

1 (25 points) Material covered Tuesday

Part 1 is all about the **consistency assumption**.

- 1.1. (5 points) Suppose there are two children, Child 1 and Child 2, who are best friends. They find a baseball in the park and each asks their parents to buy them a baseball glove. Let A_i indicate whether each child i receives a glove. Let Y_i indicate whether each child gets to play catch with their best friend. Define a set of four potential outcomes for Child 1 to correspond to the following statement: Child 1 gets to play catch if and only if both Child 1 and Child 2 receive baseball gloves.
- 1.2. (5 points) There is a campaign to vaccinate the population against COVID-19. Suppose A_i indicates whether person i is vaccinated. Suppose Y_i indicates whether that person is hospitalized in the next 12 months. A researcher says to you: $\{Y_i^1, Y_i^0\}$ will represent person i 's outcome if vaccinated and if not. In a couple sentences, discuss how interference might create problems for this definition of the potential outcomes.
- 1.3. (5 points) A medical researcher wants to understand the effect of surgery on survival. Let Y_i^0 denote unit i 's survival if given no surgery. Let $Y_i^1, Y_i^2, \dots, Y_i^k$ denote survival if given surgery from surgeon $1, 2, \dots, k$. Another researcher says this is too complicated; they propose instead to focus on the effect of surgery Y_i^{Surgery} versus no surgery Y_i^0 . In words and in math, state the treatment variation irrelevance assumption this second researcher is making to simplify the set of potential outcomes.
- 1.4. (5 points) A researcher committed to evidence-based policy recruits a set of K12 teachers who agree to try a new teaching method. After randomization, half of these highly-motivated teachers carry out the teaching method. The students in the “my teacher used the method” condition have higher test scores than those in the “my teacher did not use the method” condition. Inspired by the successful study, the district mandates that all teachers use the new method, and many of the teachers do so begrudgingly. A year later, the results are disappointing—test scores did not rise as much as expected. In a sentence or two, given an interpretation of this finding that invokes a violation of treatment variation irrelevance.
- 1.5. (5 points) An economist says “I focus on individuals, so I define my potential outcomes Y_i^a where a denotes a treatment value for person i .” A sociologist says “The outcome for each individual depends on the treatments of everyone in their network.” Suppose we are studying a network of 25 people (including person i), and the treatment is binary. For person i , how many potential outcomes does the economist define? How many does the sociologist define?

2 (25 points) Material covered Thursday

Part 2 is all about the **bounds**.

As a TA, you lead a tiny undergraduate section with 3 students this semester. Some will do the reading and some will not, and they will each be assessed at the end of the semester with a final exam score Y_i falling between 0 and 100%. You are interested in the average causal effect of doing the reading on that exam score:

$$\tau = \frac{1}{3} \sum_{i=1}^3 \left(Y_i^{\text{Reading}} - Y_i^{\text{No reading}} \right)$$

Note that this entire problem will make no assumptions of randomization or exchangeability.

- 2.1. (7 points) Before seeing any data, what are the upper and lower bounds on the possible values for τ ?

During the course, Students 1 and 2 do the reading. Student 3 does not. The observed final exam scores for students 1, 2, and 3 are $\{80\%, 100\%, 70\%\}$, depicted in the table below.

Observed Data	Student Number	Y_i^{Reading}	$Y_i^{\text{No reading}}$	Effect
	1	80%	?	?
	2	100%	?	?
	3	?	70%	?

- 2.2. (7 points) Given your data, what is the upper bound on τ ? In other words, what is the most helpful that the reading could be, on average, for the final exam score?
- 2.3. (7 points) Given your data, what is the lower bound on τ ? In other words, what is the most harmful that the reading could be, on average, for the final exam score?
- 2.4. (4 points) Using this bounding approach, will the bounds interval always contain the possibility of an effect equal to 0?