

Assignment

Write

Describe how you can determine whether there is an association between two-variable categorical data.

Remember

For categorical data in two variables, organizing and representing data in frequency distributions, marginal relative frequency distributions, conditional relative frequency distributions, and bar graphs are useful in formulating conclusions and using statistics to support your conclusions.

Practice

Men and women were surveyed to determine their favorite vehicle. The results are shown in the table.

		Favorite Vehicle			Total
		Sedan	SUV	Convertible	
Gender	Men	88	115	34	237
	Women	73	62	102	237

Suppose you were the general sales manager of a car dealership and you were planning a big car sale. Use the information in the frequency table to answer each question. Support your answers with tables and/or graphs.

1. Which type of car would you have on display in the showroom if you anticipate that most people shopping for a car would be women?
2. Which type of car would you have on display in the showroom if you anticipate that most people shopping for a car would be men?
3. Which type of car would you have on display in the showroom if you want to promote sales of cars that appeal to both men and women?

Stretch

In statistics, the frequencies in two-way tables can be used to help determine if the frequency counts are distributed identically across populations. Using probability theory, expected frequencies can be determined if the counts are distributed identically. These frequencies are calculated for each cell by multiplying the row total by the column total and dividing by the grand total of frequencies.

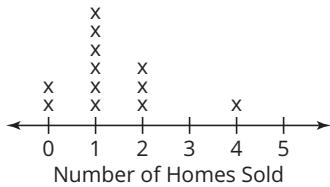
1. Complete the table of expected frequencies for gender and favorite vehicle.

Expected Frequencies	Sedan	SUV	Convertible	Row Total
Men	$\frac{(237)(161)}{474} \approx 80.5$			237
Women				237
Column Total	161	177	136	474

2. Does there seem to be a difference between the frequencies that were observed in the survey and those that would be expected if counts are distributed identically? What conclusion can you make?

Review

1. A realtor recorded the number of homes she sold each month for a year. Her numbers are shown in the dot plot.



- Describe the distribution of the dot plot.
- Calculate the mean and median. Explain what they mean in terms of the problem situation.
- Which measure of center do you think best represents these data? Explain your reasoning.

2. Two hundred residents of Grapeville City are asked in a survey where they get their prescription medicine filled. The residents are categorized by age group: young adult, adult, and senior. The results from the survey are shown in the table.

Prescription Source

		Drugstore	Grocery Store	Big Box Store	Online	Total
Age Group	Young Adult	2	4	15	44	65
	Adult	18	25	19	10	72
	Senior	26	22	10	5	63
	Total	46	51	44	59	200

- Construct a conditional relative frequency distribution of age group given source of prescription.
- A lot of young adults are currently moving to the city. Which prescription source should be most concerned? Explain your reasoning.