Lab Activity: Chi-Square Test of Independence

In this lab activity, you will conduct the chi-square tests of independence to determine whether two factors are independent.

Student Learning Outcomes

By the end of this chapter, you should be able to do the following:

- Perform a chi-square test of independence using Statcato

Preliminary

Read Chapter 11 The Chi-Square Distribution in:

Illowsky, Barbara, and Susan Dean. <u>Collaborative Statistics</u>. Connexions. 2 Mar. 2010 <http://cnx.org/content/col10522/1.37/>.

Background

This lab uses datasets from the Summary Health Statistics for U.S. Adults: National Health Interview Survey, 2008 (http://www.cdc.gov/nchs/data/series/sr_10/sr10_242.pdf). The datasets contains the frequency distributions of the degree of sadness among persons 18 years of age or over by selected characteristics in the year of 2008. You will explore whether the degree of sadness is dependent of gender and race.

Loading Data

The data files are available on the Statcato web site. You will first load the data file containing the contingency table of frequencies of doctor visits versus gender.

Loading an Online Dataset

- Go to `File > Load Dataset`. Or click the icon in the toolbar.

- Under Online Datasets, enter the web address: http://www.statcato.org/labs/data/sadness.xls. Select "Excel" in the dataset file type drop-down menu.

- Click **Load Dataset**.

The data should now be in Data window. The contingency table of degree of sadness versus gender is in columns C1 to C3. The contingency tables of sadness vs. gender for three different ethnicities (Hispanic, White, and Black) are in columns C5 – C7, C9 – C11, and C13 – C15, respectively.

Formulating the Hypothesis Test

In this test, you want to determine if degree of sadness is dependent of gender. Answer the following questions in **LR: Hypotheses**.

- State the claim that you are testing.

- State the null and alternative hypotheses.
    - H_0:
    - H_a:
- Is this a right-tailed, left-tailed, or two-tailed test?

Performing the Test of Independence

Using Statcato, you will perform calculations for the chi-square test of independence using a significance level of 0.01 ($\alpha = 0.01$).

⁘ᴛᴀᴛᴄᴀᴛᴏ **Performing Chi-Square Test of Independence**

Go to `Statistics > Multinomial Experiments > Chi-Square Contingency Table`.

- In the **Inputs** panel, you should see a list of column names under **Select a column of the table**. Select **C2 All or most of the time** and **C3 Some of the time** in the list (click **C2**, press and hold **Ctrl**, and then click **C3**), and then click the **Add to list** button. You should now see **C2** and **C3** in the **Columns** list.
- For **Significance Level**, enter **0.01**.
- Click **OK**.

Copy the computation results in the Log window to **LR: Test of Independence**.

Making Conclusions

Based on the computer-generated results, draw conclusions for the test of independence. Record your answers in **LR: Interpretation**.

Decision on Null Hypothesis

Based on the significance level α and the computed p-value, decide whether to reject H_0 and explain why. Alternatively, you may use the χ^2 test statistic and critical value.

Conclusion

Based on your decision on the null hypothesis, make a conclusion about your claim.

Discussion

Answer the following questions in **LR: Discussion**.

1. Using Statcato, determine if degree of sadness is dependent of gender in Hispanics, Whites, and Blacks. Use a 0.01 significance level.

2. Does degree of sadness appear to be dependent of gender in some ethnicities but not others? What do you think are the reasons behind such differences? Give a possible explanation.