The Chi-Square Distribution: Test for Homogeneity*

Susan Dean

This work is produced by The Connexions Project and licensed under the
Creative Commons Attribution License †

Abstract

The Test for Homogeneity is used to make 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.

The Goodness of Fit test can be used to decide whether a population fits a given distribution, but the Goodness of Fit test will not suffice to compare whether two populations follow the same unknown distribution. A different test, called the Test for Homogeneity, can be used to make 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.

NOTE: The expected value for each cell needs to be at least 5 in order to use this test.

Hypotheses
H_o: The distributions of the two populations are the same.

H_a: The distributions of the two populations are not the same.
Test Statistic
Use a χ^2 test statistic. It is computed in the same way as the test for independence.
Degrees of Freedom (df)
df = number of columns - 1
Requirements
All values in the table must be greater than or equal to 5.
Common Uses
Comparing two populations. For example: men versus women, before vs. after, east vs. west. The variable is categorical with more than two possible response values.

Example 1
Do male and female college students have the same distribution of living conditions? Use a level of significance of 0.05. Suppose that 250 randomly selected male college students and 300 randomly selected female college students were asked about their living conditions: Dormitory, Apartment, With Parents, Other. The results are shown in the table below.

*Version 1.2: Jun 29, 2012 9:08 am -0500
†http://creativecommons.org/licenses/by/3.0/

Distribution of Living Conditions for College Males and College Females

	Dormitory	Apartment	With Parents	Other
Males	72	84	49	45
Females	91	86	88	35

Table 1

Problem
Do male and female college students have the same distribution of living conditions?

Solution
H_o: The distribution of living conditions for male college students is the same as the distribution of living conditions for female college students.

H_a: The distribution of living conditions for male college students is not the same as the distribution of living conditions for female college students.

Degrees of Freedom (df):
df = number of columns - 1 = 4 - 1 = 3

Distribution for the test:χ^2_3

Calculate the test statistic:$\chi^2 = 10.1287$ (calculator or computer)

Probability statement:p-value $= P\left(\chi^2 > 10.1287\right) = 0.0175$

TI-83+ and TI-84 calculator: Press the `MATRX` key and arrow over to `EDIT`. Press `1:[A]`. Press `2 ENTER 4 ENTER`. 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`. Press `ENTER`. You should see `Observed:[A] and Expected:[B]`. Arrow down to `Calculate`. Press `ENTER`. The test statistic is 10.1287 and the p-value = 0.0175. Do the procedure a second time but arrow down to `Draw` instead of `calculate`.

Compare α and the p-value: Since no α is given, assume $\alpha = 0.05$. p-value = 0.0175. $\alpha >$ p-value.

Make a decision: Since $\alpha >$ p-value, reject H_o. This means that the distributions are not the same.

Conclusion: At a 5% level of significance, from the data, there is sufficient evidence to conclude that the distributions of living conditions for male and female college students are not the same.

Notice that the conclusion is only that the distributions are not the same. We cannot use the Test for Homogeneity to make any conclusions about how they differ.

Example 2
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.

	Perez	**Chung**	**Stevens**
Before	167	128	135
After	214	197	225

Table 2

Problem
Has there been a change in the distribution of voter preferences since the earthquake?

Solution
H_o: 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.

Degrees of Freedom (df):
df = number of columns - 1 = 3 - 1 = 2

Distribution for the test:χ_2^2

Calculate the test statistic:$\chi^2 = 3.2603$ (calculator or computer)

Probability statement:p-value $= P\left(\chi^2 > 3.2603\right) = 0.1959$

TI-83+ and TI-84 calculator: Press the `MATRX` key and arrow over to `EDIT`. Press `1:[A]`. Press `2 ENTER 3 ENTER`. 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`. 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`.

Compare α and the p-value:$\alpha = 0.05$ and the p-value = 0.1959. $\alpha <$ p-value.

Make a decision: Since $\alpha <$ p-value, do not reject H_o.

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.

** Contributed by Dr. Larry Green