

AP Statistics

Chapter 11: Inference for Distribution of Categorical Data

Day 5

HW: Finish Lesson 5 Practice Worksheet

compared random samples of 2600 U.S. and 400 Canadian heart attack patients. One key outcome was the patients' own assessment of their quality of life relative to what it had been before the heart attack. Here are the data for the patients who survived a year.

Quality of Life	Canada	United States
Much better	75	541
Somewhat better	71	498
About the same	96	779
Somewhat worse	59	282
Much worse	19	65
Total	311	2165

• Is there a significant difference between the two distributions of quality-of-life ratings? Carry out an appropriate test at the $\alpha = 0.01$ level.

H_0 : There is no difference in the distribution of quality of life in Canada and the United States.

H_a : There is a difference in the distribution of quality of life in Canada and the United States.

Conditions

- As stated, this is an SRS.
- All expected counts are at least 5.
- We can assume that there are at least 21,650 patients in the US and 3110 patients in Canada that have had heart attacks.

$$\chi^2 = \frac{(75-77.373)^2}{77.373} + \frac{(541-538.63)^2}{538.63} + \dots + \frac{(65-73.449)^2}{73.449} = 11.725$$

$$df = (5-1)(2-1) = 4 \quad p = .0195$$

Using my calculator, I performed a χ^2 -test. Since the p-value is .0195 and is greater than the significance level of $\alpha = .01$, we fail to reject H_0 . There is not enough evidence to conclude that there is a difference in the distribution of quality of life in Canada and the United States.

* Don't forget conditions

Sample surveys on sensitive issues can give different results depending on how the question is asked. A University of Wisconsin study divided 2400 respondents into 3 groups at random. All were asked if they had ever used cocaine. One group of 800 was interviewed by phone, 21% said they had used cocaine. Another 800 people were asked the same question in a one-on-one personal interview, 25% said "Yes." The remaining 800 were allowed to make an anonymous written response; 28% said "Yes." Are the differences among the three groups statistically significant? Give appropriate evidence to support your answer.

H_0 : There is no difference in the actual proportion who answer Yes based on the contact method.

H_a : There is a difference in the actual proportion who answer Yes based on the contact method.

	Observed		Expected	
	Yes	No	Yes	No
Phone	168	632	197.33	602.67
Personal	200	600	197.33	602.67
Written	224	576	197.33	602.67

$$\chi^2 = \frac{(168-197.33)^2}{197.33} + \frac{(632-602.67)^2}{602.67} + \dots + \frac{(576-602.67)^2}{602.67}$$

$$df = (3-1)(2-1) = 2 \quad p\text{-value} = .00494$$

Using my calculator, I performed a χ^2 -Test. Since the p-value is .00494 and is less than the significance level, we reject H_0 . There is sufficient evidence to conclude that there is a difference in the actual proportion who answer Yes based on the contact method.

The Chi-Square Test for Association/ Independence

The hypotheses for a chi-square test for association/independence are as follows:

H_0 : There is no association in the distribution of two categorical variables for a population.

H_a : There is an association in the distribution of two categorical variables for a population.

Do Angry People Have More Heart Attacks

A study followed a random sample of 8474 people with normal blood pressure for about four years. All the individuals were free of heart disease at the beginning of the study. Each person took the Spielberger Trait Anger Scale test, which measures how prone a person is to sudden anger. Researchers also recorded whether each individual developed coronary heart disease (CHD). This includes people who had heart attacks and those who needed medical treatment for heart disease. Do angry people have more heart disease?

	Low anger	Moderate anger	High anger	Total
CHD	53	110	27	190
No CHD	3057	4621	606	8284
Total	3110	4731	633	8474

an appropriate test to help answer this question.

	Low anger	Moderate anger	High anger	Total
CHD	53 16.9731	110 106.08	27 14.93	190
No CHD	3057 3046.3	4621 4621.9	606 618.81	8284
Total	3110	4731	633	8474

H_0 : There is no association between anger level and heart disease.
 H_a : " " is an association " "
 $\chi^2 = \frac{(53 - 69.731)^2}{69.731} + \frac{(110 - 106.08)^2}{106.08} + \dots + \frac{(606 - 618.81)^2}{618.81} = 16.07$
 $df = (2-1)(3-1) = 2$ p-value = .000323
 Using my calculator, I performed a χ^2 -test. Since the p-value is .000323 and is less than the sign. level of $\alpha = .05$ we reject H_0 . There is sufficient evidence to conclude that there is an association between anger level and heart disease.