Chapter 6 outline of discussion

This chapter has two parts: straightforward discussion of two-way tables and Simpson's Paradox.

Form a group of between 2 and 4 students to work on this.

Part 1.

Recall when you are asked for percentages, you must take a fraction, turn it into a decimal, and then convert that to a percentage. In this chapter, the main challenge is to figure out what two numbers go into the numerator and denominator of the fraction.

(5 minutes) Spend about five minutes reading questions 1-7 below. These will help you see what we are doing in these sections. You don't have to DO them yet – just think about how these are progressing to generally more complex questions. For Example 6.1, with data in Table 6.1 (pages 159-160)

1. What percent of the young adult males in this study said they were "almost certain" to be rich by aged 30?

2. What percent of the young adult females in this study said they had "some chance, but probably not" to be rich by aged 30?

3. What percent of the people in this study were male?

4. What is the distribution of opinion among all the men in the study? (Start with which of the first three questions gives you one part of the answer to this?)

5. What is the distribution of sex among all the people in this study? (Start with which of the first three questions gives you one part of this?)

6. Compare the distributions of opinion for the males and the females in this study. (Was the answer to one of the previous questions a part of the answer to this? If so, which one?)

7. What do the data tell us about whether there is a relationship between sex and their opinions about becoming rich for the people in this study?

(Five minutes) Next, spend about five minutes on these next three questions. They might be a little easier and will help you work toward answering questions 1-6. In the answers to all of these, I am not concerned so much about the answer as the process you are using to find the answer. That is – what numbers will you divide and how are you deciding which numbers to use? Usually it is finding the denominator in each case that is tricky. (There is a pattern to the way I am using different fonts here. Can you find it?)

8. What fraction of **all the males in this study** said they are "almost certain" to be rich?

9. What fraction of **all the people in this study who say they are "almost certain" to be rich** are **males**?

10. What fraction of **all the people in this study** are **males who say they are "almost certain" to be rich**?

(Five minutes) Now, spend about five minutes going back to questions 1-5 and answering those, in fraction form. You can also give the percentages, but I am most interested in your giving the fractions, so it is clear what numbers you are choosing and why.

(Five minutes) The whole class discusses the answers to questions 1-5. Include in this discussion why, in both questions 4 and 5, all the fractions (percentages) must sum to 1 or to 100%.

(Ten minutes for questions 11-14)

11. Read the definitions of marginal distribution and conditional distribution on page 163. (Section on Conditional Distribution.) Identify one of each of these in questions 1-5.

12. Read Example 6.3 in the text. Which of questions 1-7 does that example answer?

13. Perhaps the hardest part of this entire section is answering a question like number 7 if it is just presented without all of this buildup. Most students think that seems unclear which set of conditional distributions you want to compute. Go to the paragraph right at the end of the section on Conditional Distributions – at the end of page 165 and right before Exercise 6.3. Pay careful attention to the last half of that paragraph.

    a. For this table, what are the two variables?

    b. For this table, what is the response variable?

    c. So, for this table, what conditional distributions should you compute to answer our question 7?

14. Look at exercise 6.30 at the end of the chapter. Say which conditional distributions you should compute to answer the question.

Part 2. Simpson's Paradox.

(Twenty minutes) Discuss the following questions in groups.

15. In example 6.4 about evacuating people by helicopter and road, read just the first table – a two-way table – and the sentence underneath. If you had a child/friend being evacuated, what do these data suggest to you about what type of evacuation you should choose?

16. Do you think there is more to the story than that? Why?

17. The next table is three-way table, meaning that it illustrates the relationships among three variables. What are the three variables?

18. Does the new third variable seem to be relevant to the question of whether road or helicopter is a better way to be evacuated? Why?

19. Read the two paragraphs after the three-way table and explain those to the others in your group.

20. Do you see the "Paradox" – what seems like it couldn't be true but is true?

21. What is it about these data that makes the paradox happen? (Answer: An imbalance – what is imbalanced? This is rather hard, but is an important aspect.)

22. Do Exercise 6.7 about bias in the jury pool.

Quiz 5. Due Feb. 20 at the beginning of class.

Chapter 6: 6.6 Field goal shooting
Chapter 6: 6.32 Python eggs
Chapter 3: Exercise A (not from the book, but written here.)
Suppose we have a uniform distribution on the interval from 5 to 10.

a. Sketch the distribution and label the height with its numerical value.
b. What is the total area beneath the curve?
c. What proportion of the scores are less than 8?
d. What proportion of the scores are between 6 and 9?
e. What proportion of the scores are above 7?
f. What proportion of the scores are above 11?

Chapter 3: Exercise 3.26
Chapter 7. Exercise 7.38