

Questions 1 to 5: The table below shows the opinions of 953 respondents in the General Social Survey to the question “Everything considered, would you say that in general, you approve or disapprove of wiretapping?” The purpose of examining the data is to see if there is a gender difference in how people would respond to this question. The data table is on the left and the results of a chi-square test are on the right.

| Gender | Wiretapping Opinion | | Total |
|--------------|---------------------|------------|------------|
| | Approve | Disapprove | |
| Male | 134 | 272 | 406 |
| Female | 126 | 421 | 547 |
| Total | 260 | 693 | 953 |

| Expected counts are printed below observed counts | | | |
|---------------------------------------------------|---------------|---------------|-------|
| | Approve | Disapprove | Total |
| Male | 134 110.77 | 272 295.23 | 406 |
| Female | 126 149.23 | 421 397.77 | 547 |
| Total | 260 | 693 | 953 |

Chi-Sq = 11.676, DF = 1, P-Value = 0.001

1. Calculate the row percentages for the “Approve” category.

$$\text{Males: } \frac{134}{406} = .33 \text{ or } 33\% \quad \text{Females: } \frac{126}{547} = .23 \text{ or } 23\%$$

2. Based on the row percentages, does it appear that there is a difference of opinion between males and females? Explain.

Yes. While 33% of the males in the sample approved of wiretapping, only 23% of the females approved.

3. Write the null hypothesis for this situation.

There is no relationship between gender and wiretapping opinion in the population represented by this sample.

4. Show how the Expected Count in the "Male, Approve" category was calculated.

$$\text{Expected count} = \frac{(406)(260)}{953} = 110.77$$

5. Make a conclusion about the relationship based on the computer output. Be sure to state your conclusion in the context of the situation, and justify it with statistical results.

The p-value is 0.001, which is much less than 0.05. Therefore, we conclude that there is a statistically significant relationship between gender and opinion about wiretapping. You could also say that we conclude that gender and wiretapping are related in the population represented by this sample, or that opinion about wiretapping depends on gender.

6. Suppose we wanted to test the null hypothesis that taking aspirin or a placebo would be equally effective for preventing heart attacks for the population of men similar to those in Case Study 1.6. One way to express this null hypothesis is in terms of the relative risk of heart attack. Which of the following would be the correct null hypothesis in that case?
- The relative risk in the population is 0.
 - The relative risk in the population is not 0.
 - The relative risk in the population is 1.**
 - The relative risk in the population is not 1.
7. [From CyberStats Unit C10] A chi-square test for independence in a 2x2 table based on 100 observations resulted in a p-value of 0.04. If the sample size were to be increased to 1,000 observations, and the proportions in each cell remained the same, it is likely that:
- There is no way to determine whether the p-value would increase or decrease.
 - The p-value would be much smaller.**
 - The p-value would remain about the same.
 - The p-value would be much larger.
8. An article in the *Sacramento Bee* (May 8, 2000, p A12) reported that "Early results from a study of more than 2000 pregnant women in Alabama show that pregnant mothers with severe gum disease are up to eight times more likely to have underweight premature babies than moms with healthy mouths." Suppose the results of this study showed that the relationship between severe gum disease and having an underweight premature baby was statistically significant. From this, can we conclude that severe gum disease causes women to sometimes have underweight premature babies? Explain your answer.
- No, we cannot conclude that there is a cause-and-effect relationship because this was clearly an observational study. The pregnant mothers could not have been randomly assigned to have severe gum disease or a healthy mouth. The researchers simply observed which was true for each woman.*
9. Which one of the following is *not* a situation in which one or more confounding variables can cause the results to be misleading?
- Simpson's Paradox
 - Testing hypotheses about the relationship between two categorical variables using data from an observational study
 - Not reporting the baseline risk when reporting relative risk**
 - A missing third variable is not included in the analysis, but would change the results if included
10. A study on the use of seat belts versus booster seats for children ages 4 and 5 reported that "Using seat belts instead of booster seats was associated with increased risk for serious injury in an accident; the relative risk was 2.4." Based on this, it can be concluded that for this study:
- Children ages 4 and 5 in a booster seat were 2.4 times more likely to have serious injuries in an accident than were children wearing seatbelts.
 - Children ages 4 and 5 wearing seatbelts were 2.4 times more likely to have serious injuries in an accident than were children in a booster seat.**
 - The percent of children ages 4 and 5 in a booster seat was 2.4 times higher than the percent of children wearing seatbelts.
 - The percent of children ages 4 and 5 wearing seatbelts was 2.4 times higher than the percent of children in a booster seat.