Please submit your homework on Sakai sometime before 4:40pm on the due date. Homework turned in after 4:40pm on the due date will be given half credit, homework turned in after the due date will receive no credit. You may discuss these problems with each other verbally, but must write up the answers on your own, and may not share or show your answers to anyone else. Please be concise!

1. *Thyroid Levels* During a check-up, a physician finds that his patient's thyroid levels are too low. He prescribes medication at a high dose and asks her to be re-tested in a month. At the second test, the patient's thyroid levels are now too high, so the physician switches her to a low dose of medication and again asks her to be re-tested in a month. At the third test, the patient's thyroid levels are perfect, and so the doctor decides that she should stay at the low dose indefinitely, with no further testing.

   (a) (1 point) What are the units?

   (b) (1 point) What is the treatment?

   (c) (1 point) What are the potential outcomes?

   (d) (2 points) Is SUTVA plausible? If so, explain why briefly. If not, offer an assumption that, if true, would make SUTVA plausible. (Remember that SUTVA has two parts).

   (e) (1 point) Why must SUTVA be plausible in order for us to use the potential outcome framework?

   (f) (2 points) Do you agree with the physician's judgment that the patient should remain on the low dose, given the information provided here? Comment briefly and relate to concepts in the course.

   (g) (3 points) Does the physician's assignment mechanism appear to be probabilistic? Individualistic? Unconfounded? Explain in one sentence each.

2. *Gender Diversity in Teachers*
   Suppose that the federal Department of Education surveys school districts to find out how many male and female teachers are employed in each district. The Department learns that some school districts have been hiring only female teachers. Rather than intervene directly in hiring, the Department creates a workshop for school administrators that focuses on the benefits of diversity. The workshop is given at all schools districts that currently have only female teachers. At other school districts across the country, the workshop is given if requested by school administrators. Two years later, the Department again asks all school districts to indicate the number of male and female teachers employed.

   (a) (3 points) Identify the units, potential outcomes, treatment, and any observed covariates.

   (b) (2 points) Is SUTVA plausible? If so, explain why briefly. If not, offer an assumption that, if true, would make SUTVA plausible.

   (c) (4 points) Is the assignment mechanism probabilistic? Individualistic? Unconfounded? Observational or randomized? Regular? Explain in one sentence each.

   (d) (2 points) What other covariates would you collect data on, if you could? (Just listing a couple is fine... no need to list every possibility)

3. *Traffic Flow*

   Engineers in Dresden, Germany were looking at ways to improve traffic flow by enabling traffic lights to communicate information about traffic flow with nearby traffic lights ("flexible lights")[1]. The data (available here):

   (https://docs.google.com/spreadsheet/ccc?key=0AtJ5X5rxFtfqdFFyOWlnT1BXM2ZzZTNPdEd5Ni1WYWc&usp=sharing) show results of one experiment where they simulated buses moving along a street and recorded the delay time (in seconds) for both a fixed time and a flexible system of lights. The process was repeated under both conditions for a sample of 24 simulated scenarios. This is a very special situation in which we can observe both potential outcomes for each unit, because it is a computer simulation, not a real-life study, so we can ensure the conditions are identical for both runs of the simulation.

   You can use anything you want to answer this question, but future problem sets in the class will require the use of R, so I recommend doing this problem in R if possible.

     (a) (1 point) What are the units?

     (b) (1 point) What is the treatment?

     (c) (1 point) Define the unit-level causal effect (in words or notation, not a number).

     (d) (1 point) What are the potential outcomes and causal effect for the first unit?

     (e) (1 point) What is the average causal effect for this sample? Interpret in context.

     (f) (1 point) In real life, in a study involving actual traffic on real roads, explain why we could never obtain data like this.

     (g) (2 points) Let's suppose this data comes from actual traffic on real roads. Using your own assigment mechanism, assign units to treatment groups and estimate the average causal effect using only data that would be observed based on this assignment. Would this estimate make sense to use (if you hadn't already observed the truth)? Why or why not?

---

[1]Lammer and Helbing, "Self-Stabilizing decentralized signal control of realistic, saturated network traffic", Santa Fe Institute working paper # 10-09-019, September 2010