ICCPP-STATISTICS

- Test for Homogeneity

Vishal Lohchab

Scientific Assistant of
Prof. Dr. Hans-Werner Gessmann
Director ICCPP International

Karl Pearson (1900)

Test for Homogeneity

Definition

- A test of homogeneity compares the proportions of responses from two or more populations with regards to a dichotomous variable (e.g. male/female, yes/no) or variable with more than two outcome categories .

Formula

$$\sum_{i \cdot j} \frac{(O-E)^2}{2}$$

Homogeneity test statistic where

- O = observed values
- E = expected values
- i = number of rows in data contingency table
- j = number of columns in data contingency table
- df = (i −1)(j −1) degrees of freedom

Test for Homogeneity

- The chi-square test of homogeneity is the nonparametric test used in a situation where the dependent variable is categorical.
- Data can be presented using a contingency table in which populations and categories of the variable are the row and column labels.

Test for Homogeneity

- The null hypothesis states that all populations are homogeneous regarding the proportions of categories of categorical variable.

  If the null hypothesis is rejected, it is concluded that the above proportions are different in the observed populations.

Use

- Test for homogeneity, can be used to draw a conclusion about whether two populations have the same distribution.

  To calculate the test statistic for a test for homogeneity, follow the same procedure as with the test of independence.

Hypotheses

- H_0: The distributions of the two populations are the same.

- H_a: The distributions of the two populations are not the same.

Note

The expected value for each cell needs to be at least five in order for you to use this test.

- Degrees of Freedom (df)
  df = number of columns – 1

- Requirements

  All values in the table must be greater than or equal to five.

Example

Both before and after a recent earthquake, surveys were conducted asking voters which of the three candidates they planned on voting for in the upcoming city council election.

Has there been a change since the earthquake? Use a level of significance of 0.05.

The table below shows the results of the survey. Has there been a change in the distribution of voter preferences since the earthquake?

Example

	Perez	Chung	Stevens
Before	167	128	135
After	214	197	225

Solution

Step 1

H_0: The distribution of voter preferences was the same before and after the earthquake.

H_a: The distribution of voter preferences was not the same before and after the earthquake.

Solution

Step 2

Degrees of Freedom (df):

df = number of columns – 1 = 3 – 1 = 2

Solution

Step 3

Distribution for the test: χ^2_2

Solution

Step 4

Calculate the test statistic:

$\chi 2 = 3.2603$ (calculator or computer)

Solution

Step 5

Probability statement:

p-value = $P(\chi^2 > 3.2603) = 0.1959$

Press the MATRX key and arrow over to EDIT.

Press 1:[A]. Press 2 ENTER 3 ENTER.

Solution

Step 5.1

Enter the table values by row.

Press ENTER after each. Press 2nd QUIT.

Press STAT and arrow over to TESTS.

Arrow down to C: χ^2-TEST.

Solution

Step 5.2

Press ENTER. You should see
Observed: [A] and Expected: [B].

Arrow down to Calculate. Press ENTER.
The test statistic is 3.2603 and the p-value = 0.1959.

Do the procedure a second time but arrow down to Draw instead of calculate.

Solution

Step 6

Compare α and the p-value:

$\alpha = 0.05$ and the p-value = 0.1959

$\alpha <$ p-value

Solution

Step 7

Make a decision:

Since α < p-value, do not reject H_o.

Solution

Step 8

Conclusion

At a 5% level of significance, from the data, there is insufficient evidence to conclude that the distribution of voter preferences was not the same before and after the earthquake.

References

Test of Homogeneity, Chi-Square (2008). In: Kirch W. (eds) Encyclopedia of Public Health. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-5614-7_3475 (accessed Sept. 15, 2021, 16:30 pm)

Data from the Insurance Institute for Highway Safety, 2013. Available online at www.iihs.org/iihs/ratings (accessed May 24, 2013, 11:00 am).

"Parent and Family Involvement Survey of 2007 National Household Education Survey Program (NHES)," U.S. Department of Education, National Center for Education Statistics. Available online at http://nces.ed.gov/pubs2009/2009030_sup.pdf (accessed May 24, 2013, 11:30 am).

https://courses.lumenlearning.com/introstats1/chapter/test-for-homogeneity/ (accessed Sept. 15, 2021, 16:00 pm)