

Proportion Differences

A most common research question arises when an investigator has obtained two sample proportions. One asks whether or not the two sample proportions are really different considering that they are based on observations drawn randomly from a population. For example, a school nurse observes during the flu season that 13 eighth grade students are absent due to flu symptoms while only 8 of the ninth grade students are absent. The class sizes of the two grades are 110 and 121 respectively. The nurse decides to test the hypothesis that the two proportions (.118 and .066) do not differ significantly using the LazStats program. The first step is to start the Proportion Differences procedure by clicking on the Statistics menu, moving the mouse to the Comparisons option and the clicking on the Proportion Differences option. The specification form for the test then appears. We will enter the required values directly on the form and assume the samples are independent random samples from a population of eighth and ninth grade students.

Test of Equality for two Proportions

Data Entry By:

- ☒ Values Entered on this Form
- ☐ Values in the Data Grid

Test Assumptions:

- ☒ Independent Proportions
- ☐ Dependent Proportions

Sample 1 Freq. Sample Size:

Sample 2 Freq. Sample Size:

Percent Confidence Interval:

When the nurse clicks the Continue button the following results are shown in the Output form:

COMPARISON OF TWO PROPORTIONS

Test for Difference Between Two Independent Proportions

Entered Values

Sample 1: Frequency = 13 for 110 cases.

Sample 2: Frequency = 8 for 121 cases.

Proportion 1 = 0.118, Proportion 2 = 0.066, Difference = 0.052
 Standard Error of Difference = 0.038
 Confidence Level selected = 95.0
 z test statistic = 1.375 with probability = 0.0846
 z value for confidence interval = 1.960
 Confidence Interval: (-0.022, 0.126)

The nurse notices that the value of zero is within the 95% confidence interval as therefore accepts the null hypothesis that the two proportions are not different than that expected due to random sampling variability. What would the nurse conclude had the 80.0% confidence level been chosen?

If the nurse had created a data file with the above data entered into the grid such as:

CASE/VAR	FLU	GROUP
CASE 1	0	1
CASE 2	1	1
--		
CASE 110	0	1
CASE 111	0	2

CASE 231	1	2

then the option would have been to analyze data in a file.

In this case, the absence or presence of flu symptoms for the student are entered as zero (0) or one (1) and the grade is coded as 1 or 2. If the same students, say the eighth grade students, are observed at weeks 10 and 15 during the semester, then the test assumptions would be changed to Dependent Proportions. In that case the form changes again to accommodate two variables coded zero and one to reflect the observations for each student at weeks 10 and 15.

Test of Equality for Two Proportions

Data Entry By:
☐ Values Entered on This Form
☒ Values Computed from the Data Grid

Test Assumptions:
☒ Independent Proportions
☐ Dependent Proportions

Select Variables:
 Flu
 Group

First Variable:
 Flu

Group Code:
 Group

Directions: For independent groups you should have a variable (e.g. group) indicating group membership and a variable (e.g. graduated) that consists of 0's and 1's which represent not observed or observed. Use 1 and 2 for the group coding under the group variable.
 For dependent samples you should have two variables each of which contains codes of 1 or 0 for each case which

Percent Confidence Interval ? 80.0

Reset Cancel Continue