11.6: Test for Homogeneity

The goodness–of–fit test can be used to decide whether a population fits a given distribution, but it will not suffice to decide whether two populations follow the same unknown distribution. A different test, called the **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.

NOTE

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

Hypotheses

- H_0: 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)

$df = \text{number of columns} - 1$

Requirements

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

Common Uses

Comparing two populations. For example: men vs. women, before vs. after, east vs. west. The variable is categorical with more than two possible response values.

Example 11.6.1

Do male and female college students have the same distribution of living arrangements? 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 arrangements: dormitory, apartment, with parents, other. The results are shown in Table 11.6.18 Do male and female college students have the same distribution of living arrangements?

Table 11.6.18 Distribution of living arragements for college males and college females

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

Answer

Solution 11.11

H_0: The distribution of living arrangements for male college students is the same as the distribution of living arrangements for female college students.

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

Degrees of Freedom (df):
$df = \text{number of columns} - 1 = 4 - 1 = 3$

Distribution for the test: χ^2_3

Calculate the test statistic: $\chi_c^2 = 10.129$

Figure 11.6.9

The graph of the Chi-square shows the distribution and marks the critical value with three degrees of freedom at 95% level of confidence, $\alpha = 0.05$, 7.815. The graph also marks the calculated χ^2 test statistic of 10.129. Comparing the test statistic with the critical value, as we have done with all other hypothesis tests, we reach the conclusion.

Make a decision: Because the calculated test statistic is in the tail we reject H_0. 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 arrangements 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 draw any conclusions about how they differ.

Exercise 11.6.1A

Do families and singles have the same distribution of cars? Use a level of significance of 0.05. Suppose that 100 randomly selected families and 200 randomly selected singles were asked what type of car they drove: sport, sedan, hatchback, truck, van/SUV. The results are shown in Table 11.6.19. Do families and singles have the same distribution of cars? Test at a level of significance of 0.05.

Table 11.6.19

	Sport	Sedan	Hatchback	Truck	Van/SUV
Family	5	15	35	17	28
Single	45	65	37	46	7

Exercise 11.6.1B

Ivy League schools receive many applications, but only some can be accepted. At the schools listed in Table 11.6.20, two types of applications are accepted: regular and early decision.

Table 11.6.20

Application type accepted	Brown	Columbia	Cornell	Dartmouth	Penn	Yale
Regular	2,115	1,792	5,306	1,734	2,685	1,245
Early decision	577	627	1,228	444	1,195	761

We want to know if the number of regular applications accepted follows the same distribution as the number of early applications accepted. State the null and alternative hypotheses, the degrees of freedom and the test statistic, sketch the graph of the χ^2 distribution and show the critical value and the calculated value of the test statistic, and draw a conclusion about the test of homogeneity.