you used to generate the data, you should recover your data

If E is slightly wrong, is there a big effect on your estimation? Ie how robust is the estimator to slight inaccuracies in E (to answer the question: what happens when you don’t know the interference structure)

For the one with first and second, can isolate the interference from each by making matrix w for example, second but not first

# Remaining

* Figure out AIE
* Finish completely synthetic simulations
  + Plots for AIE and stuff
* Semi-synthetic with graphs & randomly generated W’s

# 12/5/2023

Just choosing three datasets rn:

* bio-grid-mouse

# Rando Ideas

**Overview**

* Plug in second degree neighbors everywhere
* Test how sensitive the misspecification is under diff coefficients
  + Ex potential conclusion: with a second deg term, the misspecification is much worse when the coeff for the second deg term is larger
  + Ex potential conclusion: in this setting, the misspecification performance doesn’t depend on the coeff’s of the second deg term, all that matters is that there is a second deg term

**Remaining Tasks**

* Email???
  + Aronow & Samii,2017; Hudgens & Halloran, 2008
  + Halloran & Struchiner (1995). Savje et al. (2021)
* Data generation
* Make plots of pi vs causal effect for each tau
* Plot pi vs expected outcome – should be p quick
* Figure out how to metric things ie how do we quantify amount of misspecification
  + Data analysis confidence interval something??? idk
* Analysis
  + How do diff betas impact misspecification, maybe make a plot of betas vs misspecification amount
  + Optional: how does noise amount impact misspecification, maybe make a plot of noise amount vs misspecification amount
* See if we can get some real graph data???
* Later
  + Write up things

**Data Generation Tasks**

* Eq 1:
  + First deg neighbor – done, poonam
    - OG: beta\_2 = ⅔, beta\_3 = ⅓
    - Need to simulate: beta\_2 = ½, beta\_3 = 1/2
  + Second deg neighbor –
  + First and second deg neighbor – done, poonam
    - OG: beta\_2 = ½, beta\_3 = ⅓, beta\_4 = ⅙
    - Need to simulate: beta\_2 = ½, beta\_3 = ¼, beta\_4 = ¼
    - Need to simulate: beta\_2 = ½, beta\_3 = ⅙, beta\_4 = ⅓
* Eq 2:
  + First deg neighbor – done, evelyn
  + Second deg neighbor – done, evelyn
  + First and second deg neighbor – done, evelyn - but need to generate
* Eq 3:
  + First deg neighbor – done, evelyn
  + Second deg neighbor – amy
  + First and second deg neighbor – amy
    - Prop of first and second combined
      * (A + A2) @ W / |A + A2|
      * Or weighted average
    - beta\_2 = ½, beta\_3 = ⅓, beta\_4 = ⅙
    - beta\_2 = ½, beta\_3 = ¼, beta\_4 = ¼
    - beta\_2 = ½, beta\_3 = ⅙, beta\_4 = ⅓

Amy Saying Shit

* Eq 2: test how misspecification varies with coeff of W\_i?
* Eq 2: what if we didn’t square and did a higher/lower power..?? Does this make sense conceptually,,,,can we willy nilly modify eq like that and have it in right range
* Eq 3: if do weighted average, can see how misspecification is impacted by the weighted average coeffs

# Notes from 11/26

* Equation 6 – maybe when we generate data with the second degree neighbors, we average among weights of second degree neighbors and then multiply by ⅓ or ⅙ or something like that???????
* **TO DO**
  + Reread and try to comprehend the Hu et al paper - everyone
  + Implement eq 2 & replicate experiment (simulate & test estimators) - evelyn
    - Includes seeing if you can vectorize
  + See if we have any interesting modifications to eq 2 - amy
    - If so, simulate & test estimators
  + Sigmoid uh see if we can come up w something to plug into the exp - idk
  + Silly little noise experiments on eq 2 and 1
    - I can do some silly noise experiment on equation 1… poonam
  + Brainstorm if there’s other misspecifications we can test - anyone
  + See if there’s more interesting test statistics to add to the 11/26 file - anyone

What dominik said:

* Most simulated data assume some probability stuff, like gaussian noise or something, etc.
* But things in reality are not always gaussian/linear
* Instead of just adding random terms to simulation, use real world datasets (Semi-synthetic generated data), get covariate x, simulate treatment, for y want some y= f(t, x) + noise
* Do semi synthetic data - take a real world dataset and extract covariates and stuff
* Take a real world dataset, get y values from it where y = f(t, x) + noise
* You can get f(t, x) from a random forest
* We can experiment with higher dimensional x's, mess w propensity score, etc.

# Summarization of Stuff

* See 11\_25\_23 and 11\_24\_23 notebooks. I did things
* Maybe some other ways to vary? Change how the graph is generated (aka not 100 regular or whatever) or keep pi constant but then vary the standard deviation of the noise/RV used for the noise to see what happens???

# Need to Replicate Paper o\_\_o

<https://arxiv.org/pdf/2104.03802.pdf> ← version of paper to work off of when no stanford wifi

* 
* Assume constant treatment assignment probabilities so pi\_i = pi\_0 for all i. They did it for pi\_i = 0.25, 0.50, 0.75
* Says that the number of neighbors is 100 for all vertices in the graph. Not sure how many vertices there are in total though (??? is this something that matters)

1. Generate adjacency matrix – it's nxn, symmetric, just 1s and 0s, each vertex has 100 neighbors
   1. <https://networkx.org/documentation/stable/reference/generated/networkx.generators.random_graphs.random_regular_graph.html#networkx.generators.random_graphs.random_regular_graph>
2. Set pi\_i
3. Generate W\_i based on Ber(pi\_i)
4. Using W, calculate Y\_i from above equation
5. ??????????????? calculate the metrics they were talking about or smthn. Idk.

Try to generate data where Y\_i is affected by neighor and neighbor's neighbors…

Second degree neighbors → adjacency matrix squared

Y\_i = (W \* A^2)/50 + (W\*A)/300 + 2W/3 + noise maybe?

Why am i dyin trying to figure out a fast way to find second degree neighbors ummm… Lol.

# Notes from Office Hours

**3 things to vary:**

1. **Nonlinear term**
2. **Second-degree neighbors**
3. **Noise????**

**Start with one paper**

**Modify to nonlinear term, maybe make it depend on not just direct neighbor but also indirect neighbors**

* **Ex: not just E\_ij**
* **Add X\_1 x\_2, e^{x\_1}, a \* x\_1 ^ 2, etc.**
* **Before experiment, maybe have like 3 points you want to vary, measure results and see what happens when the one thing is change – make result more readable and takeaway message cleaner**
* **Also see if there’s a case where there’s [i forgot what I was saying]**
* **Maybe estimator is sensitive to one particular type of misspecification but fine on another**
* **Maybe estimator performance is sensitive to noise level, like epsilon term has high noise level**
* **If high var, maybe linear term doesn’t matter too much, etc. ??**
* **Maybe if add nonlinear term, you can restore the validity**
  + **Ex: if true regression is a \* x\_1^2, beside x\_1 and x\_2, by adding x\_1^2 term, can restore the validity**
  + **Can test how to make more robust??**

[**https://www.overleaf.com/6119383926kvmrxygtprhs#1b628b**](https://www.overleaf.com/6119383926kvmrxygtprhs#1b628b)

# Project Proposal Notes

**NEED TO READ/WATCH:**

**Intro sources**

1. <https://www.youtube.com/watch?v=Oyt0hNsfpsU&>

[interference.pdf](https://drive.google.com/file/d/18NiXSQwauMM2YZGlnovZzgWP-9YbIoft/view?usp=sharing) ← roshni's notes

1. <https://www.jstor.org/stable/27640105>

Papers to read to Understand Project

* Want to have a good idea of what the problem/"interference" is, don't need to understand Every detail

1. <https://academic.oup.com/biomet/article-abstract/109/4/1165/6524620?redirectedFrom=fulltext&login=false>
   1. <https://arxiv.org/pdf/2104.03802.pdf>
2. <https://arxiv.org/abs/1404.7530>

* Read the introduction of roshni's papers: "About learning robust treatment policies" – want to understand what "robustness" is

1. <https://arxiv.org/pdf/2209.01754.pdf>
2. <https://arxiv.org/abs/2304.11735>

***PROPOSAL GUIDELINES:***

What is the problem that you will be investigating? Why is it interesting? What reading will you examine to provide context and background?

What data will you use? If you are collecting new data, how will you do it?

What method or algorithm are you proposing? If there are existing implementations, will you use them and how? How do you plan to improve or modify such implementations? You don't have to have an exact answer at this point, but you should have a general sense of how you will approach the problem you are working on.

How will you evaluate your results? Qualitatively, what kind of results do you expect (e.g. plots or figures)? Quantitatively, what kind of analysis will you use to evaluate and/or compare your results?

Evaluating the Robustness of Causal Estimators under Misspecified Interference

A lot of the estimators that we studied in class assume Stable Unit Treatment Value Assumption (SUTVA), but there are a lot of cases in the real world where SUTVA does not hold and the potential outcomes of an individual are affected by the treatment of others. For example, in the case of vaccinations, one's neighbors/coworkers/friends being vaccinated reduces one's chance of getting sick, regardless of whether or not that individual actually got vaccinated. Most estimators that account for some sort of interference assume a particular structure to the interference. For example, in Hu et al's paper "Average direct and indirect causal effects under interference", they give examples of a network model where adjacencies between different units impact the the outcome of a unit and a variable-strength ties model depending on whether the adjacency contains a close relationship with the unit or not. However, in a real world context, it may be the case that we don't know the exact structure of interference between different units, only that there may be some interference. Because of this, we are interested in determining how robust these estimators are in the case of misspecified interference. For example, the network model mentioned above leads to a linear model in Hu et al's paper; but what would happen if the interference relationship was non-linear in nature? Would the estimators of indirect and direct causal effects still be accurate?

\*\*GIVE AN EXAMPLE OF SOMETHING THAT WE COULD TRY\*\*

We will first simulate data according to some distribution and check that these estimators work well for those distributions, replicating the results in XYZ papers. Then, we will modify these distributions and test the extent to which the estimator differs from the actual treatment effect when the distribution does not match the assumptions from the papers.

We will evaluate our results by benchmarking them to the baseline results of proper distribution estimation. We will see if the difference between the estimator and the actual treatment effect, which we can calculate from the distribution that we simulate our data with, is significant.

Quantitatively,

Qualitatively,

The goal is trying to learn robust treatment policies in the presence of interference

Have data subject to interference, goal is to learn a policy from that data

Say you have data that's generated, and potential outcomes dont satisfy SUTVA, want to know who to assign treatment to maximize overall potential outcomes in population

How do you go about learning such a robust policy?

Example: vaccines – Policy maker: who to target such that overall health of population is good

If you target the right people, maybe you can improve overall population welfare in a low-cost way

Agricultural subsidy – some farmer training program, etc. if u train farmers become more productive. Maybe farmers interact with each other, etc, spread practices more effective to target farmers with high social capital → want to figure out who to target, but don't really know the social structure of the population

Could: run experiment, collect data, willing to make some assumptions about how the potential outcomes are related to each other

Notes from calling Roshni:

* partial identification for treatment effects in the presence of interference
* idea is that if there’s interference, then we don’t know…???????…???
* going to try partially identify treatment effects (get upper and lower bounds on treatment effect) assuming some structural model for the potential outcome for some interference structure
* motivation: often when people run experiments, assume SUTVA holds, that’s how we estimate treatment effect — but a lot of times there’s interference, ppl’s outcomes depend on treatment of others
* idea: study what happens when if you’re willing to assume that the potential outcomes have a simple structure but u don’t know the extent to which interference affect potential outcome
* make assumptions about how much other ppls treatment affects potential outcome
* 2^n possible potential outcomes without SUTVA
* make some assumptions about the strength of the interference, or how much someone else’s potential outcome will affect urs —> then try to get bounds on treatment effect
* don’t know exact structure, but put bounds on how strong it is — different from typical approach of assuming a certain structure
* in the presence of different interference patterns
* cite them, then say — if it ends up being hard, then we can do a simulation study like in XYZ paper