Chapter twenty two

Special Topics: Confidence Intervals

Describing a confidence interval

What is a Confidence Interval?

How is It Constructed?

Defining confidence intervals for different statistics

How is the Confidence Interval for Sample Means Calculated?

How is the Confidence Interval for Sample Proportions Calculated?

How is the Confidence Interval for a Difference of Means Calculated?

How is the Confidence Interval for a Correlation Coefficient Calculated?

How is the Confidence Interval for a Regression Coefficient Calculated?

How is the Confidence Interval for a Logistic Regression Coefficient Calculated?

D. Weisburd and C. Britt, *Statistics in Criminal Justice*, DOI 10.1007/978-1-4614-9170-5_22,

One of the main concerns of this text has been to define how we make inferences from samples to populations. This is also one of the main concerns of researchers, since in most cases they must make decisions about population parameters on the basis of sample statistics. Our approach has been to use the logic of statistical inference, which begins with the creation of a null hypothesis. Importantly, the logic of statistical inference we have reviewed so far is concerned primarily with the decision as to whether to reject or fail to reject the null hypothesis. This means in practice that we have relied on a logic that allows us to make a statement about where the population parameter is *not*.

This approach has enabled us to come to concrete decisions about population parameters on the basis of sample statistics. When we reject the null hypothesis on the basis of a statistical test, we conclude that the relationship we have examined is statistically significant. For example, when we reject the null hypothesis on the basis of our sample statistics in a statistical test of the difference of means, we have confidence that there is a difference between the means of the two populations. When we reject the null hypothesis that there is not a linear correlation between two variables, we have confidence that there is in fact a linear correlation between these two variables in the population. But the logic we have used so far does not allow us to zero in on the value of the population parameter. When we find that the relationship between two variables in a sample is statistically significant, we conclude that there is likely to be a relationship in the population from which the sample was drawn. But this decision does not provide us with an estimate of the size of that relationship in the population.

In this chapter, we turn to an approach to statistical inference that leads us to make specific statements about population parameters from sample statistics. The logic used in this approach is very similar to that described in earlier chapters. However, we do not make a single decision about the null hypothesis. Rather, we create an interval of values within which we can be fairly confident that the true parameter lies. Of course, without data on the population itself, we can never be certain of the value of the population parameter. This interval is generally called a confidence interval. In this chapter, we begin by explaining the logic behind confidence intervals and how they are used. We then illustrate how confidence intervals are constructed for the main statistics reviewed in this text.

Confidence Intervals

In the statistical tests presented in earlier chapters, we began by setting a null hypothesis. Our null hypothesis made a statement about the value of the population parameter. In practice, the null hypothesis generally stated that a statistic had a value of 0. For example, for the difference of means test, the null hypothesis generally stated that the difference between two population means was 0; for the correlation coefficient, that the population correlation had a value of 0; or for the regression coefficient, that the population regression coefficient had a value of 0. When the results of our statistical test indicated that we should reject the null hypothesis, we concluded that it was unlikely that the population parameter had the value stated in the null hypothesis. Since the null hypothesis was generally 0 or no difference, we rejected the hypothesis that the population parameter had this null value.

We can use similar logic to make a very different statement about population parameters. In this case, we ask where the population parameters are likely to be found. In statistics, the interval of values around the sample statistic within which we can be fairly confident that the true parameter lies is called a **confidence interval.** A confidence interval makes it possible for us to state where we think the population parameter is likely to be—that is, the range of values within which we feel statistically confident that the true population parameter is likely to be found. Importantly, the fact that we are confident does not mean that the population parameter actually lies in that range of values. As in tests of statistical significance, we rely on probabilities in making our decisions.

One common illustration of confidence intervals comes from newspaper articles and television news programs reporting the results from public opinion polls. In addition to stating that some percentage of the population supports a particular political candidate in an upcoming election or a particular policy, more thorough accounts of these kinds of survey results typically make reference to a range of values. For example, a poll might indicate that 60% of adults in the United States favor using the death penalty for convicted murderers, ±4% (plus or minus 4 percent).

The value of 60% is often described in statistics as a **point estimate.** Absent knowledge of the population parameter, the statistic we obtain for our sample is generally used as an estimate—in statistical terms, a point estimate—of the population parameter. The range of values represented by $\pm 4\%$ is sometimes described as the **margin of error** of a poll. In statistics, we prefer to call this margin of error a confidence interval. Based on a very specific set of statistical assumptions, it is the interval within which the true population value is expected to fall.

Confidence intervals are based on the same statistical logic as tests of statistical significance. It will be easier to understand the relationship between tests of statistical significance and the construction of confidence intervals if we start with an example that—although it is very unusual—allows us to make a straightforward link between these two concepts. Let's say that we gather data on attitudes toward the death penalty using an interval-scale measure that has both positive values, indicating support for the death penalty, and negative values, indicating opposition to the death penalty. We use an independent random sampling method to draw our sample from the population of all adult Americans. After completing our study, we find that the mean score for attitudes toward the death penalty is 0.

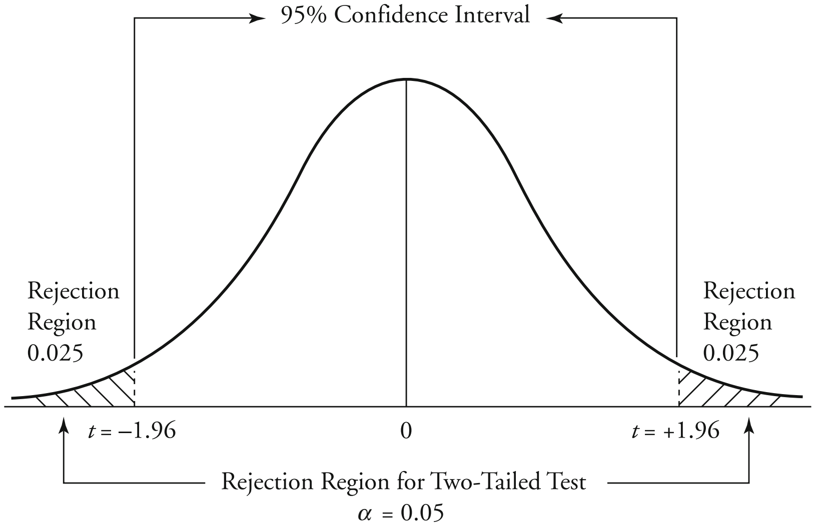
In determining a confidence interval, we rely on the same basic assumptions that we use for tests of statistical significance. If we were going to compare the mean in our sample to some hypothesized population mean, we would use a *t*-test as our test of statistical significance. This means that the *t*-test also underlies our confidence interval. Accordingly, we have to assume an interval level of measurement and make parametric assumptions regarding the population distribution. Let's assume that our sample is very large, so we can relax the assumption of a normal population distribution. We have already noted that the sampling method meets the requirements of a *t*-test.

If we intended to conduct a test of statistical significance for this example, we would have stated a null hypothesis and an alternative hypothesis and set a level of statistical significance. Let's say that the null hypothesis is that Americans are neutral regarding the death penalty. This means that H_0 for our example will be 0.0, as the scale is divided between positive attitudes greater than 0 and negative attitudes less than 0. There is no reason to posit a directional research hypothesis, so our test will be two-tailed. We will use a standard 0.05 significance level.

Figure 22.1 illustrates the *t*-test for this example. The rejection region begins at a *t*- value of 1.96 either above or below the null hypothesis of 0. In order to reject the null hypothesis, we need an observed value of t for our sample that is greater than 1.96 or less than -1.96. Since our observed value of the measure is 0, the value of t will also be 0. Clearly, we would not reject the null hypothesis in this case.

Figure 22.1

The 5% Rejection Region and 95% Confidence Interval on a Normal Frequency Distribution (where $\overline{X}$ and $H_0 = 0$)



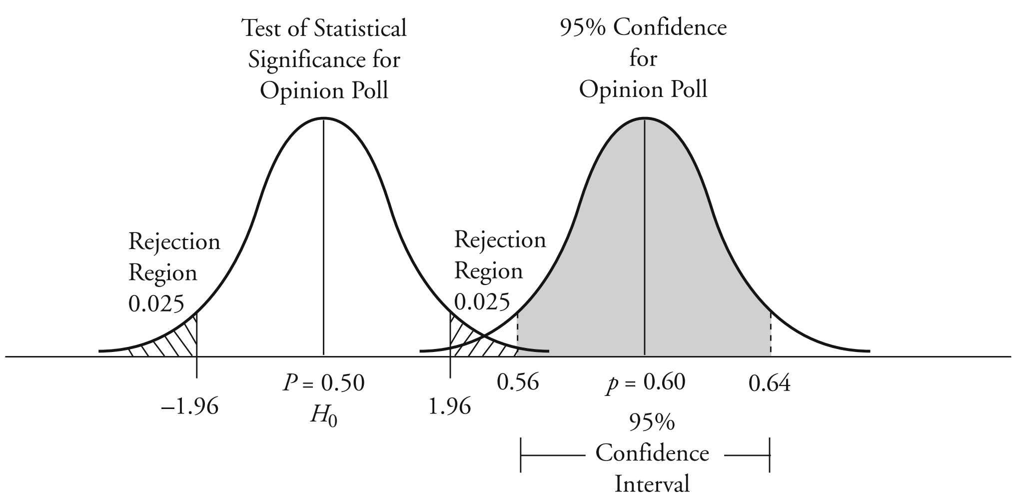
But what would a confidence interval for this example look like? With a confidence interval, we are not concerned about whether the population parameter is *not* at a specific value (for example, the null hypothesis); rather, we are concerned about specifying a range of values within which we can be fairly confident (though not certain) that the population parameter lies. How do we choose this range of values? Clearly, we want to make the interval large enough that, given the observed statistic in our sample, the population parameter is unlikely to lie outside it. As in tests of statistical significance, our choice is based on convention. With a test of statistical significance, it is common to set a threshold of 5% for the risk we are willing to take of falsely rejecting the null hypothesis. With confidence intervals, we define the width of the interval so that we can be very confident that the true population value lies within it. The confidence interval most commonly used is a 95% confidence interval. Figure 22.1 illustrates the 95% confidence interval for our example. As you can see, the confidence interval extends until the rejection region begins. It is, in this sense, the flip side of the rejection region.

Thus, a 95% confidence interval and a 5% significance level are directly related. In our example, the 5% significance rejection region represents values far enough away from the null hypothesis that we are confident in rejecting it. The 95% confidence interval represents values close enough to our observed statistic, or point estimate, that we are confident that the population parameter lies within that interval.

Of course, in practical examples it is very unlikely that our observed sample statistic will be the same as the population parameter hypothesized by the null hypothesis. A more common situation is that of the opinion poll described earlier. What would the confidence interval look like for our opinion poll? We have all the information we need to illustrate that example, except that the level of confidence of the interval was not specified. Let's assume that a 95% confidence interval was used in arriving at the margin of error. The observed statistic, or point estimate, of 60% will be the mean of the distribution. We would use a z-test rather than a t-test because we are concerned with only a single proportion. Let's assume that the other assumptions of the test were met. The margin of error of the test, or size of the confidence interval, is 4%. Figure 22.2 shows the confidence interval relative to the z distribution. As you can see, the interval ranges between 56% and 64%.

But how does this confidence interval relate to a test of statistical significance? First, we need to identify a null hypothesis. Suppose we make the null hypothesis for our test that the population is evenly divided in their attitudes toward the death penalty. In this case, the H_0 takes on a value of 0.50, meaning that about half of the population to which the sample infers are for and half against the use of the death penalty for

Figure 22.2 *95% Confidence Interval for the Public Opinion Poll Example*



convicted murderers. Note that this value is very far outside the confidence interval that we have defined for our example.

Figure 22.2 shows the sampling distribution for our example. As you can see, our point estimate of 0.60 falls much to the right of the critical value ($t = \pm 1.96$) of our test of statistical significance. As a general rule, if the null hypothesis for a test of statistical significance lies outside the confidence interval for the statistic (and the confidence interval and the significance level represent opposite parts of the same criterion—for example, 0.95 and 0.05; 0.99 and 0.01), then you may assume that the result is statistically significant. This again points to the close relationship between tests of statistical significance and confidence intervals.

While we use the logic of confidence intervals to define where a population parameter is likely to be found, the confidence interval has a very specific statistical interpretation. Were we to draw repeated samples of a specific sample size from the same population, using a 95% confidence interval, we would expect that in 95% of these samples the confidence interval would include the population parameter. That is, we would expect that in only 5 out of every 100 samples would the parameter lie outside the confidence interval. As in tests of statistical significance, we must be aware at the outset that we are only making an informed decision about the value of the population parameter. Using a 95% confidence interval, we will make the wrong decision about 5 in a 100 times.

Constructing Confidence Intervals

Confidence intervals for many different sample statistics can be constructed using the same basic equation. To illustrate how we construct a confidence interval, we use the example of a *t*-statistic. The most general equation for calculating a *t*-statistic for a sample statistic is written as follows:

$$t = \frac{\text{sample statistic} - \text{population parameter}}{\text{standard error of sampling distribution}}$$

To construct a confidence interval, we adjust this equation so that we can solve it for the population parameter. We can do this through simple algebra. Solving for the population parameter produces the following equation:

$$\text{Population parameter} = \begin{pmatrix}\text{sample} \\ \text{statistic}\end{pmatrix} \pm t \begin{pmatrix}\text{standard error of} \\ \text{sampling distribution}\end{pmatrix}$$

In setting the boundaries for the confidence interval, we will use the positive and negative values associated with a two-tailed t-test to provide the upper and lower boundaries, respectively (see Figure 22.1). After we account for the positive and negative t-values, our confidence interval is given by Equation 22.1.

$$\text{Confidence limit} = \begin{pmatrix}\text{sample}\\ \text{statistic}\end{pmatrix} \pm t\begin{pmatrix}\text{standard error of}\\ \text{sampling distribution}\end{pmatrix} \qquad \textbf{Equation 22.1}$$

The t-value in the equation coincides with the level of confidence we require (i.e., the critical t-value). Following our earlier logic, this t-value is the flip side of the significance threshold. For a 95% confidence interval, we use a t-value associated with a two-tailed 0.05 significance level. For a 99% confidence interval, we use a t-value associated with a two-tailed 0.01 significance level. In general, if α is our significance level for a two-tailed test, then we can construct a confidence interval for $100 \times (1 - \alpha)$ using the same critical t-values.

Confidence Intervals for Sample Means

Let's start by constructing a confidence interval for a sample mean ($\overline{X}$). If we rewrite Equation 22.1 to replace the general terms with the mean and the standard error, we have Equation 22.2.

$$\text{Confidence limit} = \overline{X} \pm t\left(\frac{s}{\sqrt{N-1}}\right) \qquad \textbf{Equation 22.2}$$

where $\overline{X}$ is the sample mean, s is the sample standard deviation, N is the sample size, and t is the critical t-value associated with a given significance level. To determine our critical t, we use df $= N - 1$, as in the single-sample t-test (see Chapter 10).

For an illustration of the calculation of a confidence interval for a sample mean, consider a recent study of fear of crime among Korean Americans living in the Chicago area.[1] The investigators constructed a fear of crime instrument that was measured on an interval scale and ranged in value from 11.00 to 110.00. The mean fear of crime score for the 721 respondents was 81.05, with a standard deviation of 23.41.

To calculate a 99% confidence interval for the fear of crime instrument, we use the t-value associated with a 0.01 significance level and 720 degrees of freedom (df $= 721 - 1$). Using the last line of the t distribution table in Appendix 4, we find the corresponding critical t-value to be 2.576.

---

[1]Min Sik Lee and Jeffery T. Ulmer, "Fear of Crime Among Korean Americans in Chicago Communities," *Criminology* 38:4 (2000): 1173–1206.

A 99% confidence interval for the fear of crime instrument has the following values:

Working It Out

$$\text{Confidence limit} = \overline{X} \pm t\left(\frac{s}{\sqrt{N-1}}\right)$$
$$= 81.05 \pm 2.576\left(\frac{23.41}{\sqrt{721-1}}\right)$$
$$= 81.05 \pm 2.25$$

The result of ± 2.25 indicates that the 99% confidence interval includes values ranging from a low of 78.80 (81.05 − 2.25) to a high of 83.30 (81.05 + 2.25). By using a 99% confidence interval, we can be very confident that the population mean lies somewhere between 78.80 and 83.30. In statistical terms, if we were to observe repeated samples of this size drawn from this population and calculate a confidence interval for each of them, only about 1 in 100 would fail to include the true population parameter.

Confidence Intervals for Sample Proportions We can apply the same type of logic to calculating a confidence interval for a sample proportion, modifying Equation 22.1 by replacing the critical t-value with the critical z-value. To calculate a confidence interval for a sample proportion, we use Equation 22.3:

$$\text{Confidence limit} = p \pm z\left(\sqrt{\frac{pq}{N}}\right) \qquad \textbf{Equation 22.3}$$

where p is the sample proportion, q is $1 - p$, N is the sample size, and z is the critical z-value associated with a given significance level.

In their study of fear of crime among Korean Americans in the Chicago area, the investigators also included a question about the respondent's victimization experiences. Specifically, respondents were asked whether they had experienced any kind of victimization in the past three years. Included in this global indicator of victimization were violent as well as property crime victimizations. The investigators reported that 27% of the 721 respondents had experienced some form of victimization during this time period.

Knowing that the sample proportion is 0.27 and the z-score is 1.96, we can calculate a 95% confidence interval for this proportion. We insert our values for p, q, and N into Equation 22.3.

Working It Out

$$\text{Confidence limit} = p \pm z\sqrt{\frac{pq}{N}}$$

$$= 0.27 \pm 1.960\sqrt{\frac{(0.27)(1-0.27)}{721}}$$

$$= 0.27 \pm 1.960(0.0165)$$

$$= 0.27 \pm 0.03$$

The 95% confidence interval is ±3% around the sample mean of 27%. It suggests that we can be confident that the percentage of Korean Americans living in and around Chicago who experienced some form of criminal victimization within the three-year period lies between 24% and 30%.

Confidence Intervals for a Difference of Sample Means In Chapter 11, we discussed calculating t-statistics to test for significant differences between two sample means. Another way of calculating a confidence interval for the difference of two sample means is by modifying Equation 22.1 to replace the sample mean with the difference of sample means and insert the appropriate standard error for the difference of two sample means. Recall from Chapter 11, however, that there are two methods for calculating the standard error of the sampling distribution: the separate variance method and the pooled variance method (see pages 274–279). Equations 22.4a and 22.4b present formulas for calculating a confidence interval for a difference of two sample means, using either the separate variance method or the pooled variance method.

$$\text{Confidence limit} = (\overline{X}_1 - \overline{X}_2) \pm t\sqrt{\frac{s_1^2}{N_1 - 1} + \frac{s_2^2}{N_2 - 1}}$$

Equation 22.4a Separate Variance Method

$$\text{Confidence limit} = (\overline{X}_1 - \overline{X}_2) \pm t\left(\sqrt{\frac{N_1 s_1^2 + N_2 s_2^2}{N_1 + N_2 - 2}}\sqrt{\frac{N_1 + N_2}{N_1 N_2}}\right)$$

Equation 22.4b Pooled Variance Method

In both equations, $\overline{X}_1$ and $\overline{X}_2$ represent the two sample means, s_1^2 and s_2^2 represent the two sample variances, N_1 and N_2 represent the two sample sizes, and t is the critical t-value associated with a given significance level. As with the two-sample t-test (see Chapter 11), the number of degrees of freedom for determining the critical t-value will be $\text{df} = N_1 + N_2 - 2$.

Chapter 11 presented a test for differences in bail amounts required of African American and Hispanic defendants in Los Angeles County. A sample of 1,121 African Americans were required to post a mean bail amount of \$50,841 ($s = 115{,}565$), while a sample of 1,798 Hispanics were required to post a mean bail amount of \$66,552 ($s = 190{,}801$). The difference in the two sample means is \$15,711, where Hispanics were required to post higher bail amounts, on average.

Using Equations 22.4a and 22.4b, we can calculate a 95% confidence interval for this difference of sample means. For both equations, we use the t-value associated with a 0.05 significance level and 2,917 degrees of freedom. From the last line of the t distribution table in Appendix 4, we find that critical $t = 1.960$.

Working It Out Separate Variance Method

$$\text{Confidence limit} = (\overline{X}_1 - \overline{X}_2) \pm t\sqrt{\frac{s_1^2}{N_1 - 1} + \frac{s_2^2}{N_2 - 1}}$$

$$= (50{,}841 - 66{,}552) \pm 1.960\sqrt{\frac{115{,}565^2}{1{,}121 - 1} + \frac{190{,}801^2}{1{,}798 - 1}}$$

$$= -15{,}711 \pm 1.960(5{,}673.02)$$

$$= -15{,}711 \pm 11{,}119.12$$

Working It Out Pooled Variance Method

$$\text{Confidence limit} = (\overline{X}_1 - \overline{X}_2) \pm t\left(\sqrt{\frac{N_1 s_1^2 + N_2 s_2^2}{N_1 + N_2 - 2}}\sqrt{\frac{N_1 + N_2}{N_1 N_2}}\right)$$

$$= (50{,}841 - 66{,}552)$$

$$\pm 1.960\left(\sqrt{\frac{(1{,}121)(115{,}565^2) + (1{,}798)(190{,}801^2)}{1{,}121 + 1{,}798 - 2}}\sqrt{\frac{1{,}121 + 1{,}798}{(1{,}121)(1{,}798)}}\right)$$

$$= -15{,}711 \pm 1.960(6{,}319.07)$$

$$= -15{,}711 \pm 12{,}385.38$$

Using the separate variance method, we find that the confidence interval is $\pm 11{,}119.12$ around the difference of sample means of 15,711. This interval suggests that we can be confident, based on our sample findings, that the average bail amounts posted in Los Angeles by African Americans were from \$4,591.88 to \$26,830.12 *less* than the average bail amounts required of Hispanics. The pooled variance method provides very similar results, indicating that the confidence interval is $\pm 12{,}385.38$ around the difference of sample means. Again, this interval suggests that we can be fairly confident that African Americans were required to post average bail amounts from \$3,325.62 to \$28,096.38 *less* than those required of Hispanics.

Confidence Intervals for Pearson's Correlation Coefficient, *r*

The calculation of confidence intervals for Pearson's correlation coefficient, r, relies on a similar logic, but requires an additional step. In contrast to that for sample means, sample proportions, or differences of means, the sampling distribution for Pearson's r is not normal or even approximately normal.[2] Consequently, we need to convert r into another statistic, Z^*, that does have a normal distribution. The conversion of r is known as the Fisher r-to-Z^* transformation.[3] After calculating the standard error for Z^*, we can then modify Equation 22.1 to calculate a confidence interval for Z^*. Since the values for Z^* are not directly interpretable, we will then convert the confidence limits back into values of r.

The Fisher r-to-Z^* transformation is given in Equation 22.5. In this equation, we take the natural logarithm of 1 plus r divided by 1 minus r and multiply this value by $\frac{1}{2}$.

$$Z^* = \frac{1}{2} \times \ln\left(\frac{1 + r}{1 - r}\right)$$ **Equation 22.5**

Values for Z^* for correlation coefficients ranging in value from 0.000 to 1.000 are given in Appendix 8. Note that the correlations given in the appendix are all positive. If r is negative, then Z^* will also be negative.

In Chapter 14, we reported that the correlation between unemployment rates and burglary rates for 58 counties in California was 0.491. If

[2]The sampling distribution for r will generally be normal and symmetric only for the case where $r = 0$, which is what allowed us to use the t distribution to test whether $r_p = 0$ (i.e., the null hypothesis) in Chapter 14. When $r \neq 0$, the sampling distribution is not symmetric around r, so we cannot calculate a confidence interval for r in the same way we did for sample means or the difference of sample means.

[3]Ronald A. Fisher, *Statistical Methods for Research Workers,* 14th ed. (New York: Hafner, 1970).

we locate $r = 0.491$ in Appendix 8, we find Z^* to be 0.5374. We obtain the same value for Z^* if we use Equation 22.5.

Working It Out

$$Z^* = \frac{1}{2} \times \ln\left(\frac{1+r}{1-r}\right)$$

$$= \frac{1}{2} \times \ln\left(\frac{1+0.491}{1-0.491}\right)$$

$$= 0.5374$$

The standard error of Z^*, which is based on the size of the sample (N), is presented in Equation 22.6.

$$\sigma_{sd(Z^*)} = \frac{1}{\sqrt{N-3}}$$ **Equation 22.6**

In our example concerning unemployment rates and burglary rates for California counties, we have 58 observations, so the standard error of Z^* is 0.1348.

Working It Out

$$\sigma_{sd(Z^*)} = \frac{1}{\sqrt{N-3}}$$

$$= \frac{1}{\sqrt{58-3}}$$

$$= 0.1348$$

We can now modify Equation 22.1 by inserting Z^* as the sample statistic, a critical z-value (since Z^* is approximately normally distributed), and the equation for the standard error of Z^*. The formula for the confidence interval for Z^* is given in Equation 22.7.

$$\text{Confidence limit} = Z^* \pm z\left(\frac{1}{\sqrt{N-3}}\right)$$ **Equation 22.7**

where Z^* is based on the Fisher r-to-Z^* transformation, N is the sample size, and z is the critical z-value associated with a given significance level.

Continuing our example for $Z^* = 0.5374$ and $N = 58$, we calculate a 95% confidence interval for Z^* by using critical $z = 1.960$ and inserting the values into Equation 22.7.

Working It Out

$$\text{Confidence limit} = Z^* \pm z\left(\frac{1}{\sqrt{N-3}}\right)$$

$$= 0.5374 \pm 1.960\left(\frac{1}{\sqrt{58-3}}\right)$$

$$= 0.5374 \pm 0.2642$$

The confidence interval is ± 0.2642 around $Z^* = 0.5374$, indicating that the range for Z^* is 0.2732 to 0.8016. Since we are unable to directly interpret values of Z^*, we should convert the values of Z^* back to values of r, using Appendix 8. The conversion of Z^* back to r will provide us with the confidence interval for r. For $Z^* = 0.2732$, we find that $r = 0.267$. For $Z^* = 0.8016$, we find that $r = 0.665$. For both values of Z^*, we used the closest Z^*-value reported in Appendix 8 to determine the values for r, since an exact match could not be found. These results suggest that we can be confident that the population value for the correlation coefficient falls between 0.267 and 0.665. Note that the upper and lower confidence limits are not symmetric around r—the lower limit is farther away from $r = 0.491$ than is the upper limit.

Confidence Intervals for Regression Coefficients

Confidence intervals for regression coefficients are nearly identical in form to confidence intervals for sample means. The formula for calculating confidence intervals for regression coefficients is given in Equation 22.8.

$$\text{Confidence limit} = b \pm t\left(\hat{\sigma}_b\right) \qquad \textbf{Equation 22.8}$$

where b is the regression coefficient, $\hat{\sigma}_b$ is the standard error of b, and t is the critical t-value associated with a given level of significance. The number of degrees of freedom for the critical t will be equal to $N - k - 1$, where N is the sample size and k is the number of independent variables in the regression model.

In Chapter 15, we reported that the regression coefficient representing the effect of unemployment rates on burglary rates in California was 36.7483 and the standard error for b was 8.7126 (see page 466). If we calculate a 99% confidence interval, the number of degrees of freedom will be 56 (df $= 58 - 1 - 1 = 56$), so the critical t we will use is 2.669 (see Appendix 4).

Working It Out

$$\begin{aligned}\text{Confidence limit} &= b \pm t(\hat{\sigma}_b)\\ &= 36.7483 \pm 2.669(8.7126)\\ &= 36.7483 \pm 23.2539\end{aligned}$$

The result of ± 23.2539 indicates that the 99% confidence interval includes values ranging from a low of 13.4944 to a high of 60.0022. The 99% confidence interval suggests that we can be very confident that the population value of the regression coefficient representing the effect of unemployment rates on burglary rates lies somewhere between 13.4944 and 60.0022.

Confidence Intervals for Logistic Regression Coefficients and Odds Ratios

Confidence intervals for logistic regression coefficients are calculated in exactly the same way as logistic regression coefficients (Equation 22.8). The number of degrees of freedom for determining the critical t-value also equals $N - k - 1$.

In addition to being able to calculate confidence intervals for the original logistic regression coefficients, we can also refer to confidence intervals for odds ratios. As noted in Chapter 18, we can convert our logistic regression coefficients into odds ratios by exponentiating the coefficient b. This means that we can take the lower and upper limits of our confidence interval for b and convert them to odds ratios. We can then discuss the confidence interval relative to the odds, rather than the untransformed coefficients, which are difficult to interpret.

An illustration of the use of confidence intervals for logistic regression coefficients is provided by a recent study examining the link between mental disorders and violent victimization for a sample of 747 adults.[4] The dependent variable measured whether the individual had reported a

[4] Eric Silver, "Mental Disorder and Violent Victimization: The Mediating Role of Involvement in Conflicted Social Relationships," *Criminology* 40 (2002): 191–212.

violent victimization in the preceding ten weeks. One of the nine independent variables used by the researcher was the level of neighborhood disadvantage, which was an interval-level instrument that combined economic indicators, such as poverty rate, unemployment rate, and income. The effect of neighborhood disadvantage was positive ($b = 0.33$), meaning the greater the level of neighborhood disadvantage, the more likely the individual was to have experienced a violent victimization. The standard error for b was reported to be 0.09.

To calculate a 99% confidence interval for b, we use critical $t = 2.576$, since $\text{df} = 747 - 9 - 1 = 737$, and insert the values into Equation 22.8.

Working It Out

$$\begin{aligned} \text{Confidence limit} &= b \pm t\left(\hat{\sigma}_b\right) \\ &= 0.33 \pm 2.576(0.09) \\ &= 0.33 \pm 0.23 \end{aligned}$$

The result of ± 0.23 tells us that the 99% confidence interval includes values ranging from a low of 0.10 to a high of 0.56. If we exponentiate the lower and upper limits of the confidence interval for b, we will have the lower and upper limits of the confidence interval for the odds ratio. The lower limit of the confidence interval for the odds ratio is 1.105 [Exp(0.10)], and the upper limit of the confidence interval for the odds ratio is 1.751 [Exp(0.56)]. These results suggest that we can be very confident that the population value of the odds ratio lies somewhere between 1.105 and 1.751. If we took repeated random samples of the size examined here and calculated a confidence interval for each, then in only about 1 in 100 cases would that interval fail to include the true odds ratio.

Chapter Summary

In tests of statistical significance, we make a statement about where the population parameter is *not*. In this chapter, we turned to an approach to statistical inference that leads us to make a very different type of statement about population parameters. The logic used in this approach is similar to that described in earlier chapters. However, we do not make a single decision about the null hypothesis. Rather, we create an interval of values within which we can be fairly confident that the true parameter lies—although, without data on the population itself, we can never be certain of the value of the population parameter. This interval is generally called a **confidence interval.**

A confidence interval makes it possible for us to say where we think the population parameter is likely to be—that is, the range of values within which we feel statistically confident that the true population parameter is likely to be found. A confidence interval is generally constructed around the observed statistic of interest, commonly called a **point estimate.** Absent knowledge of the population parameter, the statistic we obtain for our sample is generally used as an estimate—in statistical terms, a point estimate—of the population parameter. The size of the confidence interval is often referred to as the **margin of error.**

Confidence intervals may be constructed at any level of confidence. By convention, we use 95% and 99% confidence levels, which are based on 5% and 1% significance thresholds. While it is commonly said, when using a confidence interval, that the researcher is confident that the true parameter lies in the interval defined, confidence intervals have a specific statistical interpretation. Suppose we find, using a 95% or 99% criterion, that a confidence interval is of a certain size. If we were to draw repeated samples of the same size, using the same methods, and calculate a confidence interval for each sample, then in only 5 in 100 (for a 95% interval) or 1 in 100 (for a 99% interval) of these samples would the interval fail to include the true population parameter.

Key Terms

confidence interval An interval of values around a statistic (usually a point estimate). If we were to draw repeated samples and calculate a 95% confidence interval for each, then in only 5 in 100 of these samples would the interval fail to include the true population parameter. In the case of a 99% confidence interval, only 1 in 100 samples would fail to include the true population parameter.

margin of error The size of the confidence interval for a test. A margin of error of ±3% in an opinion poll means that the confidence interval ranged between 3% above and 3% below the point estimate or observed statistic.

point estimate An estimate of the population parameter. Absent knowledge of the population parameter, the statistic we obtain for a sample is generally used as an estimate—or, in statistical terms, a point estimate—of the population parameter.

Symbols and Formulas

To calculate the confidence interval for a sample mean:

$$\text{Confidence limit} = \overline{X} \pm t\left(\frac{s}{\sqrt{N-1}}\right)$$

To calculate the confidence interval for a sample proportion:

$$\text{Confidence limit} = p \pm z\left(\sqrt{\frac{pq}{N}}\right)$$

To calculate the confidence interval for a difference of sample means, using the separate variance method:

$$\text{Confidence limit} = (\overline{X}_1 - \overline{X}_2) \pm t\sqrt{\frac{s_1^2}{N_1 - 1} + \frac{s_2^2}{N_2 - 1}}$$

To calculate the confidence interval for a difference of sample means, using the pooled variance method:

$$\text{Confidence limit} = (\overline{X}_1 - \overline{X}_2) \pm t\left(\sqrt{\frac{N_1 s_1^2 + N_2 s_2^2}{N_1 + N_2 - 2}}\sqrt{\frac{N_1 + N_2}{N_1 N_2}}\right)$$

To convert r to Z^* (Fisher r-to-Z^* transformation):

$$Z^* = \frac{1}{2} \times \ln\left(\frac{1+r}{1-r}\right)$$

To calculate the confidence interval for Z^*:

$$\text{Confidence limit} = Z^* \pm z\left(\frac{1}{\sqrt{N-3}}\right)$$

To calculate the confidence interval for a regression or logistic regression coefficient:

$$\text{Confidence limit} = b \pm \left(t\hat{\sigma}_b\right)$$

Exercises

22.1 In a study of self-reported marijuana use, a sample of high school students were asked how many times they had smoked marijuana in the last month. Researchers reported that the average for the sample was

2.4 times, with a 95% confidence interval of ± 1.3. Explain what this result means in plain English.

22.2 Following a revolution, the new leadership of the nation of Kippax intends to hold a national referendum on whether the practice of capital punishment should be introduced. In the buildup to the referendum, a leading army general wishes to gauge how the people are likely to vote so that he can make a public statement in line with popular feeling on the issue. He commissions Greg, a statistician, to carry out a secret poll of how people plan to vote. The results of Greg's poll are as follows: The sample proportion in favor of introducing capital punishment is 52%; the sample has a 95% confidence interval of $\pm 10\%$. How should Greg explain these results to the army general?

22.3 Concerned that taxpayers were not reporting incomes honestly, a state department of revenue commissioned an independent study to estimate the number of times people had cheated on their tax returns in the last five years. The researchers interviewed a random sample of 121 adults and found that the mean number of times they had cheated on their income taxes in the last five years was 2.7, with a standard deviation of 1.1.

a. Calculate a 95% confidence interval for this sample mean.

b. Explain what this result means.

22.4 The country of Mifflin is preparing for an upcoming presidential election. A random sample of 200 likely voters in Mifflin indicates that 57% are going to vote for the Hawk Party candidate, while the remaining 43% are planning on voting for the Gopher Party candidate.

a. Calculate a 95% confidence interval for the proportion voting for the Hawk Party candidate.

b. Calculate a 99% confidence interval for the proportion voting for the Hawk Party candidate.

c. Which of the two confidence intervals provides a better indicator of who will win the election? Who do you predict will win the election?

22.5 A long-running disagreement between science and humanities professors at Big Time University focuses on which department has the smarter students. As evidence supportive of the contention that science students are smarter, a physics professor shows that the mean grade point average for a random sample of 322 recent science graduates was 3.51 ($s = 1.2$). Asserting that there is no meaningful difference, a history professor shows that the mean grade point average for a sample of 485 recent humanities graduates was 3.36 ($s = 1.6$). Construct a 99% confidence interval for this difference of means, and explain which professor appears to be more correct.

22.6 Interested in the effects of income and poverty on robbery rates, a student selected a random sample of 125 cities and correlated average income and percentage of persons living in poverty with the robbery rate. She reported the following correlations:

Income and robbery: $r = -0.215$

Poverty and robbery: $r = 0.478$

a. Calculate a 95% confidence interval for each correlation.

b. Explain what these results mean.

22.7 Delinquency researchers at DP Institute interviewed a sample of 96 adolescents about their behavior. The researchers estimated a regression model, using number of delinquent acts in the last year as the dependent variable. The table of results follows:

Variable	*b*	Standard Error
Intercept	−0.21	0.15
Age	−0.02	0.01
Number of friends arrested	2.56	0.73
Number of hours per week studying	−0.17	0.08
Number of hours per week working	0.09	0.03
Self-esteem	−1.05	0.51

a. Calculate a 95% confidence interval for each of the independent variable regression coefficients.

b. Explain what these results mean.

22.8 In a follow-up to the analysis reported in Exercise 22.7, the researchers recoded delinquency as 0 = no delinquency and 1 = one or more delinquent acts. They estimated a logistic regression model and found the following:

Variable	*b*	Standard Error
Intercept	0.05	0.04
Age	−0.12	0.05
Number of friends arrested	1.86	0.57
Number of hours per week studying	−0.23	0.09
Number of hours per week working	0.44	0.17
Self-esteem	−0.79	0.38

a. Calculate a 95% confidence interval for each of the independent variable regression coefficients.

b. Explain what these results mean.

Computer Exercises

All of the statistical packages that we are familiar with allow for the straightforward computation of confidence intervals. SPSS and Stata both allow for easy reporting of confidence intervals, as we illustrate below. Four of the confidence intervals discussed in this chapter—sample mean, difference of means, regression, and logistic regression—are the focus of our discussion. There are no options in either program for computing confidence intervals for Pearson's r. Sample syntax in both SPSS (Chapter_22.sps) and Stata (Chapter_22.do) illustrate each of the following commands.

SPSS

To obtain the confidence interval for a sample mean, use the T-TEST command, but use the /TESTVAL option to test the value specified in the null hypothesis (e.g., 0):

```
T-TEST
/TESTVAL = 0
/VARIABLES = variable_name
/CRITERIA = CI(.95).
```

where the /TESTVAL = 0 implements the null hypothesis (H_0: $\mu = 0$) and the T-TEST command will test whether the sample mean is different from 0. The /CRITERIA = CI(.95) requests a 95 % confidence interval. If the /CRITERIA line is omitted from the command, the default output will still contain the 95 % confidence interval. Where you may want to include the /CRITERIA line is in the situation where you are interested in a different confidence interval. For example, a 90 % confidence interval would be requested with /CRITERIA = CI(.90).

In Chapter 11, we discussed how to compute an independent samples t-test in SPSS with the T-TEST command. Recall the basic format for the command is

```
T-TEST GROUPS = grouping_variable(category_1 category_2)
/VARIABLES = variable_name.
```

Similar to the one-sample t-test, the default output from executing the T-TEST for independent samples will include the 95% confidence interval. Should you be interested in a confidence interval of a different size, insert the /CRITERIA = CI(.##) option as explained above.

We discussed various features and option of the linear regression command in SPSS (REGRESSION) in Chapters 15 through 17. Confidence intervals are obtained by adding the option CI(##) to the /STATISTICS line in the REGRESSION command:

```
REGRESSION
/STATISTICS COEFF CI(95) R ANOVA
/DEPENDENT dep_var
/METHOD = ENTER list_of_indep_vars.
```

where we have inserted a request for a 95 % confidence interval by adding CI(95) to the /STATISTICS line. Should you want a different confidence interval, simply change the values inside the parentheses. As we noted in previous discussions of the REGRESSION command, when you desire some additional output, it is necessary to also request all of the standard output from the command (i.e., COEFF R ANOVA). Note that the specification of the confidence interval in the REGRESSION command does not require a decimal point in the CI(##) option.

Chapter 18's Computer Exercises focused on the use of the LOGISTIC REGRESSION command. To obtain the confidence intervals for the estimated coefficients, we need to add the /PRINT = CI(##) option line to the command:

```
LOGISTIC REGRESSION VARIABLES dep_var
/METHOD = ENTER list_of_indep_vars
/PRINT = CI(95).
```

The confidence intervals will appear as the far right columns in the table of coefficients. Should you want a different confidence interval, all you need to do is change the 95 in /PRINT = CI(95) to the value of interest. NOTE: The confidence intervals computed by SPSS in the LOGISTIC REGRESSION command are for the odds ratios [Exp(B)], not the original coefficients (B). Similar to requesting confidence intervals in REGRESSION, no decimal point is used in the CI(##) option in the LOGISTIC REGRESSION command.

Stata

The **ttest** command is used to compute a one-sample ttest, where we compare a sample mean to a hypothesized value:

ttest variable_name **==** hypothesized_value

In most cases, the hypothesized value will be 0. By default, the **ttest** command will compute 95% confidence intervals. If you are interested in a different value for the confidence interval, then add the **level**(##) option to the command line:

ttest variable_name **==** hypothesized_value, **level(##)**

For example, if we were interested in testing the hypothesis that GPA in the NYS data was equal to 0 and use 90% confidence intervals, we would enter the following command:

ttest gpa **==** 0, **level(90)**

We discussed the independent samples *t*-test in Chapter 11's Computer Exercises section using the **ttest** command. The default output from the **ttest** command is a 95% confidence interval. Should we be interested in a different range, we would add the **level**(##) option:

ttest variable_name, **by**(grouping_variable) **level(##)**

In previous chapters, we have discussed the use of both the **regress** and **logit** commands to estimate linear regression and binary logistic regression models, respectively. Both of those commands report 95% confidence intervals in the table of coefficients output by default. To request a different confidence interval, add the **level**(##) option to the end of the command line (following a comma):

regress dep_var list_of_indep_vars, **level(##)**

and

logit dep_var list_of_indep_vars, **level(##)**

Consistent with all other output, the confidence intervals appear in the far right column of the coefficients table.

Problems

Open the NY data file (nys_1.sav, nys _1_student.sav, or nys_1.dta) to answer questions 1 through 4.

1. For each of the following measures of delinquency, compute a 95% confidence interval and explain what it means.
   a. Number of times the youth has stolen something valued at less than $5.
   b. Number of times the youth has cheated on exams at school.
   c. Number of times the youth has been drunk.
   d. How would your results differ if you used a 90% confidence interval? A 99% confidence interval?
2. For each of the following difference of means tests, compute a 95% confidence interval and explain what it means.
   a. Does the number of times the youth has taken something valued at less than $5 differ for males and females?
   b. Does the number of times the youth has hit his or her parents differ for whites and African Americans?
   c. Does the number of times the youth has cheated on exams differ for students earning mostly Bs and students earning mostly Cs?
   d. How would your results differ if you used a 90% confidence interval? A 99% confidence interval?

3. Rerun two of the regression models you estimated in the Computer Exercises in Chapter 16. For each model, compute 95% confidence intervals for the regression coefficients and explain what each result means. How would your results differ if you had used a 90% confidence interval? A 99% confidence interval?

4. Rerun two of the logistic regression models you estimated in the Computer Exercises in Chapter 18. For each model, compute 95% confidence intervals for the odds ratios and explain what each result means. How would your results differ if you had used a 90% confidence interval? A 99% confidence interval?

   Open the Pennsylvania Sentencing data file (pcs_98.sav or pcs_98.dta) to answer questions 5 and 6.

5. Run a regression model using length of incarceration sentence as the dependent variable and age, race, sex, offense severity score, and prior criminal history score as the independent variables. Compute 99% confidence intervals for the regression coefficients and explain what each result means.

   Run a binary logistic regression model using incarceration as the dependent variable and age, race, sex, offense severity score, and prior criminal history score as the independent variables (this is the model you estimated in the Computer Exercises in Chapter 18). Compute 99% confidence intervals for the odds ratios and explain what each result means.