

Unit 26

Estimation with Confidence Intervals

Objectives:

- To see how confidence intervals are used to estimate a population proportion, a population mean, a difference in population proportions, or a difference in population means

Before introducing more hypothesis tests, we shall consider another type of statistical analysis which we can call *estimation*. *Point estimation* refers to estimating an unknown parameter with a single observed value.

For instance, we use \bar{x} to estimate μ , we use \bar{p} to estimate λ , and we use s to estimate σ . Point estimation of a parameter is rarely satisfactory by itself, since it does not provide any sense of how much difference in value might be expected between the estimator and the parameter being estimated. *Interval estimation*, which implies estimating an unknown parameter with an interval of values, instead of just a single value, is much more useful in general, since it does provide some sense of margin of error in the estimation.

The most widely used type of interval estimator is the *confidence interval*. A confidence interval is an interval estimate which has a percentage associated with it; this percentage is the probability that the confidence interval will actually contain the parameter being estimated and is called the *level of confidence* or *confidence level*. A confidence level reflects the amount of faith we have that the interval actually will contain the parameter of interest, and it is chosen using the same type of considerations that are involved in choosing a significance level in a hypothesis test. Just as a significance level is the probability of making a Type I error in hypothesis testing, a confidence level is the probability of not making the error where the observed confidence interval does not contain the parameter of interest. The most commonly chosen confidence levels are 90%, 95%, and 99% corresponding to the most commonly chosen significance levels 0.01, 0.05, and 0.10.

Choosing a 95% confidence level implies that we are willing to tolerate a 5% (or 0.05) probability of obtaining a confidence interval which does not contain the parameter of interest. In other words, 95% of all the possible confidence intervals that we might observe actually do contain the parameter of interest, and the other 5% do not contain the parameter of interest.

Confidence intervals can be utilized as an alternative to hypothesis testing or can be used in conjunction with hypothesis testing. Suppose, for instance, that a forest ranger is interested in the proportion of trees affected by a recent blight in a wooded area. In order to perform a hypothesis test, the ranger would have to formulate null and alternative hypotheses. If the ranger has no hypothesized value for the proportion of trees affected by the blight, then a confidence interval could be utilized instead of a hypothesis test. On the other hand, suppose the manager at certain manufacturing plant is interested in the proportion of underweight packages from an assembly line. If the manager wants to know whether or not this proportion is different from 0.07, a hypothesis test could be utilized. In fact, this hypothesis test was done previously in Table 20-1; Figure 20-3 graphically displays the rejection region for the test. Since the results of the test were statistically significant at the 0.05 level, we concluded that the proportion of underweight packages is different from 0.07, and the results of the test seemed to suggest that the proportion of underweight packages is greater than 0.07. While this conclusion certainly provides the manager with some useful information, it may not be enough. The manager may be interested in just how far above 0.07 the proportion of underweight packages is. A confidence interval could be utilized to address this issue.

Finding a confidence interval for a parameter essentially implies defining a likely set of values for the parameter based on the information available in a sample. You should not be surprised that defining such a likely set of values is related to defining a rejection region, which is an unlikely set of values for a test statistic in a hypothesis test. Returning to the illustration in Table 20-1 where the manager wants to know whether or not the proportion of underweight packages is different from 0.07, we recall that the shaded area of Figure 20-3 represents unlikely values for the z test statistic if $H_0: \lambda = 0.07$ is true. Consequently, the unshaded region of Figure 20-3 represents a likely set of values for the z test statistic if $H_0: \lambda = 0.07$ is true.

Suppose the sample size n is large enough to treat the sampling distribution of \bar{p} as a normal distribution (and you should recall how to verify this!). Then, roughly speaking, we know that the difference between \bar{p} and the actual value of λ is very likely to be within two or three standard deviations, where $\sigma_{\bar{p}} = \sqrt{\lambda(1-\lambda)/n}$ is the standard deviation of the sampling distribution of \bar{p} . We can make this statement more precise by choosing a confidence level on which to base the statement. In the illustration concerning the underweight packages, the manager might choose a confidence level of 95% to correspond with the 0.05 significance level used in the hypothesis test. The unshaded area of Figure 20-3 lies between $-z_{0.025} = -1.960$ and $z_{0.025} = +1.960$, and represents the probability that \bar{p} will be within $1.96\sqrt{\lambda(1-\lambda)/n}$ of the value of λ . In other words, we can be 95% sure that \bar{p} will be within $1.960\sqrt{\lambda(1-\lambda)/n}$ of the value of λ .

Subtracting and adding $1.960\sqrt{\lambda(1-\lambda)/n}$ to \bar{p} will define limits between which we can feel 95% confident that the value of λ lies. This would give us a 95% confidence interval for λ except for one problem. In order to be able to calculate $1.960\sqrt{\lambda(1-\lambda)/n}$, we must know the value of λ ; but if we knew the value of λ , there would certainly be no reason to try to estimate λ with a confidence interval! We can easily solve this problem by estimating λ with \bar{p} in the formula $\sqrt{\lambda(1-\lambda)/n}$, that is, by using $\sqrt{\bar{p}(1-\bar{p})/n}$ to estimate $\sqrt{\lambda(1-\lambda)/n}$.

Let us now illustrate how the manager in the illustration in Table 20-1 concerning the underweight packages can obtain a 95% confidence interval for the proportion of underweight packages. In a random sample of 250 packages, 27 were found to be underweight, from which it was found that $\bar{p} = 27/250 = 0.108$. The limits for the 95% confidence interval are obtained by multiplying $z_{0.025} = 1.960$ by $\sqrt{\bar{p}(1-\bar{p})/n} = \sqrt{0.108(1-0.108)/250} = 0.01963$, and then subtracting this result from, and adding this result to, $\bar{p} = 0.108$. Specifically, the 95% confidence interval is calculated as follows:

$$0.108 - (1.960)(0.01963) < \lambda < 0.108 + (1.960)(0.01963)$$

$$0.108 - 0.0385 < \lambda < 0.108 + 0.0385$$

$$0.0695 < \lambda < 0.1465$$

We interpret this confidence interval as follows:

We are 95% confident that the proportion of underweight packages is between 0.0695 and 0.1465.

The manager concludes from this confidence interval that the percentage of underweight packages is most likely to be between 6.95% and 14.65%. Notice that the hypothesized proportion 0.07 is very close to the lower limit of the confidence interval but still within the bounds determined by the confidence interval. As a general rule, when a null hypothesis is rejected, the hypothesized value will almost never be in the corresponding confidence interval; similarly, when a null hypothesis is not rejected, the hypothesized value will almost always be included in the corresponding confidence interval.

The manager now has some idea about the percentage of underweight packages; however, the length of the confidence interval obtained here may be too long to allow the manager to decide whether or not the proportion of underweight packages should be considered clinically significant. Suppose the manager is willing to live with, say, 8% or 9% of the packages being underweight, but not willing to live with more than 10% of the packages being underweight. The limits of the confidence interval suggest that the percentage of underweight packages may be more than 10% or less than 10%. It would be desirable to have a confidence interval with a smaller length in order to make this decision.

By asking for less confidence, the length of a confidence interval will decrease. For instance, if we had used a 90% confidence level instead of a 95% confidence level in the confidence interval we just previously obtained, we would have used $z_{0.05} = 1.645$ in place of $z_{0.025} = 1.960$. This would have given us a confidence

interval with a smaller length, but the price we pay for this is that we do not feel as certain that the value of λ is within the limits of the confidence interval. The only way to decrease the length of a confidence interval without changing the confidence level is to decrease the standard error by increasing the sample size. If the manager wants a confidence interval with a smaller length while still maintaining a 95% confidence level, the sample size of 250 will have to be increased.

We can make the following statements about confidence intervals in general. With a given sample size, increasing the confidence level will increase the length of the confidence interval, and decreasing the confidence level will decrease the length of the confidence interval. With a given confidence level, increasing the sample size will tend to decrease the length of the confidence interval, and decreasing the sample size will tend to increase the length of the confidence interval. These statements should match our intuition but should also be clear from the algebraic formula for computing the confidence interval in our previous illustration.

Self-Test Problem 26-1. Self Test Problem 20-1, a hypothesis test was performed concerning the proportion of registered voters intending to vote for a candidate in an upcoming election for governor. In the random sample of 500 registered voters selected for the hypothesis test, 165 said they would vote for the candidate. The results of the test, which were statistically significant at the 0.01 level, suggested that less than 0.40 of the registered voters intended to vote for the candidate.

- (a) Considering the results of the hypothesis test in Self Test Problem 20-1, explain why a confidence interval for the proportion of registered voters intending to vote for the candidate would be of interest.
- (b) Find and interpret a 99% confidence interval for the proportion of registered voters intending to vote for the candidate.
- (c) How would a 90% or 95% confidence interval be different from the one found in part (b)?
- (d) How would the confidence interval in part (b) most likely have been different if it had been based on a random sample of 1000 voters instead of the 500 actually used?

In general, confidence intervals involving proportions and means are obtained by subtracting and adding the proper number of standard errors. Access to the proper statistical software or programmable calculator will allow you to avoid having to do much of the calculation necessary to obtain a confidence interval, although we do display the calculations for several of the confidence intervals discussed in this chapter.

Let us now consider how we might obtain a confidence interval for the difference between two proportions. Suppose that channels 4 and 8 both broadcast the same 11:00 newscast, and that the difference between the two different viewing areas in the proportion of residents who viewed last night's newscast is of interest. A hypothesis test to see if there was any evidence of a difference in these two proportions was previously done; Figure 25-1 graphically displays the rejection region for the test. Since the results of the test were statistically significant at the 0.10 level, we concluded that there is a difference between the two viewing areas in the proportion of residents who viewed last night's newscast, and the results of the test seemed to suggest that the proportion of viewers is greater in the channel 8 viewing area. A confidence interval could now be utilized to obtain some idea of just how large this difference is and to aid in deciding whether or not this difference should be considered clinically significant. We shall choose a confidence level of 90% to correspond with the 0.10 significance level used in the hypothesis test.

You can recall, or look back at the illustration to verify, that in a random sample of 175 residents in the channel 4 viewing area, 49 residents viewed the newscast, and that in a random sample of 225 residents in the channel 8 viewing area, 81 residents viewed the newscast. Also recall or verify that the standard error of the difference between sample proportions was estimated by combining these two samples together under the assumption that the null hypothesis $H_0: \lambda_4 - \lambda_8 = 0$ was correct. However, if we are going to use a confidence interval to estimate the difference between proportions, we cannot assume the proportions are equal. Consequently, our estimate of the standard error must allow for the possibility that the two proportions are different (which is in fact precisely what we concluded in our hypothesis test comparing the two viewing areas!). An estimate of the standard error of the difference between two sample proportions, without assuming that λ_1 and λ_2 are equal, is given by

$$\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}} .$$

The channel 4 sample proportion of residents who viewed the newscast is $\bar{p}_4 = 49/175 = 0.28$, and the channel 8 sample proportion of residents who viewed the newscast is $\bar{p}_8 = 81/225 = 0.36$. We then estimate the standard error of the difference between the two sample proportions to be

$$\sqrt{\frac{0.28(1-0.28)}{175} + \frac{0.36(1-0.36)}{225}} = 0.04665 .$$

As with the confidence interval for one proportion, we assume that the difference between sample proportions has a normal distribution, which will be true with sufficiently large sample sizes. The limits for the 90% confidence interval are obtained by multiplying $z_{0.05} = 1.645$ by the estimate of the standard error 0.04665, and then subtracting this result from, and adding this result to, $\bar{p}_4 - \bar{p}_8 = 0.28 - 0.36 = -0.08$.

The choice of which sample proportion is subtracted from the other is arbitrary, since we have subtracted the larger proportion from the smaller proportion, the difference $\bar{p}_4 - \bar{p}_8$ turned out to be negative. The 90% confidence interval can be calculated as follows:

$$\begin{aligned} -0.08 - (1.645)(0.04665) &< \lambda_4 - \lambda_8 < -0.08 + (1.645)(0.04665) \\ -0.08 - 0.0767 &< \lambda_4 - \lambda_8 < -0.08 + 0.0767 \\ -0.1567 &< \lambda_4 - \lambda_8 < -0.0033 \end{aligned}$$

The fact that the confidence interval lies entirely below 0 (zero) is a reflection of the fact that $H_0: \lambda_4 - \lambda_8 = 0$ was rejected, and that the data suggested that the proportion of viewers is greater in the channel 8 viewing area. When interpreting a confidence interval for the difference between two proportions, it is generally easier to think in terms of positive differences instead of negative differences. We interpret this confidence interval as follows:

We are 90% confident that the difference between the two viewing areas in the proportion of residents who viewed the newscast is between 0.0033 and 0.1567, with the Channel 8 viewing area showing a larger proportion than the Channel 4 viewing area.

If the length of this confidence interval appears to be too large to give us an accurate indication of the difference, increasing the sample sizes in order will decrease the length of the interval without changing the confidence level.

Self-Test Problem 26-2. The proportion of dissatisfied customers is being compared for two airlines, Safeway and Ricketty. In a random sample of 801 Safeway customers, 153 were dissatisfied; in a random sample of 735 Ricketty customers, 111 were dissatisfied.

- Find and interpret a 90% confidence interval for the difference in proportion of dissatisfied customers between the two airlines.
- If a hypothesis test with $\alpha = 0.10$ were to have been performed to see if there were any evidence of a difference in proportion of dissatisfied customers between the two airlines, what would most likely have been concluded?
- How would a 95% or 99% confidence interval be different from the one found in part (a)?
- How would the confidence interval in part (a) most likely have been different if it had been based on two random samples each of size 400 customers instead of the sample sizes actually used?

We have now discussed confidence intervals involving either one proportion or two proportions. Let us now consider confidence intervals involving either one mean or two means. You shall find that the general procedure is the same as with one proportion or two proportions.

Suppose we want a confidence interval for the mean amount of cereal per box from an assembly line. A hypothesis test to see if there was any evidence of that the mean is different from an advertised 12 ounces was previously done; Figure 21-1 graphically displays the rejection region for the test. Since the results of the test were statistically significant at the 0.05 level, we concluded that the mean amount of cereal per box is different from 12 ounces, and the data suggest that the mean is less than 12 ounces. A confidence interval could now be utilized to obtain some idea of just how large this difference is and to aid in deciding whether or not this difference should be considered clinically significant. We shall choose a confidence level of 95% to correspond with the 0.05 significance level used in the hypothesis test.

You can check (or recall) that this hypothesis test was done using a random sample of $n = 16$ boxes of cereal, from which we found $\bar{x} = 11.880$ ounces and $s = 0.220$ ounces. The limits for the 95% confidence interval are obtained by first multiplying $t_{15,0.025}$ by $s/\sqrt{n} = 0.220/\sqrt{16} = 0.055$, which is the estimate of the standard error that appears in the denominator of the one sample t statistic; this result is then subtracted from, and added to, \bar{x} . We will let you do these calculations of the limits of the confidence interval, and of course you can also use appropriate statistical software or a programmable calculator; then write the interpretation of the results. Your interpretation of results should be as follows:

We are 95% confident that the mean amount of cereal per box is between 11.763 and 11.997 ounces.

The fact that the confidence interval lies entirely below 12 ounces is a reflection of the fact that $H_0: \mu = 12$ was rejected, and the data suggest that the mean is less than 12 ounces.

Each of the confidence intervals in our illustrations thus far was used after a hypothesis test had been performed. As we stated earlier, a confidence interval may also be utilized in place of a hypothesis test. Suppose we are interested in the difference in mean grip strength, measured in pounds of force, between right-hand grip strength and left-hand grip strength in 21-year old, right-handed males. If our purpose is simply to estimate the difference without formulating any hypotheses, then we can utilize a confidence interval. We shall choose a 99% confidence level. Self-Test Problem 23-1(f) contains the following sample of $n = 8$ paired observations of grip strength (lbs.) which can be used to obtain a confidence interval for the mean difference:

Right	86	87	83	93	71	76	91	100
Left	84	84	76	90	67	70	87	99

You may verify that after subtracting the left-hand grip strength from the right-hand grip strength, the sample mean difference is $\bar{d} = +3.75$ pounds and the sample standard deviation of the differences is $s_d = 1.982$ pounds. The limits for the 99% confidence interval are obtained by first multiplying $t_{7,0.005}$ by $s_d/\sqrt{n} = 1.982/\sqrt{8} = 0.700743$, which is the estimate of the standard error of \bar{d} ; this result is then subtracted from, and added to, \bar{d} . We will let you do these calculations of the limits of the confidence interval, and of course you can also use appropriate statistical software or programmable calculator; then write the interpretation of the results. Your interpretation of results should be as follows:

We are 99% confident that the mean difference in grip strength between the right and left hands is between 1.30 and 6.20 pounds, with the right hand showing a larger mean.

The fact that the confidence interval lies entirely above zero suggests that $H_0: \mu_{R-L} = 0$ would be rejected if the two-sided hypothesis test were performed with $\alpha = 0.01$ (which corresponds to the selected 99% confidence level).

Finding a confidence interval for the mean of a population from one random sample and finding the confidence interval for the mean of a population of differences from one random sample of paired observations are essentially the same thing. This is because one random sample of paired observations can be treated as one random sample of observed differences between pairs. Two independent random samples would be used to find a confidence interval for the difference between two population means.

Self-Test Problem 26-3. In Self-Test Problem 21-2, a hypothesis test was performed concerning the mean lifetime of a manufacturer's light bulbs. The results of the test, which were statistically significant at the 0.10 level, suggested that the mean lifetime of the bulbs is less than the claimed 500 hours. In the random sample of 40 light bulbs selected, the sample mean was found to be 483.75 hours and the sample standard deviation was found to be 26.35 hours.

- (a) Considering the results of the hypothesis test in Self-Test Problem 21-2, explain why a confidence interval for the mean lifetime of the light bulbs would be of interest.
- (b) Find and interpret a 90% confidence interval for the mean lifetime of the manufacturer's light bulbs.
- (c) How would a 95% or 99% confidence interval be different from the one found in part (b)?
- (d) How would the confidence interval in part (b) most likely have been different if it had been based on a random sample of 20 bulbs instead of the 40 actually used?

Self-Test Problem 26-4. In Self-Test Problem 24-1, a hypothesis test was performed concerning the difference in mean time spent watching TV weekly and the mean time spent listening to the radio weekly in a particular state. The results of the test, which were statistically significant at the 0.10 level, suggested that the mean weekly radio hours is larger. In the random sample of 30 individuals selected (for which results were recorded in the SURVEY DATA, displayed as Data Set 1-1 at the end of Unit 1), the sample mean difference, subtracting weekly radio hours from weekly TV hours, was found to be -4.9 hours and the sample standard deviation of the differences was found to be 10.1755 hours.

- (a) Considering the results of the hypothesis test in Self-Test Problem 24-1, explain why a confidence interval for the mean difference in time spent watching TV weekly and time spent listening to the radio weekly would be of interest.
- (b) Find and interpret a 90% confidence interval for the mean difference in time spent watching TV weekly and time spent listening to the radio weekly.
- (c) How would a 95% or 99% confidence interval be different from the one found in part (b)?
- (d) How would it be possible to decrease the length of the 90% confidence interval in part (b) without changing the level of confidence?

Self-Test Problem 26-5. In Self-Test Problem 22-2, a hypothesis test was performed concerning the mean yearly income in a state. The results of the test were not statistically significant at the 0.01 level. In the random sample of 30 individuals selected (for which results were recorded in the SURVEY DATA, displayed as Data Set 1-1 at the end of Unit 1), the sample mean was found to be 45.4 thousand dollars and the sample standard deviation was found to be 15.895 thousand dollars.

- (a) Considering the results of the hypothesis test in Self-Test Problem 22-2, explain why a confidence interval for the mean yearly income would not be of interest.
- (b) Explain why we would expect a 99% confidence interval for the mean yearly income to contain the hypothesized 42 thousand dollars in Self Test Problem 22-2; then verify that the 99% confidence interval does contain the hypothesized mean.

Recall that when we considered a hypothesis test concerning the difference between two population means, both the pooled two sample t test and the separate two sample t test were available. However, we suggested that the separate t test always be used, since this test does not require that the variation be the same in the two populations being compared as is required by the pooled two sample t test, and when this requirement is met, both tests give close to identical results. For essentially the same reason, we suggest that a confidence interval based on the separate two sample t test statistic always be used instead of a confidence interval based on the pooled two sample t test statistic. As can be seen in Section B.2 of Appendix B, the necessary calculations for the t statistics can be rather difficult and tedious. Access to the appropriate statistical software and programmable calculators is usually essential in practice for performing the separate two sample t test or obtaining the limits of the confidence interval based on the separate t statistic.

As an example, suppose we want a 95% confidence interval for the difference in mean breaking strength (lbs.) between Deluxe brand rope and Econ brand rope. Self Test Problem 23-2(c) contains the following samples of $n_D = 12$ breaking strengths for Deluxe rope and $n_E = 10$ breaking strengths for Econ rope:

Deluxe	183	197	179	177	176	187	171	199	182	197	197	187
Econ	162	165	180	185	189	176	159	164	186	164		

With this data you will find that $\bar{x}_D = 186$ lbs. and $s_D = 9.6011$ lbs. are the respective sample mean and standard deviation for the Deluxe sample, and that $\bar{x}_E = 173$ lbs. and $s_E = 11.4018$ lbs. are the respective sample mean and standard deviation for the Econ sample. By using the formula for degrees of freedom given with the separate two sample t test statistic, one will find that $df = 18$ when rounded to the nearest integer. The

limits for the 95% confidence interval are obtained by first multiplying $t_{18;0.025} = 2.101$ by $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 4.5477$

which is the estimate of the standard error that appears in the denominator of the separate two sample t statistic; this result is then subtracted from, and added to, the difference between the two sample means (where the order of subtraction is arbitrary and will not affect the final results). We will let you do these calculations of the limits of the confidence interval, and of course you may use appropriate statistical software or a programmable calculator. Then, write the interpretation of the results. (Keep in mind that when interpreting a confidence interval for the difference between two means, it is generally easier to think in terms of positive differences instead of negative differences.) Your interpretation of results should be as follows:

We are 95% confident that the difference in mean breaking strength between Deluxe brand rope and Econ brand rope is between 3.4 and 22.6 lbs., with the Deluxe brand showing a larger mean.

The fact that the confidence interval lies entirely on one side of zero (which side depending on the order of subtraction selected) is a reflection of the fact that $H_0: \mu_D - \mu_E = 0$ would be rejected if the two-sided hypothesis test were performed with $\alpha = 0.05$ (which corresponds to the selected 95% confidence level).

(Just for the sake of comparison, we could find the 95% confidence interval based on the pooled t statistic. We find that $df = 20$, and the limits of the confidence interval are 3.7 and 22.3. These limits do not appear to be drastically different from the limits based on the separate t statistic. This is most likely a reflection of the fact that there is not a significant difference in standard deviations for the two rope brands.)

Self-Test Problem 26-6. A confidence interval for the difference in mean time spent watching TV weekly between males and females in a particular state is of interest. The sample of males selected and the sample of females selected to obtain the SURVEY DATA, displayed as Data Set 1-1 at the end of Unit 1, are treated as two independent random samples.

- Find and interpret a 95% confidence interval for the difference in mean time spent watching TV weekly between males and females; use the confidence interval based on the separate t statistic.
- If a two-sided hypothesis test with $\alpha = 0.05$ were to have been performed to see if there were any evidence of a difference in mean time spent watching TV weekly for males and females, what would have been concluded?
- How would a 90% confidence interval be different from the one found in part (a)?
- How would a 99% confidence interval be different from the one found in part (a)?
- How would the confidence interval in part (a) most likely have been different, if it had been based on random samples each of size 40 instead of the sample sizes actually used?

Self-Test Problem 26-7. In Self-Test Problem 24-3, data was observed in a study to see if the mean blood pressure was higher for employees at a Nuketown factory than for employees at a Highville factory. The results of the test were statistically significant at the 0.01 level. In a random sample of 35 blood pressures observed at the Nuketown factory, the sample mean was found to be 126.2 and the sample standard deviation was found to be 8.9; in a random sample of 40 blood pressures observed at the Highville factory, the sample mean was found to be 117.9 and the sample standard deviation was found to be 10.4.

- Considering the results of the hypothesis test in Self-Test Problem 24-3, explain why a confidence interval for the difference in mean blood pressure of employees between the Nuketown and Highville factories would be of interest.
- Find a 99% confidence interval for the difference in mean blood pressure between the two factories.
- How would a 90% or 95% confidence interval be different from the one found in part (b)?
- How would the confidence interval in part (b) most likely have been different if it had been based on two random samples each of size 80 instead of the sample sizes actually used?

We have now introduced hypothesis tests and confidence intervals concerning means and proportions for both one sample data and two sample data. Each of these hypothesis tests and confidence intervals is based on the assumption that either we are sampling from a normal distribution or we have a sufficiently large sample size. (Looking at normal probability plots, looking at box plots, and looking for outliers can be helpful in deciding whether the confidence intervals involving means will be appropriate.)

Answers to Self-Test Problems

26-1 (a) Rejecting H_0 suggests that the proportion of the registered voters intending to vote for the candidate is different from 0.40 but gives us no information about size of the difference. A confidence interval will provide information about the value of the proportion.

(b) $n = 500$, $\bar{p} = 165/500 = 0.33$.

$z_{0.005} = 2.576$ can be used to obtain the 99% confidence interval limits $0.276 < \lambda < 0.384$.

We are 99% confident that the proportion of registered voters intending to vote for the candidate is between 0.276 and 0.384.

(c) A 90% or 95% confidence interval would have a shorter length than the one in part (b).

(d) The confidence interval in part (b) would tend to have a shorter length if based on a random sample of 1000 voters.

26-2 (a) $\bar{p}_S = 153/801 = 0.191$, $\bar{p}_R = 111/735 = 0.151$.

$z_{0.05} = 1.645$ can be used to obtain the 90% confidence interval limits $0.0085 < \lambda_S - \lambda_R < 0.0715$.

We are 90% confident that the difference in proportion of dissatisfied customers between the Safeway and Ricketty airlines is between 0.0085 and 0.0715, with Safeway airlines showing the larger proportion.

(b) Since the 90% confidence interval in part (a) lies entirely above 0 (zero), $H_0: \lambda_S - \lambda_R = 0$ would have been rejected with $\alpha = 0.10$, and the results of the test would suggest that the proportion of dissatisfied customers is higher among Safeway customers.

(c) A 90% or 95% confidence interval would have a longer length than the one in part (a).

(d) The confidence interval in part (a) would tend to have a shorter length if based on two random samples each of size 400 customers.

(continued next page)

Answers to Self-Test Problems

(continued from previous page)

- 26-3** (a) Rejecting H_0 suggests that the mean lifetime of the manufacturer's light bulbs is different from 500 hours but gives us no information about the size of the difference. A confidence interval will provide information about the value of the mean.
- (b) $n = 40$, $\bar{x} = 483.75$ hours, $s = 26.35$ hours.
 $t_{39;0.05} = 1.684$ can be used to obtain the 90% confidence interval limits $476.7 < \mu < 490.8$.
We are 90% confident that the mean lifetime of the manufacturer's light bulbs is between 476.7 and 490.8 hours.
- (c) A 95% or 99% confidence interval would have a longer length than the one in part (b).
- (d) The confidence interval in part (b) would tend to have a longer length if based on a random sample of 20 bulbs.
- 26-4** (a) Rejecting H_0 suggests that there is a difference in the mean time spent watching TV weekly and mean time spent listening to the radio weekly but gives us no information about the size of the difference. A confidence interval will provide information about the size of the difference.
- (b) $n = 40$, $\bar{d} = -4.9$ hours, $s_d = 10.1755$ hours.
 $t_{29;0.05} = 1.699$ can be used to obtain the 90% confidence interval limits $-8.057 < \mu_{T-R} < -1.743$.
We are 90% confident that the mean difference in time spent watching TV weekly and time spent listening to the radio weekly is between 1.743 and 8.057 hours, with weekly radio hours showing a larger mean.
- (c) A 95% or 99% confidence interval would have a longer length than the one in part (b).
- (d) Increasing the sample size will tend to decrease the length of a confidence interval with a given confidence level.
- 26-5** (a) Since not rejecting H_0 suggests that the mean yearly income is not larger than the hypothesized 42 thousand dollars, a confidence interval for the mean difference would not be of interest.
- (b) The limits of the 99% confidence interval are 37.4 thousand dollars and 53.4 thousand dollars, between which lies the hypothesized mean 42 thousand dollars.
- 26-6** (a) $n_M = 15$, $\bar{x}_M = 14.4$ hours, $s_M = 4.9106$ hours, $n_F = 15$, $\bar{x}_F = 10.2$ hours, $s_F = 4.0391$ hours.
 $t_{27;0.025} = 2.052$ can be used to obtain the 95% confidence interval limits $0.83 < \mu_M - \mu_F < 7.57$.
We are 95% confident that the difference in mean time spent watching TV weekly between males and females is between 0.83 and 7.57 hours, with males showing a larger mean.
- (b) Since the 95% confidence interval in part (a) lies entirely above 0 (zero), $H_0: \mu_M - \mu_F = 0$ would have most likely been rejected with $\alpha = 0.05$, and the results of the test would suggest that males showing a larger mean.
- (c) A 90% confidence interval would have a shorter length than the one in part (a).
- (d) A 99% confidence interval would have a longer length than the one in part (a).
- (e) The confidence interval in part (a) would tend to have a shorter length if based on two random samples each of size 40.

(continued next page)

Answers to Self-Test Problems

(continued from previous page)

- 26-7** (a) Rejecting H_0 suggests that the mean blood pressure is higher among Nuketown factory employees than among Highville factory employees but gives us no information about the size of the difference. A confidence interval will provide information about the size of the difference.
- (b) $n_N = 35$, $\bar{x}_N = 126.2$, $s_N = 8.9$, $n_H = 40$, $\bar{x}_H = 117.9$, $s_H = 10.4$.
 $t_{73;0.005} = 2.660$ can be used to obtain the 99% confidence interval limits $2.4 < \mu_N - \mu_H < 14.2$.
We are 99% confident that the difference in mean blood pressure of employees between the males Nuketown and Highville factories is between 2.4 and 14.2, with the Nuketown factory showing a larger mean.
- (b) Since the 95% confidence interval in part (a) lies entirely above 0 (zero), $H_0: \mu_M - \mu_F = 0$ would have most likely been rejected with $\alpha = 0.05$, and the results of the test would suggest that males showing a larger mean.
- (c) A 90% or 95% confidence interval would have a shorter length than the one in part (b).
- (d) The confidence interval in part (b) would tend to have a shorter length if based on two random samples each of size 80.

Summary

Point estimation refers to estimating an unknown parameter with a single observed value. *Interval estimation*, which implies estimating an unknown parameter with an interval of values, is much more useful in general, since it does provide some sense of margin of error in the estimation.

The most widely used type of interval estimator is the *confidence interval*. A confidence interval is an interval estimate which has a percentage associated with it; this percentage is the probability that the confidence interval will actually contain the parameter being estimated and is called the *level of confidence* or *confidence level*. A confidence level reflects the amount of faith we have that the interval actually will contain the parameter of interest, and it is chosen using the same type of considerations that are involved in choosing a significance level. The most commonly chosen confidence levels are 90%, 95%, and 99%, corresponding to the most commonly chosen significance levels 0.01, 0.05, and 0.10. While hypothesis tests can be one-sided or two-sided, confidence intervals are generally always two-sided.

The limits for a confidence interval involving one or more proportions or one or more means are obtained by first multiplying an appropriate tabled z value or tabled t value by an appropriate estimate of standard error, then subtracting this result from, and adding this result to, the point estimate. The formula for the proper standard error in each case is the one in the denominator of the corresponding z or t test statistic. Access to the appropriate statistical software or programmable calculator allows one to avoid having to do much of the calculation necessary to obtain a confidence interval.