

Lesson 11: Conditional Relative Frequencies and Association

Classwork

After further discussion, the students involved in designing the superhero comic strip decided that before any decision is made, a more careful look at the data on the special powers a superhero character could possess was needed. There is an association between gender and superpower response if the superpower responses of males are not the same as the superpower responses of females. Examining each row of the table can help determine whether or not there is an association.

Exploratory Challenge 1: Conditional Relative Frequencies

Recall the two-way table from the previous lesson.

	To Fly	Freeze Time	Invisibility	Super Strength	Telepathy	Total
Females	49	60	48	1	70	228
Males	51	71	27	25	48	222
Total	100	131	75	26	118	450

A *conditional relative frequency* compares a frequency count to the marginal total that represents the condition of interest. For example, the condition of interest in the first row is females. The row conditional relative frequency of females responding “invisibility” as the favorite superpower is $\frac{48}{228}$, or approximately 0.211. This conditional relative frequency indicates that approximately 21.1% of females prefer “invisibility” as their favorite superpower. Similarly, $\frac{27}{222}$, or approximately 0.122 or 12.2%, of males prefer “invisibility” as their favorite superpower.

Exercises 1–5

- Use the frequency counts from the table in Exploratory Challenge 1 to calculate the missing row of conditional relative frequencies. Round the answers to the nearest thousandth.

	To Fly	Freeze Time	Invisibility	Super Strength	Telepathy	Total
Females			$\frac{48}{228} \approx 0.211$			
Males	$\frac{51}{222} \approx 0.230$					$\frac{222}{222} = 1.000$
Total						

2. Suppose that a student is selected at random from those who completed the survey. What do you think is the gender of the student selected? What would you predict for this student's response to the superpower question?

3. Suppose that a student is selected at random from those who completed the survey. If the selected student is male, what do you think was his response to the selection of a favorite superpower? Explain your answer.

4. Suppose that a student is selected at random from those who completed the survey. If the selected student is female, what do you think was her response to the selection of a favorite superpower? Explain your answer.

5. What superpower was selected by approximately one-third of the females? What superpower was selected by approximately one-third of the males? How did you determine each answer from the conditional relative frequency table?

Exploratory Challenge 2: Possible Association Based on Conditional Relative Frequencies

Two categorical variables are associated if the row conditional relative frequencies (or column relative frequencies) are different for the rows (or columns) of the table. For example, if the selection of superpowers selected for females is different than the selection of superpowers for males, then gender and superpower favorites are associated. This difference indicates that knowing the gender of a person in the sample indicates something about their superpower preference.

The evidence of an association is strongest when the conditional relative frequencies are quite different. If the conditional relative frequencies are nearly equal for all categories, then there is probably not an association between variables.

Exercises 6–10

Examine the conditional relative frequencies in the two-way table of conditional relative frequencies you created in Exercise 1. Note that for each superpower, the conditional relative frequencies are different for females and males.

6. For what superpowers would you say that the conditional relative frequencies for females and males are very different?

7. For what superpowers are the conditional relative frequencies nearly equal for males and females?

8. Suppose a student is selected at random from the students who completed the survey. If you had to predict which superpower this student selected, would it be helpful to know the student's gender? Explain your answer.

9. Is there evidence of an association between gender and a favorite superpower? Explain why or why not.

10. What superpower would you recommend the students at Rufus King High School select for their superhero character? Justify your choice.

Exploratory Challenge 3: Association and Cause-and-Effect

Students were given the opportunity to prepare for a college placement test in mathematics by taking a review course. Not all students took advantage of this opportunity. The following results were obtained from a random sample of students who took the placement test.

	Placed in Math 200	Placed in Math 100	Placed in Math 50	Total
Took Review Course	40	13	7	60
Did Not Take Review Course	10	15	15	40
Total	50	28	22	100

Exercises 11–16

11. Construct a row conditional relative frequency table of the above data.

	Placed in Math 200	Placed in Math 100	Placed in Math 50	Total
Took Review Course				
Did Not Take Review Course				
Total				

12. Based on the conditional relative frequencies, is there evidence of an association between whether a student takes the review course and the math course in which the student was placed? Explain your answer.

13. Looking at the conditional relative frequencies, the proportion of students who placed into Math 200 is much higher for those who took the review course than for those who did not. One possible explanation is that taking the review course caused improvement in placement test scores. What is another possible explanation?

Now consider the following statistical study:

Fifty students were selected at random from students at a large middle school. Each of these students was classified according to sugar consumption (high or low) and exercise level (high or low). The resulting data are summarized in the following frequency table.

		Exercise Level		Total
		High	Low	
Sugar Consumption	High	14	18	32
	Low	14	4	18
Total		28	22	50

14. Calculate the row conditional relative frequencies, and display them in a row conditional relative frequency table.

		Exercise Level		Total
		High	Low	
Sugar Consumption	High			
	Low			
	Total			

15. Is there evidence of an association between sugar consumption category and exercise level? Support your answer using conditional relative frequencies.

16. Do you think it is reasonable to conclude that high sugar consumption is the cause of the observed differences in the conditional relative frequencies? What other explanations could explain a difference in the conditional relative frequencies? Explain your answer.

Lesson Summary

- A conditional relative frequency compares a frequency count to the marginal total that represents the *condition* of interest.
- The differences in conditional relative frequencies are used to assess whether or not there is an association between two categorical variables.
- The greater the differences in the conditional relative frequencies, the stronger the evidence that an association exists.
- An observed association between two variables does not necessarily mean that there is a cause-and-effect relationship between the two variables.

Problem Set

Consider again the summary of data from the 100 randomly selected students in the Rufus King High School investigation of after-school activities and gender.

	Intramural Basketball	Chess Club	Jazz Band	Not Involved	Total
Females	20	10	10	20	60
Males	20	2	8	10	40
Total	40	12	18	30	100

1. Construct a row conditional relative frequency table for this data. Decimal values are given to the nearest thousandth.

	Intramural Basketball	Chess Club	Jazz Band	Not Involved	Total
Females					60
Males					40
Total					

2. For what after-school activities do you think the row conditional relative frequencies for females and males are very different? What might explain why males or females select different activities?
3. If John, a male student at Rufus King High School, completed the after-school survey, what would you predict was his response? Explain your answer.

4. If Beth, a female student at Rufus King High School, completed the after-school survey, what would you predict was her response? Explain your answer.
5. Notice that 20 female students participate in intramural basketball and that 20 male students participate in intramural basketball. Is it accurate to say that females and males are equally involved in intramural basketball? Explain your answer.
6. Do you think there is an association between gender and choice of after-school program? Explain.

Column conditional relative frequencies can also be computed by dividing each frequency in a frequency table by the corresponding column total to create a column conditional relative frequency table. Column conditional relative frequencies indicate the proportions, or relative frequencies, based on the column totals.

7. If you wanted to know the relative frequency of females surveyed who participated in chess club, would you use a row conditional relative frequency or a column conditional relative frequency?
8. If you wanted to know the relative frequency of band members surveyed who were female, would you use a row conditional relative frequency or a column conditional relative frequency?
9. For the superpower survey data, write a question that would be answered using a row conditional relative frequency.
10. For the superpower survey data, write a question that would be answered using a column conditional relative frequency.