

# THE CHI-SQUARE DISTRIBUTION: GOODNESS-OF-FIT TEST\*

Susan Dean  
Barbara Illowsky, Ph.D.

This work is produced by OpenStax-CNX and licensed under the  
Creative Commons Attribution License 3.0<sup>†</sup>

## Abstract

This module describes how the chi-square distribution is used to conduct goodness-of-fit test.

In this type of hypothesis test, you determine whether the data **"fit"** a particular distribution or not. For example, you may suspect your unknown data fit a binomial distribution. You use a chi-square test (meaning the distribution for the hypothesis test is chi-square) to determine if there is a fit or not. **The null and the alternate hypotheses for this test may be written in sentences or may be stated as equations or inequalities.**

The test statistic for a goodness-of-fit test is:

$$\sum_k \frac{(O - E)^2}{E} \quad (1)$$

where:

- $O$  = observed values (data)
- $E$  = expected values (from theory)
- $k$  = the number of different data cells or categories

**The observed values are the data values and the expected values are the values you would expect to get if the null hypothesis were true.** There are  $n$  terms of the form  $\frac{(O-E)^2}{E}$ .

The degrees of freedom are  $df = (\text{number of categories} - 1)$ .

**The goodness-of-fit test is almost always right tailed.** If the observed values and the corresponding expected values are not close to each other, then the test statistic can get very large and will be way out in the right tail of the chi-square curve.

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

## Example 1

Absenteeism of college students from math classes is a major concern to math instructors because missing class appears to increase the drop rate. Suppose that a study was done to determine if the actual student absenteeism follows faculty perception. The faculty expected that a group of 100 students would miss class according to the following chart.

---

\*Version 1.8: Jun 20, 2012 3:10 am -0500

<sup>†</sup><http://creativecommons.org/licenses/by/3.0/>

Number absences per term	Expected number of students
0 - 2	50
3 - 5	30
6 - 8	12
9 - 11	6
12+	2

Table 1

A random survey across all mathematics courses was then done to determine the actual number (**observed**) of absences in a course. The next chart displays the result of that survey.

Number absences per term	Actual number of students
0 - 2	35
3 - 5	40
6 - 8	20
9 - 11	1
12+	4

Table 2

Determine the null and alternate hypotheses needed to conduct a goodness-of-fit test.

$H_o$ : Student absenteeism **fits** faculty perception.

The alternate hypothesis is the opposite of the null hypothesis.

$H_a$ : Student absenteeism **does not fit** faculty perception.

### Problem 1

Can you use the information as it appears in the charts to conduct the goodness-of-fit test?

### Solution

**No.** Notice that the expected number of absences for the "12+" entry is less than 5 (it is 2). Combine that group with the "9 - 11" group to create new tables where the number of students for each entry are at least 5. The new tables are below.

Number absences per term	Expected number of students
0 - 2	50
3 - 5	30
6 - 8	12
9+	8

Table 3

Number absences per term	Actual number of students
0 - 2	35
3 - 5	40
6 - 8	20
9+	5

Table 4

**Problem 2**

What are the degrees of freedom (df)?

**Solution**

There are 4 "cells" or categories in each of the new tables.

$$df = \text{number of cells} - 1 = 4 - 1 = 3$$

**Example 2**

Employers particularly want to know which days of the week employees are absent in a five day work week. Most employers would like to believe that employees are absent equally during the week. Suppose a random sample of 60 managers were asked on which day of the week did they have the highest number of employee absences. The results were distributed as follows:

**Day of the Week Employees were most Absent**

	Monday	Tuesday	Wednesday	Thursday	Friday
Number of Absences	15	12	9	9	15

Table 5

**Problem**

For the population of employees, do the days for the highest number of absences occur with equal frequencies during a five day work week? Test at a 5% significance level.

**Solution**

The null and alternate hypotheses are:

- $H_0$ : The absent days occur with equal frequencies, that is, they fit a uniform distribution.
- $H_a$ : The absent days occur with unequal frequencies, that is, they do not fit a uniform distribution.

If the absent days occur with equal frequencies, then, out of 60 absent days (the total in the sample:  $15 + 12 + 9 + 9 + 15 = 60$ ), there would be 12 absences on Monday, 12 on Tuesday, 12 on Wednesday, 12 on Thursday, and 12 on Friday. These numbers are the **expected** ( $E$ ) values. The values in the table are the **observed** ( $O$ ) values or data.

This time, calculate the  $\chi^2$  test statistic by hand. Make a chart with the following headings and fill in the columns:

- Expected ( $E$ ) values (12, 12, 12, 12, 12)

- Observed ( $O$ ) values (15, 12, 9, 9, 15)
- $(O - E)$
- $(O - E)^2$
- $\frac{(O - E)^2}{E}$

The last column ( $\frac{(O - E)^2}{E}$ ) should have 0.75, 0, 0.75, 0.75, 0.75.

Now add (sum) the last column. Verify that the sum is 3. This is the  $\chi^2$  test statistic.

To find the p-value, calculate  $P(\chi^2 > 3)$ . This test is right-tailed.

(Use a computer or calculator to find the p-value. You should get p-value = 0.5578.)

The dfs are the number of cells  $- 1 = 5 - 1 = 4$ .

**TI-83+ and TI-84:** Press **2nd** **DISTR**. Arrow down to  $\chi^2$ cdf. Press **ENTER**. Enter (3,  $10^{99}$ , 4). Rounded to 4 decimal places, you should see 0.5578 which is the p-value.

Next, complete a graph like the one below with the proper labeling and shading. (You should shade the right tail.)



The decision is to not reject the null hypothesis.

**Conclusion:** At a 5% level of significance, from the sample data, there is not sufficient evidence to conclude that the absent days do not occur with equal frequencies.

NOTE: TI-83+ and some TI-84 calculators do not have a special program for the test statistic for the goodness-of-fit test. The next example (Example 11-3) has the calculator instructions. The newer TI-84 calculators have in **STAT TESTS** the test **Chi2 GOF**. To run the test, put the observed values (the data) into a first list and the expected values (the values you expect if the null hypothesis is true) into a second list. Press **STATTESTS** and **Chi2 GOF**. Enter the list names for the Observed list and the Expected list. Enter the degrees of freedom and press **calculate** or **draw**. Make sure you clear any lists before you start. See below.

NOTE: **To Clear Lists in the calculators:** Go into **STAT EDIT** and arrow up to the list name area of the particular list. Press **CLEAR** and then arrow down. The list will be cleared. Or, you can press **STAT** and press 4 (for **ClrList**). Enter the list name and press **ENTER**.

### Example 3

One study indicates that the number of televisions that American families have is distributed (this is the **given** distribution for the American population) as follows:

Number of Televisions	Percent
0	10
1	16
2	55
3	11
over 3	8

Table 6

The table contains expected ( $E$ ) percents.

A random sample of 600 families in the far western United States resulted in the following data:

Number of Televisions	Frequency
0	66
1	119
2	340
3	60
over 3	15
	<b>Total = 600</b>

Table 7

The table contains observed ( $O$ ) frequency values.

### Problem

At the 1% significance level, does it appear that the distribution "number of televisions" of far western United States families is different from the distribution for the American population as a whole?

### Solution

This problem asks you to test whether the far western United States families distribution fits the distribution of the American families. This test is always right-tailed.

The first table contains expected percentages. To get expected ( $E$ ) frequencies, multiply the percentage by 600. The expected frequencies are:

Number of Televisions	Percent	Expected Frequency
0	10	$(0.10) \cdot (600) = 60$
1	16	$(0.16) \cdot (600) = 96$
2	55	$(0.55) \cdot (600) = 330$
3	11	$(0.11) \cdot (600) = 66$
over 3	8	$(0.08) \cdot (600) = 48$

Table 8

Therefore, the expected frequencies are 60, 96, 330, 66, and 48. In the TI calculators, you can let the calculator do the math. For example, instead of 60, enter  $.10 \cdot 600$ .

$H_o$ : The "number of televisions" distribution of far western United States families is the same as the "number of televisions" distribution of the American population.

$H_a$ : The "number of televisions" distribution of far western United States families is different from the "number of televisions" distribution of the American population.

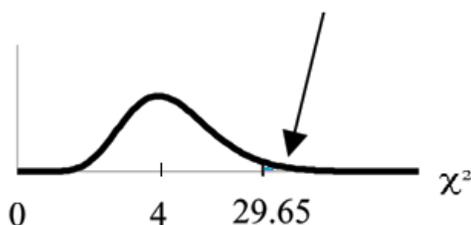
Distribution for the test:  $\chi^2_4$  where  $df = (\text{the number of cells}) - 1 = 5 - 1 = 4$ .

NOTE:  $df \neq 600 - 1$

Calculate the test statistic:  $\chi^2 = 29.65$

Graph:

p-value = 0.000006 (almost 0)



**Probability statement:**  $p\text{-value} = P(\chi^2 > 29.65) = 0.000006$ .

**Compare  $\alpha$  and the p-value:**

- $\alpha = 0.01$
- $p\text{-value} = 0.000006$

So,  $\alpha > p\text{-value}$ .

**Make a decision:** Since  $\alpha > p\text{-value}$ , reject  $H_o$ .

This means you reject the belief that the distribution for the far western states is the same as that of the American population as a whole.

**Conclusion:** At the 1% significance level, from the data, there is sufficient evidence to conclude that the "number of televisions" distribution for the far western United States is different from the "number of televisions" distribution for the American population as a whole.

NOTE: TI-83+ and some TI-84 calculators: Press STAT and ENTER. Make sure to clear lists L1, L2, and L3 if they have data in them (see the note at the end of Example 11-2). Into L1, put the observed frequencies 66, 119, 349, 60, 15. Into L2, put the expected frequencies  $.10 \cdot 600$ ,  $.16 \cdot 600$ ,  $.55 \cdot 600$ ,  $.11 \cdot 600$ ,  $.08 \cdot 600$ . Arrow over to list L3 and up to the name area "L3". Enter  $(L1-L2)^2/L2$  and ENTER. Press 2nd QUIT. Press 2nd LIST and arrow over to MATH. Press 5. You should see "sum" (Enter L3). Rounded to 2 decimal places, you should see 29.65. Press 2nd DISTR. Press 7 or Arrow down to 7:  $\chi^2cdf$  and press ENTER. Enter (29.65, 1E99, 4). Rounded to 4 places, you should see  $5.77E-6 = .000006$  (rounded to 6 decimal places) which is the p-value.

The newer TI-84 calculators have in STAT TESTS the test Chi2 GOF. To run the test, put the observed values (the data) into a first list and the expected values (the values you expect if the null hypothesis is true) into a second list. Press STATTESTS and Chi2 GOF. Enter the list names for the Observed list and the Expected list. Enter the degrees of freedom and press calculate or draw. Make sure you clear any lists before you start.

**Example 4**

Suppose you flip two coins 100 times. The results are 20 HH, 27 HT, 30 TH, and 23 TT. Are the coins fair? Test at a 5% significance level.

**Solution**

This problem can be set up as a goodness-of-fit problem. The sample space for flipping two fair coins is {HH, HT, TH, TT}. Out of 100 flips, you would expect 25 HH, 25 HT, 25 TH, and 25 TT. This is the expected distribution. The question, "Are the coins fair?" is the same as saying, "Does the distribution of the coins (20 HH, 27 HT, 30 TH, 23 TT) fit the expected distribution?"

**Random Variable:** Let  $X$  = the number of heads in one flip of the two coins.  $X$  takes on the value 0, 1, 2. (There are 0, 1, or 2 heads in the flip of 2 coins.) Therefore, the **number of cells is 3**. Since  $X$  = the number of heads, the observed frequencies are 20 (for 2 heads), 57 (for 1 head), and 23 (for 0 heads or both tails). The expected frequencies are 25 (for 2 heads), 50 (for 1 head), and 25 (for 0 heads or both tails). This test is right-tailed.

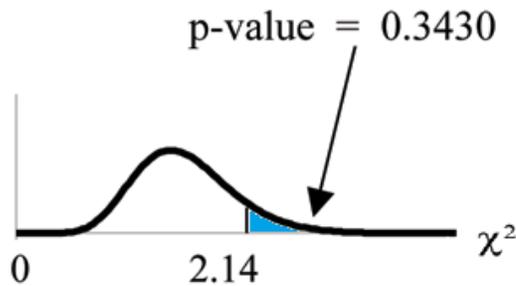
$H_o$ : The coins are fair.

$H_a$ : The coins are not fair.

**Distribution for the test:**  $\chi^2$  where  $df = 3 - 1 = 2$ .

**Calculate the test statistic:**  $\chi^2 = 2.14$

**Graph:**



**Probability statement:**  $p\text{-value} = P(\chi^2 > 2.14) = 0.3430$

**Compare  $\alpha$  and the p-value:**

- $\alpha = 0.05$
- $p\text{-value} = 0.3430$

So,  $\alpha < p\text{-value}$ .

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

**Conclusion:** There is insufficient evidence to conclude that the coins are not fair.

NOTE: TI-83+ and some TI- 84 calculators: Press STAT and ENTER. Make sure you clear lists L1, L2, and L3 if they have data in them. Into L1, put the observed frequencies 20, 57, 23. Into L2, put the expected frequencies 25, 50, 25. Arrow over to list L3 and up to the name area "L3". Enter  $(L1-L2)^2/L2$  and ENTER. Press 2nd QUIT. Press 2nd LIST and arrow over to MATH. Press 5. You should see "sum". Enter L3. Rounded to 2 decimal places, you should see 2.14. Press 2nd DISTR. Arrow down to 7:  $\chi^2cdf$  (or press 7). Press ENTER. Enter 2.14, 1E99, 2). Rounded to 4 places, you should see .3430 which is the p-value.

The newer TI-84 calculators have in STAT TESTS the test Chi2 GOF. To run the test, put the

observed values (the data) into a first list and the expected values (the values you expect if the null hypothesis is true) into a second list. Press **STATTESTS** and **Chi2 GOF**. Enter the list names for the Observed list and the Expected list. Enter the degrees of freedom and press **calculate** or **draw**. Make sure you clear any lists before you start.