causal relationship, we need to evaluate how well our theory has performed in terms of all four of the causal hurdles from Chapter 3.

7.3.4 The Null Hypothesis and *p*-Values

In Chapter 1 we introduced the concept of the null hypothesis. Our definition was "A null hypothesis is also a theory-based statement but it is about what we would expect to observe if our theory were incorrect." Thus, following the logic that we previously outlined, if our theory-driven hypothesis is that there is covariation between X and Y, then the corresponding null hypothesis is that there is no covariation between X and Y. In this context, another interpretation of the p-value is that it conveys the level of confidence with which we can reject the null hypothesis.

7.4 THREE BIVARIATE HYPOTHESIS TESTS

We now turn to three specific bivariate hypothesis tests. In each case, we are testing for whether there is a relationship between X and Y. We are doing this with sample data, and then, based on what we find, making inferences about the underlying population.

7.4.1 Example 1: Tabular Analysis

Tabular presentations of data on two variables are still used quite widely. In the more recent political science literature, scholars use them as stepping-stones on the way to multivariate analyses. It is worth noting at this point in the process that, in tables, most of the time the dependent variable is displayed in the rows whereas the independent variable is displayed in the columns. Any time that you see a table, it is very important to take some time to make sure that you understand what is being conveyed. We can break this into the following three-step process:

1. Figure out what the variables are that define the rows and columns of the table.
2. Figure out what the individual cell values represent. Sometimes they will be the number of cases that take on the particular row and column values; other times the cell values will be proportions (ranging from 0 to 1.0) or percentages (ranging from 0 to 100). If this is the case, it is critical that you figure out whether the researcher calculated the percentages or proportions for the entire table or for each column or row.
3. Figure out what, if any, general patterns you see in the table.

Table 7.2. Union households and vote in the 2008 U.S. presidential election

Candidate	Not from a union household	From a union household	Total
McCain	47.1	33.4	45.0
Obama	52.9	66.6	55.0
Total	100.0	100.0	100.0

Note: Cell entries are column percentages.

Let's go through these steps with Table 7.2. In this table we are testing the theory that affiliation with trade unions makes people more likely to support left-leaning candidates.[4] We can tell from the title and the column and row headings that this table is comparing the votes of people from union households with those not from union households in the 2008 U.S. presidential election. We can use the information in this table to test the hypothesis that voters from union households were more likely to support Democratic Party presidential candidate Barack Obama.[5] As the first step in reading this table, we determine that the columns indicate values for the independent variable (whether or not the individual was from a union household) and that the rows indicate values for the dependent variable (presidential vote). The second step is fairly straightforward; the table contains a footnote that tells us that the "cell entries are column percentages." This is the easiest format for pursuing step 3, because the column percentages correspond to the comparison that we want to make. We want to compare the presidential votes of people from union households with the presidential votes of people not from union households. The pattern is fairly clear: People from the union households overwhelmingly supported Obama (66.6 for Obama and 33.4 for McCain), whereas people from the nonunion households only marginally favored Obama (52.9 for Obama and 47.1 for McCain). If we think in terms of independent (X) and dependent (Y) variables, the comparison that we have made is between the distribution of the dependent variable ($Y =$ Presidential Vote) across values of the independent variable ($X =$ Type of Household).

[4] Take a moment to assess this theory in terms of the first two of the four hurdles that we discussed in Chapter 3. The causal mechanism is that left-leaning candidates tend to support policies favored by trade unions. Is this credible? What about hurdle 2? Can we rule out the possibility that support for left-leaning candidates make one more likely to be affiliated with a trade union?

[5] What do you think about the operationalization of these two variables? How well does it stand up to what we discussed in Chapter 5?

Table 7.3. Gender and vote in the 2008 U.S. presidential election: Hypothetical scenario

Candidate	Male	Female	Row total
McCain	?	?	45.0
Obama	?	?	55.0
Column total	100.0	100.0	100.0

Note: Cell entries are column percentages.

In Table 7.2, we follow the simple convention of placing the values of the independent variable in the columns and the values of the dependent variable in the rows. Then, by calculating column percentages for the cell values, this makes comparing across the columns straightforward. It is wise to adhere to these norms, because it is the easiest way to make the comparison that we want, and because it is the way many readers will expect to see the information.

In our next example we are going to go step-by-step through a bivariate test of the hypothesis that gender (X) is related to vote (Y) in U.S. presidential elections. To test this hypothesis about gender and presidential vote, we are going to use data from the 2008 National Annenberg Election Survey (NAES from here on). This is an appropriate set of data for testing this hypothesis because these data are from a randomly selected sample of cases from the underlying population of interest (U.S. adults). Before we look at results obtained by using actual data, think briefly about the measurement of the variables of interest and what we would expect to find if there was no relationship between the two variables.

Table 7.3 shows partial information from a hypothetical example in which we know that 45.0% of our sample respondents report having voted for John McCain and 55.0% of our sample respondents report having voted for Barack Obama. But, as the question marks inside this table indicate, we do not know how voting breaks down in terms of gender. If there were no relationship between gender and presidential voting in 2008, consider what we would expect to see given what we know from Table 7.3. In other words, what values should replace the question marks in Table 7.3 if there were no relationship between our independent variable (X) and dependent variable (Y)?

If there is not a relationship between gender and presidential vote, then we should expect to see no major differences between males and females in terms of how they voted for John McCain and Barack Obama. Because we know that 45.0% of our cases voted for McCain and 55.0% for Obama,

Table 7.4. Gender and vote in the 2008 U.S. presidential election: Expectations for hypothetical scenario if there were no relationship

Candidate	Male	Female	Row total
McCain	45.0	45.0	45.0
Obama	55.0	55.0	55.0
Column total	100.0	100.0	100.0

Note: Cell entries are column percentages.

Table 7.5. Gender and vote in the 2008 U.S. presidential election

Candidate	Male	Female	Row total
McCain	?	?	1,434
Obama	?	?	1,755
Column total	1,379	1,810	3,189

Note: Cell entries are number of respondents.

what should we expect to see for males and for females? We should expect to see the same proportions of males and females voting for each candidate. In other words, we should expect to see the question marks replaced with the values in Table 7.4. This table displays the expected cell values for the null hypothesis that there is no relationship between gender and presidential vote.

Table 7.5 shows the total number of respondents who fit into each column and row from the 2008 NAES. If we do the calculations, we can see that the numbers in the rightmost column of Table 7.5 correspond with the percentages from Table 7.3. We can now combine the information from Table 7.5 with our expectations from Table 7.4 to calculate the number of respondents that we would expect to see in each cell if gender and presidential vote were unrelated. We display these calculations in Table 7.6. In Table 7.7, we see the actual number of respondents that fell into each of the four cells.

Finally, in Table 7.8, we compare the observed number of cases in each cell (O) with the number of cases that we would expect to see if there were no relationship between our independent and dependent variables (E).

We can see a pattern. Among males, the proportion observed voting for Obama is lower than what we would expect if there were no relationship

Table 7.6. Gender and vote in the 2008 U.S. presidential election: Calculating the expected cell values if gender and presidential vote are unrelated

Candidate	Male	Female
McCain	(45% of 1,379) $= 0.45 \times 1,379 = 620.55$	(45% of 1,810) $= 0.45 \times 1,810 = 814.5$
Obama	(55% of 1,379) $= 0.55 \times 1,379 = 758.45$	(55% of 1,810) $= 0.55 \times 1,810 = 995.5$

Note: Cell entries are expectation calculations if these two variables are unrelated.

Table 7.7. Gender and vote in the 2008 U.S. presidential election

Candidate	Male	Female	Row total
McCain	682	752	1,434
Obama	697	1,058	1,755
Column total	1,379	1,810	3,189

Note: Cell entries are number of respondents.

Table 7.8. Gender and vote in the 2008 U.S. presidential election

Candidate	Male	Female
McCain	$O = 682; E = 620.55$	$O = 752; E = 814.5$
Obama	$O = 697; E = 758.45$	$O = 1,058; E = 995.5$

Note: Cell entries are the number observed (O); the number expected if there were no relationship (E).

between the two variables. Also, among men, the proportion voting for McCain is higher than what we would expect if there were no relationship. For females this pattern is reversed – the proportion voting for Obama (McCain) is higher (lower) than we would expect if there were no relationship between gender and vote for U.S. president. The pattern of these differences is in line with the theory that women support Democratic Party candidates more than men do. Although these differences are present, we have not yet determined that they are of such a magnitude that we should now have increased confidence in our theory. In other words, we want to know whether or not these differences are statistically significant.

To answer this question, we turn to the **chi-squared (χ^2) test for tabular association**. Karl Pearson originally developed this test when he was testing theories about the influence of nature versus nurture at the beginning of the 20th century. His formula for the χ^2 statistic is

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

The summation sign in this formula signifies that we sum over each cell in the table; so a 2×2 table would have four cells to add up. If we think about an individual cell's contribution to this formula, we can see the underlying logic of the χ^2 test. If the value observed, O, is exactly equal to the expected value if there were no relationship between the two variables, E, then we would get a contribution of zero from that cell to the overall formula (because $O - E$ would be zero). Thus, if all observed values were exactly equal to the values that we expect if there were no relationship between the two variables, then $\chi^2 = 0$. The more the O values differ from the E values, the greater the value will be for χ^2. Because the numerator on the right-hand side of the χ^2 formula $(O - E)$ is squared, any difference between O and E will contribute positively to the overall χ^2 value.

Here are the calculations for χ^2 made with the values in Table 7.8:

$$\begin{aligned}
\chi^2 &= \sum \frac{(O - E)^2}{E} \\
&= \frac{(682 - 620.55)^2}{620.55} + \frac{(752 - 814.5)^2}{814.5} + \frac{(697 - 758.45)^2}{758.45} + \frac{(1{,}058 - 995.5)^2}{995.5} \\
&= \frac{3{,}776.1}{620.55} + \frac{3{,}906.25}{814.5} + \frac{3{,}776.1}{758.45} + \frac{3906.25}{995.5} \\
&= 6.09 + 4.8 + 4.98 + 3.92 = 19.79.
\end{aligned}$$

So our calculated value of χ^2 is 19.79 based on the observed data. What do we do with this? We need to compare that 19.79 with some predetermined standard, called a **critical value**, of χ^2. If our calculated value is greater than the critical value, then we conclude that there is a relationship between the two variables; and if the calculated value is less than the critical value, we cannot make such a conclusion.

How do we obtain this critical value? First, we need a piece of information known as the **degrees of freedom** (df) for our test.[6] In this case, the df calculation is very simple: $\text{df} = (r - 1)(c - 1)$, where r is the number of rows

[6] We define degrees of freedom in the next section.

in the table, and c is the number of columns in the table. In the example in Table 7.8, there are two rows and two columns, so $(2-1)(2-1) = 1$.

You can find a table with critical values of χ^2 in Appendix A. If we adopt the standard p-value of .05, we see that the critical value of χ^2 for df = 1 is 3.841. Therefore a calculated χ^2 value of 19.79 is well over the minimum value needed to achieve a p-value of .05. In fact, continuing out in this table, we can see that we have exceeded the critical value needed to achieve a p-value of .001.

At this point, we have established that the relationship between our two variables meets a conventionally accepted standard of statistical significance (i.e., $p < .05$). Although this result is supportive of our hypothesis, we have not yet established a causal relationship between gender and presidential voting. To see this, think back to the four hurdles along the route to establishing causal relationships that we discussed in Chapter 3. Thus far, we have cleared the third hurdle, by demonstrating that X (gender) and Y (vote) covary. From what we know about politics, we can easily cross hurdle 1, "Is there a credible causal mechanism that links X to Y?" Women might be more likely to vote for candidates like Obama because, among other things, women depend on the social safety net of the welfare state more than men do. If we turn to hurdle 2, "Can we rule out the possibility that Y could cause X?," we can pretty easily see that we have met this standard through basic logic. We know with confidence that changing one's vote does not lead to a change in one's gender. We hit the most serious bump in the road to establishing causality for this relationship when we encounter hurdle 4, "Have we controlled for all confounding variables Z that might make the association between X and Y spurious?" Unfortunately, our answer here is that we do not yet know. In fact, with a bivariate analysis, we cannot know whether some other variable Z is relevant because, by definition, there are only two variables in such an analysis. So, until we see evidence that Z variables have been controlled for, our scorecard for this causal claim is [*y y y n*].

7.4.2 Example 2: Difference of Means

In our second example, we examine a situation in which we have a continuous dependent variable and a categorical independent variable. In this type of bivariate hypothesis test, we are looking to see if the means are different across the values of the independent variable. We follow the basic logic of hypothesis testing: comparing our real-world data with what we would expect to find if there were no relationship between our independent and dependent variables. We use the sample means and standard deviations to make inferences about the unobserved population.