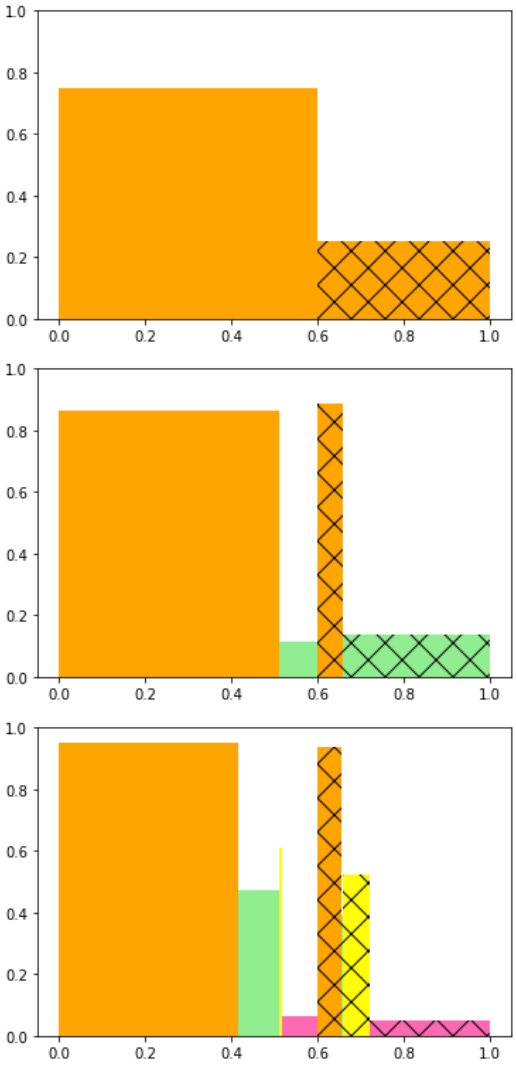
Philosophy of Statistics: Homework 4

due on Gradescope by 11am on Thursday May 20

Guidelines. Some questions ask you to justify your answers. For these questions, credit will be based on how well you justify your answers, not on whether your answers are correct. (There's often no consensus on the correct answers, even among statisticians.) However, that doesn't mean that anything goes: some answers will be hard to justify well. I give suggested word counts but these are just ballpark numbers. Don't sweat them too much. Collaboration is encouraged, but make sure to write up your answers by yourself and list your collaborators.

Problem 1 (15 points). We can represent Simpson reversals in pictures, such as the following:



What do the pictures mean? More specifically:

(A) What does the width of a column represent?

(B) What does the height of a column represent?

(C) What does the area of a column represent?

(D) What does no hatching v. hatching represent?

(E) Within a particular picture, what do the colors represent?

(F) What is the relation between the total unhatched areas in the first, second, and third pictures, and similarly for the hatched areas?

Problem 2 (20 points). As discussed in class, and represented in the pictures above, your data could exhibit not just one but indefinitely many Simpson reversals. For example, comparing the treatment group and control group in a drug trial, there might be:

- overall, a *higher* recovery rate
- for men and for women, *lower* recovery rates
- for men over 50, men under 50, women over 50, and women under 50, *higher* recovery rates

Can you come up with such a case, where there are two Simpson reversals in the data, by filling in the table below with suitable numbers?

	T	T	T	T	C	C	C	C
	M	M	W	W	M	M	W	W
	≥ 50	< 50	≥ 50	< 50	≥ 50	< 50	≥ 50	< 50
recovered								
didn't recover								
recovery rate								

Your answer should show the calculations of the other relevant recovery rates too: i.e. for men, for women, and overall, in both the treatment and control groups.

Problem 3 (5 points). In a problem from a previous homework, I asked you to come up with an example of a Simpson reversal, though I didn't use that terminology. Which problem?

Problem 4 (15 points). Do Simpson reversals arise in practice? Yes, sometimes: the kidney stones and baseball examples from class are genuine cases. Search online to find another real-life example of a Simpson reversal. In your answer you should state your source, explain any necessary background, and present the data and relevant calculations clearly.

Problem 5 (15 points). In a *binary* Simpson reversal, the population is split into *two* sub-populations. The examples we looked at in class were all binary Simpson reversals. (In, say, the kidney stones study, the patients were split into those with large kidney stones and those with small kidney stones.) But a Simpson reversal can occur when the population is split into any number of sub-populations, not just two. For example, comparing the treatment group and control group in a drug trial, there might be a *lower* recovery rate overall, but a *higher* recovery rate among men over 50, men under 50, women over 50 and women under 50.

Can you come up with such a case, by filling in the tables below with suitable numbers?

	T	T	T	T
	men over 50	men under 50	women over 50	women under 50
recovered				
didn't recover				
recovery rate				

	C	C	C	C
	men over 50	men under 50	women over 50	women under 50
recovered				
didn't recover				
recovery rate				

You should also show the calculation of the overall recovery rate, both in the treatment group and in the control group.

Problem 6 (10 points). Draw a picture, similar to the ones above, to represent the data you came up with in Problem 5. Your picture needn't be exactly to scale, but should be close enough that it's easy to check that it corresponds to your data. (You might find it easier to do this problem before the previous one: i.e. draw the picture first and come up with data based on your picture.)

Problem 7 (20 points). Judea Pearl argues that when the data exhibit a Simpson reversal, in some cases we should base our conclusion (e.g. about whether the drug is beneficial) on the aggregated data and in other cases on the segregated data, depending on the *causal story* behind the data.

Can you come up with two cases of your own such that we should go by the aggregated data in one and the segregated data in the other? Describe the two cases, including the data exhibiting the reversal and the causal story behind the data, and explain why you think we should go by the aggregated data or segregated data.

Your cases needn't be realistic: just focus on making your claims about them as convincing as possible.