Chapter 900

Chi-Square Effect Size Estimator

Introduction

This procedure calculates the effect size of the Chi-square test for use in power and sample size calculations. Based on your input, the procedure provides effect size estimates for Chi-square goodness-of-fit tests and for Chi-square tests of independence.

The *Chi-square test* is often used to test whether sets of frequencies or proportions follow certain patterns. The two most common cases are in tests of goodness of fit and tests of independence in contingency tables.

The *Chi-square goodness-of-fit* test is used to test whether a set of data follows a particular distribution. For example, you might want to test whether a set of data comes from the normal distribution.

The *Chi-square test for independence* in a contingency table is another common application of this test. Here individuals (people, animals, or things) are classified by two (nominal or ordinal) classification variables into a two-way contingency table. This table contains the counts of the number of individuals in each combination of the row categories and column categories. The Chi-square test determines if there is dependence (association) between the two classification variables.

Effect Size

For each cell of a table containing m cells, there are two proportions considered: one specified by a null hypothesis and the other specified by the alternative hypothesis. Usually, the proportions specified by the alternative hypothesis are those occurring in the data. Define p_{0i} to be the proportion in cell i given by the null hypothesis and p_{1i} to be the proportion in cell i according to the alternative hypothesis. The effect size, W, is calculated using the following formula

$$W = \sqrt{\sum_{i=1}^{m} \frac{(p_{0i} - p_{1i})^2}{p_{0i}}}.$$

The formula for computing the Chi-square value, χ^2, is

$$\chi^2 = \sum_{i=1}^{m} \frac{(O_i - E_i)^2}{E_i} = N \sum_{i=1}^{m} \frac{(p_{0i} - p_{1i})^2}{p_{0i}},$$

where N is the total count in all the cells. Hence, the relationship between W and χ^2 is

$$\chi^2 = NW^2$$

or

$$W = \sqrt{\frac{\chi^2}{N}}$$

Contingency Table Tab

This window allows you to enter up to an eight-by-eight contingency table. You can enter percentages or counts. If you enter counts, the Chi-Square and Prob Level values are correct and may be used to test the independence of the row and column variables. If you enter percentages, you should ignore the Chi-Square and Prob Level values.

Note that if you are entering percentages, it does not matter whether you enter table percentages or row (or column) percentages as long as you are consistent.

Example

Suppose you are planning a survey with the primary purpose of testing whether marital status is related to gender. You decide to adopt four marital status categories: never married, married, divorced, widowed. In the population you are studying, previous studies have found the following percentages in each of these categories:

Never Married	27%
Married	39%
Divorced	23%
Widowed	11%

You decide that you want to calculate the effect size when the individual percentages for males and females are

Gender	Male	Female
Never Married	22%	32%
Married	46%	33%
Divorced	22%	24%
Widowed	10%	11%

To complete this example, you would load the Chi-Square Effect Size Estimator procedure from the PASS-Other menu and enter "22 46 22 10" across the top row and "32 33 24 11" across the next row. The value of W turns out to be 0.143626.

Note that even though a Chi-square value (4.13) and probability level (0.248) are displayed, you would ignore them since you have entered percentages, not counts, into the table. If you had entered counts, these results could be used to test the hypothesis of independence.

Multinomial Test Tab

This window allows you to enter a multinomial table with up to fourteen cells. You can enter percentages or counts. If you enter counts, the Chi-Square and Prob Level values are correct and may be used to test the statistical significance of the table. If you enter percentages, you should ignore the Chi-Square and Prob Level values.

Note that if you are using the window to perform a goodness-of-fit test on a set of data, you will need to adjust the degrees of freedom for the number of parameters you are estimating. For example, if you are testing whether the data are distributed normally and you estimate the mean and standard deviation from the data, you will need to reduce the degrees of freedom by two.

Example

Suppose you are going to use the Chi-square goodness-of-fit statistic calculated from a multinomial table to test whether a set of exponential data follow the normal distribution. That is, you want to find a reasonable effect size for comparing exponentially distributed data to the normal distribution.

You decide to divide the data into five groups: 5 or less, 5-10, 10-15, 15-20, 20+

Using tables for the normal and exponential distributions, you find that the probabilities for each group are

Category	Normal	Exponential
5 or Less	11%	39%
5 to 10	20%	26%
10 to 15	38%	18%
15 to 20	20%	11%
Above 20	11%	6%

To complete this example, you would select the Multinomial Test tab and enter "39 26 18 11 6" in the Data Values column and "11 20 38 20 11" in the Hypothesized Proportions column. The calculated value of W is 0.948271. You would enter this value into the Effect Size option of the Chi-Square Test window in **PASS** to determine the necessary sample size.