Work Problems Chapter 14

Suppose I want to know whether there are differences in the likelihood of being diagnosed with depression for people who live in different types of communities (urban, suburban, rural). I collect data from people from these three different types of communities and get the data summarized in Table 14.11. Use this data to answer the following set of questions.

Table 14.11. Raw data of depressed and not depressed people living in urban, rural, and suburban areas.

	Urban	*Rural*	*Suburban*	*Row Total*
Depressed	120	90	100	310
Not Depressed	600	300	400	1300
Column Total	720	390	500	1610

1. Calculate the expected values for each of the 6 cells in the table.

	Urban	*Rural*	*Suburban*
Depressed	$\frac{(310)(720)}{1610} = 138.63$	$\frac{(310)(390)}{1610} = 75.09$	$\frac{(310)(500)}{1610} = 96.27$
Not Depressed	$\frac{(1300)(720)}{1610} = 581.37$	$\frac{(1300)(390)}{1610} = 314.91$	$\frac{(1300)(500)}{1610} = 403.73$

2. Calculate the sum of squared differences between the observed and expected values to find the observed chi-square value.

	Urban	*Rural*	*Suburban*

Depressed	$\frac{(120-138.63)^2}{138.63}=2.50$	$\frac{(90-75.09)^2}{75.09}=2.96$	$\frac{(100-96.27)^2}{96.27}=.14$
Not Depressed	$\frac{(600-581.37)^2}{581.37}=.60$	$\frac{(300-314.91)^2}{314.91}=.71$	$\frac{(400-403.73)^2}{403.73}=.03$

2.50 + .60 + 2.96 + .71 + .14 + .03 = 6.94. This is the chi-square value (χ^2 = 6.94.)

3. Report the degrees of freedom (*df*) for this problem.

   ***R* = 2 and *C* = 3, so *df* = (2 – 1)(3 – 1) = 2.**

4. Using the *df* you just calculated, and an alpha level of .05, find the critical value for the chi-square statistic in Appendix E.

   The critical χ^2 = 5.99 with *df* = 2 and alpha level of .05.

5. Compare the critical value from Appendix E with the observed chi-square value that you calculated in question #2 and decide whether your observed value is statistically significant.

   The observed χ^2 = 6.94 and the critical χ^2 = 5.99. Because the observed value is larger than the critical value, our chi-square statistic is statistically significant.

6. What does the chi-square statistic that you calculated tell you? What doesn't it tell you?

   Because our observed chi-square statistic is statistically significant, we know that some of our observed frequencies differ from the expected frequencies. In suburban areas, it appears that the proportion of depressed and non-depressed people is about what would be expected by chance. In urban areas, it appears that there are more depressed and fewer non-depressed people than we would expect by chance. This pattern is reversed in the rural areas.

Here are two more questions that are not based on the data presented above:

7. Explain when you would use a non-parametric test rather than a parametric test.

   Non-parametric tests are better than parametric tests when the data do not form a normal distribution and, sometimes, when the scales of measurement on the variables are nominal and/or ordinal rather than interval/ratio.

8. Suppose that in a large company, there is an allegation of gender bias in who

receives promotions and who does not. Explain how the chi-square test of independence compares observed and expected frequencies to determine whether this allegation is true.

In this example, there would be four groups, or cells of a table: Women who received promotions, women who did not, men who received promotions, and men who did not. Using the actual data, we would first determine the observed frequencies for each of these four groups. Then, using the total number of each gender and the total number of promoted vs. non-promoted, we could calculate the expected frequencies for each cell. For example, if half of all employees were men and half of all employees received a promotion, we would expect, by chance alone, that one half of all women received promotions and the other half did not. By comparing the observed frequencies with the expected frequencies, and using the differences between the two to calculate a chi-square statistic, we can determine whether one gender is more or less likely than chance to have received a promotion.