Nominal Data and the Chi-Square Tests

10

Jupiterimages/Stockbyte/Thinkstock

Chapter Learning Objectives

After reading this chapter, you should be able to do the following:

1. Describe nominal data.
2. Complete and explain the chi-square goodness-of-fit-test.
3. Complete and explain the chi-square test of independence.

Introduction

When an important development in statistical analysis took place in the early part of the 20th century, more often than not Karl Pearson was associated with it. As the text previously noted, many of those who made important contributions were members of the department that Pearson founded at University College London. Those who gravitated to Pearson's department included William Sealy Gosset, who developed the *t* tests; R. A. Fisher, who developed analysis of variance; and Charles Spearman, who did the early work on factor analysis. Although social relations among these men were not always harmonious, they were enormously productive scholars, and this was particularly true of Pearson. Besides the correlation coefficient named for him, Pearson developed an analytical approach related to Spearman's factor analysis called principal components analysis, as well as the procedures that are the subjects of this chapter, the chi-square tests. (The Greek letter *chi* [χ] is pronounced "kye" and rhymes with *sky*. Chi is the Greek equivalent of the letter *c*, rather than the letter *x*, which it resembles.)

10.1 Nominal Data

With the exception of Spearman's rho in Chapter 8, Chapters 1 through 9 have focused on procedures designed for interval or ratio data. Sometimes, however, the data are neither interval scale nor the ordinal-scale data that Spearman's rho accommodates. When the data are nominal scale, researchers often use one of the chi-square tests.

Because our focus has been so much on interval- and ratio-scale data, it might be helpful to review what makes data nominal scale. Nominal data either fit a category or do not, which is why they are sometimes referred to as "categorical data." Because of this presence-or-absence quality, analyses of nominal data are based on counting how frequently they occur, and for that reason they are also called "count data." Compared to ratio, interval, and even ordinal data, nominal data provide relatively little information. They reveal only the presence or absence of a characteristic, not how much of the characteristic, or how the individual's possession of the characteristic compares to others in the category. To illustrate: when people are classified according to whether they are

1. left-handed or right-handed, or
2. Buddhist, Jewish, Muslim, or
3. African American, Hispanic, or Native American, or
4. blue-eyed or brown-eyed, or
5. introverted or extroverted,

then the resulting data are nominal scale.

Parameters and Tests for Nominal Data

Because data of different scales provide different kinds of information, the statistical procedure used in their analyses is tailored accordingly. Because nominal data concerns itself with *frequency*, the related analytical procedures—in this instance, the chi-square tests—are based on how many individuals are in a particular category. To put it simply, the measurement procedure for chi-square is counting.

Recall from Chapter 8 that tests for nominal data are nonparametric tests. The "no parameters" element means that employing these tests does not obligate the researcher to meet most of the traditional parameters, or requirements, for statistical tests. The *t* tests and ANOVA, for example, require that the dependent variable be normally distributed in its population. The Pearson correlation and ordinary least-squares regression upon which it is based (Chapter 9) also require that the *x* and *y* variables be normally distributed. Like Spearman's rho (Chapter 8), which is also a nonparametric test, the chi-square tests set normality and homogeneity requirements aside; they are "distribution free" tests. However, in the statistical equivalent of no such thing as a free lunch, all of this analytical flexibility has a cost. The chi-square's drawback has to do with the power of the test, which the chapter will later discuss.

When working with nominal data, most of the descriptive statistics used to this point are irrelevant. As the most frequently occurring value, the mode, of course, can still be calculated, but the means and medians to which we compared the mode in order to determine skew require at least interval data. Nominal data offer no standard deviation or range values to examine to evaluate kurtosis. It is just as well that the chi-square tests are nonparametric since most of the values needed to determine normality are unavailable in any case.

10.2 The Chi-Square Tests

This chapter explains two chi-square tests. The analysis in both tests is based on comparing the frequency (count) with which something actually occurs, compared to the frequency with which it is expected to occur.

The first test is called the $1 \times k$ ("one by kay"), or the **goodness-of-fit chi-square test**. Like the independent variable in the one-way ANOVA, this test accommodates just one variable, but that one variable can have any number of categories greater than one. For instance, a psychologist could analyze whether those participating in court-ordered group therapy sessions for drug addiction represent some vocations more than others. In that case, the variable is vocation. It can have any number of manifestations (clerical workers, laborers, the unemployed, educators, and so on), but the only variable is vocation.

The second chi-square test the chapter takes up is called the $r \times k$ ("are by kay"), or the **chi-square test of independence**. This test accommodates two variables. Each of the two variables can be further divided into any number of categories. A researcher might be interested in whether marital status (single never-married, married, divorced) is related to graduating on-time among university students (graduated within four years, did not graduate within four years).

The Goodness-of-Fit or $1 \times k$ Chi-Square Test

This test asks whether an outcome is different enough from an initial hypothesis that research should conclude that the difference is not likely to have occurred by chance. The focus on whether an outcome might be expected to have occurred by chance makes the $1 \times k$ like all significance tests. The important difference is that it accommodates a nominal-scale, dependent variable. For some illustrations of problems that might involve the $1 \times k$ chi-square, consider the following:

Fuse/Thinkstock

Do women and men pursue psychology majors in equal numbers? A 1 × *k* chi-square test will provide an answer.

Those responsible for recruitment in the university's college of social sciences wonder whether opting for a psychology major relates to the potential students' gender. The variable is the gender of the student, with two categories: female and male. The research questions whether, in a randomly selected group of psychology majors, male or female students occur with significantly different frequencies.

This problem is similar to an independent groups *t* test in that it has two independent categories. The difference in the two tests is whether the count or frequency with which subjects occur in each category significantly strays from a pre-determined hypothesis, rather than whether the groups' means, which nominal data cannot provide, are significantly different from each other.

In a second example, a military psychologist wants to know whether recruits represent urban, suburban, semi-rural, and rural backgrounds in similar proportions. The psychologist selects a random sample of 50 recent recruits and determines their demographic origins. The variable is the population characteristics of the recruits' origins. In the absence of information to the contrary, the researcher's hypothesis is probably that recruits come from different areas of the country in equal proportions. If the psychologist determines that twice as many people live in suburban areas as in semi-rural areas, however, perhaps the corresponding hypothesis is that recruits from suburban areas will be twice as numerous as those from rural areas. The psychologist might also hypothesize that patriotism, which may affect the individual's desire to join the military, runs higher in rural than in urban populations, so that the expectation is that rural recruits will occur in greater proportions than those from urban environments. With multiple groups represented in this hypothetical problem, it bears some similarity to a one-way ANOVA, but without any sums of squares to analyze.

Try It!: #1

How many variables will the 1 × *k* chi-square accommodate?

Without wishing to belabor the point, the independent *t* test and analysis of variance divide subjects into two or more categories, with each category characterized by a different level, or manifestation, of the independent variable. The study analyzes how the different levels affect some other variable, the dependent variable. The chi-square similarly has two or more categories, but it analyzes the frequency with which individuals are distributed into those different categories.

Observed and Expected Frequencies

To restate our approach, then, the measurement involved in chi-square analysis is simply counting. Researchers who use this analysis are interested in the frequency with which something occurs in a category. More specifically, rather than comparing sample means to population means, or sample means to each other, chi square examines differences between the frequency with which individuals occur in a particular category (symbolized by f_o), and the frequency with which they were expected to occur (symbolized by f_e).

The f_e and f_o values are simply the number of observations in each category; they are frequency counts. When the expected number varies sufficiently from the observed number, the result is statistically significant.

The Chi-Square Test Statistic

The test statistic for the chi-square test is as follows:

Formula 10.1

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

where

χ^2 = the value of the chi-square statistic

f_o = the frequency observed in the particular category

f_e = the frequency expected in the particular category

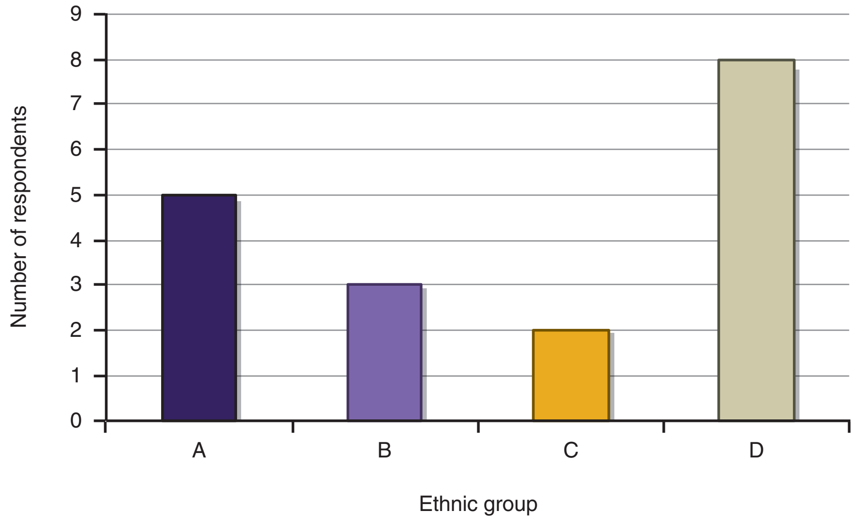
Studying the test statistic for the chi-square test is quite revealing. To calculate the value of this statistic, start with these steps:

1. Count the number in each category (f_o).
2. Determine the number expected in each category (f_e). When the assumption is that all categories are equal, this will be the total number of subjects divided by the number of categories.
3. As a quick check before continuing, note that the sum of the f_e categories must equal the sum of the f_o categories. Then, perform the following mathematical operations:
   a. Subtract f_e from f_o.
   b. Square the difference.
   c. Divide the squared difference by f_e.
   d. Sum the squared differences divided by f_e across the categories.
   e. Compare to the critical value of chi-square for the number of categories, minus 1 degree of freedom. (The critical values of chi-square appear in Table 10.2.)

A Goodness-of-Fit (1 × *k* Chi-Square) Problem

Using the ethnic diversity of voters as an example, a psychologist who has examined voting patterns and ethnicity perhaps wishes to test the assumption that voting in a general election is unrelated to voters' ethnic group membership. On election day, the psychologist journeys to a polling place in an ethnically diverse part of the city and administers a brief survey to those who have just voted. One question concerns the respondents' ethnic group. Figure 10.1 shows the data for the 18 people who completed the survey.

Figure 10.1: Voter participation data



Although the calculations are not difficult, determining the value of chi-square involves some arithmetic. An easy way to keep track of the calculations is to arrange the data into a table like Table 10.1. The rows are numbered to be consistent with the numbered steps listed after Formula 10.1 for calculating the chi-square statistic. The results from the survey are the frequency-observed values in the first line of the table. The frequency-expected values are n divided by the number of categories: $18 \div 4 = 4.50$. That value indicates that if the ethnic group membership of the voters in this group is exactly equivalent, 4.50 of the respondents will declare for each group. Do not let the .50 value in each f_e distract you. Although the f_o numbers have no chance of any such value, that f_e value is the same for all groups; the issue is whether the $f_o - f_e$ differences are significantly different from category to category.

Table 10.1: A goodness-of-fit chi-square problem for voting patterns

Value	Ethnic group A	Ethnic group B	Ethnic group C	Ethnic group D
1. f_o	5	3	2	8
2. f_e	4.50	4.50	4.50	4.50
3a. $f_o - f_e$	0.50	−1.50	−2.50	3.50
3b. $f_o - f_e^2$	0.25	2.25	6.25	12.25
3c. $f_o - f_e^2 / f_e$	0.06	0.50	1.39	2.72
3d. $\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} = 0.06 + 0.50 + 1.39 + 2.72 = 4.67$				

Determining Significance

For this problem, the value of chi-square is $\chi^2 = 4.67$. Having calculated the statistic, the researcher needs something with which to compare it, a critical value, and—as with other tests—the critical value is indexed to degrees of freedom for the problem.

- The degrees of freedom for a goodness-of-fit problem are the number of categories in the problem, minus 1.
- With subjects in the voting participation problem divided into four different ethnic groups, there are $4 - 1 = 3$ *df*.

Try It!: #2

Why can chi-square values never be negative?

The critical values for chi-square in Table 10.2 (also table B.7 in Appendix B) are arranged by degrees of freedom down the left side, and the level at which the test is conducted across the top.

Table 10.2: The critical values of chi-squared

df	$p = 0.05$	$p = 0.01$	$p = 0.001$
1	3.84	6.64	10.83
2	5.99	9.21	13.82
3	7.82	11.35	16.27
4	9.49	13.28	18.47
5	11.07	15.09	20.52
6	12.59	16.81	22.46
7	14.07	18.48	24.32
8	15.51	20.09	26.13
9	16.92	21.67	27.88
10	18.31	23.21	29.59
11	19.68	24.73	31.26
12	21.03	26.22	32.91
13	22.36	27.69	34.53
14	23.69	29.14	36.12
15	25.00	30.58	37.70
16	26.30	32.00	39.25
17	27.59	33.41	40.79
18	28.87	34.81	42.31
19	30.14	36.19	43.82
20	31.41	37.57	45.32
21	32.67	38.93	46.80
22	33.92	40.29	48.27
23	35.17	41.64	49.73
24	36.42	42.98	51.18
25	37.65	44.31	52.62
26	38.89	45.64	54.05
27	40.11	46.96	55.48
28	41.34	48.28	56.89
29	42.56	49.59	58.30
30	43.77	50.89	59.70

Source: Virginia Tech, Quantitative Population Ecology. (n.d.). Table of chi-square statistics. Retrieved from https://web.archive.org/web/20150930232540/http://alexei.nfshost.com/PopEcol/tables/chisq.html

To keep the size of the table manageable, the values are carried to just two decimals. For consistency, the final values of chi-square will be also rounded to two decimal places.

The critical value for a chi-square problem with $df = 3$ and $p = 0.05$ is 7.82. To distinguish the calculated value of chi-square from the critical value, follow the same pattern adopted for the other tests. First, calculate the value from the test results:

$\chi^2 = 4.667$ for the calculated value

This value is compared to the critical value, which is indicated by the subscripts for the level of probability of alpha error for the test (0.05) and its degrees of freedom.

$$\chi^2_{0.05(3)} = 7.82$$

With a calculated value less than the critical value from the table, the differences in the ethnicity of the voters in these four groups are *not* statistically significant; the researcher attributes the differences to chance. That may seem like a strange conclusion when the differences in the f_o values are so substantial. The explanation goes back to the heart of what a goodness-of-fit test is designed to analyze. Pearson focused not on the differences (in this case) between ethnic groups, but on the differences between what was observed and what could be expected to occur if the initial hypothesis is valid. The comparison is not how ethnic group C compares to ethnic group D, for example, but how the f_o and f_e values *within each category* differ. The difference in ethnic group C between 2 (f_o) and 4.5 (f_e) is a different matter than the difference between 2 (ethnic group C) and 8 (ethnic group D). The result indicates that across the four groups, the difference between f_o and f_e does not vary enough for the result to be significant. This much difference could have occurred by chance.

The Hypotheses in a Goodness-of-Fit Test

Consistent with the other tests of significant differences (*z, t, F*), the null hypothesis in the chi-square tests is the hypothesis of no difference, symbolized by $H_0: f_o = f_e$. However, as the symbols indicate, the difference between what is observed (f_o) and what is expected (f_e) is what is at issue. When the f_o and f_e for a particular category show an approximate equivalence, we fail to reject the null hypothesis. The alternate hypothesis is that what is observed is significantly different from the expected, $H_A: f_o \neq f_e$. Literally, the frequency observed does not equal the frequency expected.

In the case of ethnic-group voting behavior, the statistical decision is to fail to reject H_0. The differences between what was observed and what was expected across the four groups were not great enough to be statistically significant.

A Goodness-of-Fit Problem with Nonequivalent Frequencies Expected

In the ethnicity and voting problem, the researcher tested the assumption that what could be expected (f_e) did not differ from group-to-group among the four ethnic groups. However, researchers do not always assume equivalent f_e values. When the hypothesis is that the frequencies will vary from category to category, the different f_e values must be calculated for the categories in order to reflect the hypothesis.

Perhaps a psychologist working with the military observes that service personnel exposed to combat situations for more than six months appear to experience post-traumatic stress disorder (PTSD) about three times more frequently than those with less than six months of combat exposure. To test this hypothesis, the f_e values will need to indicate the different expectations. Gathering data for a group of service personnel, the psychologist has the following:

Of 429 service personnel, 154 were exposed to combat situations for less than six months and the other 275 had six months or more of combat exposure.

Those 154 and 275 numbers indicate the f_o values for the problem. As always with chi-square problems, the f_e values must sum to the same 429 value, but the f_e numbers must also reflect the 3-to-1 hypothesis. To determine the f_e values, follow these steps:

1. Take the ratio, 3 to 1 in this example.
2. Add the elements of the ratio together: $3 + 1 = 4.3$.
3. Divide the total number of subjects, n, by the sum of the ratio elements:
   $429 \div 4 = 107.25$

The f_e value for those exposed to combat situations for less than six months will be 1×107.25. The f_e value for those exposed to combat situations for six months or more will be $3 \times 107.25 = 321.75$.

The balance of the problem involves the same procedure used in Table 10.1 except that there are only two categories. The problem is completed in Table 10.3.

Table 10.3: A goodness-of-fit chi-square with unequal frequencies

Combat experience		
Value	**Less than 6 months**	**6 months or more**
f_o	154	275
f_e	107.25	321.75
$f_o - f_e$	46.75	−46.75
$f_o - f_e^2$	2185.56	2185.56
$f_o - f_e^2 / f_e$	20.38	6.79
$\sum\frac{(f_o - f_e)^2}{f_e} = \chi 2 = 27.17$		

Note that the null hypothesis reflects the assumption that there will be no difference between what was expected and what was observed. In this particular problem, the hypothesis is that $f_o \neq f_e$. What the psychologist expected was a PTSD rate that was three times higher among service personnel who had been exposed to combat situations for six months or more than among personnel who had less than six months of exposure. The value calculated is $\chi^2 = 27.17$, and the associated critical value from the table for $p = 0.05$ and one degree of freedom is $\chi^2_{0.05(1)} = 3.84$. With a calculated value of chi-square higher than the critical value from the table, the result is statistically significant. What does that mean in terms of what the psychologist expected to occur? It means that among these 429 service personnel, H_0: $f_o = f_e$ must be rejected. The rate of PTSD is *not* three times higher among those with six months or more of combat exposure. Just examining the f_o values indicates that the PTSD rate is about double for those with six months or more of combat experience compared to those with less than six months of exposure. The psychologist's expectation does not hold for these personnel.

The Chi-Square and Statistical Power

Before a chi-square result is significant, the difference between what is expected and what actually occurs must be substantial. Nominal data cannot match the sophistication of ratio, interval, or even ordinal data, because the data used in a chi-square problem reflect only frequency. They do not contain the information that can indicate the subtle differences in measured qualities that data of the other scales reflect. The analytical price paid for relying exclusively on nominal data is power. Recall that in statistical terms, power refers to the probability of detecting significance.

Users of distribution-free tests like chi-square gain great flexibility. They need not make any judgments about normality or linearity, but as the chapter earlier stated, such flexibility comes at a cost. The flexibility's ever-present companion is an increased probability of a type II error. The failure to detect significance is higher with these distribution-free tests than with the procedures in the earlier chapters. The departures from the $f_o = f_e$ assumption must be quite extreme before they can be chalked up to anything except sampling variability. That was the situation in the first problem on voter turnout, when it appeared that there were substantial differences in the voting behavior of people of different ethnic backgrounds, but they were nonsignificant nevertheless.

Remember that type I and II errors are related, however. When the likelihood of failing to detect a statistically significant difference is higher than usual (a type II error), the probability of finding significant difference in error (a type I error) is correspondingly reduced. Although the chi-square tests have a relatively high incidence of type II error, at least the probability of type I error is lower than with many of the parametric alternatives.

These characteristics notwithstanding, the loss of power from using a chi-square test is often a nonissue. If the data are nominal scale to begin with, researchers can make no decision about what kind of test to use; their only choice is to use one that accommodates nominal data. Power becomes an issue when data are ordinal scale or higher but some requirement such as normality is suspect. In that case the analyst must make a decision about the best course: rely on a traditional parametric test in spite of suspect normality, or adopt a nonparametric test with relaxed requirements but also consequent loss of power.

Try It!: #3

The risk of which type of decision error increases with chi-square problems?

For example, suppose that the voting-survey researcher also asked respondents how many times in the last 15 years they had participated in elections. The researcher may have intended to use analysis of variance to determine whether ethnic-group differences exist in the level of participation in elections. However, 18 respondents divided among four ethnic groups is a very small sample size. The two people in ethnic group C provide little basis for completing an ANOVA; the sample is simply too small. With groups so small, just one or two extremely low or extremely high scores will skew results, making normality an issue. In such a case, a shift to a nonparametric test like the goodness-of-fit test, where neither the normality of the data nor the sample size is central, is likely to be more appropriate.

10.3 The Chi-Square Test of Independence

Both of the chi-square problems we have worked in this chapter have been goodness-of-fit ($1 \times k$) tests. Like all goodness-of-fit tests, the first problem involved just one variable, although it was divided into four categories to reflect the ethnicity of the voter. The second problem's one variable—the incidence of post-traumatic stress disorder among service personnel—was divided into two categories: those deployed to combat situations for less than six months and those deployed for six months or more. The goodness-of-fit test works well for any number of data categories related to a single, nominal-scale, variable.

Sometimes the question is more complex. Maybe the question involves the ethnicity of the respondent *and* whether the individual voted in the last election. Or perhaps the PTSD problem looks at the incidence among service personnel of different deployment periods *and* whether the service personnel were men or women. Both of those examples involve two variables. In any statistical analysis, researchers add variables to be able to explain the scoring variability more completely. Although z, t, and one-way ANOVA procedures are extremely important, they, like the goodness-of-fit test, are all restricted to a single independent variable. Relatively few outcomes, particularly related to human subjects, can be adequately explained by a single variable. People are too complicated.

Both the chi-square tests in this chapter compare what is observed to what is expected, but in the goodness-of-fit test, f_o to f_e differences test a hypothesis about frequencies in categories. The chi-square test of independence uses the f_o to f_e differences to test whether the two variables being examined, as the name suggests, operate independently of each other. This second chi-square test is also known as the $r \times k$ chi-square for reasons that will become clear below.

The Hypotheses in the Chi-Square Test of Independence

The null and alternate hypotheses look the same as they do in the $1 \times k$:

- $H_0: f_o = f_e$
- $H_A: f_o \neq f_e$

The hypotheses are reminders that the problem seeks to resolve how the frequencies observed compare to the frequencies expected. As before, H_0 is rejected for calculated values of chi-square that are larger than the table value. However, in an $r \times k$ chi-square problem, the null hypothesis also indicates that the two variables are unrelated: the frequency with which one variable occurs does not affect the frequency of the other. If the null hypothesis is rejected (indicating that the two variables *are* related), the analysis has another step: determining the strength of the relationship between the variables, as the following example demonstrates.

A Chi-Square Test of Independence Problem

Let us return to the ethnicity and voting behavior problem. The researcher now decides to expand the study to gain a more comprehensive view of how ethnicity and the tendency to vote might be related. With a list of registered voters in hand, the researcher sends out several questionnaires asking, among other things, the individual's ethnicity and whether the person voted in the last national election. With 36 responses, the researcher has gathered the following data:

Ethnic Group A: Of the 12 respondents, 8 voted

Ethnic Group B: Of the 8 respondents, 2 voted

Ethnic Group C: Of the 8 respondents, 3 voted

Ethnic Group D: Of the 8 respondents, 7 voted

The Contingency Table

In this two-variable chi-square test, a table called a **contingency table** helps to keep the data organized. The subsets of one variable are reflected in the rows of the table (the r in the $r \times k$), and the subsets of the other variable are listed in the table columns or categories (the k in the $r \times k$). Table 10.4, an example of a contingency table, shows the breakdown of ethnicity and voting behavior data results.

Table 10.4: Contingency table

	Voted in last election		
Ethnic group	Yes	No	Total number of respondents
A	8 a	4 b	12
B	2 c	6 d	8
C	3 e	5 f	8
D	7 g	1 h	8
Totals	20	16	36

The subject's ethnicity is indicated in the rows, which end with a row for column totals. The columns indicate how many voted and how many did not, as well as the total number in each ethnic group. Each of the 8 cells is identified with a letter, which the researcher will use to calculate the chi-square value. Cell a, for example, indicates that eight of the people in ethnic group A voted.

Calculating the Frequency-Expected Values, f_e

As it was with the $1 \times k$ chi-square test, the frequency-observed (f_o) values reflect what actually occurred. The frequencies expected (f_e) are calculated differently than they were in the one-variable test, however. Because each value reflects the influence of *two* variables (each cell in the contingency table is at the intersection of a row *and* a column), a researcher cannot just divide the number of subjects by the number of cells and use the same f_e value for each cell. The fact that the two variables might have a different impact on some combinations than on others disallows such an approach. The f_e value must reflect the impact that both variables have on the outcome in each combination. The f_e value for cell a in the $r \times k$ chi-square test is completed this way:

> The f_e value for a particular cell is the row total for that cell times the column total for that cell, divided by the total number of subjects.

The f_e value for cell a, for example, is the row total for cell a (12) times the total for the column in which cell a is found (20), divided by the total number of subjects (36): $(12 \times 20) \div 36 = 6.67$.

The f_e calculations for cells b through h follow:

b: $(12 \times 16) \div 36 = 5.33$

c: $(8 \times 20) \div 36 = 4.44$

d: $(8 \times 16) \div 36 = 3.56$

e: $(8 \times 20) \div 36 = 4.44$

f: $(8 \times 16) \div 36 = 3.56$

g: $(8 \times 20) \div 36 = 4.44$

h: $(8 \times 16) \div 36 = 3.56$

Using the frequency-observed values in the cells of the contingency table and the calculated frequency-expected values, the researcher can create the same table used in the goodness-of-fit problems earlier:

For each of the eight cells,

1. subtract f_e from f_o,
2. square the difference,
3. divide the squared difference by f_e, and
4. sum the results from each of the cells, which is the value of chi-square.

Table 10.5 completes the ethnicity and voting behavior problem.

Table 10.5: The chi-square test of independence: Ethnicity and voting behavior

Value	a	b	c	d	e	f	g	h
f_o	8	4	2	6	3	5	7	1
f_e	6.67	5.33	4.44	3.56	4.44	3.56	4.44	3.56
$f_o - f_e$	1.33	−1.33	−2.44	2.44	−1.44	1.44	2.56	−2.56
$f_o - f_e^2$	1.77	1.77	5.95	5.95	2.07	2.07	6.55	6.55
$f_o - f_e^2 / f_e$	0.27	0.33	1.34	1.67	0.47	0.58	1.48	4.84
$\Sigma \frac{(f_o - f_e)^2}{f_e} = \chi 2 = 7.98$								

Degrees of Freedom in the Chi-Square Test of Independence

For a chi-square test of independence, the number of degrees of freedom is determined by the number of categories of one variable, minus one, times the number of categories in the other variable, minus one. For this problem, which has four rows and two columns in the contingency table, the number of degrees of freedom is $(4 - 1) \times (2 - 1) = 3$.

From the table for critical values of chi-square (Table 10.2), the value for 3 degrees of freedom and testing for alpha error at $p = 0.05$ is $\chi^2_{0.05\,(3)} = 7.82$.

Interpreting the *r* × *k* Result

By conducting the chi-square test of independence, the researcher is asking, "Is ethnicity related to whether the individual votes in a national election?" As with the first test, Pearson compared what actually occurs in a particular situation (f_o) to what can be expected, but with the test of independence, what is expected is based on the hypothesis that the variables involved are unrelated, uncorrelated. The null hypothesis for this test is based on that uncorrelated hypothesis, so the f_e values are calculated to indicate what to expect when the variables are independent of each other. The substantial variations of f_o from f_e prompt larger values of chi-square. If the variations between f_o and f_e are great enough that they meet or exceed the critical value, the statistical decision is to reject the null hypothesis and conclude that the variables are not independent of each other; they are correlated.

Try It!: #4

How many categories in either variable can the $r \times k$ chi-square accommodate?

The psychologist's data on ethnicity and voting behavior produced a calculated value of chi-square which exceeds the critical value from Table 10.2 for $p = 0.05$ and three degrees of freedom. It is statistically significant. The lack of independence indicates that voting behavior for some ethnic groups is different than it is for those of other ethnic groups.

Classifying the *r* × *k* Test

Earlier chapters organized statistical tests according to whether they addressed the hypothesis of difference or the hypothesis of association. Tests like *z*, *t*, and ANOVA (*F*) are analyses of significant differences between samples and populations, or differences between samples. The Pearson and Spearman correlation procedures quantified the strength of the relationship between two variables; they addressed the hypothesis of association. The chi-square test of independence does not fit this either-or classification. The researcher initially questioned whether there are significant differences in voting behavior among the different ethnic groups, which makes the $r \times k$ sound a lot like an ANOVA. But the analysis is based on whether ethnicity and voting behavior are related, a question that makes the test more of a correlation analysis. The $r \times k$ test addresses both of those main hypotheses. It straddles the ground between the hypotheses of difference and association.

Phi Coefficient and Cramér's V

Because the researcher's results indicate that ethnicity and voting behavior are not independent, a supplementary question follows: How related are the two variables? This is a correlation question; whenever the $r \times k$ chi-square value is statistically significant, the strength of the relationship must be determined.

Research uses several correlation procedures for nominal data. Depending upon circumstances, the **phi coefficient** (symbolized by the Greek letter *phi*, φ), or a variation of the phi coefficient called **Cramér's V** are the correlation procedures we'll use. Phi coefficient is appropriate when at least one of the variables has only two levels. In the ethnicity and voting behavior problem, the individual voted or did not—these are the only two categories. If both variables have three or more levels, the researcher must choose Cramér's V.

Both Cramér's V and phi coefficient employ the previously determined value of chi-square in their calculations, which means that most of the work has been done before the strength-of-the-correlation question is resolved. In addition, the correlation values require no separate significance tests. Since these values are calculated only when the chi-square value is statistically significant, any related φ or V coefficient will likewise be significant. For the same reason, the phi coefficient and Cramér's V may also be considered effect sizes. Like all significance tests, the initial $r \times k$ result indicates only whether the chi-square value is or is not statistically significant. Phi and Cramér's V provide information about the magnitude of the relationship.

The formula for the phi coefficient follows:

Formula 10.2

$$\varphi = \sqrt{\left(\frac{\chi^2}{n}\right)}$$

where

χ^2 = the value already available from the $r \times k$ problem

n = the total number of observations (the sum of all the f_o values)

The formula for Cramér's V follows:

Formula 10.3

$$V = \sqrt{\frac{\chi^2}{n(k-1)}}$$

where

χ^2 = the value already available from the $r \times k$ problem

n = the total number of observations

$k - 1$ = the number of rows or columns—whichever is less—minus 1

When there are only two levels of either or both of the variables, then $V = \varphi$ because the denominator in V would be $n \times 1$.

For the ethnicity and voting behavior problem, $\chi^2 = 7.98$, and it is statistically significant. The phi coefficient is

$$\varphi = \sqrt{\left(\frac{\chi^2}{n}\right)}$$
$$= \sqrt{\left(\frac{7.98^2}{36}\right)}$$
$$= \mathbf{0.47}$$

Note that the chi-square value is symbolized by χ^2. Sometimes students note the exponent and proceed to square the value, 7.98 in this case. Keep in mind that the exponent refers to the way the value was initially calculated $(f_o - f_e)^2$ and that to square it again as part of the phi or Cramér's V coefficient would be an error.

The phi coefficient value is interpreted like any other correlation value, except for two things: 1. A researcher does not need to check for significance because that was already established for the chi-square value upon which it is based. 2. Because the coefficient is reached by calculating a square root, phi can never be negative. Correlations closer to 1.0 are stronger; values closer to 0 are weaker. The correlation here between ethnicity and the tendency to vote might be described as a moderate correlation.

Apply It!

Chi-Square Test for Locating Group Homes

Because working with the mentally ill in group homes rather than in secure facilities is more economical and in some cases more effective, a group of mental health professionals has decided to gauge whether public opinion favors outpatient treatment of mental health inmates convicted of nonviolent crimes. The group assumes that there will be differences of opinion in rural, semi-rural, suburban, and urban locations, since some of those locations are more likely to house group homes than others. Researchers sent to people in different locations 3,000 questionnaires, asking residents whether they favor releasing certain classifications of the mentally ill to the custody of group homes for treatment and rehabilitation into society. The health professionals received 1,434 completed questionnaires. The null hypothesis (H_0: $f_o = f_e$) holds that whether people favor group homes rather than assignment to a secure facility will be unrelated to where they live. The alternate hypothesis (H_A: $f_o \neq f_e$) is that the two variables are related. Table 10.6 summarizes the survey data results.

Table 10.6: Questionnaire results concerning release of mentally ill to group homes

	Favor group homes		
Type of neighborhood	**Yes**	**No**	**Totals**
Urban	248	128	**376**
Suburban	212	143	**355**
Small town	198	95	**293**
Semi-rural	33	67	**100**
Rural	133	177	**310**
Totals	**824**	**610**	**1434**

Because this 5 × 2 chi-square involves so many repetitious calculations, this would be a good problem to complete using Excel. The related calculations are demonstrated below, but in the meantime, note that $\chi^2 = 75.47$. By comparison, the critical value is $\chi^2_{0.05(4)} = 9.49$. With a statistically significant value of chi-square, the psychologists reject the null hypothesis and proceed to calculate phi:

$$\varphi = \sqrt{(\frac{\chi^2}{n})}$$

Inserting the values for χ^2 and for n, they discover that

$$\varphi = \sqrt{(\frac{75.47}{1{,}434})}$$

$$\varphi = \mathbf{0.23}$$

Researchers initially chose to break down responses by area of residence because they assumed that there might be differences of opinion about the value of establishing group homes depending upon the likelihood of one near respondents' places of residence. The initial $r \times k$ results support this explanation. The two variables (where one lives and whether one supports) do not operate independently. However, the phi coefficient value makes it clear that the relationship is not a particularly strong one. A value of $\varphi = 0.23$ suggest that where one lives and whether one supports group homes as an alternative to secure institutions are only modestly related.

Apply It! boxes written by Shawn Murphy

A 3 × 3 Problem

An earthquake and an accompanying tsunami have wreaked havoc in a particular country. A psychologist is analyzing the appeal of three different outreach programs (A, B, and C) to people who describe their losses as primarily material, primarily emotional, or a combination. The three programs offer counseling regarding insurance and financial recovery (A), therapy focused on coping with change and the loss of loved ones (B), and therapy that includes elements of both financial recovery and emotional disruption (C). The data are as follows:

- Among 8 people describing their losses as primarily material, 6 opt for A, 2 for B, and 0 for C.
- Among 6 people describing their losses as primarily emotional, 0 opt for A, 2 for B, and 4 for C.
- Among 6 people describing their losses as a combination, 0 opt for A, 5 opt for B, and 1 for C.

The question is whether people who have different kinds of losses choose to be involved in different kinds of counseling programs. Table 10.7 depicts the contingency table for these data.

Table 10.8 shows the solution for this problem.

Table 10.7: Contingency table for a 3 × 3 problem

Type of Loss	Program			Totals
	A	B	C	
Material	6a	2b	0c	**8**
Emotional	0d	2e	4f	**6**
Material and Emotional	0g	5h	1i	**6**
Totals	**6**	**9**	**5**	**20**

Table 10.8: Another $r \times k$ problem: Outreach programs and the type of loss suffered

Value	a	b	c	d	e	f	g	h	i
f_o	6	2	0	0	2	4	0	5	1
f_e	2.4	3.6	2	1.8	2.7	1.5	1.8	2.7	1.5
$f_o - f_e$	3.6	−1.6	−2	−1.8	−0.7	2.5	−1.8	2.3	−0.5
$f_o - f_e^2$	12.96	2.56	4	3.24	0.49	6.25	3.24	5.29	0.25
$f_o - f_e^2 / f_e$	5.4	0.71	2	1.8	0.18	4.17	1.8	1.96	0.17
$\Sigma\frac{(f_o - f_e)^2}{f_e} = \chi 2 = 18.19$									

Completing the calculations produces a chi-square value of $\chi^2 = 18.19$. With four degrees of freedom $(3 - 1) \times (3 - 1)$, the critical value for testing for alpha error at $p = 0.05$ $({\chi^2}_{0.05(4)})$ is $9.49 = 18.19_{0.05(4)}$. The type of loss the individual suffered is related to the kind of counseling the individual chooses. Because both variables have more than three levels, researchers use Cramér's V as the correlation procedure to determine the strength of the relationship between type of loss and kind of counseling.

With $n = 20$ and the number of either the rows or columns $= 3$, $k - 1 = 2$. Cramér's V is calculated as follows:

$$V = \sqrt{\frac{\chi^2}{n(k-1)}} = \sqrt{\frac{18.19}{20(2)}}$$

$$= \mathbf{0.67}$$

The $r \times k$ result indicated that there are significant differences between what was observed and what would have been expected had there been no relationship between the type of counseling program and the kind of loss victims suffered. Cramér's V quantified the strength of that relationship. Besides being statistically significant (which the significant χ^2 value establishes), the strength of the relationship appears to be substantial.

Apply It!
Public Policy Research

Mark Massel/iStock/Thinkstock

Psychologists associated with a particular national psychological association are interested in how people view the insanity defense in criminal trials. Their research is based on anecdotal data that suggest that the viability of an insanity defense has something to do with the region of the country in which the respondent lives. The psychologists develop a questionnaire on mental health issues where one of the items is in Likert-type form and is posed as follows:

> At certain times, circumstances make otherwise rational people unaccountable for their actions.

The response choices are the following:

- disagree
- neither agree nor disagree
- agree

The researchers solicit responses from five different areas of the country identified as these:

- the Northeast
- the South
- the Midwest
- the Northwest
- the West

This is a 5 (regions) × 3 (type of response) chi-square test of independence. The question the procedure will answer is whether where the respondent lives is independent of the respondent's response to this particular item. The null hypothesis associated with such a question is that the two are unrelated, indicated by no statistically significant differences between f_o and f_e. Results for the survey are shown in Table 10.9.

Table 10.9: Agreement with insanity defense according to region

Campaign	Disagree	Agree	Strongly Agree	Totals
Northwest	1	4	13	**18**
Northeast	2	5	9	**16**
Midwest	10	4	2	**16**
West	11	5	1	**17**
South	13	3	0	**16**
Totals	**37**	**21**	**25**	**83**

(continued)

(continued)

Having completed the analysis, the psychologists find that

$$\chi^2 = 42.04.$$

The number of degrees of freedom is $(5 - 1) \times (3 - 1) = 8$, so

$$\chi^2_{0.05(8)} = 15.51.$$

The calculated value of chi-square is larger than the critical value of chi-square from Table 10.2, indexed by the probability of a type I error and the degrees of freedom. For these data at least, where respondents live does relate to how they answer the item about accountability. To determine the strength of the relationship between the region of residence and the response type, the psychologists will need Cramér's V, since both variables have more than two levels.

In this case, the response type (agree, neither agree . . .) represents the variable with the fewest number of levels (3), so $k - 1 = 2$, and $n =$ the total number of responses (83). Therefore,

$$V = \sqrt{\frac{\chi^2}{n(k-1)}}$$

$$= \sqrt{\frac{42.04}{(83 \times 2)}}$$

$$= \mathbf{0.50}$$

This is another moderate correlation. Psychologists who are called as expert witnesses in trials involving an insanity defense will have a more difficult time convincing the jury in some areas of the country than in others.

Apply It! boxes written by Shawn Murphy

10.4 Completing the $r \times k$ with Excel

The Excel statistical package focuses on analyses for interval and ratio data. Excel offers no procedure for either of the chi-square tests, but they are not difficult to set up and do not require elaborate programming. To illustrate, here is how to complete the second $r \times k$ problem on Excel. To produce Figure 10.2, follow these steps:

1. Leave cell A1 blank, and then label cells B1–F1 as follows:

   B1: f_o

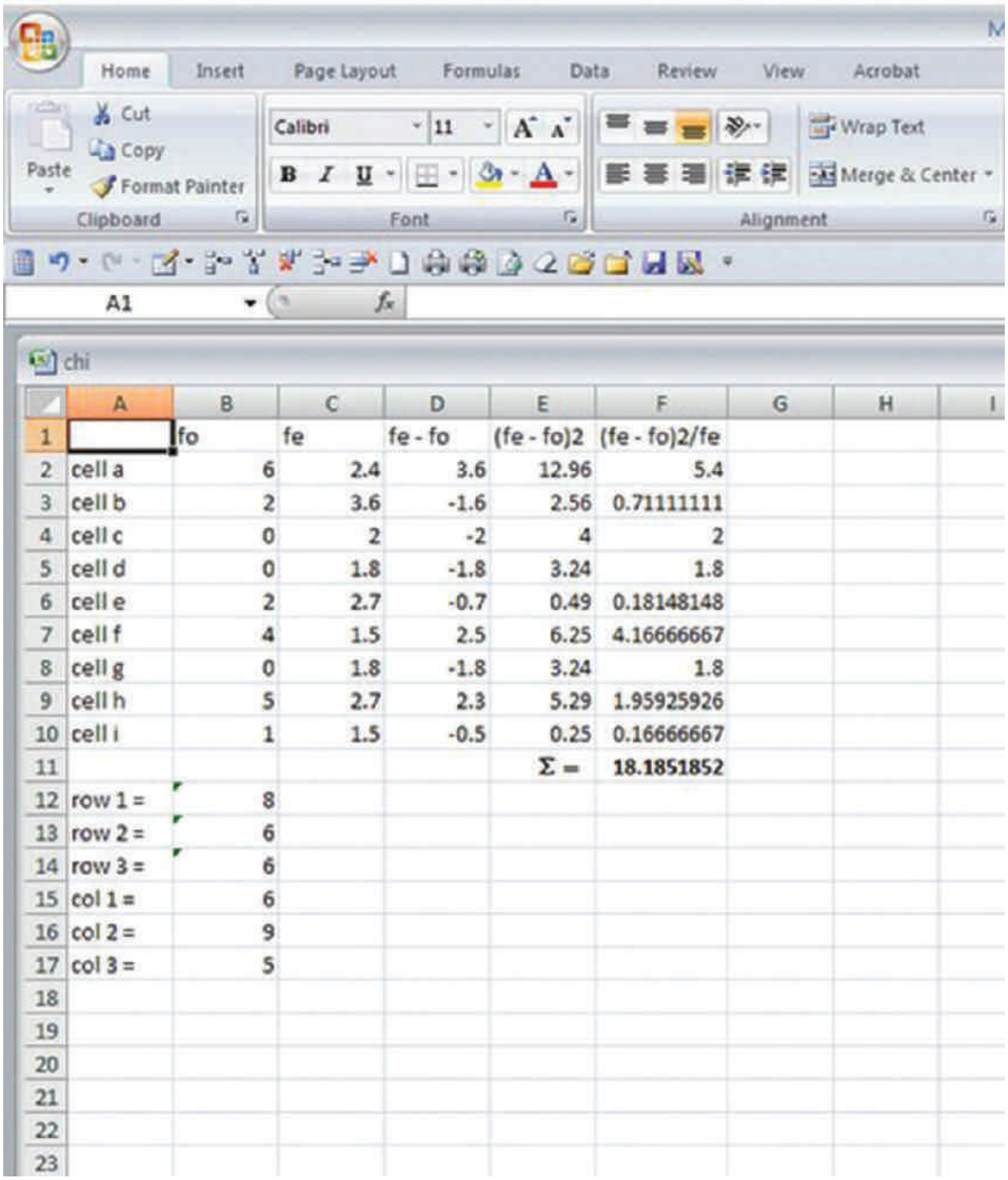
   C1: f_e

   D1: $f_o - f_e$

   E1: $(f_o - f_e)^2$

   F1: $(f_o - f_e)^2/f_e$

Figure 10.2: Completing the chi-square test of independence using Excel



	A	B	C	D	E	F
1		fo	fe	fe - fo	(fe - fo)2	(fe - fo)2/fe
2	cell a	6	2.4	3.6	12.96	5.4
3	cell b	2	3.6	-1.6	2.56	0.71111111
4	cell c	0	2	-2	4	2
5	cell d	0	1.8	-1.8	3.24	1.8
6	cell e	2	2.7	-0.7	0.49	0.18148148
7	cell f	4	1.5	2.5	6.25	4.16666667
8	cell g	0	1.8	-1.8	3.24	1.8
9	cell h	5	2.7	2.3	5.29	1.95925926
10	cell i	1	1.5	-0.5	0.25	0.16666667
11					Σ =	18.1851852
12	row 1 =	8				
13	row 2 =	6				
14	row 3 =	6				
15	col 1 =	6				
16	col 2 =	9				
17	col 3 =	5				

Source: Microsoft Excel. Used with permission from Microsoft.

2. Beginning in cell A2 and continuing to cell A10, enter the labels, as shown in Figure 10.1.
3. In column B under "f_o," enter the f_o values for cells a through i from the contingency table.
4. Leave cell A11 blank, and then beginning in cell A12, enter the label "row 1 =." The label "row 2 =" will go in cell A13.

5. Starting in cell B12, enter the formula for determining the sum of row 1 (as if you were adding cells a, b, and c in the contingency table): **=sum(B2:B4)**.

   - For the row 2 total in cell B13, the command will be **=sum(B5:B7)**. Click **Enter**.
   - For the row 3 total in cell B14 the command will be **=sum(B8:B10)**. Click **Enter**.
   - For the column 1 total in cell B15 the command will be **=sum(B2+B5+B8)**. Click **Enter**.
   - For the column 2 total in cell B16 the command will be **=sum(B3+B6+B9)**. Click **Enter**.
   - For the column 3 total in cell B17 the command will be **=sum(B4+B7+B10)**. Click **Enter**.

6. In cell C2, enter the formula for determining the f_e value for what is cell a in the contingency table, with the formula **=(B12*B15)/20**. For cells b–i in the contingency table, the commands are the following:

   - In cell C3, enter the command **=(B12*B16)/20**.
   - In cell C4, enter the command **=(B12*B17)/20**.
   - In cell C5, enter the command **=(B13*B15)/20**.
   - In cell C6, enter the command **=(B13*B16)/20**.
   - In cell C7, enter the command **=(B13*B17)/20**.
   - In cell C8, enter the command **=(B14*B15)/20**.
   - In cell C9, enter the command **=(B14*B16)/20**.
   - In cell C10, enter the command **=(B14*B17)/20**.

7. In cell D2, enter the command **=B2-C2**.

   - Repeat this command for cells D3–D10 by placing the cursor on the result in cell D2 and dragging the cursor down to cell D10 so that all the cells from D2 to D10 are highlighted.
   - From the **Home** tab, select the **small down arrow** to the right of the larger down arrow at the top of the page on the extreme right just under the summation (∑) sign, and then select **down**.

8. In cell E2, enter the command **=D2^2** which will enter the square of the cell D2 contents in cell E2.

   - Repeat this command for cells E3–E10 by placing the cursor on the result in cell E2 and dragging it down to cell E10 so that all the cells from E2 to E10 are highlighted.
   - From the **Home** tab, select the **small down arrow** to the right of the larger down arrow at the top of the page on the extreme right just under the summation (∑) sign, and then select **down**.

9. In cell F2, enter the command **=E2/C2**.

   - Repeat this command for cells F3–F10 by placing the cursor on the result in cell F2 and dragging it down to cell F10 so that all the cells from F2 to F10 are highlighted.

- From the **Home** tab, select the **small down arrow** to the right of the larger down arrow at the top of the page on the extreme right just below the summation (∑) sign, then select **down**.

10. Sum the values from cells F2 to F10 by entering the command **=sum(F2:F10)** in cell F11. The result is the chi-square value, which may modestly differ from a long-hand calculation because Excel carries several more decimals than the two in the longhand solution.

Writing Up Statistics

Habersham (2014) used the chi-square test of independence to study the relationship between college persistence and student characteristics such as dual enrollment (concurrent college and high school enrollment), gender, and ethnicity. Of the three variables, only gender was significantly related to college persistence. The students' persistence in post-secondary studies was dependent upon neither their prior experience with dual enrollment nor their ethnicity. Interestingly, the author does not report the strength of the relationship between gender and persistence.

Summary and Resources

Chapter Summary

Nineteenth-century British prime minister and apparent skeptic Benjamin Disraeli observed that what we expect usually does not occur and what we do not anticipate typically happens. Karl Pearson, who happened to be one of Disraeli's contemporaries, would have disagreed. In fact, Pearson's chi-square tests are based on the assumption that under normal circumstances, it is the expected thing that happens, and when the outcome differs significantly from what is expected, that is noteworthy.

The chi-square tests are the only procedures in this book that are designed for nominal-scale data, data assigned to categories rather than quantities (Objective 1). For that reason, they stand apart from the others. There are, however, some important similarities as well. Both the goodness-of-fit ($1 \times k$) test and the test of independence ($r \times k$) are tests of significant differences like the z, t, and ANOVA tests. In addition, a significant finding with the $r \times k$ test leads to calculating either the phi coefficient or, when both variables have more than two categories, Cramér's V (Objectives 2 and 3). The phi coefficient and Cramér's V are both correlation procedures, which places them in the same category as the Pearson and Spearman correlations. Although this chapter may seem very different from the first nine chapters, the differences are confined to the scale of the data involved and the kinds of values that are calculated as a result, not to the kinds of decisions that are made.

Data of any scale can be reduced to nominal-scale data and subjected to chi-square analysis. It is possible to reduce ratio-scale data into nominal-scale data. For example, this occurs if a psychologist who has collected data on the number of times clients manifest compulsive behaviors elects to ignore the number of behaviors (ratio data) and focus only on whether or not people from different groups manifest compulsive behavior (yes/no responses constitute nominal data—Objective 1) according to significantly different frequencies (Objective 2). Such research provides for a simpler analysis, but any opportunity to examine the *degree* of compulsivity is forfeit. The differences between the client who is minimally compulsive and the one who is extremely compulsive are lost. Reductions in data scale result in lost data, and when researchers have a choice, most avoid reducing data scale, in spite of the opportunity to use chi-square analysis with all its flexibility.

Chi-square tests' appeal, on the other hand, is that they do not require that the data meet any of the normality requirements that parametric tests impose. Very small data sets, with their inherent risks to normality, are more acceptable in chi-square tests than they are with *t* test, for example. The chi-square tests should have a place in every behavioral scientist's analytical repertoire.

To some degree, the chi-square tests in this chapter represent a nominal data equivalent to the one-way ANOVA and the factorial ANOVA tests used for interval and ratio data. Like the one-way ANOVA, the goodness-of-fit chi-square involves just one independent variable (political party affiliation, for example), but it can be divided into any number of categories. Like the factorial ANOVA, the chi-square test of independence involves two independent variables (party affiliation and gender, for example), although in the case of the chi-square test, two variables are the limit. (Objective 3). Although the coverage of statistical procedures in this book has not been exhaustive, it has been representative. The 10 chapters introduced some of the most important concepts in data analysis and some of the statistical procedures that provide the foundation for nearly all quantitative analysis. Someone with a grasp of the material in these chapters has a solid footing from which to conduct research, complete the data analysis that many reports require, and read the scholarly research.

A parting comment: data-analysis concepts do not often come up in day-to-day conversation, but they probably should provide at least the backdrop for those conversations. You now have the tools for summarizing data and presenting them to others and for answering a variety of important questions. Do the people in your community differ from a national population in voter participation? Are graduation rates at your university significantly different from those at others you considered attending? Is the relationship between marital status and salary merely a random relationship? People frequently make snap judgments about such things, but you need not. You now have the tools that will allow you, with some work in data collection, to answer such questions definitively. So it is worth working to keep your data-analysis tools well-oiled. Think of the understanding you have gained of these topics as elements of a cognitive muscle: Do not let it atrophy. Occasional practice and frequent review will keep your understanding active. The author wishes each student the very best of luck. Happy analyzing!

Key Terms

chi-square test of independence, or $r \times k$ chi-square A test of whether two nominal variables operate independently of each other.

contingency table The arrangement of data from a chi-square test of independence where the categories of one nominal variable are in rows and the categories of the other variable are in columns.

Cramér's V Variation of the phi coefficient used to determine the strength of the relationship between two variables with more than two categories.

goodness-of-fit chi-square test, or 1 × *k* chi-square A test for significant differences in the frequency with which nominal data occur in distinct categories.

phi coefficient Used to determine the strength of the relationship when there are two categories of each variable in the chi-square test of independence. When there are more than two categories of both variables, the measure is Cramér's V.

Review Questions

Answers to the odd-numbered questions are provided in Appendix A.

1. A market researcher for an advertising agency wants to determine whether people listen to three local radio stations in similar proportions. The researcher conducts a telephone survey to find out which station people prefer. In two hours, 52 of the respondents indicated preference for one of the three. The data are as follows:

   Station A: 22
   Station B: 18
   Station C: 12

   Are the differences among their preferences statistically significant?

2. What does it mean when an $r \times k$ problem is statistically significant?

3. Why are significant $r \times k$ problems followed by either phi coefficient or Cramér's V? What are the circumstances under which each procedure is used?

4. A counselor working with people who are developmentally disabled has read research relating accurate responses to the type of reinforcement they receive. For three weeks, the counselor provides verbal praise every time a client completes a simple task correctly and then totals the number of simple problem-solving tasks that are completed successfully. For the next three weeks, the counselor provides a small piece of candy for each successfully completed task. The data for the group are as follows:

   After 3 weeks of verbal praise—17 tasks completed successfully in one hour.
   After 3 weeks of tangible rewards—27 tasks completed successfully in one hour.

   Are the differences statistically significant?

5. An armed-services psychologist believes that among those who have been in a combat area, ground forces experience post-traumatic stress more frequently than Navy or Air Force personnel. Data from a number of armed-services personnel who have been in combat areas reveal the following:

   Of 40 Army personnel, 22 have experienced some post-traumatic stress.
   Of 30 Air Force personnel, 3 have experience some post-traumatic stress.
   Of 30 Navy personnel, 12 have experienced some post-traumatic stress.

a. Are the branch of service and post-traumatic stress related?
b. If so, what is the strength of the relationship?

6. A university counselor has a theory that students' employment and their grades may be related. Among 20 employed students, 8 have grades above 3.0. Among 15 non-employed students, 12 have grades better than 3.0. Is the counselor correct? If so, what is the strength of the relationship?

7. A university administrator believes that undergraduate students of different majors attend the writing lab in different proportions. Test this with the following data:

   Of 20 English majors, 2 attend the writing lab.
   Of 18 engineering majors, 10 attend the lab.
   Of 15 history majors, 6 attend the lab.

   Do major and lab attendance operate independently?

8. Juvenile offenders in court-ordered treatment can choose between community-service activities and a series of group-counseling sessions. The therapist believes that they will choose therapy over community service by a ratio of two to one. Among 35 offenders, 20 opt for therapy. Is the counselor correct?

9. In an effort to assist students more effectively, a high school counselor wonders whether students who opt for college choose one institution with significantly greater frequency than the others. Data are as follows for 70 graduating seniors who are going to college:

   37 apply to the local community college
   23 apply to the state university
   10 apply to the private, liberal-arts college

   Are the differences in institutional preference statistically significant?

10. A car salesperson attempts to determine whether age and the type of car purchased are related.

    Of 15 people in their 20s, 3 opt for sports cars, 8 for economy cars, and 4 for sedans.
    Of 18 people in their 30s, 7 opt for sports cars, 4 for economy cars, and 7 for sedans.
    Of 12 people in their 40s, 2 opt for sports cars, 4 for economy cars, and 6 for sedans.

    Are customers' ages and the type of car they select related?

Answers to Try It! Questions

1. The $1 \times k$ chi-square will accommodate just one nominal-scale variable, but it may have any number of categories.

2. The fact that chi-square values are squared as they are calculated does away with any possibility of a negative value.

3. Type II errors are typically more of a problem with chi-square than type I errors are.

4. The $r \times k$ can accommodate just two variables, but theoretically, those two variables can have any number of categories.