UNIVERSITY OF
SAN FRANCISCO

CHANGE THE WORLD FROM HERE

Tutorial 26: Chi Square Tests - Goodness of Fit

Description

In this tutorial we will learn how to carry out tests using the Chi Square statistic and distribution. As a reminder the Chi Square distribution is the distribution that is used to describe data drawn from population variables that are not normally distributed. These variables tend to be categorical (nominal or ordinal/text) as opposed to being continuous (interval or ratio). This tutorial will focus on a Chi Square test of Goodness of Fit. Typically, this test examines the observed frequencies/proportions of categories for a single variable and measures them against an expected or predicted distribution of frequencies/proportions.

So it is possible to test two sets of hypotheses:
H_0: All population proportions are equal
H_1: Not H_0
Or
H_0: There are no differences in a specified population from the proportions in another known population.
H_1: Not H_0

For this tutorial we will focus on the variable **parentedu** which asks participants to state the educational level of their parents as a proxy for socioeconomic status. Participants could select one of 4 parental education categories:
1) Both parents have college level degree or higher
2) One parent has college level degree or higher
3) Neither parent has college level degree or higher
4) unknown

a) We will first test the hypothesis that the proportion of the four categories are equal and there is no difference in the proportions of education level.
b) Second we will test the hypothesis that the proportions of the four categories follow an expected distribution or model for example we will test whether the proportions in the dataset (observed proportions) are similar to the following proportions:
    i)   Both parents have college level degree or higher = 40%
    ii)   One parent has college level degree or higher = 25%
   iii) Neither parent has college level degree or higher = 25%
    iv)   unknown = 10%

Note: we are using the datafile ARMF2020_wave1andwave2.omv

Content
1. **Conduct a Chi Square Goodness of Fit Test - All proportions are equal hypothesis**
2. **Interpret Output for Chi Square Goodness of Fit Test - All proportions are equal hypothesis**
3. **Conduct a Chi Square Goodness of Fit Test - proportions are similar to a expected distribution of proportions**
4. **Interpret Output for Chi Square Goodness of Fit Test - proportions are similar to a expected distribution of proportions**
5. **APA Format describing the findings**

STEPS

1. **Conduct a Chi Square Goodness of Fit Test - All proportions are equal hypothesis**
   a. Go to the 'Analyses' tab.
   b. Click on 'Frequencies'.
   c. Select 'N Outcomes/ χ^2 Goodness of fit':
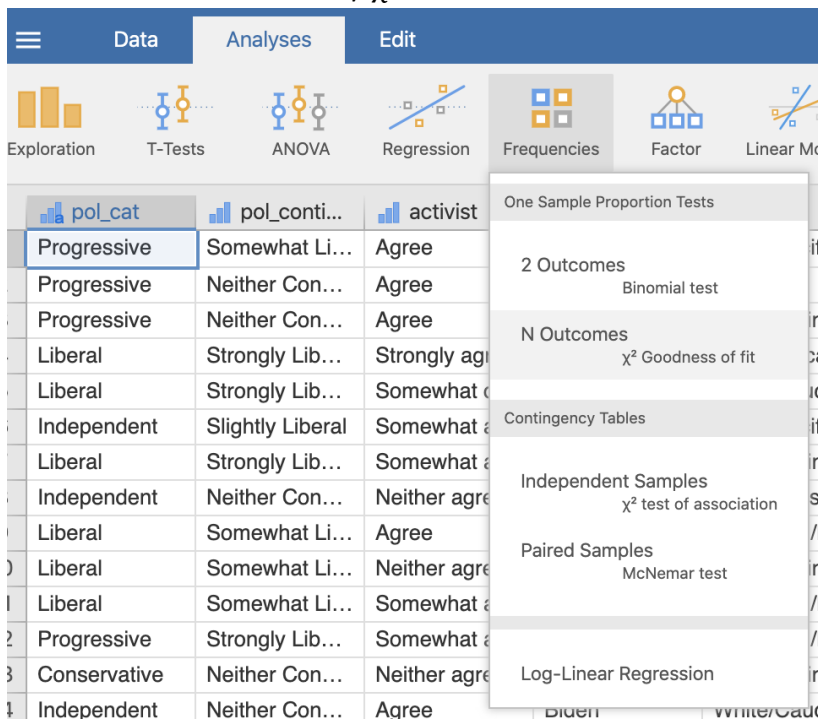
   d. Move the variable of interest **parentedu** into the variable slot on the right.

Proportion Test (N Outcomes)

e. The default is to check that all proportions are equal. You may check this default by clicking on the arrow next to 'Expected Proportions' and examine the proportions column.

Level	Ratio	Proportion
both parents have college level degree or higher	1	0.250
one parent has college level degree or higher	1	0.250
neither parent has college level degree or higher	1	0.250
unknown	1	0.250

2. **Interpret Output for Chi Square Goodness of Fit Test - All proportions are equal hypothesis**

Proportion Test (N Outcomes)

Proportions – parentedu

Level	Count	Proportion
both parents have college level degree or higher	75	0.4167
one parent has college level degree or higher	51	0.2833
neither parent has college level degree or higher	50	0.2778
unknown	4	0.0222

χ^2 Goodness of Fit

χ^2	df	p
58.7	3	<.001

    a. Proportions - parentedu table: provides a description of the observed data in the sample
1. Count column: Provides the exact count of participants who selected each category
2. Proportion column: Provides the proportion of participants who selected each category

  b. χ^2 Goodness of fit table: Provides a test of the hypotheses that all proportions reported in the 'Proportions' table are equal.
    i. χ^2 column: provides the χ^2 test statistic
    ii. df column: provides the df of freedom for the test
    iii. p column: provides the p-value for the test and the p-value is less than .05 (alpha level) which indicates that the test is statistically significant and the hypotheses that all proportions are equal is rejected. Thus the proportions are statistically significantly not equal to one another.

3. **Conduct a Chi Square Goodness of Fit Test - proportions are similar to a expected distribution of proportions**

Recall the expected distribution that we are testing is as follows:
    i. Both parents have college level degree or higher = 40%
    ii. One parent has college level degree or higher = 25%
    iii. Neither parent has college level degree or higher = 25%
    iv. unknown = 10%
  a. Follow steps a to d in step 1
  b. Click on the arrow next to 'Expected Proportions' and change the entries under the 'Ratio' column to support the distribution that we want to test:

Expected Proportions		
Level	**Ratio**	**Proportion**
both parents have college level degree or higher	0.4	0.400
one parent has college level degree or higher	0.25	0.250
neither parent has college level degree or higher	0.25	0.250
unknown	0.1	0.100

4. **Interpret Output for Chi Square Goodness of Fit Test - proportions are similar to a expected distribution of proportions**

Proportion Test (N Outcomes)

Proportions – parentedu

Level	Count	Proportion
both parents have college level degree or higher	75	0.4167
one parent has college level degree or higher	51	0.2833
neither parent has college level degree or higher	50	0.2778
unknown	4	0.0222

χ^2 Goodness of Fit

χ^2	df	p
12.4	3	0.006

  a. Proportions - parentedu table: provides a description of the observed data in the sample
       i. Count column: Provides the exact count of participants who selected each category
       ii. Proportion column: Provides the proportion of participants who selected each category
  b. χ^2 Goodness of fit table: Provides a test of the hypotheses that the observed proportions in the table match the expected proportions entered (.40, .25, .25, .10)
       iv. χ^2 column: provides the χ^2 test statistic
       v. df column: provides the df of freedom for the test
       vi. p column: provides the p-value for the test and the p-value is less than .05 (alpha level) which indicates that the test is statistically significant and the hypothesis that proportions follow the expected distribution (.40, .25, .25, .10) is rejected. Thus the observed proportions are statistically

significantly not equal to those expected proportions. Although this test is not definitively pointing to the source of significance - we can see that the proportion of participants reporting they do not know their parents' income was less than expected (expected was .10 and actual result was .02).
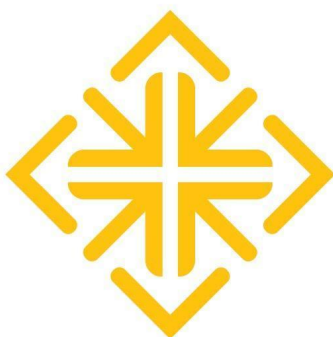
c. APA Format describing the findings

a) For the equal proportions hypothesis: Categories of parental education were shown to be statistically significantly different from one another χ^2(3, n = 180) = 58.7, $p < .05$.
b) For the expected distribution hypothesis: Categories of parental education were shown to be statistically significantly different from the expected proportion χ^2(3, n = 180) = 12.4, $p < .05$.

--END TUTORIAL--

This Jamovi tutorial is a companion to a video tutorial and these materials were developed by:

Aline Hitti, Ph.D., University of San Francisco

Saera Khan, Ph.D., University of San Francisco

Sage Stefanick, University of San Francisco

Erica Divinagracia, University of San Francisco

Adolfo Barrales, University of San Francisco

UNIVERSITY OF SAN FRANCISCO

CHANGE THE WORLD FROM HERE