11.3: 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 alternative 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:

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

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 number of degrees of freedom is df = (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 number of expected values inside each cell needs to be at least five in order to use this test.

Exercise 11.3.1

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 rate follows faculty perception. The faculty expected that a group of 100 students would miss class according to Table $11.3.1$.

Table 11.3.1

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

A random survey across all mathematics courses was then done to determine the actual number **(observed)** of absences in a course. The chart in Table $11.3.2$ displays the results of that survey.

Table 11.3.2

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

Number of absences per term	Actual number of students
9–11	1
12+	4

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

$\mathbf{H_a}$**:** Student absenteeism **fits** faculty perception.

The alternative hypothesis is the opposite of the null hypothesis.

$\mathbf{H_a}$**:** Student absenteeism **does not fit** faculty perception.

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

Answer

Solution 11.4

a. **No.** Notice that the expected number of absences for the "12+" entry is less than five (it is two). Combine that group with the "9–11" group to create new tables where the number of students for each entry are at least five. The new results are in Table $11.3.3$and Table $11.3.4$

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

Table 11.3

Table 11.3.4

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

b. What is the number of degrees of freedom (df)?

Answer

Solution 11.4

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

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

? Example 11.3.1

A factory manager needs to understand how many products are defective versus how many are produced. The number of expected defects is listed in Table $11.3.5$.

Table 11.3.5

Number produced	Number defective
0–100	5
101–200	6
201–300	7
301–400	8
401–500	10

A random sample was taken to determine the actual number of defects. Table $11.3.6$ shows the results of the survey.

Table 11.3.6

Number produced	Number defective
0–100	5
101–200	7
201–300	8
301–400	9
401–500	11

State the null and alternative hypotheses needed to conduct a goodness-of-fit test, and state the degrees of freedom.

Exercise 11.3.2

Employers 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 they had the highest number of employee absences. The results were distributed as in Table $11.3.7$. 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.

Table 11.3.7 Day of the Week Employees were Most Absent

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

Answer

Solution 11.5

The null and alternative 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 \chi2 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 $(15, 12, 9, 9, 15)$
- $(O – E)$

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

Now add (sum) the last column. The sum is three. This is the χ^2 test statistic.

The calculated test statistics is 3 and the critical value of the χ^2 distribution at 4 degrees of freedom the 0.05 level of confidence is 9.48. This value is found in the χ^2 table at the 0.05 column on the degrees of freedom row 4.

The degrees of freedom are the number of cells $- 1 = 5 - 1 = 4$

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

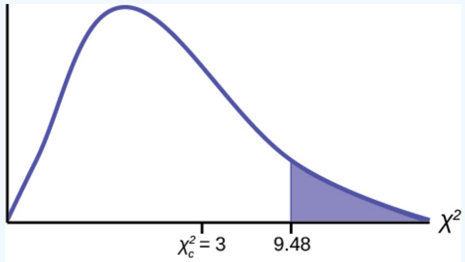

Figure 11.3.5

$$\chi^2_c = \sum_k \frac{(\mathbf{O} - \mathbf{E})^2}{\mathbf{E}} = \mathbf{3}$$

The decision is not to reject the null hypothesis because the calculated value of the test statistic is not in the tail of the distribution.

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.

? Example 11.3.2

Teachers want to know which night each week their students are doing most of their homework. Most teachers think that students do homework equally throughout the week. Suppose a random sample of 56 students were asked on which night of the week they did the most homework. The results were distributed as in Table 11.3.8.

Table 11.3.8

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Number of students	11	8	10	7	10	5	5

From the population of students, do the nights for the highest number of students doing the majority of their homework occur with equal frequencies during a week? What type of hypothesis test should you use?

? Exercise 11.3.3

One study indicates that the number of televisions that American families have is distributed (this is the **given** distribution for the American population) as in Table 11.3.9.

Table 11.3.9

Number of Televisions	Percent
0	10
1	16
2	55

Number of Televisions	Percent
3	11
4+	8

The table contains expected (E) percents.

A random sample of 600 families in the far western United States resulted in the data in Table $11.3.10$

Table 11.3.10

Number of Televisions	Frequency
0	66
1	119
2	340
3	60
4+	15
	Total = 600

The table contains observed (O) frequency values.

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?

Answer

Solution 11.6

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 shown in Table $11.3.11$

Table 11.3.11

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

Therefore, the expected frequencies are 60, 96, 330, 66, and 48.

H_0: 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: χ_4^2 where $df = (\text{ the number of cells }) - 1 = 5 - 1 = 4$.

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

Graph:

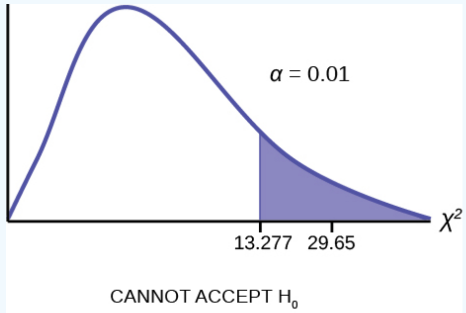

Figure 11.3.6

The graph of the Chi-square shows the distribution and marks the critical value with four degrees of freedom at 99% level of confidence, α = .01, 13.277. The graph also marks the calculated chi squared test statistic of 29.65. 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 test statistic is in the tail of the distribution we cannot accept the null hypothesis.

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.

Example 11.3.3

The expected percentage of the number of pets students have in their homes is distributed (this is the given distribution for the student population of the United States) as in Table $11.3.12$

Table 11.3.12

Number of pets	Percent
0	18
1	25
2	30
3	18
4+	9

A random sample of 1,000 students from the Eastern United States resulted in the data in Table $11.3.13$

Table 11.3.13

Number of pets	Frequency
0	210
1	240
2	320
3	140
4+	90

At the 1% significance level, does it appear that the distribution “number of pets” of students in the Eastern United States is different from the distribution for the United States student population as a whole?

Exercise 11.3.4

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

Answer

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 $HH, 25HT, 25TH$, and $25TT$. This is the expected distribution from the binomial probability distribution. The question, "Are the coins fair?" is the same as saying, "Does the distribution of the coins $(20HH, 27HT, 30TH, 23TT)$ fit the expected distribution?"

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

$\mathbf{H_0}$: The coins are fair.

$\mathbf{H_a}$: The coins are not fair.

Distribution for the test: χ^2_2 where $df = 3 - 1 = 2$.

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

Graph:

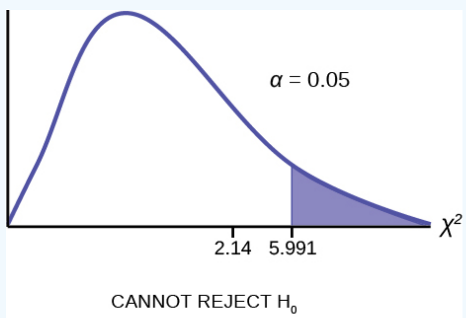

Figure 11.3.7

The graph of the Chi-square shows the distribution and marks the critical value with two degrees of freedom at 95% level of confidence, $\alpha = 0.05$, 5.991. The graph also marks the calculated χ^2 test statistic of 2.14. Comparing the test statistic with the critical value, as we have done with all other hypothesis tests, we reach the conclusion.

Conclusion: There is insufficient evidence to conclude that the coins are not fair: we cannot reject the null hypothesis that the coins are fair.