Lesson 13: Summarizing Bivariate Categorical Data in a Two-Way Table

Classwork

Exercises 1–8

On an upcoming field day at school, the principal wants to provide ice cream during lunch. She will offer three flavors: chocolate, strawberry, and vanilla. She selected your class to complete a survey to help her determine how much of each flavor to buy.

1. Answer the following question. Wait for your teacher to count how many students selected each flavor. Then, record the class totals for each flavor in the chart below.

    "Which of the following three ice cream flavors is your favorite: chocolate, strawberry, or vanilla?"

Ice Cream Flavor	Chocolate	Strawberry	Vanilla	Total
Number of Students				

2. Which ice cream flavor do most students prefer?

3. Which ice cream flavor do the fewest students prefer?

4. What percentage of students preferred each flavor? Round to the nearest tenth of a percent.

5. Do the numbers in the chart above summarize data on a categorical variable or a numerical variable?

6.   Do the students in your class represent a random sample of all students in your school? Why or why not? Discuss this with your neighbor.

7.   Is your class representative of all the other classes at your school? Why or why not? Discuss this with your neighbor.

8.   Do you think the principal will get an accurate estimate of the proportion of students that prefer each ice cream flavor for the whole school using only your class? Why or why not? Discuss this with your neighbor.

Example 1

Students in a different class were asked the same question about their favorite ice cream flavor. The table below shows the ice cream flavors and the number of students who chose each flavor for that particular class. This table is called a one-way frequency table because it shows the counts of a univariate categorical variable.

This is the univariate categorical variable. ⟶

These are the counts for each category. ⟶

Ice Cream Flavor	Chocolate	Strawberry	Vanilla	Total
Number of Students	11	4	10	25

We compute the relative frequency for each ice cream flavor by dividing the count by the total number of observations.

$$\text{relative frequency} = \frac{\text{count for a category}}{\text{total number of observations}}$$

Since 11 out of 25 students answered chocolate, the relative frequency would be $\frac{11}{25} = 0.44$. This relative frequency shows that 44% of the class prefers chocolate ice cream. In other words, the relative frequency is the proportional value that each category is of the whole.

Exercises 9–10

Use the table for the preferred ice cream flavors from the class in Example 1 to answer the following questions.

9. What is the relative frequency for the category strawberry?

10. Write a sentence interpreting the relative frequency value in the context of strawberry ice cream preference.

Example 2

The principal also wondered if boys and girls have different favorite ice cream flavors. She decided to redo the survey by taking a random sample of students from the school and recording both their favorite ice cream flavor and their gender. She asked the following two questions:

- "Which of the following ice cream flavors is your favorite: chocolate, strawberry, or vanilla?"
- "What is your gender: male or female?"

The results of the survey are as follows:

- Of the 30 students who prefer chocolate ice cream, 22 are males.
- Of the 25 students who prefer strawberry ice cream, 15 are females.
- Of the 27 students who prefer vanilla ice cream, 13 are males.

The values of two variables, which were ice cream flavor and gender, were recorded in this survey. Since both of the variables are categorical, the data are bivariate categorical data.

Exercises 11–17

11. Can we display these data in a one-way frequency table?   Why or why not?

12. Summarize the results of the second survey of favorite ice cream flavors in the following table:

		Favorite Ice Cream Flavor			
		Chocolate	Strawberry	Vanilla	Total
Gender	Male				
	Female				
	Total				

13. Calculate the relative frequencies for the table above and write them in the table.

		Favorite Ice Cream Flavor			
		Chocolate	Strawberry	Vanilla	Total
Gender	Male				
	Female				
	Total				

Use the relative frequency values in the table to answer the following questions:

14. What is the proportion of the students that prefer chocolate ice cream?

15. What is the proportion of students that are female and prefer vanilla ice cream?

16. Write a sentence explaining the meaning of the approximate relative frequency 0.55.

17. Write a sentence explaining the meaning of the approximate relative frequency 0.10.

Example 3

In the previous exercises, you used the total number of students to calculate relative frequencies. These relative frequencies were the proportion of the whole group who answered the survey a certain way. Sometimes we use row or column totals to calculate relative frequencies. We call these *row relative frequencies* or *column relative frequencies*.

Below is the two-way frequency table for your reference. To calculate "the proportion of male students that prefer chocolate ice cream," divide the 22 male students who preferred chocolate ice cream by the total of 45 male students. This proportion is $\frac{22}{45} = 0.49$. Notice that you used the row total to make this calculation. This is a row relative frequency.

		Favorite Ice Cream Flavor			
		Chocolate	Strawberry	Vanilla	Total
Gender	Male	22	10	13	45
	Female	8	15	14	37
	Total	30	25	27	82

Exercises 18–22

In Exercise 13, you used the total number of students to calculate relative frequencies. These relative frequencies were the proportion of the whole group who answered the survey a certain way.

18. Suppose you are interested in the proportion of male students that prefer chocolate ice cream. How is this value different from "the proportion of students that are male and prefer chocolate ice cream"? Discuss this with your neighbor.

19. Use the table provided in Example 3 to calculate the following relative frequencies.

    a. What proportion of students that prefer vanilla ice cream is female?

b.   What proportion of male students prefers strawberry ice cream? Write a sentence explaining the meaning of this proportion in context of this problem.

c.   What proportion of female students prefers strawberry ice cream?

d.   What proportion of students who prefer strawberry ice cream is female?

20. A student is selected at random from this school. What would you predict this student's favorite ice cream to be? Explain why you choose this flavor.

21. Suppose the randomly selected student is male. What would you predict his favorite flavor of ice cream to be? Explain why you choose this flavor.

22. Suppose the randomly selected student is female. What would you predict her favorite flavor of ice cream to be? Explain why you choose this flavor.

Lesson Summary

- Univariate categorical data are displayed in a one-way frequency table.
- Bivariate categorical data are displayed in a two-way frequency table.
- *Relative frequency* is the frequency divided by a total (frequency/total).
- A *cell relative frequency* is a cell frequency divided by the total number of observations.
- A *row relative frequency* is a cell frequency divided by the row total.
- A *column relative frequency* is a cell frequency divided by the column total.

Problem Set

Every student at Abigail Douglas Middle School is enrolled in exactly one extracurricular activity. The school counselor recorded data on extracurricular activity and gender for all 254 eighth-grade students at the school .

The counselor's findings for the 254 eighth-grade students are the following:

- Of the 80 students enrolled in band, 42 are male.
- Of the 21 students enrolled in art, 9 are female.
- Of the 65 students enrolled in choir, 20 are male.
- Of the 88 students enrolled in sports, 30 are female.

1. Complete the table below.

		Extracurricular Activities				
		Band	Choir	Sports	Art	Total
Gender	Female					
	Male					
	Total					

2. Write a sentence explaining the meaning of the frequency 38 in this table.

3. What proportion of students is male and enrolled in choir?

4. What proportion of students is enrolled in a musical extracurricular activity (i.e., band or choir)?

5. What proportion of male students is enrolled in sports?

6. What proportion of students enrolled in sports is male?

    A pregnant woman will often undergo ultrasound tests to monitor her baby's health. These tests can also be used to predict the gender of the baby, but these predictions are not always accurate. Data on the gender predicted by ultrasound and the actual gender of the baby for 1,000 babies are summarized in the two-way table below.

		Predicted Gender	
		Female	Male
Actual Gender	Female	432	48
	Male	130	390

7. Write a sentence explaining the meaning of the frequency 130 in this table.

8. What is the proportion of babies predicted to be male but were actually female?

9. What is the proportion of incorrect ultrasound gender predictions?

10. For babies predicted to be female, what proportion of the predictions was correct?

11. For babies predicted to be male, what proportion of the predictions was correct?