

In the situations in previous chapters, the scores have all been *numerical values* on some dimension, such as a score on a standard achievement test, length of time in a relationship, an employer's rating of an employee's job effectiveness on a nine-point scale, and so forth. By contrast, relationship style of a man's partner is an example of what in Chapter 1 we called a *nominal variable* (or a *categorical variable*). A nominal variable is one in which the information is the number of people in each category. These are called nominal variables because the different categories or levels of the variable have names instead of numbers. Hypothesis testing with nominal variables uses *chi-square tests*.

THE CHI-SQUARE STATISTIC AND THE CHI-SQUARE TEST FOR GOODNESS OF FIT

The basic idea of any chi-square test is that you compare how well an *observed breakdown* of people over various categories fits some *expected breakdown*. In this chapter you will learn about two types of chi-square tests: First, you will learn about the **chi-square test for goodness of fit**, which is a chi-square test involving levels of a *single nominal variable*. Later in the chapter, you will learn about the **chi-square test for independence**, which is used when there are *two nominal variables*, each with several categories.

In the relationship style example—in which there is a single nominal variable with three categories—you are comparing the observed breakdown of 50, 26, and 25 to the expected breakdown of about 34 (33.67) for each style. A breakdown of numbers of people expected in each category is actually a frequency distribution, as you learned in Chapter 1. Thus, a chi-square test is more formally described as comparing an **observed frequency** distribution to an **expected frequency** distribution. Overall, what the hypothesis testing involves is first figuring a number for the amount of mismatch between the observed frequency and the expected frequency and then seeing whether that number indicates a greater mismatch than you would expect by chance.

Let's start with how you would come up with that mismatch number for the observed versus expected frequencies. The mismatch between observed and expected for any one category is just the observed frequency minus the expected frequency. For example, consider again the Harter et al. study. For self-focused men with an other-focused partner, the observed frequency of 50 is 16.33 more than the expected frequency of 33.67 (recall the expected frequency is 1/3 of the 101 total). For the second category, the difference is -7.67 . For the third, -8.67 . These differences are shown in the Difference column of Table 13-1.

You do not use these differences directly. One reason is that some differences are positive and some are negative. Thus, they would cancel each other out. To get around this, you square each difference. (This is the same strategy we saw in Chapter 2 to deal with difference scores in figuring the variance.) In the relationship-style example, the squared difference for self-focused men with other-focused partners is 16.33 squared, or 266.67. For those men with self-focused partners, it is 58.83. For those with mutuality style partners, 75.17. These squared differences are shown in the Difference Squared column of Table 13-1.

In the Harter et al. (1997) example, the expected frequencies are the same in each category. But in other research situations, expected frequencies for the different categories may not be the same. A particular amount of difference between observed and expected has a different importance according to the size of the expected

chi-square test for goodness of fit

chi-square test for independence

observed frequency

expected frequency

frequency. For example, a difference of eight people between observed and expected is a much bigger mismatch if the expected frequency is 10 than if the expected frequency is 1,000. If the expected frequency is 10, a difference of 8 would mean that the observed frequency was 18 or 2, frequencies that are dramatically different from 10. But if the expected frequency is 1,000, a difference of 8 is only a slight mismatch. This would mean that the observed frequency was 1,008 or 992, frequencies that are only slightly different from 1,000.

How can you adjust the mismatch (the squared difference) between observed and expected for a particular category? What you need to do is adjust or weight the mismatch to take into account the expected frequency for that category. You can do this by dividing your squared difference for a category by the expected frequency for that category. Thus, if the expected frequency for a particular category is 10, you divide the squared difference by 10. If the expected frequency for the category is 1,000, you divide the squared difference by 1,000. In this way, you weight each squared difference by the expected frequency. This weighting puts the squared difference onto a more appropriate scale of comparison. In our example, for men with an other-focused partner, you would weight the mismatch by dividing the squared difference of 266.67 by 33.67, giving 7.92. For those with a self-focused partner, 58.83 divided by 33.67 gives 1.75. For those with a mutuality style partner, 75.17 divided by 33.67 gives 2.23. These adjusted mismatches (squared differences divided by expected frequencies) are shown in the rightmost column of Table 13-1.

What remains is to get an overall figure for the mismatch between observed and expected frequencies. This final step is done by adding up the mismatch for all the categories. In the Harter et al. example, this would be 7.92 plus 1.75 plus 2.23, for a total of 11.90. This final number (the sum of the weighted squared differences) is an overall indication of the amount of mismatch between the expected and observed frequencies. It is called the **chi-square statistic**. In terms of a formula,

$$(13-1) \quad \chi^2 = \sum \frac{(O - E)^2}{E}$$

In this formula, χ^2 is the chi-square statistic. Σ is the summation sign, telling you to sum over all the different categories. O is the observed frequency for a category (the number of people actually found in that category in the study). E is the expected frequency for a category (in the Harter et al. example, it is based on what you would expect if there were equal numbers in each category).

Applying the formula to the example,

$$\begin{aligned} \chi^2 &= \sum \frac{(O - E)^2}{E} = \frac{(50 - 33.67)^2}{33.67} + \frac{(26 - 33.67)^2}{33.67} + \frac{(25 - 33.67)^2}{33.67} \\ &= 11.90. \end{aligned}$$

STEPS FOR FIGURING THE CHI-SQUARE STATISTIC

Here is a summary of what we've said so far in terms of steps:

- ➊ Determine the actual, observed frequencies in each category.
- ➋ Determine the expected frequencies in each category.
- ➌ In each category, take observed minus expected frequencies.
- ➍ Square each of these differences.

chi-square statistic

Chi-square is the sum, over all the categories, of the squared difference between observed and expected frequencies divided by the expected frequency.

Tip for Success

Notice in the chi-square formula that, for each category, you first divide the squared difference between observed and expected frequencies by the expected frequency, and then you sum the resulting values for all the categories. This is a slightly different procedure than you are used to from previous chapters (in which you often first summed a series of squared values in the numerator and then divided by a denominator value), so be sure to follow the formula carefully.

- Ⓔ Divide each squared difference by the expected frequency for its category.
- Ⓕ Add up the results of Step Ⓔ for all the categories.

THE CHI-SQUARE DISTRIBUTION

Now we turn to the question of whether the chi-square statistic you have figured is a bigger mismatch than you would expect by chance. To answer that, you need to know how likely it is to get chi-square statistics of various sizes by chance. As long as you have a reasonable number of people in the study, the distribution of the chi-square statistic that would arise by chance follows quite closely a known mathematical distribution—the **chi-square distribution**.

chi-square distribution

The exact shape of the chi-square distribution depends on the degrees of freedom. For a chi-square test, the degrees of freedom are the number of categories that are free to vary, given the totals. In the partners' relationship style example, there are three categories. If you know the total number of people and you know the number in any two categories, you could figure out the number in the third category—so only two are free to vary. That is, in a study like this one (which uses a chi-square test for goodness of fit), if there are three categories, there are two degrees of freedom. In terms of a formula,

The degrees of freedom for the chi-square test for goodness of fit are the number of categories minus 1.

$$df = N_{\text{Categories}} - 1 \tag{13-2}$$

Web Link

<http://www.stat.sc.edu/~west/applets/chisqdemo1.html>. This web page nicely illustrates the shape of the chi-square distribution for different degrees of freedom.

Chi-square distributions for several different degrees of freedom are shown in Figure 13-1. Notice that the distributions are all skewed to the right. This is because the chi-square statistic cannot be less than 0 but can have very high values. (Chi-square must be positive because it is figured by adding a group of fractions in each of which the numerator and denominator both have to be positive. The numerator has to be positive because it is squared. The denominator has to be positive because the number of people expected in a given category can't be a negative number—you can't expect less than no one!)

THE CHI-SQUARE TABLE

What matters most about the chi-square distribution for hypothesis testing is the cutoff for a chi-square to be extreme enough to reject the null hypothesis. A **chi-square table** gives the cutoff chi-squares for different significance levels and various degrees of freedom. Table 13-2 shows a portion of a chi-square table like the

chi-square table

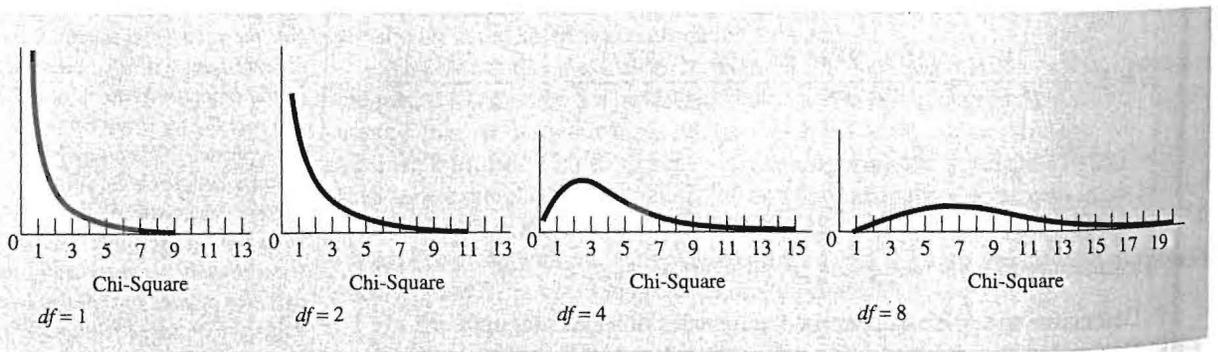


FIGURE 13-1 Examples of chi-square distributions for different degrees of freedom.

one in the Appendix (Table A-4). For our example, where there were two degrees of freedom, the table shows that the cutoff chi-square for the .05 level is 5.992.

In the Harter et al. (1997) example, we figured a chi-square of 11.90. This is clearly larger than the chi-square cutoff (using the .05 significance level) of 5.992 (see Figure 13-2). Thus, the researchers rejected the null hypothesis. That is, they rejected as too unlikely that the mismatch they observed could have come about if in the population of self-focused men there were an equal number of partners of each relationship style. It seemed more reasonable to conclude that there truly were different proportions of relationship styles of the partners of such men.

STEPS OF HYPOTHESIS TESTING

Let us review the chi-square test for goodness of fit using the same example, but this time systematically following the standard steps of hypothesis testing. In the process we also consider some fine points.

1 Restate the question as a research hypothesis and a null hypothesis about the populations. There are two populations:

Population 1: Self-focused men like those in the study.

Population 2: Self-focused men whose partners are equally of the three relationship styles.

The research hypothesis is that the distribution of people over categories in the two populations is different; the null hypothesis is that they are the same.

2 Determine the characteristics of the comparison distribution. The comparison distribution is a chi-square distribution with 2 degrees of freedom. (Once you know the total, there are only two category numbers still free to vary.)

3 Determine the cutoff on the comparison distribution at which the null hypothesis should be rejected. You do this by looking up the cutoff on the chi-square table for your significance level and the study's degrees of freedom. In this example, we used the .05 significance level, and we determined in Step 2 that there were 2 degrees of freedom. Based on the chi-square table, this gives a cutoff chi-square of 5.992 (see Figure 13-2).

4 Determine your sample's score on the comparison distribution. Your sample's score is the chi-square figured from the sample. In other words, this is where you do all the figuring.

A Determine the actual, observed frequencies in each category. These are shown in the first column of Table 13-1.

B Determine the expected frequencies in each category. We figured these each to be 33.67 based on expecting an equal distribution of the 101 partners.

TABLE 13-2 Portion of a Chi-Square Table

df	Significance Level		
	.10	.05	.01
1	2.706	3.841	6.635
2	4.605	5.992	9.211
3	6.252	7.815	11.345
4	7.780	9.488	13.277
5	9.237	11.071	15.087

Note: Full table is Table A-4 in the Appendix.

Tip for Success

It is important not to be confused by the terminology here. The comparison distribution is the distribution to which we compare the number that summarizes the whole pattern of the result. With a t test, this number is the t score and we use a t distribution. With an analysis of variance, it is the F ratio and we use an F distribution. Accordingly, with a chi-square test, our comparison distribution is a distribution of the chi-square statistic. This can be confusing because when preparing to use the chi-square distribution, you compare a distribution of observed frequencies to a distribution of expected frequencies. Yet the distribution of expected frequencies is not a comparison distribution in the sense that we use this term in Step 2 of hypothesis testing.

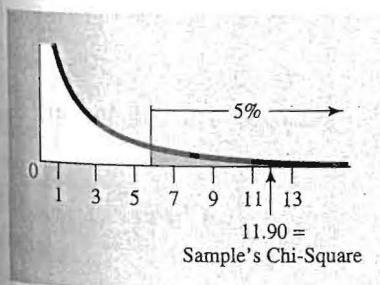


FIGURE 13-2 For the Harter et al. (1997) example, the chi-square distribution ($df = 2$) showing the cutoff for rejecting the null hypothesis at the .05 level and the sample's chi-square.

④ **In each category, take observed minus expected frequencies.** These are shown in the third column of Table 13-1.

⑤ **Square each of these differences.** These are shown in the fourth column of Table 13-1.

⑥ **Divide each squared difference by the expected frequency for its category.** These are shown in the fifth column of Table 13-1.

⑦ **Add up the results of Step ⑥ for all the categories.** The result we figured earlier (11.90) is the chi-square statistic for the sample. It is shown in Figure 13-2.

⑧ **Decide whether to reject the null hypothesis.** The chi-square cutoff to reject the null hypothesis (from Step ③) is 5.992 and the chi-square of the sample (from Step ⑦) is 11.90. Thus, you can reject the null hypothesis. The research hypothesis that the two populations are different is supported. That is, Harter et al. could conclude that the partners of self-focused men are not equally likely to be of the three relationship styles.

ANOTHER EXAMPLE

A fictional research team of clinical psychologists want to test a theory that mental health is affected by the level of a certain mineral in the diet, "mineral Q." The research team has located a region of the United States where mineral Q is found in very high concentrations in the soil. As a result, it is in the water people drink and in locally grown food. The researchers carry out a survey of older people who have lived in this area their whole life, focusing on mental health disorders. Of the 1,000 people surveyed, 134 had at some point in their life experienced an anxiety disorder, 160 had suffered from alcohol or drug abuse, 97 from mood disorders (such as major chronic depression), and 12 from schizophrenia; 597 had never experienced any of these problems. (In this example, we ignore the problem of what happens when a person had more than one of these problems.)

The psychologists then compared their results to what would be expected based on large surveys of the U.S. public in general. In these surveys, 14.6% of adults at some point in their lives suffer from an anxiety disorder, 16.4% from alcohol or drug abuse, 8.3% from mood disorders, and 1.5% from schizophrenia; 59.2% do not experience any of these conditions (Regier et al., 1984). If their sample of 1,000 is not different from the general U.S. population, 14.6% of them (146) should have had anxiety disorders, 16.4% of them (164) should have suffered from alcohol and drug abuse, and so forth. The question the clinical psychologists posed is this: On the basis of the sample we have studied, can we conclude that the rates of various mental health problems among people in this region are different from those of the general U.S. population?

Table 13-3 shows the observed and expected frequencies and the figuring for the chi-square test.

① **Restate the question as a research hypothesis and a null hypothesis about the populations.** There are two populations:

Population 1: People in the U.S. region with a high level of mineral Q.

Population 2: The U.S. population.

Tip for Success

Note in this example that the expected frequencies are figured based on what would be expected in the U.S. population. This is quite different from the situations we have considered before where the expected frequencies were based on an even division.

TABLE 13-3 Observed and Expected Frequencies and the Chi-Square Goodness of Fit Test for Types of Mental Health Disorders in a U.S. Region High in Mineral Q Compared to the General U.S. Population (Fictional Data)

Condition	Observed ^A	Expected ^B
Anxiety disorder	134	146 (14.6% × 1,000)
Alcohol and drug abuse	160	164 (16.4% × 1,000)
Mood disorders	97	83 (8.3% × 1,000)
Schizophrenia	12	15 (1.5% × 1,000)
None of these conditions	597	592 (59.2% × 1,000)

Degrees of freedom = $N_{\text{Categories}} - 1 = 5 - 1 = 4$ ^C

Chi-square needed, $df = 4$, .05 level: 9.488 ^D

$$\begin{aligned} \chi^2 &= \sum \frac{(O - E)^2}{E} = \frac{(134 - 146)^2}{146} + \frac{(160 - 164)^2}{164} + \frac{(97 - 83)^2}{83} + \frac{(12 - 15)^2}{15} + \frac{(597 - 592)^2}{592} \\ &= \frac{-12^2}{146} + \frac{-4^2}{164} + \frac{14^2}{83} + \frac{-3^2}{15} + \frac{5^2}{592} \\ &= \frac{144}{146} + \frac{16}{164} + \frac{196}{83} + \frac{9}{15} + \frac{25}{592} \\ &= .99 + .10 + 2.36 + .60 + .04 = 4.09 \end{aligned}$$

Decision: Do not reject the null hypothesis. ^E

The research hypothesis is that the distribution of numbers of people over the five mental health categories is different in the two populations; the null hypothesis is that it is the same.

^C **Determine the characteristics of the comparison distribution.** The comparison distribution is a chi-square distribution with 4 degrees of freedom (that is, $df = N_{\text{Categories}} - 1 = 5 - 1 = 4$). See Figure 13-3.

^D **Determine the cutoff sample score on the comparison distribution at which the null hypothesis should be rejected.** We will use the standard 5% significance level and we have just seen that there are 4 degrees of freedom. Thus, Table 13-2 (or Table A-4 in the Appendix) shows that the clinical psychologists need a chi-square of at least 9.488 to reject the null hypothesis. This is shown in Figure 13-3.

^E **Determine your sample's score on the comparison distribution.** The chi-square figuring is shown in Table 13-3.

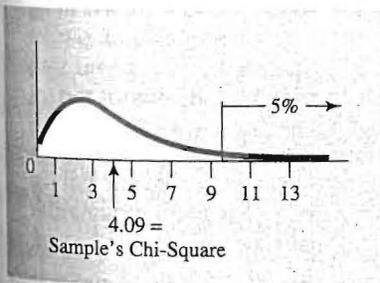


FIGURE 13-3 For the mineral Q example, the chi-square distribution ($df = 4$) showing the cutoff for rejecting the null hypothesis at the .05 level and the sample's chi square.

④ **Determine the actual, observed frequencies in each category.** These are shown in the first column of Table 13-3.

⑤ **Determine the expected frequencies in each category.** These are figured by multiplying the expected percentage by the total number. For example, with 14.6% expected to have anxiety disorders, the actual expected number to have anxiety disorders is 146 (that is, 14.6% of 1000). All of the expected frequencies are shown in Table 13-3.

⑥ **In each category, take observed minus expected frequencies.** The result of these subtractions are shown in the numerators of the second formula line on Table 13-3.

⑦ **Square each of these differences.** The results of these squarings are shown in the numerators of the third formula line on Table 13-3.

⑧ **Divide each squared difference by the expected frequency for its category.** The result of these divisions are shown in the fourth formula line on Table 13-3.

⑨ **Add up the results of Step ⑧ for all the categories.** The sum comes out to 4.09. The addition is shown on Table 13-3; the location on the chi-square distribution is shown in Figure 13-3.

⑩ **Decide whether to reject the null hypothesis.** The sample's chi-square (from Step ⑨) of 4.09 is less extreme than the cutoff (from Step ⑩) of 9.488. The researchers cannot reject the null hypothesis; the study is inconclusive. (Having failed to reject the null hypothesis with such a large sample, it is reasonable to suppose that if there is any difference between the populations, it is quite small.)

HOW ARE YOU DOING?

1. In what situation do you use a chi-square test for goodness of fit?
2. List the steps for figuring the chi-square statistic and explain the logic behind each step.
3. Write the formula for the chi-square statistic and define each of the symbols.
4. (a) What is a chi-square distribution? (b) What is its shape? (c) Why does it have that shape?
5. Carry out a chi-square test for goodness of fit (using the .05 significance level) for a sample in which one category has 15 people, the other has 35 people, and the categories are expected to have equal frequencies. (a) Use the steps of hypothesis testing and (b) sketch the distribution involved.

ANSWERS

1. When you want to test whether a sample's distribution of people across categories represents a population that is significantly different from a population with an expected distribution of people across categories.
2. ④ **Determine the actual, observed frequencies in each category.** This is the key information for the sample studied.
⑤ **Determine the expected frequencies in each category.** Having these numbers makes it possible to make a direct comparison of what is expected to the observed frequencies.
⑥ **In each category, take observed minus expected frequencies.** This is the direct comparison of the distribution for the sample versus the distribution representing the expected population.

① **Square each of these differences.** This gets rid of the direction of the difference (since the interest is only in how much difference there is).

② **Divide each squared difference by the expected frequency for its category.** This adjusts the degree of difference for the absolute size of the expected frequencies.

③ **Add up the results of Step ② for all the categories.** This gives you a statistic for the overall degree of discrepancy.

$$3. \chi^2 = \sum \frac{(O - E)^2}{E}$$

χ^2 is the chi-square statistic; Σ tells you to sum over all the different categories; O is the observed frequency for a category; E is the expected frequency for a category.

4. (a) For any particular number of categories, the distribution you would expect if you figured a very large number of chi-square statistics for samples from a population in which the distribution of people over categories is the expected distribution.
- (b) It is skewed to the right.
- (c) It has this shape because a chi-square statistic can't be less than 0 (since the numerator, a squared score, has to be positive, and its denominator, an expected number of individuals, also has to be positive), but there is no limit to how large it can be.
5. (a) ① **Restate the question as a research hypothesis and a null hypothesis about the populations.** There are two populations:

Population 1: People like those in the sample.

Population 2: People who have an equal distribution of the two categories.

The research hypothesis is that the distribution of numbers of people over categories is different in the two populations; the null hypothesis is that it is the same.

② **Determine the characteristics of the comparison distribution.** The comparison distribution is a chi-square distribution with 1 degree of freedom (that is, $df = N_{\text{categories}} - 1 = 2 - 1 = 1$).

③ **Determine the cutoff sample score on the comparison distribution at which the null hypothesis should be rejected.** At the .05 level with $df = 1$, cutoff is 3.841.

④ **Determine your sample's score on the comparison distribution.**

④ **Determine the actual, observed frequencies in each category.** As given in the problem, these are 15 and 35.

⑤ **Determine the expected frequencies in each category.** With 50 people total and expecting an even breakdown, the expected frequencies are 25 and 25.

⑥ **In each category, take observed minus expected frequencies.** These come out to -10 (that is, $15 - 25 = -10$) and 10 (that is, $35 - 25 = 10$).

⑦ **Square each of these differences.** Both come out to 100 (that is, $-10^2 = 100$ and $10^2 = 100$).

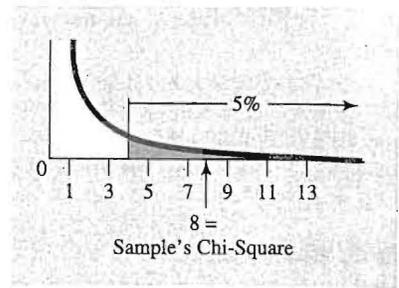
⑧ **Divide each squared difference by the expected frequency for its category.** These both come out to 4 (that is, $100/25 = 4$).

⑨ **Add up the results of Step ⑧ for all the categories.** $4 + 4 = 8$.

⑩ **Decide whether to reject the null hypothesis.** The sample's chi-square of 8 is more extreme than the cutoff of 3.841. Reject the null hypothesis; people like those in the sample are different from the expected even breakdown.

(b) See Figure 13-4.

FIGURE 13-4 For "How Are You Doing?" question 5, the chi-square distribution ($df = 1$) showing the cutoff for rejecting the null hypothesis at the .05 level and the sample's chi-square.



THE CHI-SQUARE TEST FOR INDEPENDENCE

So far, we have looked at the distribution of *one nominal variable* with several categories, such as the relationship style of men's partners. In fact, this kind of situation is fairly rare in research. We began with an example of this kind mainly as a stepping-stone to a more common actual research situation, to which we now turn.

The most common use of chi-square is when there are *two nominal variables*, each with several categories. For example, Harter et al. (1997) might have studied whether the breakdown of partners of self-focused men was the same as the breakdown of partners of other-focused men. If that were their purpose, we would have had two nominal variables. Relationship styles of partners would be the first nominal variable. Men's own relationship styles would be the second nominal variable. Hypothesis testing in this kind of situation is called a *chi-square test for independence*. You learn shortly why it has this name.

Suppose researchers at a large university survey 200 staff members who commute to work about the kind of transportation they use to get to work as well as whether they are "morning people" (prefer to go to bed early and awaken early) or "night people" (go to bed late and awaken late). Table 13-4 shows the results. Notice the two nominal variables: *types of transportation* (with three levels) and *sleep tendency* (with two levels).

CONTINGENCY TABLES

contingency table

Table 13-4 is a **contingency table**—a table in which the distributions of two nominal variables are set up so that you have the frequencies of their combinations as well as the totals. Thus, in Table 13-4, the 60 in the bus-morning combination is how many morning people ride the bus. (A contingency table is similar to the tables

TABLE 13-4 Contingency Table of Observed Frequencies of Morning and Night People Using Different Types of Transportation (Fictional Data)

		Transportation			Total
		Bus	Carpool	Own Car	
Sleep Tendency	Morning	60	30	30	120 (60%)
	Night	20	20	40	80 (40%)
	Total	80	50	70	200 (100%)

in factorial analysis of variance that you learned about in Chapter 10; but in a contingency table, the number in each cell is a *number of people*, not a *mean*.)

Table 13-4 is a 3×2 contingency table because it has three levels of one variable crossed with two levels of the other. (Which dimension is named first does not matter.) It is also possible to have larger contingency tables, such as a 4×7 or a 6×18 table. Smaller tables, 2×2 contingency tables, are especially common.

INDEPENDENCE

The question in this example is whether there is any relation between the type of transportation people use and whether they are morning or night people. If there is no relation, the *proportion* of morning and night people is the same among bus riders, carpoolers, and those who drive their own cars. Or, to put it the other way, if there is no relation, the *proportion* of bus riders, carpoolers, and own car drivers is the same for morning and night people. However you describe it, the situation of no relation between the variables in a contingency table is called **independence**.¹

independence

SAMPLE AND POPULATION

In the observed survey results in the example, the proportions of night and morning people in the sample vary with different types of transportation. For example, the bus riders are split 60–20, so three-fourths of the bus riders are morning people. Among people who drive their own cars, the split is 30–40. Thus, a slight majority are night people. Still, the sample is only of 200. It is possible that in the larger population, the type of transportation a person uses is independent of the person's being a morning or a night person. The big question is whether the lack of independence in the sample is large enough to reject the null hypothesis of independence in the population. That is, you need to do a chi-square test.

DETERMINING EXPECTED FREQUENCIES

One thing that is new in a chi-square test for independence is that you now have to figure differences between observed and expected for each *combination* of categories—that is, for each **cell** of the contingency table. (When there was only one nominal variable, you figured these differences just for each category of that single nominal variable.) Table 13-5 is the contingency table for the example survey. This time, we have shown the expected frequency (in parentheses) next to each observed frequency.

cell

Tip for Success

Always ensure that you have the same number of expected frequencies as observed frequencies. For example, with a 2×3 contingency table, there will be 6 observed frequencies and 6 corresponding expected frequencies.

The key consideration in figuring expected frequencies in a contingency table is that “expected” is based on the two variables being independent. If they are independent, then the proportions up and down the cells of each column should be the same. In the example, overall, there are 60% morning people and 40% night people; thus, if transportation method is independent of being a morning or night person, this 60%–40% split should hold for each column (each transportation type). First, the 60%–40% overall split should hold for the bus group. This would make an expected frequency in the bus cell for morning people of 60% of 80, which comes out to 48 people. The expected frequency for the bus riders who are night people is 32 (that is, 40% of 80 is 32). The same principle holds for the other columns: The

¹Independence is usually used to talk about a lack of relation between two nominal variables. However, if you have studied Chapter 11, it may be helpful to think of independence as roughly the same as the situation of no correlation ($r = 0$).

TABLE 13-5 Contingency Table of Observed (and Expected) Frequencies of Morning and Night People Using Different Types of Transportation (Fictional Data)

		Transportation			Total
		Bus	Carpool	Own Car	
Sleep Tendency	Morning	60 (48) ^a	30 (30)	30 (42)	120 (60%)
	Night	20 (32)	20 (20)	40 (28)	80 (40%)
	Total	80	50	70	200 (100%)

^aExpected frequencies are in parentheses.

Tip for Success

Be sure to check that you are selecting the correct row percentage and column total for each cell. Selecting from the wrong row or column is a common source of errors in the figuring for chi-square.

A cell's expected frequency is the number in its row divided by the total, multiplied by the number in its column.

Tip for Success

As a check on your arithmetic, it is a good idea to make sure that the expected and observed frequencies add up to the same row and column totals.

50 carpool people should have a 60%–40% split, giving an expected frequency of 30 morning people who carpool (that is, 60% of 50 is 30) and 20 night people who carpool (that is, 40% of 50 is 20), and the 70 own-car people should have a 60%–40% split, giving expected frequencies of 42 and 28.

Summarizing what we have said in terms of steps,

- ❶ Find each row's percentage of the total.
- ❷ For each cell, multiply its row's percentage by its column's total.

Applying these steps to the top left cell (morning persons who ride the bus),

❶ Find each row's percentage of the total. The 120 in the morning person row is 60% of the overall total of 200 (that is, $120/200 = 60\%$).

❷ For each cell, multiply its row's percentage by its column's total. The column total for the bus riders is 80; 60% of 80 comes out to 48 (that is, $.60 \times 80 = 48$).

These steps can also be stated as a formula,

$$E = \left(\frac{R}{N}\right)(C) \quad (13-3)$$

In this formula, E is the expected frequency for a particular cell, R is the number of people observed in this cell's row, N is the number of people total, and C is the number of people observed in this cell's column. (If you reverse cells and columns, the expected frequency still comes out the same.)

Applying the formula to the same top left cell,

$$E = \left(\frac{R}{N}\right)(C) = \left(\frac{120}{200}\right)(80) = (.60)(80) = 48.$$

Looking at the entire Table 13-5, notice that the expected frequencies add up to the same totals as the observed frequencies. For example, in the first column (bus), the expected frequencies of 48 and 32 add up to 80, just as the observed frequencies in that column of 60 and 20 do. Similarly, in the top row (morning), the expected frequencies of 48, 30, and 42 add up to 120, the same total as for the observed frequencies of 60, 30, and 30.

FIGURING CHI-SQUARE

You figure chi-square the same way as in the chi-square test for goodness of fit, except that you now figure the weighted squared difference for each cell and add these up. Here is how it works for our survey example:

$$\begin{aligned} \chi^2 &= \sum \frac{(O - E)^2}{E} = \frac{(60 - 48)^2}{48} + \frac{(30 - 30)^2}{30} + \frac{(30 - 42)^2}{42} \\ &\quad + \frac{(20 - 32)^2}{32} + \frac{(20 - 20)^2}{20} + \frac{(40 - 28)^2}{28} \\ &= 3 + 0 + 3.43 + 4.5 + 0 + 5.14 = 16.07. \end{aligned}$$

DEGREES OF FREEDOM

A contingency table with many cells may have relatively few degrees of freedom. In our example, there are six cells but only 2 degrees of freedom. Recall that the degrees of freedom are the number of categories free to vary once the totals are known. With a chi-square test for independence, the number of categories is the number of cells; the totals include the row and column totals—and if you know the row and column totals, you have a lot of information.

Consider the sleep tendency and transportation example. Suppose you know the first two cell frequencies across the top, for example, and all the row and column totals. You could then figure all the other cell frequencies just by subtraction. Table 13-6 shows the contingency table for this example with just the row and column totals and these two cell frequencies. Let's start with the Morning/Own Car cell. There is a total of 120 morning people and the other two morning-person cells have 90 in them (60 + 30). Thus, only 30 remain for the Morning/Own Car cell. Now consider the three night person cells. You know the frequencies for all the morning people cells and the column totals for each type of transportation. Thus, each cell frequency for the night people is its column's total minus the morning people in that column. For example, there are 80 bus riders and 60 are morning people. Thus, the remaining 20 must be night people.

What you can see in all this is that with knowledge of only two of the cells, you could figure out the frequencies in each of the other cells. Thus, although there are six cells, there are only 2 degrees of freedom—only two cells whose frequencies are really free to vary once we have all the row and column totals.

TABLE 13-6 Contingency Table Showing Marginal and Two Cells' Observed Frequencies to Illustrate Figuring of Degrees of Freedom

		Transportation			Total
		Bus	Carpool	Own Car	
Sleep Tendency	Morning	60	30	—	120 (60%)
	Night	—	—	—	80 (40%)
	Total	80	50	70	200 (100%)

The degrees of freedom for the chi-square test for independence is the number of columns minus 1 multiplied by the number of rows minus 1.

In a chi-square test for independence, the degrees of freedom is the number of columns minus 1 multiplied by the number of rows minus 1. Put as a formula,

$$df = (N_{\text{Columns}} - 1)(N_{\text{Rows}} - 1) \quad (13-4)$$

N_{Columns} is the number of columns and N_{Rows} is the number of rows.

Using this formula for our survey example,

$$df = (N_{\text{Columns}} - 1)(N_{\text{Rows}} - 1) = (3 - 1)(2 - 1) = (2)(1) = 2.$$

HYPOTHESIS TESTING

With 2 degrees of freedom, Table 13-2 (or Table A-4) shows that the chi-square you need for significance at the .01 level is 9.211. The chi-square of 16.07 for our example is larger than this cutoff. Thus, you can reject the null hypothesis that the two variables are independent in the population.

STEPS OF HYPOTHESIS TESTING

Now let's go through the survey example again, this time following the steps of hypothesis testing.

❶ Restate the question as a research hypothesis and a null hypothesis about the populations. There are two populations:

Population 1: People like those surveyed.

Population 2: People for whom being a night or a morning person is independent of the kind of transportation they use to commute to work.

The null hypothesis is that the two populations are the same—that, in general the proportions using different types of transportation are the same for morning and night people. The research hypothesis is that the two populations are different, that among people in general the proportions using different types of transportation are different for morning and night people.

Put another way, the null hypothesis is that the two variables are independent (they are unrelated to each other). The research hypothesis is that they are not independent (that they are related to each other).

❷ Determine the characteristics of the comparison distribution. The comparison distribution is a chi-square distribution with 2 degrees of freedom (the number of columns minus 1 multiplied by the number of rows minus 1).

❸ Determine the cutoff sample score on the comparison distribution at which the null hypothesis should be rejected. You use the same table as for any chi-square test. In the example, setting a .01 significance level with 2 degrees of freedom, you need a chi-square of 9.211.

❹ Determine your sample's score on the comparison distribution.

Ⓐ Determine the actual, observed frequencies in each cell. These are the results of the survey, as given in Table 13-4.

Ⓑ Determine the expected frequencies in each cell. These are shown in Table 13-5. For example, for the bottom right cell (Night Person/Own Car cell),

① **Find each row's percentage of the total.** The 80 people in the night person's row are 40% of the overall total of 200 (that is, $80/200 = 40\%$).

② **For each cell, multiply its row's percentage by its column's total.** The column total for those with their own car is 70; 40% of 70 comes out to 28 (that is, $.40 \times 70 = 28$).

③ **In each cell, take observed minus expected frequencies.** For example, for the Night Person/Own Car cell, this comes out to 12 (that is, $40 - 28 = 12$).

④ **Square each of these differences.** For example, for the Night Person/Own Car cell, this comes out to 144 (that is, $12^2 = 144$).

⑤ **Divide each squared difference by the expected frequency for its cell.** For example, for the Night Person/Own Car cell, this comes out to 5.14 (that is, $144/28 = 5.14$).

⑥ **Add up the results of Step ⑤ for all the cells.** As we saw, this came out to 16.07.

⑦ **Decide whether to reject the null hypothesis.** The chi-square needed to reject the null hypothesis is 9.211 and the chi-square for our sample is 16.07 (see Figure 13-5). Thus, you can reject the null hypothesis. The research hypothesis that the two variables are not independent in the population is supported. That is, the proportions of type of transportation used to commute to work are different for morning and night people.

A SECOND EXAMPLE

Riehl (1994) studied the college experience of students who were the first generation in their family to attend college. These students were compared to other students who were not the first generation in their family to go to college. (All students in the study were from Indiana University.) One of the variables Riehl measured was whether or not students dropped out during their first semester.

Table 13-7 shows the results along with the expected frequencies (shown in parentheses). Below the contingency table is the figuring for the chi-square test for independence.

① **Restate the question as a null hypothesis and a research hypothesis about the populations.** There are two populations:

Population 1: Students like those surveyed.

Population 2: Students whose dropping out or staying in college their first semester is independent of whether or not they are the first generation in their family to go to college.

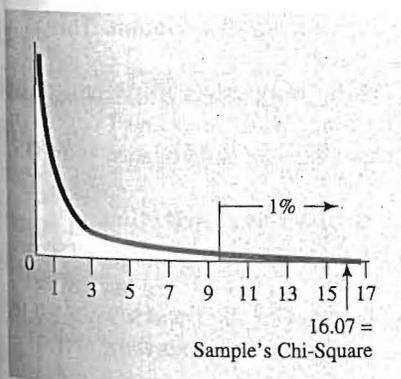


FIGURE 13-5 For the sleep tendency and transportation example, chi-square distribution ($df = 2$) showing the cutoff for rejecting the null hypothesis at the .01 level and the sample's chi-square.

TABLE 13-7 Results and Figuring of the Chi-Square Test for Independence Comparing Whether First-Generation College Students Differ from Others in First Semester Dropouts

	Generation to Go to College		Total
	First	Other	
Dropped Out	73 (57.7) [Ⓐ] [Ⓑ]	89 (103.9)	162 (7.9%)
Did Not Drop Out	657 (672.3)	1,226 (1,211.1)	1,883 (92.1%)
	730	1,315	2,045

$$df = (N_{\text{Columns}} - 1)(N_{\text{Rows}} - 1) = (2 - 1)(2 - 1) = (1)(1) = 1. \text{ ②}$$

Chi-square needed, $df = 1$, .01 level: 6.635. [Ⓒ]

$$\chi^2 = \sum \frac{(O - E)^2}{E} = \frac{(73 - 57.7)^2}{57.7} + \frac{(89 - 103.9)^2}{103.9} + \frac{(657 - 672.3)^2}{672.3} + \frac{(1,226 - 1,211.1)^2}{1,211.1}$$

$$\text{③} = \frac{15.3^2}{57.7} + \frac{-14.9^2}{103.9} + \frac{-15.3^2}{672.3} + \frac{14.9^2}{1,211.1}$$

$$\text{④} = \frac{234.1}{57.7} + \frac{222}{103.9} + \frac{234.1}{672.3} + \frac{222}{1,211.1}$$

$$\text{⑤} = 4.06 + 2.14 + .35 + .18$$

$$= 6.73 \text{ ⑥}$$

Decision: Reject the null hypothesis. ^⑤

Notes. 1. With a 2×2 analysis, the differences and squared differences (numerators) are the same for all four cells. In this example, the cells are a little different due to rounding error. 2. Data from Riehl (1994). The exact chi-square (6.73) is slightly different from that reported in the article (7.2), due to rounding error.

The null hypothesis is that the two populations are the same—that, in general, whether or not students drop out of college is independent of whether or not they are the first generation in their family to go to college. The research hypothesis is that the populations are not the same. In other words, the research hypothesis is that students like those surveyed are different from the hypothetical population in which dropping out is unrelated to whether or not you are first generation.

^② **Determine the characteristics of the comparison distribution.** This is a chi-square distribution with 1 degree of freedom.

^③ **Determine the cutoff sample score on the comparison distribution at which the null hypothesis should be rejected.** Using the .01 level and 1 degree of freedom, Table A-4 shows that you need a chi-square for significance is 6.635. This is shown in Figure 13-6.

^④ **Determine your sample's score on the comparison distribution.**

[Ⓐ] **Determine the actual, observed frequencies in each cell.** These are the results of the survey, as given in Table 13-7.

[Ⓑ] **Determine the expected frequencies in each cell.** These are shown in parentheses in Table 13-7.

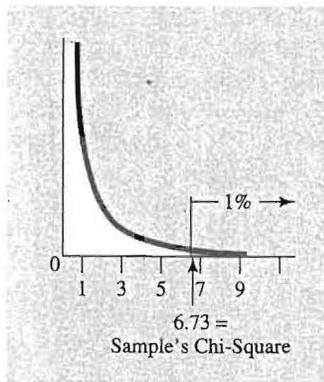


FIGURE 13-6 For the example from Riehl (1994), chi-square distribution ($df = 1$) showing the cutoff for rejecting the null hypothesis at the .01 level and the sample's chi square.

For example, for the top left cell (First Generation/Dropped Out cell),

1 Find each row's percentage of the total. The Dropped Out row's 162 is 7.9% of the overall total of 2,045 (that is, $162/2,045 = 7.9\%$).

2 For each cell, multiply its row's percentage by its column's total. The column total for the First Generation students is 730; 7.9% of 730 comes out to 57.7 (that is, $.079 \times 730 = 57.7$).

3 In each cell, take observed minus expected frequencies. These are shown in Table 13-7. For example, for the First Generation/Dropped Out cell, this comes out to 15.3 (that is, $73 - 57.7 = 15.3$).

4 Square each of these differences. These are also shown in Table 13-7. For example, for the First Generation/Dropped Out cell, this comes out to 234.1 (that is, $15.3^2 = 234.1$).

5 Divide each squared difference by the expected frequency for its cell. Once again, these are shown in Table 13-7. For example, for the First Generation/Dropped Out cell, this comes out to 4.06 (that is, $234.1/57.7 = 4.06$).

6 Add up the results of Step 5 for all the cells. As shown in Table 13-7, this comes out to 6.73. Its location on the chi-square distribution is shown in Figure 13-6.

7 Decide whether to reject the null hypothesis. Your chi-square of 6.73 is larger than the cutoff of 6.635. Thus, you can reject the null hypothesis. That is, judging from a sample of 2,045 Indiana University students, first generation students are somewhat more likely to drop out during their first semester than are other students. (Remember, of course, that there could be many reasons for this result.)

HOW ARE YOU DOING?

- (a) In what situation do you use a chi-square test for independence? (b) How is this different from the situation in which you would use a chi-square test for goodness of fit?
- (a) List the steps for figuring the expected frequencies in a contingency table. (b) Write the formula for expected frequencies in a contingency table and define each of its symbols.
- (a) Write the formula for figuring degrees of freedom in a chi-square test for independence and define each of its symbols. (b) Explain the logic behind this formula.

ANSWERS

1. (a) When you have the number of people in the various combinations of levels of two nominal variables and want to test whether the difference from independence in the sample is sufficiently great to reject the null hypothesis of independence in the population. (b) The focus is on the independence of *two nominal variables*, whereas in a chi-square test for goodness of fit the focus is on the distribution of people over categories of a *single nominal variable*.

2. (a) **1 Find each row's percentage of the total.**

1 For each cell, multiply its row's percentage by its column's total.

(b) $E = \left(\frac{R}{N}\right)(C)$.

E is the expected frequency for a particular cell, R is the number of people observed in this cell's row, N is the total number of people, and C is the number of people observed in this cell's column.

3. (a) $df = (N_{\text{Columns}} - 1)(N_{\text{Rows}} - 1)$.

df are the degrees of freedom, N_{Columns} is the number of columns, and N_{Rows} is the number of rows.

(b) df are the number of cell totals free to vary given you know the column and row totals. If you know the totals in all the columns but one, you can figure the total in the cells in the remaining column by subtraction. Similarly, if you know the total in all the rows but one, you can figure the total in the cells in the remaining row by subtraction.

4. (a) **1 Restate the question as a null hypothesis and a research hypothesis about the populations.** There are two populations:

Population 1: People like those studied.

Population 2: People whose being in a particular category of Nominal Variable A is independent of their being in a particular category of Nominal Variable B.

The null hypothesis is that the two populations are the same; the research hypothesis is that the populations are not the same.

2 Determine the characteristics of the comparison distribution. This is a chi-square distribution with 1 degree of freedom. That is, $df = (N_{\text{Columns}} - 1)(N_{\text{Rows}} - 1) = (2 - 1)(2 - 1) = 1$.

3 Determine the cutoff sample score on the comparison distribution at which the null hypothesis should be rejected. From Table A-4 (in the Appendix), for the .10 level and 1 degree of freedom, the needed chi-square is 2.706.

4 Determine your sample's score on the comparison distribution.

A Determine the actual, observed frequencies in each cell. These are shown in the contingency table for the problem.

B Determine the expected frequencies in each cell.

1 Find each row's percentage of the total. For the top row $20/80 = 25\%$; for the second row, $60/80 = 75\%$.

Nominal Variable A		Category 1		Category 2	
		10	50	10	50
Nominal Variable B		Category 1		Category 2	
10	10	10	10	10	10

4. Carry out a chi-square test for independence for the observed scores below (using the .10 significance level). (a) Use the steps of hypothesis testing, and (b) sketch the distribution involved.

