CHAPTER 12

Name:

Narrative: Aspirin and polyps

A study done by the Ohio State University Medical Center examined whether or not taking an aspirin a day would help colon cancer patients reduce the chance of getting subsequent colon polyps. 635 patients with colon cancer participated; 317 of them were randomly assigned to the aspirin group, and the other 318 patients were assigned to a placebo (non-aspirin) group. 54 patients in the aspirin group developed subsequent polyps, compared to 86 patients in the non-aspirin group. Many confounding variables were controlled for in this well-designed randomized experiment.

1. {Aspirin and polyps narrative} Which variable is the explanatory variable and which is the response variable in this study?
   
   Answer: Explanatory variable - aspirin/placebo; response variable - developed subsequent polyps/did not develop subsequent polyps.

2. {Aspirin and polyps narrative} Make a contingency table summarizing the data from this study, making the explanatory variable the row variable and the response variable the column variable.
   
   Answer:
   
<table>
<thead>
<tr>
<th></th>
<th>Polyps</th>
<th>No Polyps</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin</td>
<td>54</td>
<td>263</td>
<td>317</td>
</tr>
<tr>
<td>Placebo</td>
<td>86</td>
<td>232</td>
<td>318</td>
</tr>
<tr>
<td>Total</td>
<td>140</td>
<td>495</td>
<td>635</td>
</tr>
</tbody>
</table>

3. {Aspirin and polyps narrative} How many cells will the contingency table have for the data presented by this study?
   
   Answer: 4

4. {Aspirin and polyps narrative} Does taking aspirin appear to be related to the chance of developing subsequent polyps in colon cancer patients? Justify your answer using percentages.
   
   Answer: Yes. The percentage in the aspirin group suffering subsequent polyps was 17%, compared to 27% for the placebo group.

5. {Aspirin and polyps narrative} For colon cancer patients who take an aspirin a day, what is the expected rate for developing subsequent polyps?
   a. Approximately 170 occurrences per 1,000 patients.
   b. Approximately 54 occurrences per 635 patients.
   c. Approximately 54 occurrences per 140 patients.
   d. None of the above.
   
   Answer: A

6. {Aspirin and polyps narrative} Based on the results of this study, what can you conclude?
   a. Taking an aspirin a day can help prevent subsequent polyps from developing in colon cancer patients.
   b. Taking an aspirin a day has a causal effect on subsequent polyp development in colon cancer patients.
   c. Taking an aspirin a day reduces the chances of developing subsequent polyps in colon cancer patients.
   d. All of the above.
   
   Answer: D

7. {Aspirin and polyps narrative} What percentage of all the patients in this study developed subsequent polyps?
   a. $\frac{140}{635} = 22\%$
   b. $\frac{54}{140} = 39\%$
   c. $\frac{140}{495} = 28\%$
   d. None of the above.
   
   Answer: A
8. {Aspirin and polyps narrative} What percentage of patients on the placebo did not develop subsequent polyps?
   a. $\frac{232}{495} = 47\%$
   b. $\frac{232}{318} = 73\%$
   c. $\frac{232}{635} = 37\%$
   d. None of the above.
   ANSWER: B

9. Each row and column combination in a contingency table is called a(n) __________
   ANSWER: CELL

10. A table that displays the number of individuals who fall into each combination of categorical variables is called a(n) __________ table.
    ANSWER: CONTINGENCY

11. {Aspirin and polyps narrative} What is the baseline risk of developing subsequent polyps? Interpret your answer.
    ANSWER: 27%; THE RISK OF HAVING SUBSEQUENT POLYPS IF YOU DO NOT TAKE AN ASPIRIN A DAY IS 27%.

12. {Aspirin and polyps narrative} What is the relative risk of developing polyps for colon cancer patients who don’t take an aspirin a day (compared to those who do)? Interpret your answer.
    ANSWER: 1.59. THE RISK OF DEVELOPING SUBSEQUENT POLYPS IS 1.59 TIMES GREATER FOR PATIENTS WHO DO NOT TAKE AN ASPIRIN A DAY, COMPARED TO THOSE THAT DO.

13. {Aspirin and polyps narrative} What are the odds of developing subsequent polyps for colon cancer patients who don’t take an aspirin a day?
    a. 86 to 318
    b. 86 to 54
    c. 86 to 232
    d. None of the above
    ANSWER: C

14. Which of the following is not one of the basic ways to express the chances that a randomly selected individual will fall into a particular category for a categorical variable?
    a. The proportion of the total that falls into that group.
    b. The percentage of the total that falls into that group.
    c. Comparing one category to another category in the form of relative odds.
    d. All of the above are appropriate.
    ANSWER: D

Narrative: Genetic defect
Suppose a population contains 100,000 individuals, of which 500 are carriers of a certain genetic defect.

15. {Genetic defect narrative} Which of the following is not an appropriate way to express this result as a portion of the total?
    a. One half of one percent (0.5%) of all individuals in this population are carriers.
    b. The proportion in this population who are carriers is 0.005.
    c. The probability that a person randomly selected from this population is a carrier is 0.5%
    d. All of the above are appropriate.
    ANSWER: D
16. \{Genetic defect narrative\} Which of the following is an appropriate way to express this result in terms of odds?
   a. The odds of carrying the genetic defect are 5 to 995.
   b. The odds of carrying the genetic defect are 500 out of 100,000.
   c. The odds of carrying the genetic defect are 500 to 1.
   d. None of the above
   ANSWER: A

Narrative: Audition qualifications
Suppose an audition process requires applicants to first meet a series of qualifications before they can be accepted for an audition. Suppose 88% of the women met these qualifications, and 86% of the men met these qualifications. Also assume that there were 100 men and 100 women who applied.

17. \{Audition qualifications narrative\} What is the selection ratio for men compared to women? Interpret your answer.
   ANSWER: .98; MEN ARE .98 TIMES AS LIKELY TO MAKE IT THROUGH THIS ROUND OF THE AUDITION PROCESS AS WOMEN.

18. \{Audition qualifications narrative\} What is the odds ratio for being accepted (versus rejected) for women compared to men?
   ANSWER: 1.19; THE ODDS OF A WOMAN BEING ACCEPTED (VERSUS REJECTED) ARE 1.19 TIMES THOSE FOR A MAN.

19. \{Audition qualifications narrative\} What is the odds ratio for being accepted (versus rejected) for men compared to women?
   ANSWER: .84; THE ODDS OF A MAN BEING ACCEPTED (VERSUS REJECTED) ARE .84 TIMES THOSE FOR A WOMAN.

Narrative: Talent and gender
Suppose a talent contest is looking for people who can either sing or act. 900 people tried out for the first round of auditions, 300 women and 600 men. Of the men who tried out, 55% of them were allowed to go on to the next round, compared to 42% of the women. However, if you break the data down by singing versus acting auditions, the relationship changes. For acting, 20% of the women made it through, compared to 15% of the men; for singing, 85% of the women made it through, compared to 75% of the men.

20. \{Talent and gender narrative\} What do we call this phenomenon, and how could this have happened?
   ANSWER: SIMPSON’S PARADOX. MORE MEN APPLIED, AND MORE MEN MAY HAVE TRIED OUT FOR SINGING, WHICH WAS EASIER FOR BOTH GENDERS TO RECEIVE A ‘PASS THROUGH’ (COMARED TO ACTING).

21. \{Talent and gender narrative\} What is the statistical phenomena that occurred in this situation?
   a. A statistical error; these results cannot be correct.
   b. Simpson’s Paradox.
   c. The Empirical Rule.
   d. Selection Bias.
   ANSWER: B

22. What does Simpson’s Paradox teach us?
   a. The nature of an association can reverse direction when data from several groups are combined to form a single group.
   b. Omitting a third variable can mask the true relationship between two categorical variables.
   c. It is dangerous to summarize information over groups, especially if the individuals were not randomly assigned to those groups.
   d. All of the above.
   ANSWER: D
23. When you see a group of individuals summarized into a single contingency table containing only one row variable and one column variable, what should your reaction be?
   a. Great; this data is clear and easy to understand.
   b. I need to find out if the data have been collapsed over some important third variable.
   c. This is garbage; there is always more to a statistical relationship than just two simple variables.
   d. None of the above.
   ANSWER: B

24. Suppose you are on a jury in a trial someday. How could you encounter Simpson’s Paradox?
   a. You could see data that were collected from two different studies, giving you two different results.
   b. One side could present the data using two variables, and the other side could break the same data down by a third variable that reverses the direction of the results.
   c. One side could use counts to summarize the data, and the other side could use percentages or rates, reversing the direction of the relationship.
   d. All of the above.
   ANSWER: B

25. When omitting a third variable masks the relationship between two categorical variables, this phenomenon is called
   ANSWER: SIMPSON’S PARADOX

26. Simpson’s Paradox makes it clear that it is dangerous to ________ information across groups, especially if the subjects or experimental units were not ________ assigned to those groups.
   ANSWERS (RESPECTIVELY): COMBINE; RANDOMLY