

## 11.2: Chi-Square Goodness of Fit

In probability, you calculated probabilities using both experimental and theoretical methods. There are times when it is important to determine how well the experimental values match the theoretical values. An example of this is if you wish to verify if a die is fair. To determine if observed values fit the expected values, you want to see if the difference between observed values and expected values is large enough to say that the test statistic is unlikely to happen if you assume that the observed values fit the expected values. The test statistic in this case is also the chi-square. The process is the same as for the chi-square test for independence.

### Hypothesis Test for Goodness of Fit Test

1. State the null and alternative hypotheses and the level of significance

$H_o$ : The data are consistent with a specific distribution

$H_A$ : The data are not consistent with a specific distribution

Also, state your  $\alpha$  level here.

2. State and check the assumptions for the hypothesis test

a. A random sample is taken.

b. Expected frequencies for each cell are greater than or equal to 5 (The expected frequencies,  $E$ , will be calculated later, and this assumption means  $E \geq 5$ ).

3. Find the test statistic and p-value

Finding the test statistic involves several steps. First the data is collected and counted, and then it is organized into a table (in a table each entry is called a cell). These values are known as the observed frequencies, which the symbol for an observed frequency is  $O$ . The table is made up of  $k$  entries. The total number of observed frequencies is  $n$ . The expected frequencies are calculated by multiplying the probability of each entry,  $p$ , times  $n$ .

$$\text{Expected frequency( entry } i) = E = n * p$$

#### Test Statistic:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where  $O$  is the observed frequency and  $E$  is the expected frequency.

Again, the test statistic involves squaring the differences, so the test statistics are all positive. Thus a chi-squared test for goodness of fit is always right tailed.

p-value:

Using the TI-83/84:  $\chi$  cdf ( lower limit, 1E99,  $df$ )

Using R:  $1 - \text{pchisq}(\chi^2, df)$

Where the degrees of freedom is  $df = k - 1$

4. Conclusion

This is where you write reject  $H_o$  or fail to reject  $H_o$ . The rule is: if the p-value  $< \alpha$ , then reject  $H_o$ . If the p-value  $\geq \alpha$ , then fail to reject  $H_o$ .

5. Interpretation

This is where you interpret in real world terms the conclusion to the test. The conclusion for a hypothesis test is that you either have enough evidence to show  $H_A$  is true, or you do not have enough evidence to show  $H_A$  is true.

#### Example 11.2.1 goodness of fit test using the formula

Suppose you have a die that you are curious if it is fair or not. If it is fair then the proportion for each value should be the same. You need to find the observed frequencies and to accomplish this you roll the die 500 times and count how often each side comes up. The data is in Example 11.2.1. Do the data show that the die is fair? Test at the 5% level.

Table 11.2.1: Observed Frequencies of Die

Die values	1	2	3	4	5	6	Total

Observed Frequency	78	87	87	76	85	87	100
--------------------	----	----	----	----	----	----	-----

### Solution

1. State the null and alternative hypotheses and the level of significance

$H_o$ : The observed frequencies are consistent with the distribution for fair die (the die is fair)

$H_A$ : The observed frequencies are not consistent with the distribution for fair die (the die is not fair)

$\alpha = 0.05$

2. State and check the assumptions for the hypothesis test

- A random sample is taken since each throw of a die is a random event.
- Expected frequencies for each cell are greater than or equal to 5. See step 3.

3. Find the test statistic and p-value

First you need to find the probability of rolling each side of the die. The sample space for rolling a die is  $\{1, 2, 3, 4, 5, 6\}$ .

Since you are assuming that the die is fair, then  $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$ .

Now you can find the expected frequency for each side of the die. Since all the probabilities are the same, then each expected frequency is the same.

$$\text{Expected Frequency} = E = n * p = 500 * \frac{1}{6} \approx 83.33$$

Test Statistic:

It is easier to calculate the test statistic using a table.

Table 11.2.2: Calculation of the Chi-Square Test Statistic

$O$	$E$	$O-E$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
78	83.33	-5.22	28.4089	0.340920437
87	83.33	3.67	13.4689	0.161633265
87	83.33	3.67	13.4689	0.161633265
76	83.33	-7.33	53.7289	0.644772591
85	83.33	1.67	2.7889	0.033468139
87	83.33	3.67	13.4689	0.161633265
Total		0.02		$\chi^2 \approx 1.504060962$

The test statistic is  $\chi^2 \approx 1.504060962$

The degrees of freedom are  $df = k - 1 = 6 - 1 = 5$

Using TI-83/84:  $p\text{-value} = \chi^2 \text{cdf}(1.50406096, 1E99, 5) \approx 0.913$

Using R:  $p\text{-value} = 1 - \text{pchisq}(1.50406096, 5) \approx 0.9126007$

4. Conclusion

Fail to reject  $H_o$  since the p-value is greater than 0.05.

5. Interpretation

There is not enough evidence to show that the die is not consistent with the distribution for a fair die. There is not enough evidence to show that the die is not fair.

### Example 11.2.2 goodness of fit test using technology

Suppose you have a die that you are curious if it is fair or not. If it is fair then the proportion for each value should be the same. You need to find the observed frequencies and to accomplish this you roll the die 500 times and count how often each side comes up. The data is in Example 11.2.1. Do the data show that the die is fair? Test at the 5% level.

#### Solution

1. State the null and alternative hypotheses and the level of significance

$H_0$ : The observed frequencies are consistent with the distribution for fair die (the die is fair)

$H_A$ : The observed frequencies are not consistent with the distribution for fair die (the die is not fair)

$\alpha = 0.05$

2. State and check the assumptions for the hypothesis test

- A random sample is taken since each throw of a die is a random event.
- Expected frequencies for each cell are greater than or equal to 5. See step 3.

3. Find the test statistic and p-value

Using the TI-83/84 calculator:

#### Using the TI-83:

To use the TI-83 calculator to compute the test statistic, you must first put the data into the calculator. Type the observed frequencies into L1 and the expected frequencies into L2. Then you will need to go to L3, arrow up onto the name, and type in  $(L1 - L2)^2 / L2$ . Now you use 1-Var Stats L3 to find the total. See *Figure 11.2.1* for the initial setup, *Figure 11.2.2* for the results of that calculation, and *Figure 11.2.3* for the result of the 1-Var Stats L3.

Figure 11.2.1: Input into TI-83

Figure 11.2.2: Result for L3 on TI-83

Figure 11.2.3: 1-Var Stats L3 Result on TI-83

The total is the chi-square value,  $\chi^2 = \sum x \approx 1.50406$ .

The p-value is found using  $p\text{-value} = \chi^2 \text{cdf}(1.50406096, 1E99, 5) \approx 0.913$  where the degrees of freedom is  $df = k - 1 = 6 - 1 = 5$ .

#### Using the TI-84:

To run the test on the TI-84, type the observed frequencies into L1 and the expected frequencies into L2, then go into STAT, move over to TEST and choose  $\chi^2$  GOF-Test from the list. The setup for the test is in *Figure 11.2.4*

Figure 11.2.4: Setup for Chi-Square Goodness of Fit Test on TI-84

Once you press ENTER on Calculate you will see the results in *Figure 11.2.5*

Figure 11.2.5: Results for Chi-Square Test on TI-83/84

The test statistic is  $\chi^2 \approx 1.504060962$

The  $p\text{-value} \approx 0.913$

The CNTRB represent the  $\frac{(O - E)^2}{E}$  for each die value. You can see the values by pressing the right arrow.

Using R:

Type in the observed frequencies. Call it something like observed.

`observed<- c(type in data with commas in between)`

Type in the probabilities that you are comparing to the observed frequencies.

Call it something like null.probs.

`null.probs <- c(type in probabilities with commas in between)`

`chisq.test(observed, p=null.probs)` – the command for the hypothesis test

For this example (Note since you are looking to see if the die is fair, then the probability of each side of a fair die coming up is  $1/6$ .)

```
observed<-c(78, 87, 87, 76, 85, 87)
```

```
null.probs<-c(1/6, 1/6, 1/6, 1/6, 1/6, 1/6)
```

```
chisq.test(observed, p=null.probs)
```

Output:

Chi-squared test for given probabilities

data: observed

X-squared = 1.504, df = 5, p-value = 0.9126

The test statistic is  $\chi^2 = 1.504$  and the p-value = 0.9126.

#### 4. Conclusion

Fail to reject  $H_o$  since the p-value is greater than 0.05.

#### 5. Interpretation

There is not enough evidence to show that the die is not consistent with the distribution for a fair die. There is not enough evidence to show that the die is not fair.

## Homework

### Exercise 11.2.1

In each problem show all steps of the hypothesis test. If some of the assumptions are not met, note that the results of the test may not be correct and then continue the process of the hypothesis test.

1. According to the M&M candy company, the expected proportion can be found in Example 11.2.3 In addition, the table contains the number of M&M's of each color that were found in a case of candy (Madison, 2013). At the 5% level, do the observed frequencies support the claim of M&M?

Table 11.2.3: M&M Observed and Proportions

	Blue	Brown	Green	Orange	Red	Yellow	Total
Observed Frequencies	481	371	483	544	372	369	2620
Expected Proportion	0.24	0.13	0.16	0.20	0.13	0.14	

2. Eyeglassomatic manufactures eyeglasses for different retailers. They test to see how many defective lenses they made the time period of January 1 to March 31. Example 11.2.4 gives the defect and the number of defects. Do the data support the notion that each defect type occurs in the same proportion? Test at the 10% level.

Table 11.2.4: Number of Defective Lenses

Defect type	Number of defects
Scratch	5865
Right shaped - small	4613
Flaked	1992
Wrong axis	1838
Chamfer wrong	1596
Crazing, cracks	1546
Wrong shape	1485

Defect type	Number of defects
Wrong PD	1398
Spots and bubbles	1371
Wrong height	1130
Right shape - big	1105
Lost in lab	976
Spots/bubble - intern	976

3. On occasion, medical studies need to model the proportion of the population that has a disease and compare that to observed frequencies of the disease actually occurring. Suppose the end-stage renal failure in south-west Wales was collected for different age groups. Do the data in Example 11.2.5 show that the observed frequencies are in agreement with proportion of people in each age group (Boyle, Flowerdew & Williams, 1997)? Test at the 1% level.

Table 11.2.5: Renal Failure Frequencies

Age Group	16-29	30-44	45-59	60-75	75+	Total
Observed Frequency	32	66	132	218	91	539
Expected Proportion	0.23	0.25	0.22	0.21	0.09	

4. In Africa in 2011, the number of deaths of a female from cardiovascular disease for different age groups are in Example 11.2.6 ("Global health observatory," 2013). In addition, the proportion of deaths of females from all causes for the same age groups are also in Example 11.2.6 Do the data show that the death from cardiovascular disease are in the same proportion as all deaths for the different age groups? Test at the 5% level.

Table 11.2.6: Deaths of Females for Different Age Groups

Age	5-14	15-29	30-49	50-69	Total
Cardiovascular Frequency	9	16	56	433	513
All Cause Proportion	0.10	0.12	0.26	0.52	

5. In Australia in 1995, there was a question of whether indigenous people are more likely to die in prison than non-indigenous people. To figure out, the data in Example 11.2.7 was collected. ("Aboriginal deaths in," 2013). Do the data show that indigenous people die in the same proportion as non-indigenous people? Test at the 1% level.

Table 11.2.7: Death of Prisoners

	Prisoner Dies	Prisoner Did Not Die	Total
Indigenous Prisoner Frequency	17	2890	2907
Frequency of Non-Indigenous Prisoner	42	14459	14501

6. A project conducted by the Australian Federal Office of Road Safety asked people many questions about their cars. One question was the reason that a person chooses a given car, and that data is in Example 11.2.8 ("Car preferences," 2013).

Table 11.2.8: Reason for Choosing a Car

Safety	Reliability	Cost	Performance	Comfort	Looks
--------	-------------	------	-------------	---------	-------

84

62

46

34

47

27

**Answer**

For all hypothesis tests, just the conclusion is given. See solutions for the entire answer.

1. Reject  $H_0$
3. Reject  $H_0$
5. Reject  $H_0$

This page titled [11.2: Chi-Square Goodness of Fit](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Kathryn Kozak](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.