



This is “Two-Sample Problems”, chapter 9 from the book [Beginning Statistics \(index.html\)](#) (v. 1.0).

This book is licensed under a [Creative Commons by-nc-sa 3.0](http://creativecommons.org/licenses/by-nc-sa/3.0/) license. See the license for more details, but that basically means you can share this book as long as you credit the author (but see below), don't make money from it, and do make it available to everyone else under the same terms.

This content was accessible as of December 29, 2012, and it was downloaded then by [Andy Schmitz](#) (<http://lardbucket.org>) in an effort to preserve the availability of this book.

Normally, the author and publisher would be credited here. However, the publisher has asked for the customary Creative Commons attribution to the original publisher, authors, title, and book URI to be removed. Additionally, per the publisher's request, their name has been removed in some passages. More information is available on this project's [attribution page](http://2012books.lardbucket.org/attribution.html?utm_source=header).

For more information on the source of this book, or why it is available for free, please see [the project's home page](http://2012books.lardbucket.org/). You can browse or download additional books there.

Chapter 9

Two-Sample Problems

The previous two chapters treated the questions of estimating and making inferences about a parameter of a single population. In this chapter we consider a comparison of parameters that belong to two different populations. For example, we might wish to compare the average income of all adults in one region of the country with the average income of those in another region, or we might wish to compare the proportion of all men who are vegetarians with the proportion of all women who are vegetarians.

We will study construction of confidence intervals and tests of hypotheses in four situations, depending on the parameter of interest, the sizes of the samples drawn from each of the populations, and the method of sampling. We also examine sample size considerations.

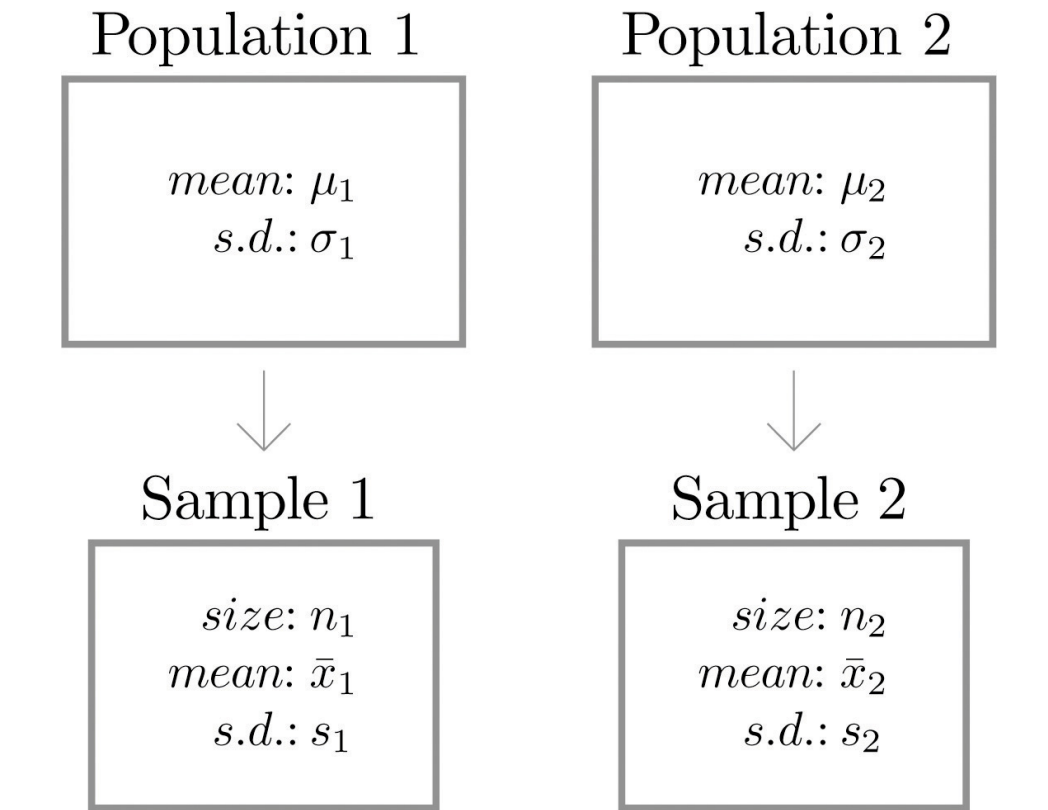
9.1 Comparison of Two Population Means: Large, Independent Samples

LEARNING OBJECTIVES

1. To understand the logical framework for estimating the difference between the means of two distinct populations and performing tests of hypotheses concerning those means.
2. To learn how to construct a confidence interval for the difference in the means of two distinct populations using large, independent samples.
3. To learn how to perform a test of hypotheses concerning the difference between the means of two distinct populations using large, independent samples.

Suppose we wish to compare the means of two distinct populations. [Figure 9.1 "Independent Sampling from Two Populations"](#) illustrates the conceptual framework of our investigation in this and the next section. Each population has a mean and a standard deviation. We arbitrarily label one population as Population 1 and the other as Population 2, and subscript the parameters with the numbers 1 and 2 to tell them apart. We draw a random sample from Population 1 and label the sample statistics it yields with the subscript 1. Without reference to the first sample we draw a sample from Population 2 and label its sample statistics with the subscript 2.

Figure 9.1 Independent Sampling from Two Populations



Definition

Samples from two distinct populations are **independent** if each one is drawn without reference to the other, and has no connection with the other.

Our goal is to use the information in the *samples* to estimate the difference $\mu_1 - \mu_2$ in the means of the two *populations* and to make statistically valid inferences about it.

Confidence Intervals

Since the mean \bar{x}_1 of the sample drawn from Population 1 is a good estimator of μ_1 and the mean \bar{x}_2 of the sample drawn from Population 2 is a good estimator of μ_2 , a reasonable point estimate of the difference $\mu_1 - \mu_2$ is $\bar{x}_1 - \bar{x}_2$. In order to widen this point estimate into a confidence interval, we first suppose that both samples

are large, that is, that both $n_1 \geq 30$ and $n_2 \geq 30$. If so, then the following formula for a confidence interval for $\mu_1 - \mu_2$ is valid. The symbols s_1^2 and s_2^2 denote the squares of s_1 and s_2 . (In the relatively rare case that both population standard deviations σ_1 and σ_2 are known they would be used instead of the sample standard deviations.)

100 (1 - α) % Confidence Interval for the Difference Between Two Population Means: Large, Independent Samples

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

The samples must be independent, and *each* sample must be large: $n_1 \geq 30$ and $n_2 \geq 30$.

EXAMPLE 1

To compare customer satisfaction levels of two competing cable television companies, 174 customers of Company 1 and 355 customers of Company 2 were randomly selected and were asked to rate their cable companies on a five-point scale, with 1 being least satisfied and 5 most satisfied. The survey results are summarized in the following table:

Company 1	Company 2
$n_1 = 174$	$n_2 = 355$
$\bar{x}_1 = 3.51$	$\bar{x}_2 = 3.24$
$s_1 = 0.51$	$s_2 = 0.52$

Construct a point estimate and a 99% confidence interval for $\mu_1 - \mu_2$, the difference in average satisfaction levels of customers of the two companies as measured on this five-point scale.

Solution:

The point estimate of $\mu_1 - \mu_2$ is

$$\bar{x}_1 - \bar{x}_2 = 3.51 - 3.24 = 0.27.$$

In words, we estimate that the average customer satisfaction level for Company 1 is 0.27 points higher on this five-point scale than it is for Company 2.

To apply the formula for the confidence interval, proceed exactly as was done in [Chapter 7 "Estimation"](#). The 99% confidence level means that $\alpha = 1 - 0.99 = 0.01$ so that $z_{\alpha/2} = z_{0.005}$. From [Figure 12.3 "Critical Values of "](#) we read directly that $z_{0.005} = 2.576$. Thus

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0.27 \pm 2.576 \sqrt{\frac{0.51^2}{174} + \frac{0.52^2}{355}} = 0.27 \pm$$

We are 99% confident that the difference in the population means lies in the interval $[0.15, 0.39]$, in the sense that in repeated sampling 99% of all intervals constructed from the sample data in this manner will contain

$\mu_1 - \mu_2$. In the context of the problem we say we are 99% confident that the average level of customer satisfaction for Company 1 is between 0.15 and 0.39 points higher, on this five-point scale, than that for Company 2.

Hypothesis Testing

Hypotheses concerning the relative sizes of the means of two populations are tested using the same critical value and p -value procedures that were used in the case of a single population. All that is needed is to know how to express the null and alternative hypotheses and to know the formula for the standardized test statistic and the distribution that it follows.

The null and alternative hypotheses will always be expressed in terms of the difference of the two population means. Thus the null hypothesis will always be written

$$H_0 : \mu_1 - \mu_2 = D_0$$

where D_0 is a number that is deduced from the statement of the situation. As was the case with a single population the alternative hypothesis can take one of the three forms, with the same terminology:

Form of H_a	Terminology
$H_a : \mu_1 - \mu_2 < D_0$	Left-tailed
$H_a : \mu_1 - \mu_2 > D_0$	Right-tailed
$H_a : \mu_1 - \mu_2 \neq D_0$	Two-tailed

As long as the samples are independent and both are large the following formula for the standardized test statistic is valid, and it has the standard normal distribution. (In the relatively rare case that both population standard deviations σ_1 and σ_2 are known they would be used instead of the sample standard deviations.)

**Standardized Test Statistic for Hypothesis Tests
Concerning the Difference Between Two Population
Means: Large, Independent Samples**

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

The test statistic has the standard normal distribution.

The samples must be independent, and *each* sample must be large: $n_1 \geq 30$ and $n_2 \geq 30$.

EXAMPLE 2

Refer to [Note 9.4 "Example 1"](#) concerning the mean satisfaction levels of customers of two competing cable television companies. Test at the 1% level of significance whether the data provide sufficient evidence to conclude that Company 1 has a higher mean satisfaction rating than does Company 2. Use the critical value approach.

Solution:

- Step 1. If the mean satisfaction levels μ_1 and μ_2 are the same then $\mu_1 = \mu_2$, but we always express the null hypothesis in terms of the difference between μ_1 and μ_2 , hence H_0 is $\mu_1 - \mu_2 = 0$. To say that the mean customer satisfaction for Company 1 is higher than that for Company 2 means that $\mu_1 > \mu_2$, which in terms of their difference is $\mu_1 - \mu_2 > 0$. The test is therefore

$$H_0 : \mu_1 - \mu_2 = 0$$

$$\text{vs. } H_a : \mu_1 - \mu_2 > 0 \quad @ \alpha = 0.01$$

- Step 2. Since the samples are independent and both are large the test statistic is

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

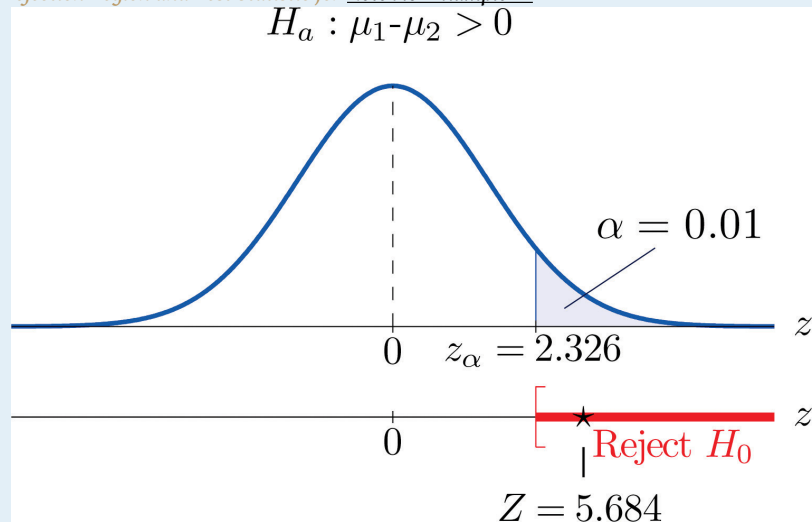
- Step 3. Inserting the data into the formula for the test statistic gives

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(3.51 - 3.24) - 0}{\sqrt{\frac{0.51^2}{174} + \frac{0.52^2}{355}}} = 5.684$$

- Step 4. Since the symbol in H_a is “>” this is a right-tailed test, so there is a single critical value, $z_\alpha = z_{0.01}$, which from the last line in [Figure 12.3 "Critical Values of "](#) we read off as 2.326. The rejection region is $[2.326, \infty)$.

Figure 9.2

Rejection Region and Test Statistic for [Note 9.6 "Example 2"](#)



- Step 5. As shown in [Figure 9.2 "Rejection Region and Test Statistic for "](#) the test statistic falls in the rejection region. The decision is to reject H_0 . In the context of the problem our conclusion is:

The data provide sufficient evidence, at the 1% level of significance, to conclude that the mean customer satisfaction for Company 1 is higher than that for Company 2.

EXAMPLE 3

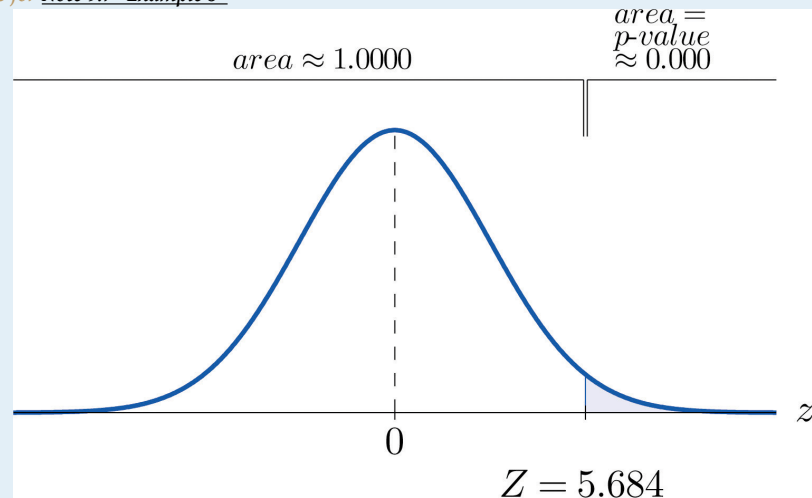
Perform the test of [Note 9.6 "Example 2"](#) using the p -value approach.

Solution:

The first three steps are identical to those in [Note 9.6 "Example 2"](#).

- Step 4. The observed significance or p -value of the test is the area of the right tail of the standard normal distribution that is cut off by the test statistic $Z = 5.684$. The number 5.684 is too large to appear in [Figure 12.2 "Cumulative Normal Probability"](#), which means that the area of the left tail that it cuts off is 1.0000 to four decimal places. The area that we seek, the area of the right tail, is therefore $1 - 1.0000 = 0.0000$ to four decimal places. See [Figure 9.3](#). That is, $p\text{-value} = 0.0000$ to four decimal places. (The actual value is approximately 0.000 000 007.)

Figure 9.3
P-Value for Note 9.7 "Example 3"



- Step 5. Since $0.0000 < 0.01$, $p\text{-value} < \alpha$ so the decision is to reject the null hypothesis:

The data provide sufficient evidence, at the 1% level of significance, to conclude that the mean customer satisfaction for Company 1 is higher than that for Company 2.

KEY TAKEAWAYS

- A point estimate for the difference in two population means is simply the difference in the corresponding sample means.
- In the context of estimating or testing hypotheses concerning two population means, “large” samples means that *both* samples are large.
- A confidence interval for the difference in two population means is computed using a formula in the same fashion as was done for a single population mean.
- The same five-step procedure used to test hypotheses concerning a single population mean is used to test hypotheses concerning the difference between two population means. The only difference is in the formula for the standardized test statistic.

EXERCISES

BASIC

1. Construct the confidence interval for $\mu_1 - \mu_2$ for the level of confidence and the data from independent samples given.
 - a. 90% confidence,
 $n_1 = 45, \bar{x}_1 = 27, s_1 = 2$
 $n_2 = 60, \bar{x}_2 = 22, s_2 = 3$
 - b. 99% confidence,
 $n_1 = 30, \bar{x}_1 = -112, s_1 = 9$
 $n_2 = 40, \bar{x}_2 = -98, s_2 = 4$
2. Construct the confidence interval for $\mu_1 - \mu_2$ for the level of confidence and the data from independent samples given.
 - a. 95% confidence,
 $n_1 = 110, \bar{x}_1 = 77, s_1 = 15$
 $n_2 = 85, \bar{x}_2 = 79, s_2 = 21$
 - b. 90% confidence,
 $n_1 = 65, \bar{x}_1 = -83, s_1 = 12$
 $n_2 = 65, \bar{x}_2 = -74, s_2 = 8$
3. Construct the confidence interval for $\mu_1 - \mu_2$ for the level of confidence and the data from independent samples given.
 - a. 99.5% confidence,
 $n_1 = 130, \bar{x}_1 = 27.2, s_1 = 2.5$
 $n_2 = 155, \bar{x}_2 = 38.8, s_2 = 4.6$
 - b. 95% confidence,
 $n_1 = 68, \bar{x}_1 = 215.5, s_1 = 12.3$

$$n_2 = 84, \bar{x}_2 = 287.8, s_2 = 14.1$$

4. Construct the confidence interval for $\mu_1 - \mu_2$ for the level of confidence and the data from independent samples given.
- a. 99.9% confidence,
 $n_1 = 275, \bar{x}_1 = 70.2, s_1 = 1.5$
 $n_2 = 325, \bar{x}_2 = 63.4, s_2 = 1.1$
- b. 90% confidence,
 $n_1 = 120, \bar{x}_1 = 35.5, s_1 = 0.75$
 $n_2 = 146, \bar{x}_2 = 29.6, s_2 = 0.80$
5. Perform the test of hypotheses indicated, using the data from independent samples given. Use the critical value approach. Compute the p -value of the test as well.
- a. Test $H_0 : \mu_1 - \mu_2 = 3$ vs. $H_a : \mu_1 - \mu_2 \neq 3$ @ $\alpha = 0.05$,
 $n_1 = 35, \bar{x}_1 = 25, s_1 = 1$
 $n_2 = 45, \bar{x}_2 = 19, s_2 = 2$
- b. Test $H_0 : \mu_1 - \mu_2 = -25$ vs. $H_a : \mu_1 - \mu_2 < -25$ @ $\alpha = 0.10$,
 $n_1 = 85, \bar{x}_1 = 188, s_1 = 15$
 $n_2 = 62, \bar{x}_2 = 215, s_2 = 19$
6. Perform the test of hypotheses indicated, using the data from independent samples given. Use the critical value approach. Compute the p -value of the test as well.
- a. Test $H_0 : \mu_1 - \mu_2 = 45$ vs. $H_a : \mu_1 - \mu_2 > 45$ @ $\alpha = 0.001$,
 $n_1 = 200, \bar{x}_1 = 1312, s_1 = 35$
 $n_2 = 225, \bar{x}_2 = 1256, s_2 = 28$
- b. Test $H_0 : \mu_1 - \mu_2 = -12$ vs. $H_a : \mu_1 - \mu_2 \neq -12$ @ $\alpha = 0.10$,
 $n_1 = 35, \bar{x}_1 = 121, s_1 = 6$
 $n_2 = 40, \bar{x}_2 = 135, s_2 = 7$

7. Perform the test of hypotheses indicated, using the data from independent samples given. Use the critical value approach. Compute the p -value of the test as well.
- Test $H_0 : \mu_1 - \mu_2 = 0$ vs. $H_a : \mu_1 - \mu_2 \neq 0$ @ $\alpha = 0.01$,
 $n_1 = 125, \bar{x}_1 = -46, s_1 = 10$
 $n_2 = 90, \bar{x}_2 = -50, s_2 = 13$
 - Test $H_0 : \mu_1 - \mu_2 = 20$ vs. $H_a : \mu_1 - \mu_2 > 20$ @ $\alpha = 0.05$,
 $n_1 = 40, \bar{x}_1 = 142, s_1 = 11$
 $n_2 = 40, \bar{x}_2 = 118, s_2 = 10$
8. Perform the test of hypotheses indicated, using the data from independent samples given. Use the critical value approach. Compute the p -value of the test as well.
- Test $H_0 : \mu_1 - \mu_2 = 13$ vs. $H_a : \mu_1 - \mu_2 < 13$ @ $\alpha = 0.01$,
 $n_1 = 35, \bar{x}_1 = 100, s_1 = 2$
 $n_2 = 35, \bar{x}_2 = 88, s_2 = 2$
 - Test $H_0 : \mu_1 - \mu_2 = -10$ vs. $H_a : \mu_1 - \mu_2 \neq -10$ @ $\alpha = 0.10$,
 $n_1 = 146, \bar{x}_1 = 62, s_1 = 4$
 $n_2 = 120, \bar{x}_2 = 73, s_2 = 7$
9. Perform the test of hypotheses indicated, using the data from independent samples given. Use the p -value approach.
- Test $H_0 : \mu_1 - \mu_2 = 57$ vs. $H_a : \mu_1 - \mu_2 < 57$ @ $\alpha = 0.10$,
 $n_1 = 117, \bar{x}_1 = 1309, s_1 = 42$
 $n_2 = 133, \bar{x}_2 = 1258, s_2 = 37$
 - Test $H_0 : \mu_1 - \mu_2 = -1.5$ vs. $H_a : \mu_1 - \mu_2 \neq -1.5$ @ $\alpha = 0.20$,
 $n_1 = 65, \bar{x}_1 = 16.9, s_1 = 1.3$
 $n_2 = 57, \bar{x}_2 = 18.6, s_2 = 1.1$
10. Perform the test of hypotheses indicated, using the data from independent samples given. Use the p -value approach.

- a. Test $H_0 : \mu_1 - \mu_2 = -10.5$ vs. $H_a : \mu_1 - \mu_2 > -10.5$ @ $\alpha = 0.01$,
 $n_1 = 64, \bar{x}_1 = 85.6, s_1 = 2.4$
 $n_2 = 50, \bar{x}_2 = 95.3, s_2 = 3.1$
- b. Test $H_0 : \mu_1 - \mu_2 = 110$ vs. $H_a : \mu_1 - \mu_2 \neq 110$ @ $\alpha = 0.02$,
 $n_1 = 176, \bar{x}_1 = 1918, s_1 = 68$
 $n_2 = 241, \bar{x}_2 = 1782, s_2 = 146$
11. Perform the test of hypotheses indicated, using the data from independent samples given. Use the p -value approach.
- a. Test $H_0 : \mu_1 - \mu_2 = 50$ vs. $H_a : \mu_1 - \mu_2 > 50$ @ $\alpha = 0.005$,
 $n_1 = 72, \bar{x}_1 = 272, s_1 = 26$
 $n_2 = 103, \bar{x}_2 = 213, s_2 = 14$
- b. Test $H_0 : \mu_1 - \mu_2 = 7.5$ vs. $H_a : \mu_1 - \mu_2 \neq 7.5$ @ $\alpha = 0.10$,
 $n_1 = 52, \bar{x}_1 = 94.3, s_1 = 2.6$
 $n_2 = 38, \bar{x}_2 = 88.6, s_2 = 8.0$
12. Perform the test of hypotheses indicated, using the data from independent samples given. Use the p -value approach.
- a. Test $H_0 : \mu_1 - \mu_2 = 23$ vs. $H_a : \mu_1 - \mu_2 < 23$ @ $\alpha = 0.20$,
 $n_1 = 314, \bar{x}_1 = 198, s_1 = 12.2$
 $n_2 = 220, \bar{x}_2 = 176, s_2 = 11.5$
- b. Test $H_0 : \mu_1 - \mu_2 = 4.4$ vs. $H_a : \mu_1 - \mu_2 \neq 4.4$ @ $\alpha = 0.05$,
 $n_1 = 32, \bar{x}_1 = 40.3, s_1 = 0.5$
 $n_2 = 30, \bar{x}_2 = 35.5, s_2 = 0.7$

APPLICATIONS

13. In order to investigate the relationship between mean job tenure in years among workers who have a bachelor's degree or higher and those who do not, random samples of each type of worker were taken, with the following results.

	n	\bar{x}	s
Bachelor's degree or higher	155	5.2	1.3
No degree	210	5.0	1.5

- Construct the 99% confidence interval for the difference in the population means based on these data.
 - Test, at the 1% level of significance, the claim that mean job tenure among those with higher education is greater than among those without, against the default that there is no difference in the means.
 - Compute the observed significance of the test.
14. Records of 40 used passenger cars and 40 used pickup trucks (none used commercially) were randomly selected to investigate whether there was any difference in the mean time in years that they were kept by the original owner before being sold. For cars the mean was 5.3 years with standard deviation 2.2 years. For pickup trucks the mean was 7.1 years with standard deviation 3.0 years.
- Construct the 95% confidence interval for the difference in the means based on these data.
 - Test the hypothesis that there is a difference in the means against the null hypothesis that there is no difference. Use the 1% level of significance.
 - Compute the observed significance of the test in part (b).
15. In previous years the average number of patients per hour at a hospital emergency room on weekends exceeded the average on weekdays by 6.3 visits per hour. A hospital administrator believes that the current weekend mean exceeds the weekday mean by fewer than 6.3 hours.
- Construct the 99% confidence interval for the difference in the population means based on the following data, derived from a study in which 30 weekend and 30 weekday one-hour periods were randomly selected and the number of new patients in each recorded.

	n	\bar{x}	s
Weekends	30	13.8	3.1

	n	\bar{x}	s
Weekdays	30	8.6	2.7

- b. Test at the 5% level of significance whether the current weekend mean exceeds the weekday mean by fewer than 6.3 patients per hour.
 - c. Compute the observed significance of the test.

16. A sociologist surveys 50 randomly selected citizens in each of two countries to compare the mean number of hours of volunteer work done by adults in each. Among the 50 inhabitants of Lilliput, the mean hours of volunteer work per year was 52, with standard deviation 11.8. Among the 50 inhabitants of Blefuscu, the mean number of hours of volunteer work per year was 37, with standard deviation 7.2.
 - a. Construct the 99% confidence interval for the difference in mean number of hours volunteered by all residents of Lilliput and the mean number of hours volunteered by all residents of Blefuscu.
 - b. Test, at the 1% level of significance, the claim that the mean number of hours volunteered by all residents of Lilliput is more than ten hours greater than the mean number of hours volunteered by all residents of Blefuscu.
 - c. Compute the observed significance of the test in part (b).

17. A university administrator asserted that upperclassmen spend more time studying than underclassmen.
 - a. Test this claim against the default that the average number of hours of study per week by the two groups is the same, using the following information based on random samples from each group of students. Test at the 1% level of significance.

	n	\bar{x}	s
Upperclassmen	35	15.6	2.9
Underclassmen	35	12.3	4.1

- b. Compute the observed significance of the test.

18. An kinesiologist claims that the resting heart rate of men aged 18 to 25 who exercise regularly is more than five beats per minute less than that of men who do not exercise regularly. Men in each category were selected at random and their resting heart rates were measured, with the results shown.

	n	\bar{x}	s
Regular exercise	40	63	1.0

	n	\bar{x}	s
No regular exercise	30	71	1.2

- Perform the relevant test of hypotheses at the 1% level of significance.
- Compute the observed significance of the test.

19. Children in two elementary school classrooms were given two versions of the same test, but with the order of questions arranged from easier to more difficult in Version A and in reverse order in Version B. Randomly selected students from each class were given Version A and the rest Version B. The results are shown in the table.

	n	\bar{x}	s
Version A	31	83	4.6
Version B	32	78	4.3

- Construct the 90% confidence interval for the difference in the means of the populations of all children taking Version A of such a test and of all children taking Version B of such a test.
- Test at the 1% level of significance the hypothesis that the A version of the test is easier than the B version (even though the questions are the same).
- Compute the observed significance of the test.

20. The Municipal Transit Authority wants to know if, on weekdays, more passengers ride the northbound blue line train towards the city center that departs at 8:15 a.m. or the one that departs at 8:30 a.m. The following sample statistics are assembled by the Transit Authority.

	n	\bar{x}	s
8:15 a.m. train	30	323	41
8:30 a.m. train	45	356	45

- Construct the 90% confidence interval for the difference in the mean number of daily travellers on the 8:15 train and the mean number of daily travellers on the 8:30 train.
- Test at the 5% level of significance whether the data provide sufficient evidence to conclude that more passengers ride the 8:30 train.
- Compute the observed significance of the test.

21. In comparing the academic performance of college students who are affiliated with fraternities and those male students who are unaffiliated, a random sample of students was drawn from each of the two populations on a university campus. Summary statistics on the student GPAs are given below.

	n	\bar{x}	s
Fraternity	645	2.90	0.47
Unaffiliated	450	2.88	0.42

Test, at the 5% level of significance, whether the data provide sufficient evidence to conclude that there is a difference in average GPA between the population of fraternity students and the population of unaffiliated male students on this university campus.

22. In comparing the academic performance of college students who are affiliated with sororities and those female students who are unaffiliated, a random sample of students was drawn from each of the two populations on a university campus. Summary statistics on the student GPAs are given below.

	n	\bar{x}	s
Sorority	330	3.18	0.37
Unaffiliated	550	3.12	0.41

Test, at the 5% level of significance, whether the data provide sufficient evidence to conclude that there is a difference in average GPA between the population of sorority students and the population of unaffiliated female students on this university campus.

23. The owner of a professional football team believes that the league has become more offense oriented since five years ago. To check his belief, 32 randomly selected games from one year's schedule were compared to 32 randomly selected games from the schedule five years later. Since more offense produces more points per game, the owner analyzed the following information on points per game (ppg).

	n	\bar{x}	s
ppg previously	32	20.62	4.17
ppg recently	32	22.05	4.01

Test, at the 10% level of significance, whether the data on points per game provide sufficient evidence to conclude that the game has become more offense oriented.

24. The owner of a professional football team believes that the league has become more offense oriented since five years ago. To check his belief, 32 randomly selected games from one year's schedule were compared to 32 randomly selected games from the schedule five years later. Since more offense produces

more offensive yards per game, the owner analyzed the following information on offensive yards per game (oypg).

	n	\bar{x}	s
oypg previously	32	316	40
oypg recently	32	336	35

Test, at the 10% level of significance, whether the data on offensive yards per game provide sufficient evidence to conclude that the game has become more offense oriented.

LARGE DATA SET EXERCISES

25. Large Data Sets 1A and 1B list the SAT scores for 1,000 randomly selected students. Denote the population of all male students as Population 1 and the population of all female students as Population 2.
- <http://www.gone.2012books.lardbucket.org/sites/all/files/data1A.xls>
- <http://www.gone.2012books.lardbucket.org/sites/all/files/data1B.xls>
- Restricting attention to just the males, find n_1 , \bar{x}_1 , and s_1 . Restricting attention to just the females, find n_2 , \bar{x}_2 , and s_2 .
 - Let μ_1 denote the mean SAT score for all males and μ_2 the mean SAT score for all females. Use the results of part (a) to construct a 90% confidence interval for the difference $\mu_1 - \mu_2$.
 - Test, at the 5% level of significance, the hypothesis that the mean SAT scores among males exceeds that of females.
26. Large Data Sets 1A and 1B list the GPAs for 1,000 randomly selected students. Denote the population of all male students as Population 1 and the population of all female students as Population 2.
- <http://www.gone.2012books.lardbucket.org/sites/all/files/data1A.xls>
- <http://www.gone.2012books.lardbucket.org/sites/all/files/data1B.xls>
- Restricting attention to just the males, find n_1 , \bar{x}_1 , and s_1 . Restricting attention to just the females, find n_2 , \bar{x}_2 , and s_2 .
 - Let μ_1 denote the mean GPA for all males and μ_2 the mean GPA for all females. Use the results of part (a) to construct a 95% confidence interval for the difference $\mu_1 - \mu_2$.
 - Test, at the 10% level of significance, the hypothesis that the mean GPAs among males and females differ.

27. Large Data Sets 7A and 7B list the survival times for 65 male and 75 female laboratory mice with thymic leukemia. Denote the population of all such male mice as Population 1 and the population of all such female mice as Population 2.

<http://www.gone.2012books.lardbucket.org/sites/all/files/data7A.xls>

<http://www.gone.2012books.lardbucket.org/sites/all/files/data7B.xls>

- Restricting attention to just the males, find n_1 , \bar{x}_1 , and s_1 . Restricting attention to just the females, find n_2 , \bar{x}_2 , and s_2 .
- Let μ_1 denote the mean survival for all males and μ_2 the mean survival time for all females. Use the results of part (a) to construct a 99% confidence interval for the difference $\mu_1 - \mu_2$.
- Test, at the 1% level of significance, the hypothesis that the mean survival time for males exceeds that for females by more than 182 days (half a year).
- Compute the observed significance of the test in part (c).

ANSWERS

1.
 - a. $(4.20, 5.80)$,
 - b. $(-18.54, -9.46)$
3.
 - a. $(-12.81, -10.39)$,
 - b. $(-76.50, -68.10)$
5.
 - a. $Z = 8.753, \pm z_{0.025} = \pm 1.960$, reject H_0 , p -value = 0.0000;
 - b. $Z = -0.687, -z_{0.10} = -1.282$, do not reject H_0 , p -value = 0.2451
7.
 - a. $Z = 2.444, \pm z_{0.005} = \pm 2.576$, do not reject H_0 , p -value = 0.0146.
 - b. $Z = 1.702, z_{0.05} = 1.645$, reject H_0 , p -value = 0.0446
9.
 - a. $Z = -1.19$, p -value = 0.1170, do not reject H_0 ;
 - b. $Z = -0.92$, p -value = 0.3576, do not reject H_0
11.
 - a. $Z = 2.68$, p -value = 0.0037, reject H_0 ;
 - b. $Z = -1.34$, p -value = 0.1802, do not reject H_0
13.
 - a. 0.2 ± 0.4 ,
 - b. $Z = 1.360, z_{0.01} = 2.326$, do not reject H_0 (not greater)
 - c. p -value = 0.0869
15.
 - a. 5.2 ± 1.9 ,
 - b. $Z = -1.466, -z_{0.050} = -1.645$, do not reject H_0 (exceeds by 6.3 or more)
 - c. p -value = 0.0708
17.
 - a. $Z = 3.888, z_{0.01} = 2.326$, reject H_0 (upperclassmen study more)
 - b. p -value = 0.0001
19.
 - a. 5 ± 1.8 ,
 - b. $Z = 4.454, z_{0.01} = 2.326$, reject H_0 (Test A is easier)
 - c. p -value = 0.0000
21. $Z = 0.738, \pm z_{0.025} = \pm 1.960$, do not reject H_0 (no difference)
23. $Z = -1.398, -z_{0.10} = -1.282$, reject H_0 (more offense oriented)
25.
 - a. $n_1 = 419, \bar{x}_1 = 1540.33, s_1 = 205.40, n_2 = 581, \bar{x}_2 = 1520.38$, and $s_2 = 217.34$.
 - b. $(-2.24, 42.15)$

- c. $H_0 : \mu_1 - \mu_2 = 0$ vs. $H_a : \mu_1 - \mu_2 > 0$. Test Statistic: $Z = 1.48$.
 Rejection Region: $[1.645, \infty)$. Decision: Fail to reject H_0 .
27. a. $n_1 = 65, \bar{x}_1 = 665.97, s_1 = 41.60, n_2 = 75, \bar{x}_2 = 455.89$,
 and $s_2 = 63.22$.
- b. $(187.06, 233.09)$
- c. $H_0 : \mu_1 - \mu_2 = 182$ vs. $H_a : \mu_1 - \mu_2 > 182$. Test Statistic: $Z = 3.14$. Rejection Region: $[2.33, \infty)$. Decision: Reject H_0 .
- d. $p - value = 0.0008$

9.2 Comparison of Two Population Means: Small, Independent Samples

LEARNING OBJECTIVES

1. To learn how to construct a confidence interval for the difference in the means of two distinct populations using small, independent samples.
2. To learn how to perform a test of hypotheses concerning the difference between the means of two distinct populations using small, independent samples.

When one or the other of the sample sizes is small, as is often the case in practice, the Central Limit Theorem does not apply. We must then impose conditions on the population to give statistical validity to the test procedure. We will assume that both populations from which the samples are taken have a normal probability distribution and that their standard deviations are equal.

Confidence Intervals

When the two populations are normally distributed and have equal standard deviations, the following formula for a confidence interval for $\mu_1 - \mu_2$ is valid.

100 (1 - α) % Confidence Interval for the Difference Between Two Population Means: Small, Independent Samples

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \quad \text{where} \quad s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

The number of degrees of freedom is $df = n_1 + n_2 - 2$.

The samples must be independent, the populations must be normal, and the population standard deviations must be equal. “Small” samples means that either $n_1 < 30$ or $n_2 < 30$.

The quantity s_p^2 is called the **pooled sample variance**. It is a weighted average of the two estimates s_1^2 and s_2^2 of the common variance $\sigma_1^2 = \sigma_2^2$ of the two populations.

EXAMPLE 4

A software company markets a new computer game with two experimental packaging designs. Design 1 is sent to 11 stores; their average sales the first month is 52 units with sample standard deviation 12 units. Design 2 is sent to 6 stores; their average sales the first month is 46 units with sample standard deviation 10 units. Construct a point estimate and a 95% confidence interval for the difference in average monthly sales between the two package designs.

Solution:

The point estimate of $\mu_1 - \mu_2$ is

$$\bar{x}_1 - \bar{x}_2 = 52 - 46 = 6$$

In words, we estimate that the average monthly sales for Design 1 is 6 units more per month than the average monthly sales for Design 2.

To apply the formula for the confidence interval, we must find $t_{\alpha/2}$. The 95% confidence level means that $\alpha = 1 - 0.95 = 0.05$ so that $t_{\alpha/2} = t_{0.025}$. From [Figure 12.3 "Critical Values of "](#), in the row with the heading $df = 11 + 6 - 2 = 15$ we read that $t_{0.025} = 2.131$. From the formula for the pooled sample variance we compute

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(10)(12)^2 + (5)(10)^2}{15} = 129.\bar{3}$$

Thus

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} = 6 \pm (2.131) \sqrt{129.\bar{3} \left(\frac{1}{11} + \frac{1}{6} \right)} \approx$$

We are 95% confident that the difference in the population means lies in the interval $[-6.3, 18.3]$, in the sense that in repeated sampling 95% of all intervals constructed from the sample data in this manner will contain $\mu_1 - \mu_2$. Because the interval contains both positive and negative values the statement in the context of the problem is that we are 95% confident

that the average monthly sales for Design 1 is between 18.3 units higher and 6.3 units lower than the average monthly sales for Design 2.

Hypothesis Testing

Testing hypotheses concerning the difference of two population means using small samples is done precisely as it is done for large samples, using the following standardized test statistic. The same conditions on the populations that were required for constructing a confidence interval for the difference of the means must also be met when hypotheses are tested.

Standardized Test Statistic for Hypothesis Tests Concerning the Difference Between Two Population Means: Small, Independent Samples

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad \text{where} \quad s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

The test statistic has Student's t -distribution with $df = n_1 + n_2 - 2$ degrees of freedom.

The samples must be independent, the populations must be normal, and the population standard deviations must be equal. "Small" samples means that either $n_1 < 30$ or $n_2 < 30$.

EXAMPLE 5

Refer to [Note 9.11 "Example 4"](#) concerning the mean sales per month for the same computer game but sold with two package designs. Test at the 1% level of significance whether the data provide sufficient evidence to conclude that the mean sales per month of the two designs are different. Use the critical value approach.

Solution:

- Step 1. The relevant test is

$$H_0 : \mu_1 - \mu_2 = 0$$

$$\text{vs. } H_a : \mu_1 - \mu_2 \neq 0 \quad @ \alpha = 0.01$$

- Step 2. Since the samples are independent and at least one is less than 30 the test statistic is

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

which has Student's t -distribution with $df = 11 + 6 - 2 = 15$ degrees of freedom.

- Step 3. Inserting the data and the value $D_0 = 0$ into the formula for the test statistic gives

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{(52 - 46) - 0}{\sqrt{129.3 \left(\frac{1}{11} + \frac{1}{6} \right)}} = 1.040$$

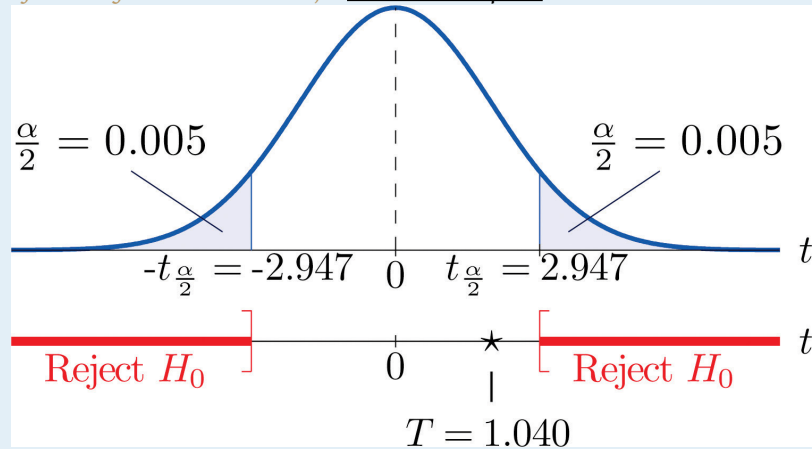
$$\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \quad \sqrt{129.3 \left(\frac{1}{11} + \frac{1}{6} \right)}$$

- Step 4. Since the symbol in H_a is “ \neq ” this is a two-tailed test, so there are two critical values, $\pm t_{\alpha/2} = \pm t_{0.005}$. From the row

in Figure 12.3 "Critical Values of " with the heading $df = 15$ we read off $t_{0.005} = 2.947$. The rejection region is $(-\infty, -2.947] \cup [2.947, \infty)$.

Figure 9.4

Rejection Region and Test Statistic for Note 9.13 "Example 5"



- Step 5. As shown in Figure 9.4 "Rejection Region and Test Statistic for " the test statistic does not fall in the rejection region. The decision is not to reject H_0 . In the context of the problem our conclusion is:

The data do not provide sufficient evidence, at the 1% level of significance, to conclude that the mean sales per month of the two designs are different.

EXAMPLE 6

Perform the test of [Note 9.13 "Example 5"](#) using the p -value approach.

Solution:

The first three steps are identical to those in [Note 9.13 "Example 5"](#).

- Step 4. Because the test is two-tailed the observed significance or p -value of the test is the double of the area of the right tail of Student's t -distribution, with 15 degrees of freedom, that is cut off by the test statistic $T = 1.040$. We can only approximate this number. Looking in the row of [Figure 12.3 "Critical Values of "](#) headed $df = 15$, the number 1.040 is between the numbers 0.866 and 1.341, corresponding to $t_{0.200}$ and $t_{0.100}$.

The area cut off by $t = 0.866$ is 0.200 and the area cut off by $t = 1.341$ is 0.100. Since 1.040 is between 0.866 and 1.341 the area it cuts off is between 0.200 and 0.100. Thus the p -value (since the area must be doubled) is between 0.400 and 0.200.

- Step 5. Since $p > 0.200 > 0.01$, $p > \alpha$, so the decision is not to reject the null hypothesis:

The data do not provide sufficient evidence, at the 1% level of significance, to conclude that the mean sales per month of the two designs are different.

KEY TAKEAWAYS

- In the context of estimating or testing hypotheses concerning two population means, "small" samples means that *at least one* sample is small. In particular, even if one sample is of size 30 or more, if the other is of size less than 30 the formulas of this section must be used.
- A confidence interval for the difference in two population means is computed using a formula in the same fashion as was done for a single population mean.

EXERCISES

BASIC

In all exercises for this section assume that the populations are normal and have equal standard deviations.

1. Construct the confidence interval for $\mu_1 - \mu_2$ for the level of confidence and the data from independent samples given.
 - a. 95% confidence,
 $n_1 = 10, \bar{x}_1 = 120, s_1 = 2$
 $n_2 = 15, \bar{x}_2 = 101, s_2 = 4$
 - b. 99% confidence,
 $n_1 = 6, \bar{x}_1 = 25, s_1 = 1$
 $n_2 = 12, \bar{x}_2 = 17, s_2 = 3$
2. Construct the confidence interval for $\mu_1 - \mu_2$ for the level of confidence and the data from independent samples given.
 - a. 90% confidence,
 $n_1 = 28, \bar{x}_1 = 212, s_1 = 6$
 $n_2 = 23, \bar{x}_2 = 198, s_2 = 5$
 - b. 99% confidence,
 $n_1 = 14, \bar{x}_1 = 68, s_1 = 8$
 $n_2 = 20, \bar{x}_2 = 43, s_2 = 3$
3. Construct the confidence interval for $\mu_1 - \mu_2$ for the level of confidence and the data from independent samples given.
 - a. 99.9% confidence,
 $n_1 = 35, \bar{x}_1 = 6.5, s_1 = 0.2$
 $n_2 = 20, \bar{x}_2 = 6.2, s_2 = 0.1$
 - b. 99% confidence,

$$n_1 = 18, \bar{x}_1 = 77.3, s_1 = 1.2$$

$$n_2 = 32, \bar{x}_2 = 75.0, s_2 = 1.6$$

4. Construct the confidence interval for $\mu_1 - \mu_2$ for the level of confidence and the data from independent samples given.

- a. 99.5% confidence,

$$n_1 = 40, \bar{x}_1 = 85.6, s_1 = 2.8$$

$$n_2 = 20, \bar{x}_2 = 73.1, s_2 = 2.1$$

- b. 99.9% confidence,

$$n_1 = 25, \bar{x}_1 = 215, s_1 = 7$$

$$n_2 = 35, \bar{x}_2 = 185, s_2 = 12$$

5. Perform the test of hypotheses indicated, using the data from independent samples given. Use the critical value approach.

- a. Test $H_0 : \mu_1 - \mu_2 = 11$ vs. $H_a : \mu_1 - \mu_2 > 11$ @ $\alpha = 0.025$,

$$n_1 = 6, \bar{x}_1 = 32, s_1 = 2$$

$$n_2 = 11, \bar{x}_2 = 19, s_2 = 1$$

- b. Test $H_0 : \mu_1 - \mu_2 = 26$ vs. $H_a : \mu_1 - \mu_2 \neq 26$ @ $\alpha = 0.05$,

$$n_1 = 17, \bar{x}_1 = 166, s_1 = 4$$

$$n_2 = 24, \bar{x}_2 = 138, s_2 = 3$$

6. Perform the test of hypotheses indicated, using the data from independent samples given. Use the critical value approach.

- a. Test $H_0 : \mu_1 - \mu_2 = 40$ vs. $H_a : \mu_1 - \mu_2 < 40$ @ $\alpha = 0.10$,

$$n_1 = 14, \bar{x}_1 = 289, s_1 = 11$$

$$n_2 = 12, \bar{x}_2 = 254, s_2 = 9$$

- b. Test $H_0 : \mu_1 - \mu_2 = 21$ vs. $H_a : \mu_1 - \mu_2 \neq 21$ @ $\alpha = 0.05$,

$$n_1 = 23, \bar{x}_1 = 130, s_1 = 6$$

$$n_2 = 27, \bar{x}_2 = 113, s_2 = 8$$

7. Perform the test of hypotheses indicated, using the data from independent samples given. Use the critical value approach.
- a. Test $H_0 : \mu_1 - \mu_2 = -15$ vs. $H_a : \mu_1 - \mu_2 < -15$ @ $\alpha = 0.10$,
 $n_1 = 30, \bar{x}_1 = 42, s_1 = 7$
 $n_2 = 12, \bar{x}_2 = 60, s_2 = 5$
- b. Test $H_0 : \mu_1 - \mu_2 = 103$ vs. $H_a : \mu_1 - \mu_2 \neq 103$ @ $\alpha = 0.10$,
 $n_1 = 17, \bar{x}_1 = 711, s_1 = 28$
 $n_2 = 32, \bar{x}_2 = 598, s_2 = 21$
8. Perform the test of hypotheses indicated, using the data from independent samples given. Use the critical value approach.
- a. Test $H_0 : \mu_1 - \mu_2 = 75$ vs. $H_a : \mu_1 - \mu_2 > 75$ @ $\alpha = 0.025$,
 $n_1 = 45, \bar{x}_1 = 674, s_1 = 18$
 $n_2 = 29, \bar{x}_2 = 591, s_2 = 13$
- b. Test $H_0 : \mu_1 - \mu_2 = -20$ vs. $H_a : \mu_1 - \mu_2 \neq -20$ @ $\alpha = 0.005$,
 $n_1 = 30, \bar{x}_1 = 137, s_1 = 8$
 $n_2 = 19, \bar{x}_2 = 166, s_2 = 11$
9. Perform the test of hypotheses indicated, using the data from independent samples given. Use the p -value approach. (The p -value can be only approximated.)
- a. Test $H_0 : \mu_1 - \mu_2 = 12$ vs. $H_a : \mu_1 - \mu_2 > 12$ @ $\alpha = 0.01$,
 $n_1 = 20, \bar{x}_1 = 133, s_1 = 7$
 $n_2 = 10, \bar{x}_2 = 115, s_2 = 5$
- b. Test $H_0 : \mu_1 - \mu_2 = 46$ vs. $H_a : \mu_1 - \mu_2 \neq 46$ @ $\alpha = 0.10$,
 $n_1 = 24, \bar{x}_1 = 586, s_1 = 11$
 $n_2 = 27, \bar{x}_2 = 535, s_2 = 13$
10. Perform the test of hypotheses indicated, using the data from independent samples given. Use the p -value approach. (The p -value can be only approximated.)

- a. Test $H_0 : \mu_1 - \mu_2 = 38$ vs. $H_a : \mu_1 - \mu_2 < 38$ @ $\alpha = 0.01$,
 $n_1 = 12, \bar{x}_1 = 464, s_1 = 5$
 $n_2 = 10, \bar{x}_2 = 432, s_2 = 6$
- b. Test $H_0 : \mu_1 - \mu_2 = 4$ vs. $H_a : \mu_1 - \mu_2 \neq 4$ @ $\alpha = 0.005$,
 $n_1 = 14, \bar{x}_1 = 68, s_1 = 2$
 $n_2 = 17, \bar{x}_2 = 67, s_2 = 3$
11. Perform the test of hypotheses indicated, using the data from independent samples given. Use the p -value approach. (The p -value can be only approximated.)
- a. Test $H_0 : \mu_1 - \mu_2 = 50$ vs. $H_a : \mu_1 - \mu_2 > 50$ @ $\alpha = 0.01$,
 $n_1 = 30, \bar{x}_1 = 681, s_1 = 8$
 $n_2 = 27, \bar{x}_2 = 625, s_2 = 8$
- b. Test $H_0 : \mu_1 - \mu_2 = 35$ vs. $H_a : \mu_1 - \mu_2 \neq 35$ @ $\alpha = 0.10$,
 $n_1 = 36, \bar{x}_1 = 325, s_1 = 11$
 $n_2 = 29, \bar{x}_2 = 286, s_2 = 7$
12. Perform the test of hypotheses indicated, using the data from independent samples given. Use the p -value approach. (The p -value can be only approximated.)
- a. Test $H_0 : \mu_1 - \mu_2 = -4$ vs. $H_a : \mu_1 - \mu_2 < -4$ @ $\alpha = 0.05$,
 $n_1 = 40, \bar{x}_1 = 80, s_1 = 5$
 $n_2 = 25, \bar{x}_2 = 87, s_2 = 5$
- b. Test $H_0 : \mu_1 - \mu_2 = 21$ vs. $H_a : \mu_1 - \mu_2 \neq 21$ @ $\alpha = 0.01$,
 $n_1 = 15, \bar{x}_1 = 192, s_1 = 12$
 $n_2 = 34, \bar{x}_2 = 180, s_2 = 8$

APPLICATIONS

13. A county environmental agency suspects that the fish in a particular polluted lake have elevated mercury level. To confirm that suspicion, five striped bass in that lake were caught and their tissues were tested for mercury. For the purpose of comparison, four striped bass in an unpolluted lake were also caught and tested. The fish tissue mercury levels in mg/kg are given below.

Sample 1 (from polluted lake)	Sample 2 (from unpolluted lake)
0.580	0.382
0.711	0.276
0.571	0.570
0.666	0.366
0.598	

- Construct the 95% confidence interval for the difference in the population means based on these data.
 - Test, at the 5% level of significance, whether the data provide sufficient evidence to conclude that fish in the polluted lake have elevated levels of mercury in their tissue.
14. A genetic engineering company claims that it has developed a genetically modified tomato plant that yields on average more tomatoes than other varieties. A farmer wants to test the claim on a small scale before committing to a full-scale planting. Ten genetically modified tomato plants are grown from seeds along with ten other tomato plants. At the season's end, the resulting yields in pound are recorded as below.

Sample 1 (genetically modified)	Sample 2 (regular)
20	21
23	21
27	22
25	18
25	20
25	20
27	18
23	25
24	23
22	20

- a. Construct the 99% confidence interval for the difference in the population means based on these data.
 - b. Test, at the 1% level of significance, whether the data provide sufficient evidence to conclude that the mean yield of the genetically modified variety is greater than that for the standard variety.
15. The coaching staff of a professional football team believes that the rushing offense has become increasingly potent in recent years. To investigate this belief, 20 randomly selected games from one year's schedule were compared to 11 randomly selected games from the schedule five years later. The sample information on rushing yards per game (rypg) is summarized below.

	n	\bar{x}	s
rypg previously	20	112	24
rypg recently	11	114	21

- a. Construct the 95% confidence interval for the difference in the population means based on these data.
 - b. Test, at the 5% level of significance, whether the data on rushing yards per game provide sufficient evidence to conclude that the rushing offense has become more potent in recent years.
16. The coaching staff of professional football team believes that the rushing offense has become increasingly potent in recent years. To investigate this belief, 20 randomly selected games from one year's schedule were compared to

11 randomly selected games from the schedule five years later. The sample information on passing yards per game (pypg) is summarized below.

	n	\bar{x}	s
pypg previously	20	203	38
pypg recently	11	232	33

- Construct the 95% confidence interval for the difference in the population means based on these data.
 - Test, at the 5% level of significance, whether the data on passing yards per game provide sufficient evidence to conclude that the passing offense has become more potent in recent years.
17. A university administrator wishes to know if there is a difference in average starting salary for graduates with master's degrees in engineering and those with master's degrees in business. Fifteen recent graduates with master's degree in engineering and 11 with master's degrees in business are surveyed and the results are summarized below.

	n	\bar{x}	s
Engineering	15	68,535	1627
Business	11	63,230	2033

- Construct the 90% confidence interval for the difference in the population means based on these data.
 - Test, at the 10% level of significance, whether the data provide sufficient evidence to conclude that the average starting salaries are different.
18. A gardener sets up a flower stand in a busy business district and sells bouquets of assorted fresh flowers on weekdays. To find a more profitable pricing, she sells bouquets for 15 dollars each for ten days, then for 10 dollars each for five days. Her average daily profit for the two different prices are given below.

	n	\bar{x}	s
\$15	10	171	26
\$10	5	198	29

- Construct the 90% confidence interval for the difference in the population means based on these data.
- Test, at the 10% level of significance, whether the data provide sufficient evidence to conclude the gardener's average daily profit will be higher if the bouquets are sold at \$10 each.

ANSWERS

1.
 - a. $(16.16, 21.84)$,
 - b. $(4.28, 11.72)$
3.
 - a. $(0.13, 0.47)$,
 - b. $(1.14, 3.46)$
5.
 - a. $T = 2.787, t_{0.025} = 2.131$, reject H_0 ,
 - b. $T = 1.831, \pm t_{0.025} = \pm 2.023$, do not reject H_0
7.
 - a. $T = -1.349, -t_{0.10} = -1.303$, reject H_0 ,
 - b. $T = 1.411, \pm t_{0.05} = \pm 1.678$, do not reject H_0
9.
 - a. $T = 2.411, df = 28, p\text{-value} > 0.01$, do not reject H_0 ,
 - b. $T = 1.473, df = 49, p\text{-value} < 0.10$, reject H_0
11.
 - a. $T = 2.827, df = 55, p\text{-value} < 0.01$, reject H_0 .
 - b. $T = 1.699, df = 63, p\text{-value} < 0.10$, reject H_0
13.
 - a. 0.2267 ± 0.2182 ,
 - b. $T = 3.635, df = 7, t_{0.05} = 1.895$, reject H_0 (elevated levels)
15.
 - a. -2 ± 17.7 ,
 - b. $T = -0.232, df = 29, -t_{0.05} = -1.699$, do not reject H_0 (not more potent)
17.
 - a. 5305 ± 1227 ,
 - b. $T = 7.395, df = 24, \pm t_{0.05} = \pm 1.711$, reject H_0 (different)

9.3 Comparison of Two Population Means: Paired Samples

LEARNING OBJECTIVES

1. To learn the distinction between independent samples and paired samples.
2. To learn how to construct a confidence interval for the difference in the means of two distinct populations using paired samples.
3. To learn how to perform a test of hypotheses concerning the difference in the means of two distinct populations using paired samples.

Suppose chemical engineers wish to compare the fuel economy obtained by two different formulations of gasoline. Since fuel economy varies widely from car to car, if the mean fuel economy of two independent samples of vehicles run on the two types of fuel were compared, even if one formulation were better than the other the large variability from vehicle to vehicle might make any difference arising from difference in fuel difficult to detect. Just imagine one random sample having many more large vehicles than the other. Instead of independent random samples, it would make more sense to select pairs of cars of the same make and model and driven under similar circumstances, and compare the fuel economy of the two cars in each pair. Thus the data would look something like [Table 9.1 "Fuel Economy of Pairs of Vehicles"](#), where the first car in each pair is operated on one formulation of the fuel (call it Type 1 gasoline) and the second car is operated on the second (call it Type 2 gasoline).

Table 9.1 Fuel Economy of Pairs of Vehicles

Make and Model	Car 1	Car 2
Buick LaCrosse	17.0	17.0
Dodge Viper	13.2	12.9
Honda CR-Z	35.3	35.4
Hummer H 3	13.6	13.2
Lexus RX	32.7	32.5
Mazda CX-9	18.4	18.1
Saab 9-3	22.5	22.5

Make and Model	Car 1	Car 2
Toyota Corolla	26.8	26.7
Volvo XC 90	15.1	15.0

The first column of numbers form a sample from Population 1, the population of all cars operated on Type 1 gasoline; the second column of numbers form a sample from Population 2, the population of all cars operated on Type 2 gasoline. It would be incorrect to analyze the data using the formulas from the previous section, however, since the samples were not drawn independently. What is correct is to compute the difference in the numbers in each pair (subtracting in the same order each time) to obtain the third column of numbers as shown in [Table 9.2 "Fuel Economy of Pairs of Vehicles"](#) and treat the differences as the data. At this point, the new sample of differences $d_1 = 0.0, \dots, d_9 = 0.1$ in the third column of [Table 9.2 "Fuel Economy of Pairs of Vehicles"](#) may be considered as a random sample of size $n = 9$ selected from a population with mean $\mu_d = \mu_1 - \mu_2$. This approach essentially transforms the paired two-sample problem into a one-sample problem as discussed in the previous two chapters.

Table 9.2 Fuel Economy of Pairs of Vehicles

Make and Model	Car 1	Car 2	Difference
Buick LaCrosse	17.0	17.0	0.0
Dodge Viper	13.2	12.9	0.3
Honda CR-Z	35.3	35.4	-0.1
Hummer H 3	13.6	13.2	0.4
Lexus RX	32.7	32.5	0.2
Mazda CX-9	18.4	18.1	0.3
Saab 9-3	22.5	22.5	0.0
Toyota Corolla	26.8	26.7	0.1
Volvo XC 90	15.1	15.0	0.1

Note carefully that although it does not matter what order the subtraction is done, it must be done in the *same* order for all pairs. This is why there are both positive and negative quantities in the third column of numbers in [Table 9.2 "Fuel Economy of Pairs of Vehicles"](#).

Confidence Intervals

When the population of differences is normally distributed the following formula for a confidence interval for $\mu_d = \mu_1 - \mu_2$ is valid.

100 (1 - α) % Confidence Interval for the Difference Between Two Population Means: Paired Difference Samples

$$\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}}$$

where there are n pairs, \bar{d} is the mean and s_d is the standard deviation of their differences.

The number of degrees of freedom is $df = n - 1$.

The population of differences must be normally distributed.

EXAMPLE 7

Using the data in [Table 9.1 "Fuel Economy of Pairs of Vehicles"](#) construct a point estimate and a 95% confidence interval for the difference in average fuel economy between cars operated on Type 1 gasoline and cars operated on Type 2 gasoline.

Solution:

We have referred to the data in [Table 9.1 "Fuel Economy of Pairs of Vehicles"](#) because that is the way that the data are typically presented, but we emphasize that with paired sampling one immediately computes the differences, as given in [Table 9.2 "Fuel Economy of Pairs of Vehicles"](#), and uses the differences as the data.

The mean and standard deviation of the differences are

$$\bar{d} = \frac{\Sigma d}{n} = \frac{1.3}{9} = 0.1\bar{4} \text{ and } s_d = \sqrt{\frac{\Sigma d^2 - \frac{1}{n} (\Sigma d)^2}{n-1}} = \sqrt{\frac{0.41 - \frac{1}{9} (1.3)^2}{8}}$$

The point estimate of $\mu_1 - \mu_2 = \mu_d$ is

$$\bar{d} = 0.14$$

In words, we estimate that the average fuel economy of cars using Type 1 gasoline is 0.14 mpg greater than the average fuel economy of cars using Type 2 gasoline.

To apply the formula for the confidence interval, we must find $t_{\alpha/2}$. The 95% confidence level means that $\alpha = 1 - 0.95 = 0.05$ so that $t_{\alpha/2} = t_{0.025}$. From [Figure 12.3 "Critical Values of "](#), in the row with the heading $df = 9 - 1 = 8$ we read that $t_{0.025} = 2.306$. Thus

$$\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}} = 0.14 \pm 2.306 \left(\frac{0.1\bar{6}}{\sqrt{9}} \right) \approx 0.14 \pm 0.13$$

We are 95% confident that the difference in the population means lies in the interval $[0.01, 0.27]$, in the sense that in repeated sampling 95% of all intervals constructed from the sample data in this manner will contain $\mu_d = \mu_1 - \mu_2$. Stated differently, we are 95% confident that mean fuel economy is between 0.01 and 0.27 mpg greater with Type 1 gasoline than with Type 2 gasoline.

Hypothesis Testing

Testing hypotheses concerning the difference of two population means using paired difference samples is done precisely as it is done for independent samples, although now the null and alternative hypotheses are expressed in terms of μ_d instead of $\mu_1 - \mu_2$. Thus the null hypothesis will always be written

$$H_0 : \mu_d = D_0$$

The three forms of the alternative hypothesis, with the terminology for each case, are:

Form of H_a	Terminology
$H_a : \mu_d < D_0$	Left-tailed
$H_a : \mu_d > D_0$	Right-tailed
$H_a : \mu_d \neq D_0$	Two-tailed

The same conditions on the population of differences that was required for constructing a confidence interval for the difference of the means must also be met when hypotheses are tested. Here is the standardized test statistic that is used in the test.

**Standardized Test Statistic for Hypothesis Tests
Concerning the Difference Between Two Population
Means: Paired Difference Samples**

$$T = \frac{\bar{d} - D_0}{s_d / \sqrt{n}}$$

where there are n pairs, \bar{d} is the mean and s_d is the standard deviation of their differences.

The test statistic has Student's t -distribution with $df = n - 1$ degrees of freedom.

The population of differences must be normally distributed.

EXAMPLE 8

Using the data of [Table 9.2 "Fuel Economy of Pairs of Vehicles"](#) test the hypothesis that mean fuel economy for Type 1 gasoline is greater than that for Type 2 gasoline against the null hypothesis that the two formulations of gasoline yield the same mean fuel economy. Test at the 5% level of significance using the critical value approach.

Solution:

The only part of the table that we use is the third column, the differences.

- Step 1. Since the differences were computed in the order Type 1 mpg – Type 2 mpg, better fuel economy with Type 1 fuel corresponds to $\mu_d = \mu_1 - \mu_2 > 0$. Thus the test is

$$H_0 : \mu_d = 0$$

$$\text{vs. } H_a : \mu_d > 0 \quad @ \alpha = 0.05$$

(If the differences had been computed in the opposite order then the alternative hypotheses would have been $H_a : \mu_d < 0$.)

- Step 2. Since the sampling is in pairs the test statistic is

$$T = \frac{\bar{d} - D_0}{s_d / \sqrt{n}}$$

- Step 3. We have already computed \bar{d} and s_d in the previous example. Inserting their values and $D_0 = 0$ into the formula for the test statistic gives

$$T = \frac{\bar{d} - D_0}{s_d / \sqrt{n}} = \frac{0.14}{0.16 / 3} = 2.600$$

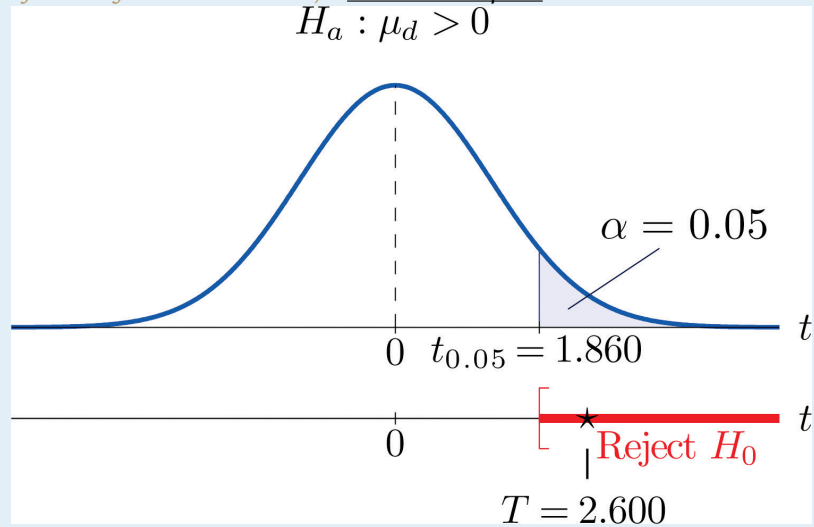
- Step 4. Since the symbol in H_a is ">" this is a right-tailed test, so there is a single critical value, $t_\alpha = t_{0.05}$ with 8 degrees of freedom, which from

the row labeled $df = 8$ in [Figure 12.3 "Critical Values of "](#) we read off as 1.860. The rejection region is $[1.860, \infty)$.

- Step 5. As shown in [Figure 9.5 "Rejection Region and Test Statistic for "](#) the test statistic falls in the rejection region. The decision is to reject H_0 . In the context of the problem our conclusion is:

Figure 9.5

Rejection Region and Test Statistic for [Note 9.20 "Example 8"](#)



The data provide sufficient evidence, at the 5% level of significance, to conclude that the mean fuel economy provided by Type 1 gasoline is greater than that for Type 2 gasoline.

EXAMPLE 9

Perform the test of [Note 9.20 "Example 8"](#) using the p -value approach.

Solution:

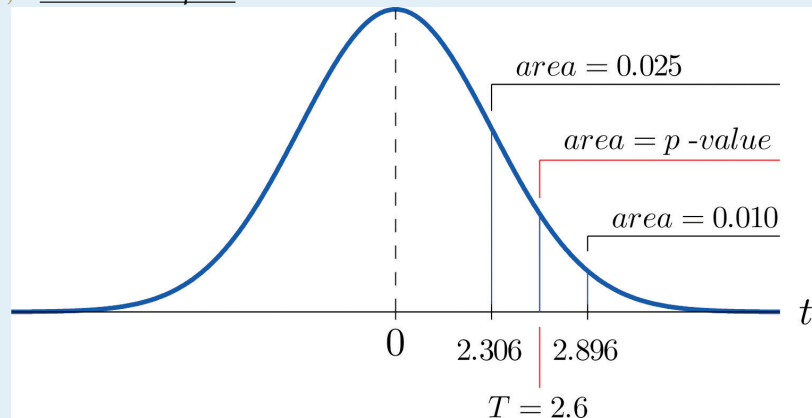
The first three steps are identical to those in [Note 9.20 "Example 8"](#).

- Step 4. Because the test is one-tailed the observed significance or p -value of the test is just the area of the right tail of Student's t -distribution, with 8 degrees of freedom, that is cut off by the test statistic $T = 2.600$. We can only approximate this number.

Looking in the row of [Figure 12.3 "Critical Values of "](#) headed $df = 8$, the number 2.600 is between the numbers 2.306 and 2.896, corresponding to $t_{0.025}$ and $t_{0.010}$.

The area cut off by $t = 2.306$ is 0.025 and the area cut off by $t = 2.896$ is 0.010. Since 2.600 is between 2.306 and 2.896 the area it cuts off is between 0.025 and 0.010. Thus the p -value is between 0.025 and 0.010. In particular it is less than 0.025. See [Figure 9.6](#).

Figure 9.6
P-Value for [Note 9.21 "Example 9"](#)



- Step 5. Since $0.025 < 0.05$, $p < \alpha$ so the decision is to reject the null hypothesis:

The data provide sufficient evidence, at the 5% level of significance, to conclude that the mean fuel economy provided by Type 1 gasoline is greater than that for Type 2 gasoline.

The paired two-sample experiment is a very powerful study design. It bypasses many unwanted sources of “statistical noise” that might otherwise influence the outcome of the experiment, and focuses on the possible difference that might arise from the one factor of interest.

If the sample is large (meaning that $n \geq 30$) then in the formula for the confidence interval we may replace $t_{\alpha/2}$ by $z_{\alpha/2}$. For hypothesis testing when the number of pairs is at least 30, we may use the same statistic as for small samples for hypothesis testing, except now it follows a standard normal distribution, so we use the last line of [Figure 12.3 "Critical Values of "](#) to compute critical values, and p -values can be computed exactly with [Figure 12.2 "Cumulative Normal Probability"](#), not merely estimated using [Figure 12.3 "Critical Values of "](#).

KEY TAKEAWAYS

- When the data are collected in pairs, the differences computed for each pair are the data that are used in the formulas.
- A confidence interval for the difference in two population means using paired sampling is computed using a formula in the same fashion as was done for a single population mean.
- The same five-step procedure used to test hypotheses concerning a single population mean is used to test hypotheses concerning the difference between two population means using pair sampling. The only difference is in the formula for the standardized test statistic.

EXERCISES

BASIC

In all exercises for this section assume that the population of differences is normal.

1. Use the following paired sample data for this exercise.

Population 1 35 32 35 35 36 35 36

Population 2 28 26 27 26 29 27 29

- Compute \bar{d} and s_d .
- Give a point estimate for $\mu_1 - \mu_2 = \mu_d$.
- Construct the 95% confidence interval for $\mu_1 - \mu_2 = \mu_d$ from these data.
- Test, at the 10% level of significance, the hypothesis that $\mu_1 - \mu_2 > 7$ as an alternative to the null hypothesis that $\mu_1 - \mu_2 = 7$.

2. Use the following paired sample data for this exercise.

Population 1 103 127 96 110

Population 2 81 106 73 88

Population 1 90 118 130 106

Population 2 70 95 109 83

- Compute \bar{d} and s_d .
- Give a point estimate for $\mu_1 - \mu_2 = \mu_d$.
- Construct the 90% confidence interval for $\mu_1 - \mu_2 = \mu_d$ from these data.
- Test, at the 1% level of significance, the hypothesis that $\mu_1 - \mu_2 < 24$ as an alternative to the null hypothesis that $\mu_1 - \mu_2 = 24$.

3. Use the following paired sample data for this exercise.

Population 1 40 27 55 34

Population 2 53 42 68 50

- Compute \bar{d} and s_d .
- Give a point estimate for $\mu_1 - \mu_2 = \mu_d$.
- Construct the 99% confidence interval for $\mu_1 - \mu_2 = \mu_d$ from these data.

- d. Test, at the 10% level of significance, the hypothesis that $\mu_1 - \mu_2 \neq -12$ as an alternative to the null hypothesis that $\mu_1 - \mu_2 = -12$.
4. Use the following paired sample data for this exercise.
- | | | | | | |
|--------------|-----|-----|-----|-----|-----|
| Population 1 | 196 | 165 | 181 | 201 | 190 |
| Population 2 | 212 | 182 | 199 | 210 | 205 |
- a. Compute \bar{d} and s_d .
- b. Give a point estimate for $\mu_1 - \mu_2 = \mu_d$.
- c. Construct the 98% confidence interval for $\mu_1 - \mu_2 = \mu_d$ from these data.
- d. Test, at the 2% level of significance, the hypothesis that $\mu_1 - \mu_2 \neq -20$ as an alternative to the null hypothesis that $\mu_1 - \mu_2 = -20$.

APPLICATIONS

5. Each of five laboratory mice was released into a maze twice. The five pairs of times to escape were:

Mouse	1	2	3	4	5
First release	129	89	136	163	118
Second release	113	97	139	85	75

- a. Compute \bar{d} and s_d .
- b. Give a point estimate for $\mu_1 - \mu_2 = \mu_d$.
- c. Construct the 90% confidence interval for $\mu_1 - \mu_2 = \mu_d$ from these data.
- d. Test, at the 10% level of significance, the hypothesis that it takes mice less time to run the maze on the second trial, on average.
6. Eight golfers were asked to submit their latest scores on their favorite golf courses. These golfers were each given a set of newly designed clubs. After playing with the new clubs for a few months, the golfers were again asked to submit their latest scores on the same golf courses. The results are summarized below.

Golfer	1	2	3	4	5	6	7	8
Own clubs	77	80	69	73	73	72	75	77
New clubs	72	81	68	73	75	70	73	75

- a. Compute \bar{d} and s_d .
 - b. Give a point estimate for $\mu_1 - \mu_2 = \mu_d$.
 - c. Construct the 99% confidence interval for $\mu_1 - \mu_2 = \mu_d$ from these data.
 - d. Test, at the 1% level of significance, the hypothesis that on average golf scores are lower with the new clubs.
7. A neighborhood home owners association suspects that the recent appraisal values of the houses in the neighborhood conducted by the county government for taxation purposes is too high. It hired a private company to appraise the values of ten houses in the neighborhood. The results, in thousands of dollars, are

House	County Government	Private Company
1	217	219
2	350	338
3	296	291
4	237	237
5	237	235
6	272	269
7	257	239
8	277	275
9	312	320
10	335	335

- a. Give a point estimate for the difference between the mean private appraisal of all such homes and the government appraisal of all such homes.
 - b. Construct the 99% confidence interval based on these data for the difference.
 - c. Test, at the 1% level of significance, the hypothesis that appraised values by the county government of all such houses is greater than the appraised values by the private appraisal company.
8. In order to cut costs a wine producer is considering using duo or 1 + 1 corks in place of full natural wood corks, but is concerned that it could affect buyers's perception of the quality of the wine. The wine producer shipped eight pairs of bottles of its best young wines to eight wine experts. Each pair includes one

bottle with a natural wood cork and one with a duo cork. The experts are asked to rate the wines on a one to ten scale, higher numbers corresponding to higher quality. The results are:

Wine Expert	Duo Cork	Wood Cork
1	8.5	8.5
2	8.0	8.5
3	6.5	8.0
4	7.5	8.5
5	8.0	7.5
6	8.0	8.0
7	9.0	9.0
8	7.0	7.5

- Give a point estimate for the difference between the mean ratings of the wine when bottled are sealed with different kinds of corks.
 - Construct the 90% confidence interval based on these data for the difference.
 - Test, at the 10% level of significance, the hypothesis that on the average duo corks decrease the rating of the wine.
9. Engineers at a tire manufacturing corporation wish to test a new tire material for increased durability. To test the tires under realistic road conditions, new front tires are mounted on each of 11 company cars, one tire made with a production material and the other with the experimental material. After a fixed period the 11 pairs were measured for wear. The amount of wear for each tire (in mm) is shown in the table:

Car	Production	Experimental
1	5.1	5.0
2	6.5	6.5
3	3.6	3.1
4	3.5	3.7
5	5.7	4.5
6	5.0	4.1
7	6.4	5.3

Car	Production	Experimental
8	4.7	2.6
9	3.2	3.0
10	3.5	3.5
11	6.4	5.1

- Give a point estimate for the difference in mean wear.
 - Construct the 99% confidence interval for the difference based on these data.
 - Test, at the 1% level of significance, the hypothesis that the mean wear with the experimental material is less than that for the production material.
10. A marriage counselor administered a test designed to measure overall contentment to 30 randomly selected married couples. The scores for each couple are given below. A higher number corresponds to greater contentment or happiness.

Couple	Husband	Wife
1	47	44
2	44	46
3	49	44
4	53	44
5	42	43
6	45	45
7	48	47
8	45	44
9	52	44
10	47	42
11	40	34
12	45	42
13	40	43
14	46	41

Couple	Husband	Wife
15	47	45
16	46	45
17	46	41
18	46	41
19	44	45
20	45	43
21	48	38
22	42	46
23	50	44
24	46	51
25	43	45
26	50	40
27	46	46
28	42	41
29	51	41
30	46	47

- Test, at the 1% level of significance, the hypothesis that on average men and women are not equally happy in marriage.
- Test, at the 1% level of significance, the hypothesis that on average men are happier than women in marriage.

LARGE DATA SET EXERCISES

- Large Data Set 5 lists the scores for 25 randomly selected students on practice SAT reading tests before and after taking a two-week SAT preparation course. Denote the population of all students who have taken the course as Population 1 and the population of all students who have not taken the course as Population 2.

<http://www.gone.2012books.lardbucket.org/sites/all/files/data5.xls>

- a. Compute the 25 differences in the order after – before, their mean \bar{d} , and their sample standard deviation s_d .
 - b. Give a point estimate for $\mu_d = \mu_1 - \mu_2$, the difference in the mean score of all students who have taken the course and the mean score of all who have not.
 - c. Construct a 98% confidence interval for μ_d .
 - d. Test, at the 1% level of significance, the hypothesis that the mean SAT score increases by at least ten points by taking the two-week preparation course.
12. Large Data Set 12 lists the scores on one round for 75 randomly selected members at a golf course, first using their own original clubs, then two months later after using new clubs with an experimental design. Denote the population of all golfers using their own original clubs as Population 1 and the population of all golfers using the new style clubs as Population 2.

<http://www.gone.2012books.lardbucket.org/sites/all/files/data12.xls>

- a. Compute the 75 differences in the order original clubs – new clubs, their mean \bar{d} , and their sample standard deviation s_d .
 - b. Give a point estimate for $\mu_d = \mu_1 - \mu_2$, the difference in the mean score of all golfers using their original clubs and the mean score of all golfers using the new kind of clubs.
 - c. Construct a 90% confidence interval for μ_d .
 - d. Test, at the 1% level of significance, the hypothesis that the mean golf score decreases by at least one stroke by using the new kind of clubs.
13. Consider the previous problem again. Since the data set is so large, it is reasonable to use the standard normal distribution instead of Student's t -distribution with 74 degrees of freedom.
- a. Construct a 90% confidence interval for μ_d using the standard normal distribution, meaning that the formula is $\bar{d} \pm z_{\alpha/2} \frac{s_d}{\sqrt{n}}$. (The computations done in part (a) of the previous problem still apply and need not be redone.) How does the result obtained here compare to the result obtained in part (c) of the previous problem?
 - b. Test, at the 1% level of significance, the hypothesis that the mean golf score decreases by at least one stroke by using the new kind of clubs, using the standard normal distribution. (All the work done in part (d) of the previous problem applies, except the critical value is now z_α instead of t_α (or the p -value can be computed exactly instead of only approximated, if you used the p -value approach.) How does the result obtained here compare to the result obtained in part (c) of the previous problem?

- c. Construct the 99% confidence intervals for μ_d using both the t - and z -distributions. How much difference is there in the results now?

ANSWERS

1.
 - a. $\bar{d} = 7.4286, s_d = 0.9759,$
 - b. $\bar{d} = 7.4286,$
 - c. $(6.53, 8.33),$
 - d. $T = 1.162, df = 6, t_{0.10} = 1.44,$ do not reject H_0
3.
 - a. $\bar{d} = -14.25, s_d = 1.5,$
 - b. $\bar{d} = -14.25,$
 - c. $(-18.63, -9.87),$
 - d. $T = -3.000, df = 3, \pm t_{0.05} = \pm 2.353,$ reject H_0
5.
 - a. $\bar{d} = 25.2, s_d = 35.6609,$
 - b. 25.2,
 - c. 25.2 ± 34.0
 - d. $T = 1.580, df = 4, t_{0.10} = 1.533,$ reject H_0 (takes less time)
7.
 - a. 3.2,
 - b. 3.2 ± 7.5
 - c. $T = 1.392, df = 9, t_{0.10} = 2.821,$ do not reject H_0 (government appraisals not higher)
9.
 - a. 0.65,
 - b. $0.65 \pm 0.69,$
 - c. $T = 3.014, df = 10, t_{0.01} = 2.764,$ reject H_0 (experimental material wears less)
11.
 - a. $\bar{d} = 16.68$ and $s_d = 10.77$
 - b. $\bar{d} = 16.68$
 - c. $(11.31, 22.05)$
 - d. $H_0 : \mu_1 - \mu_2 = 10$ vs. $H_a : \mu_1 - \mu_2 > 10$. Test Statistic: $T = 3.1014$. $d.f. = 24$. Rejection Region: $[2.492, \infty)$. Decision: Reject H_0 .
13.
 - a. $(1.6266, 2.6401)$. Endpoints change in the third decimal place.
 - b. $H_0 : \mu_1 - \mu_2 = 1$ vs. $H_a : \mu_1 - \mu_2 > 1$. Test Statistic: $Z = 3.6791$. Rejection Region: $[2.33, \infty)$. Decision: Reject H_0 . The decision is the same as in the previous problem.
 - c. Using the t -distribution, $(1.3188, 2.9478)$. Using the z -distribution, $(1.3401, 2.9266)$. There is a difference.

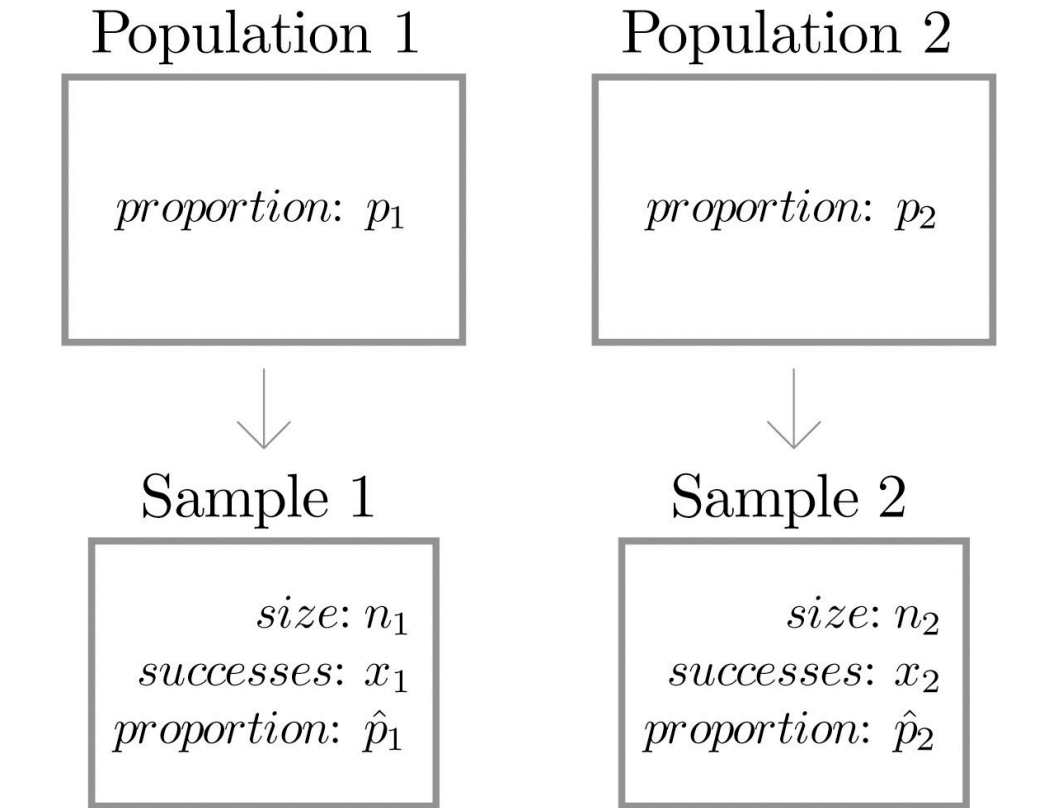
9.4 Comparison of Two Population Proportions

LEARNING OBJECTIVES

1. To learn how to construct a confidence interval for the difference in the proportions of two distinct populations that have a particular characteristic of interest.
2. To learn how to perform a test of hypotheses concerning the difference in the proportions of two distinct populations that have a particular characteristic of interest.

Suppose we wish to compare the proportions of two populations that have a specific characteristic, such as the proportion of men who are left-handed compared to the proportion of women who are left-handed. [Figure 9.7 "Independent Sampling from Two Populations In Order to Compare Proportions"](#) illustrates the conceptual framework of our investigation. Each population is divided into two groups, the group of elements that have the characteristic of interest (for example, being left-handed) and the group of elements that do not. We arbitrarily label one population as Population 1 and the other as Population 2, and subscript the proportion of each population that possesses the characteristic with the number 1 or 2 to tell them apart. We draw a random sample from Population 1 and label the sample statistic it yields with the subscript 1. Without reference to the first sample we draw a sample from Population 2 and label its sample statistic with the subscript 2.

Figure 9.7 Independent Sampling from Two Populations In Order to Compare Proportions



Our goal is to use the information in the *samples* to estimate the difference $p_1 - p_2$ in the two *population* proportions and to make statistically valid inferences about it.

Confidence Intervals

Since the sample proportion \hat{p}_1 computed using the sample drawn from Population 1 is a good estimator of population proportion p_1 of Population 1 and the sample proportion \hat{p}_2 computed using the sample drawn from Population 2 is a good estimator of population proportion p_2 of Population 2, a reasonable point estimate of the difference $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2$. In order to widen this point estimate into a confidence interval we suppose that both samples are large, as described in [Section 7.3 "Large Sample Estimation of a Population Proportion"](#) in [Chapter 7 "Estimation"](#) and repeated below. If so, then the following formula for a confidence interval for $p_1 - p_2$ is valid.

100 (1 - α) % Confidence Interval for the Difference Between Two Population Proportions

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2 (1 - \hat{p}_2)}{n_2}}$$

The samples must be independent, and *each* sample must be large: each of the intervals

$$\left[\hat{p}_1 - 3 \sqrt{\frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1}}, \hat{p}_1 + 3 \sqrt{\frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1}} \right]$$

and

$$\left[\hat{p}_2 - 3 \sqrt{\frac{\hat{p}_2 (1 - \hat{p}_2)}{n_2}}, \hat{p}_2 + 3 \sqrt{\frac{\hat{p}_2 (1 - \hat{p}_2)}{n_2}} \right]$$

must lie wholly within the interval $[0, 1]$.

EXAMPLE 10

The department of code enforcement of a county government issues permits to general contractors to work on residential projects. For each permit issued, the department inspects the result of the project and gives a “pass” or “fail” rating. A failed project must be re-inspected until it receives a pass rating. The department had been frustrated by the high cost of re-inspection and decided to publish the inspection records of all contractors on the web. It was hoped that public access to the records would lower the re-inspection rate. A year after the web access was made public, two samples of records were randomly selected. One sample was selected from the pool of records before the web publication and one after. The proportion of projects that passed on the first inspection was noted for each sample. The results are summarized below. Construct a point estimate and a 90% confidence interval for the difference in the passing rate on first inspection between the two time periods.

No public web access	$n_1 = 500 \quad \hat{p}_1 = 0.67$
Public web access	$n_2 = 100 \quad \hat{p}_2 = 0.80$

Solution:

The point estimate of $p_1 - p_2$ is

$$\hat{p}_1 - \hat{p}_2 = 0.67 - 0.80 = -0.13$$

Because the “No public web access” population was labeled as Population 1 and the “Public web access” population was labeled as Population 2, in words this means that we estimate that the proportion of projects that passed on the first inspection increased by 13 percentage points after records were posted on the web.

The sample sizes are sufficiently large for constructing a confidence interval since for sample 1:

$$3 \sqrt{\frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1}} = 3 \sqrt{\frac{(0.67) (0.33)}{500}} = 0.06$$

so that

$$\begin{aligned} & \left[\hat{p}_1 - 3 \sqrt{\frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1}}, \hat{p}_1 + 3 \sqrt{\frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1}} \right] \\ & = [0.67 - 0.06, 0.67 + 0.06] = [0.61, 0.73] \subset [0, 1] \end{aligned}$$

and for sample 2:

$$3 \sqrt{\frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1}} = 3 \sqrt{\frac{(0.8)(0.2)}{100}} = 0.12$$

so that

$$\begin{aligned} & \left[\hat{p}_2 - 3 \sqrt{\frac{\hat{p}_2 (1 - \hat{p}_2)}{n_2}}, \hat{p}_2 + 3 \sqrt{\frac{\hat{p}_2 (1 - \hat{p}_2)}{n_2}} \right] \\ & = [0.8 - 0.12, 0.8 + 0.12] = [0.68, 0.92] \subset [0, 1] \end{aligned}$$

To apply the formula for the confidence interval, we first observe that the 90% confidence level means that $\alpha = 1 - 0.90 = 0.10$ so that $z_{\alpha/2} = z_{0.05}$. From [Figure 12.3 "Critical Values of "](#) we read directly that $z_{0.05} = 1.645$. Thus the desired confidence interval is

$$\begin{aligned} (\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2 (1 - \hat{p}_2)}{n_2}} \\ & = -0.13 \pm 1.645 \sqrt{\frac{(0.67)(0.33)}{500} + \frac{(0.8)(0.2)}{100}} \\ & = -0.13 \pm 0.07 \end{aligned}$$

The 90% confidence interval is $[-0.20, -0.06]$. We are 90% confident that the difference in the population proportions lies in the interval $[-0.20, -0.06]$, in the sense that in repeated sampling 90% of all intervals constructed from the sample data in this manner will contain $p_1 - p_2$. Taking into account the labeling of the two populations, this means that we are 90% confident that the proportion of projects that pass on

the first inspection is between 6 and 20 percentage points higher after public access to the records than before.

Hypothesis Testing

In hypothesis tests concerning the relative sizes of the proportions p_1 and p_2 of two populations that possess a particular characteristic, the null and alternative hypotheses will always be expressed in terms of the difference of the two population proportions. Hence the null hypothesis is always written

$$H_0 : p_1 - p_2 = D_0$$

The three forms of the alternative hypothesis, with the terminology for each case, are:

Form of H_a	Terminology
$H_a : p_1 - p_2 < D_0$	Left-tailed
$H_a : p_1 - p_2 > D_0$	Right-tailed
$H_a : p_1 - p_2 \neq D_0$	Two-tailed

As long as the samples are independent and both are large the following formula for the standardized test statistic is valid, and it has the standard normal distribution.

Standardized Test Statistic for Hypothesis Tests Concerning the Difference Between Two Population Proportions

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - D_0}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

The test statistic has the standard normal distribution.

The samples must be independent, and *each* sample must be large: each of the intervals

$$\left[\hat{p}_1 - 3 \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}}, \hat{p}_1 + 3 \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}} \right]$$

and

$$\left[\hat{p}_2 - 3 \sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}, \hat{p}_2 + 3 \sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \right]$$

must lie wholly within the interval $[0,1]$.

EXAMPLE 11

Using the data of [Note 9.25 "Example 10"](#), test whether there is sufficient evidence to conclude that public web access to the inspection records has increased the proportion of projects that passed on the first inspection by more than 5 percentage points. Use the critical value approach at the 10% level of significance.

Solution:

- Step 1. Taking into account the labeling of the populations an increase in passing rate at the first inspection by more than 5 percentage points after public access on the web may be expressed as $p_2 > p_1 + 0.05$, which by algebra is the same as $p_1 - p_2 < -0.05$. This is the alternative hypothesis. Since the null hypothesis is always expressed as an equality, with the same number on the right as is in the alternative hypothesis, the test is

$$H_0 : p_1 - p_2 = -0.05$$

$$\text{vs. } H_a : p_1 - p_2 < -0.05 \quad @ \alpha = 0.10$$

- Step 2. Since the test is with respect to a difference in population proportions the test statistic is

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - D_0}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

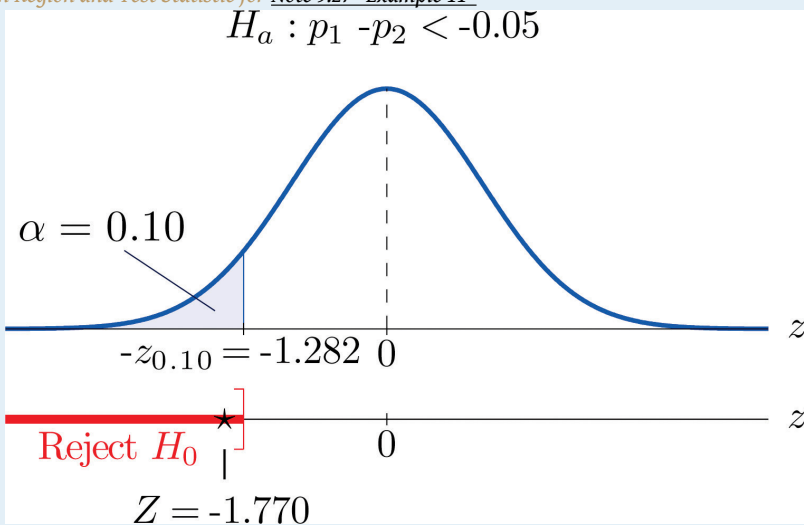
- Step 3. Inserting the values given in [Note 9.25 "Example 10"](#) and the value $D_0 = -0.05$ into the formula for the test statistic gives

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - D_0}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} = \frac{(-0.13) - (-0.05)}{\sqrt{\frac{(0.67)(0.33)}{500} + \frac{(0.8)(0.2)}{100}}} = -1.770$$

- Step 4. Since the symbol in H_a is "<" this is a left-tailed test, so there is a single critical value, $z_\alpha = -z_{0.10}$. From the last row in [Figure 12.3 "Critical Values of "](#) $z_{0.10} = 1.282$, so $-z_{0.10} = -1.282$. The rejection region is $(-\infty, -1.282]$.
- Step 5. As shown in [Figure 9.8 "Rejection Region and Test Statistic for "](#) the test statistic falls in the rejection region. The decision is to reject H_0 . In the context of the problem our conclusion is:

The data provide sufficient evidence, at the 10% level of significance, to conclude that the rate of passing on the first inspection has increased by more than 5 percentage points since records were publicly posted on the web.

Figure 9.8
Rejection Region and Test Statistic for [Note 9.27 "Example 11"](#)



EXAMPLE 12

Perform the test of [Note 9.27 "Example 11"](#) using the p -value approach.

Solution:

The first three steps are identical to those in [Note 9.27 "Example 11"](#).

- Step 4. Because the test is left-tailed the observed significance or p -value of the test is just the area of the left tail of the standard normal distribution that is cut off by the test statistic $Z = -1.770$. From [Figure 12.2 "Cumulative Normal Probability"](#) the area of the left tail determined by -1.77 is 0.0384. The p -value is 0.0384.
- Step 5. Since the p -value 0.0384 is less than $\alpha = 0.10$, the decision is to reject the null hypothesis: The data provide sufficient evidence, at the 10% level of significance, to conclude that the rate of passing on the first inspection has increased by more than 5 percentage points since records were publicly posted on the web.

Finally a common misuse of the formulas given in this section must be mentioned. Suppose a large pre-election survey of potential voters is conducted. Each person surveyed is asked to express a preference between, say, Candidate A and Candidate B. (Perhaps “no preference” or “other” are also choices, but that is not important.) In such a survey, estimators \hat{p}_A and \hat{p}_B of p_A and p_B can be calculated. It is important to realize, however, that these two estimators were not calculated from two independent samples. While $\hat{p}_A - \hat{p}_B$ may be a reasonable estimator of $p_A - p_B$, the formulas for confidence intervals and for the standardized test statistic given in this section are not valid for data obtained in this manner.

KEY TAKEAWAYS

- A confidence interval for the difference in two population proportions is computed using a formula in the same fashion as was done for a single population mean.
- The same five-step procedure used to test hypotheses concerning a single population proportion is used to test hypotheses concerning the difference between two population proportions. The only difference is in the formula for the standardized test statistic.

EXERCISES

BASIC

1. Construct the confidence interval for $p_1 - p_2$ for the level of confidence and the data given. (The samples are sufficiently large.)
 - a. 90% confidence,
 $n_1 = 1670, \hat{p}_1 = 0.42$
 $n_2 = 900, \hat{p}_2 = 0.38$
 - b. 95% confidence,
 $n_1 = 600, \hat{p}_1 = 0.84$
 $n_2 = 420, \hat{p}_2 = 0.67$

2. Construct the confidence interval for $p_1 - p_2$ for the level of confidence and the data given. (The samples are sufficiently large.)
 - a. 98% confidence,
 $n_1 = 750, \hat{p}_1 = 0.64$
 $n_2 = 800, \hat{p}_2 = 0.51$
 - b. 99.5% confidence,
 $n_1 = 250, \hat{p}_1 = 0.78$
 $n_2 = 250, \hat{p}_2 = 0.51$

3. Construct the confidence interval for $p_1 - p_2$ for the level of confidence and the data given. (The samples are sufficiently large.)
 - a. 80% confidence,
 $n_1 = 300, \hat{p}_1 = 0.255$
 $n_2 = 400, \hat{p}_2 = 0.193$
 - b. 95% confidence,

$$n_1 = 3500, \hat{p}_1 = 0.147$$

$$n_2 = 3750, \hat{p}_2 = 0.131$$

4. Construct the confidence interval for $p_1 - p_2$ for the level of confidence and the data given. (The samples are sufficiently large.)

- a. 99% confidence,

$$n_1 = 2250, \hat{p}_1 = 0.915$$

$$n_2 = 2525, \hat{p}_2 = 0.858$$

- b. 95% confidence,

$$n_1 = 120, \hat{p}_1 = 0.650$$

$$n_2 = 200, \hat{p}_2 = 0.505$$

5. Perform the test of hypotheses indicated, using the data given. Use the critical value approach. Compute the p -value of the test as well. (The samples are sufficiently large.)

- a. Test $H_0 : p_1 - p_2 = 0$ vs. $H_a : p_1 - p_2 > 0$ @ $\alpha = 0.10$,

$$n_1 = 1200, \hat{p}_1 = 0.42$$

$$n_2 = 1200, \hat{p}_2 = 0.40$$

- b. Test $H_0 : p_1 - p_2 = 0$ vs. $H_a : p_1 - p_2 \neq 0$ @ $\alpha = 0.05$,

$$n_1 = 550, \hat{p}_1 = 0.61$$

$$n_2 = 600, \hat{p}_2 = 0.67$$

6. Perform the test of hypotheses indicated, using the data given. Use the critical value approach. Compute the p -value of the test as well. (The samples are sufficiently large.)

- a. Test $H_0 : p_1 - p_2 = 0.05$ vs. $H_a : p_1 - p_2 > 0.05$ @ $\alpha = 0.05$,

$$n_1 = 1100, \hat{p}_1 = 0.57$$

$$n_2 = 1100, \hat{p}_2 = 0.48$$

- b. Test $H_0 : p_1 - p_2 = 0$ vs. $H_a : p_1 - p_2 \neq 0$ @ $\alpha = 0.05$,

$$n_1 = 800, \hat{p}_1 = 0.39$$

$$n_2 = 900, \hat{p}_2 = 0.43$$

7. Perform the test of hypotheses indicated, using the data given. Use the critical value approach. Compute the p -value of the test as well. (The samples are sufficiently large.)

- a. Test $H_0 : p_1 - p_2 = 0.25$ vs. $H_a : p_1 - p_2 < 0.25$ @ $\alpha = 0.005$,

$$n_1 = 1400, \hat{p}_1 = 0.57$$

$$n_2 = 1200, \hat{p}_2 = 0.37$$

- b. Test $H_0 : p_1 - p_2 = 0.16$ vs. $H_a : p_1 - p_2 \neq 0.16$ @ $\alpha = 0.02$,

$$n_1 = 750, \hat{p}_1 = 0.43$$

$$n_2 = 600, \hat{p}_2 = 0.22$$

8. Perform the test of hypotheses indicated, using the data given. Use the critical value approach. Compute the p -value of the test as well. (The samples are sufficiently large.)

- a. Test $H_0 : p_1 - p_2 = 0.08$ vs. $H_a : p_1 - p_2 > 0.08$ @ $\alpha = 0.025$,

$$n_1 = 450, \hat{p}_1 = 0.67$$

$$n_2 = 200, \hat{p}_2 = 0.52$$

- b. Test $H_0 : p_1 - p_2 = 0.02$ vs. $H_a : p_1 - p_2 \neq 0.02$ @ $\alpha = 0.001$,

$$n_1 = 2700, \hat{p}_1 = 0.837$$

$$n_2 = 2900, \hat{p}_2 = 0.854$$

9. Perform the test of hypotheses indicated, using the data given. Use the p -value approach. (The samples are sufficiently large.)

- a. Test $H_0 : p_1 - p_2 = 0$ vs. $H_a : p_1 - p_2 < 0$ @ $\alpha = 0.005$,

$$n_1 = 1100, \hat{p}_1 = 0.22$$

$$n_2 = 1300, \hat{p}_2 = 0.27$$

- b. Test $H_0 : p_1 - p_2 = 0$ vs. $H_a : p_1 - p_2 \neq 0$ @ $\alpha = 0.01$,

$$n_1 = 650, \hat{p}_1 = 0.35$$

$$n_2 = 650, \hat{p}_2 = 0.41$$

10. Perform the test of hypotheses indicated, using the data given. Use the p -value approach. (The samples are sufficiently large.)

- a. Test $H_0 : p_1 - p_2 = 0.15$ vs. $H_a : p_1 - p_2 > 0.15$ @ $\alpha = 0.10$,

$$n_1 = 950, \hat{p}_1 = 0.41$$

$$n_2 = 500, \hat{p}_2 = 0.23$$

- b. Test $H_0 : p_1 - p_2 = 0.10$ vs. $H_a : p_1 - p_2 \neq 0.10$ @ $\alpha = 0.10$,

$$n_1 = 220, \hat{p}_1 = 0.92$$

$$n_2 = 160, \hat{p}_2 = 0.78$$

11. Perform the test of hypotheses indicated, using the data given. Use the p -value approach. (The samples are sufficiently large.)

- a. Test $H_0 : p_1 - p_2 = 0.22$ vs. $H_a : p_1 - p_2 > 0.22$ @ $\alpha = 0.05$,

$$n_1 = 90, \hat{p}_1 = 0.72$$

$$n_2 = 75, \hat{p}_2 = 0.40$$

- b. Test $H_0 : p_1 - p_2 = 0.37$ vs. $H_a : p_1 - p_2 \neq 0.37$ @ $\alpha = 0.02$,

$$n_1 = 425, \hat{p}_1 = 0.772$$

$$n_2 = 425, \hat{p}_2 = 0.331$$

12. Perform the test of hypotheses indicated, using the data given. Use the p -value approach. (The samples are sufficiently large.)

- a. Test $H_0 : p_1 - p_2 = 0.50$ vs. $H_a : p_1 - p_2 < 0.50$ @ $\alpha = 0.10$,

$$n_1 = 40, \hat{p}_1 = 0.65$$

$$n_2 = 55, \hat{p}_2 = 0.24$$

- b. Test $H_0 : p_1 - p_2 = 0.30$ vs. $H_a : p_1 - p_2 \neq 0.30$ @ $\alpha = 0.10$,
 $n_1 = 7500, \hat{p}_1 = 0.664$
 $n_2 = 1000, \hat{p}_2 = 0.319$

APPLICATIONS

In all the remaining exercises the samples are sufficiently large (so this need not be checked).

13. Voters in a particular city who identify themselves with one or the other of two political parties were randomly selected and asked if they favor a proposal to allow citizens with proper license to carry a concealed handgun in city parks. The results are:

	Party A	Party B
Sample size, n	150	200
Number in favor, x	90	140

- Give a point estimate for the difference in the proportion of all members of Party A and all members of Party B who favor the proposal.
 - Construct the 95% confidence interval for the difference, based on these data.
 - Test, at the 5% level of significance, the hypothesis that the proportion of all members of Party A who favor the proposal is less than the proportion of all members of Party B who do.
 - Compute the p -value of the test.
14. To investigate a possible relation between gender and handedness, a random sample of 320 adults was taken, with the following results:

	Men	Women
Sample size, n	168	152
Number of left-handed, x	24	9

- Give a point estimate for the difference in the proportion of all men who are left-handed and the proportion of all women who are left-handed.
- Construct the 95% confidence interval for the difference, based on these data.

- c. Test, at the 5% level of significance, the hypothesis that the proportion of men who are left-handed is greater than the proportion of women who are.
 - d. Compute the p -value of the test.
15. A local school board member randomly sampled private and public high school teachers in his district to compare the proportions of National Board Certified (NBC) teachers in the faculty. The results were:

	Private Schools	Public Schools
Sample size, n	80	520
Proportion of NBC teachers, \hat{p}	0.175	0.150

- a. Give a point estimate for the difference in the proportion of all teachers in area public schools and the proportion of all teachers in private schools who are National Board Certified.
 - b. Construct the 90% confidence interval for the difference, based on these data.
 - c. Test, at the 10% level of significance, the hypothesis that the proportion of all public school teachers who are National Board Certified is less than the proportion of private school teachers who are.
 - d. Compute the p -value of the test.
16. In professional basketball games, the fans of the home team always try to distract free throw shooters on the visiting team. To investigate whether this tactic is actually effective, the free throw statistics of a professional basketball player with a high free throw percentage were examined. During the entire last season, this player had 656 free throws, 420 in home games and 236 in away games. The results are summarized below.

	Home	Away
Sample size, n	420	236
Free throw percent, \hat{p}	81.5%	78.8%

- a. Give a point estimate for the difference in the proportion of free throws made at home and away.
- b. Construct the 90% confidence interval for the difference, based on these data.
- c. Test, at the 10% level of significance, the hypothesis that there exists a home advantage in free throws.
- d. Compute the p -value of the test.

17. Randomly selected middle-aged people in both China and the United States were asked if they believed that adults have an obligation to financially support their aged parents. The results are summarized below.

	China	USA
Sample size, n	1300	150
Number of yes, x	1170	110

Test, at the 1% level of significance, whether the data provide sufficient evidence to conclude that there exists a cultural difference in attitude regarding this question.

18. A manufacturer of walk-behind push mowers receives refurbished small engines from two new suppliers, A and B. It is not uncommon that some of the refurbished engines need to be lightly serviced before they can be fitted into mowers. The mower manufacturer recently received 100 engines from each supplier. In the shipment from A, 13 needed further service. In the shipment from B, 10 needed further service. Test, at the 10% level of significance, whether the data provide sufficient evidence to conclude that there exists a difference in the proportions of engines from the two suppliers needing service.

LARGE DATA SET EXERCISES

19. Large Data Sets 6A and 6B record results of a random survey of 200 voters in each of two regions, in which they were asked to express whether they prefer Candidate A for a U.S. Senate seat or prefer some other candidate. Let the population of all voters in region 1 be denoted Population 1 and the population of all voters in region 2 be denoted Population 2. Let p_1 be the proportion of voters in Population 1 who prefer Candidate A, and p_2 the proportion in Population 2 who do.

<http://www.gone.2012books.lardbucket.org/sites/all/files/data6A.xls>

<http://www.gone.2012books.lardbucket.org/sites/all/files/data6B.xls>

- Find the relevant sample proportions \hat{p}_1 and \hat{p}_2 .
- Construct a point estimate for $p_1 - p_2$.
- Construct a 95% confidence interval for $p_1 - p_2$.
- Test, at the 5% level of significance, the hypothesis that the same proportion of voters in the two regions favor Candidate A, against the alternative that a larger proportion in Population 2 do.

20. Large Data Set 11 records the results of samples of real estate sales in a certain region in the year 2008 (lines 2 through 536) and in the year 2010 (lines 537 through 1106). Foreclosure sales are identified with a 1 in the second column. Let all real estate sales in the region in 2008 be Population 1 and all real estate sales in the region in 2010 be Population 2.

<http://www.gone.2012books.lardbucket.org/sites/all/files/data11.xls>

- a. Use the sample data to construct point estimates \hat{p}_1 and \hat{p}_2 of the proportions p_1 and p_2 of all real estate sales in this region in 2008 and 2010 that were foreclosure sales. Construct a point estimate of $p_1 - p_2$.
- b. Use the sample data to construct a 90% confidence for $p_1 - p_2$.
- c. Test, at the 10% level of significance, the hypothesis that the proportion of real estate sales in the region in 2010 that were foreclosure sales was greater than the proportion of real estate sales in the region in 2008 that were foreclosure sales. (The default is that the proportions were the same.)

ANSWERS

1.
 - a. $(0.0068, 0.0732)$,
 - b. $(0.1163, 0.2237)$
3.
 - a. $(0.0210, 0.1030)$,
 - b. $(0.0001, 0.0319)$
5.
 - a. $Z = 0.996, z_{0.10} = 1.282, p\text{-value} = 0.1587$, do not reject H_0 ,
 - b. $Z = -2.120, \pm z_{0.025} = \pm 1.960, p\text{-value} = 0.0340$, reject H_0
7.
 - a. $Z = -2.602, -z_{0.005} = -2.576, p\text{-value} = 0.0047$, reject H_0 ,
 - b. $Z = 2.020, \pm z_{0.01} = \pm 2.326, p\text{-value} = 0.0434$, do not reject H_0
9.
 - a. $Z = -2.85, p\text{-value} = 0.0022$, reject H_0 ,
 - b. $Z = -2.23, p\text{-value} = 0.0258$, do not reject H_0
11.
 - a. $Z = 1.36, p\text{-value} = 0.0869$, do not reject H_0 ,
 - b. $Z = 2.32, p\text{-value} = 0.0204$, do not reject H_0
13.
 - a. -0.10 ,
 - b. -0.10 ± 0.101 ,
 - c. $Z = -1.943, -z_{0.05} = -1.645$, reject H_0 (fewer in Party A favor),
 - d. $p\text{-value} = 0.0262$
15.
 - a. 0.025 ,
 - b. 0.025 ± 0.0745 ,
 - c. $Z = 0.552, z_{0.10} = 1.282$, do not reject H_0 (as many public school teachers are certified),
 - d. $p\text{-value} = 0.2912$
17. $Z = 4.498, \pm z_{0.005} = \pm 2.576$, reject H_0 (different)
19.
 - a. $\hat{p}_1 = 0.355$ and $\hat{p}_2 = 0.41$
 - b. $\hat{p}_1 - \hat{p}_2 = -0.055$
 - c. $(-0.1501, 0.0401)$
 - d. $H_0 : p_1 - p_2 = 0$ vs. $H_a : p_1 - p_2 < 0$. Test Statistic:
 $Z = -1.1335$. Rejection Region: $(-\infty, -1.645]$. Decision: Fail to reject H_0 .

9.5 Sample Size Considerations

LEARNING OBJECTIVE

1. To learn how to apply formulas for estimating the size samples that will be needed in order to construct a confidence interval for the difference in two population means or proportions that meets given criteria.

As was pointed out at the beginning of [Section 7.4 "Sample Size Considerations"](#) in [Chapter 7 "Estimation"](#), sampling is typically done with definite objectives in mind. For example, a physician might wish to estimate the difference in the average amount of sleep gotten by patients suffering a certain condition with the average amount of sleep got by healthy adults, at 90% confidence and to within half an hour. Since sampling costs time, effort, and money, it would be useful to be able to estimate the smallest size samples that are likely to meet these criteria.

Estimating $\mu_1 - \mu_2$ with Independent Samples

Assuming that large samples will be required, the confidence interval formula for estimating the difference $\mu_1 - \mu_2$ between two population means using independent samples is $(\bar{x}_1 - \bar{x}_2) \pm E$, where

$$E = z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

To say that we wish to estimate the mean to within a certain number of units means that we want the margin of error E to be no larger than that number. The number $z_{\alpha/2}$ is determined by the desired level of confidence.

The numbers s_1 and s_2 are estimates of the standard deviations σ_1 and σ_2 of the two populations. In analogy with what we did in [Section 7.4 "Sample Size Considerations"](#) in [Chapter 7 "Estimation"](#) we will assume that we either know or can reasonably approximate σ_1 and σ_2 .

We cannot solve for both n_1 and n_2 , so we have to make an assumption about their relative sizes. We will specify that they be equal. With these assumptions we obtain

the minimum sample sizes needed by solving the equation displayed just above for $n_1 = n_2$.

Minimum Equal Sample Sizes for Estimating the Difference in the Means of Two Populations Using Independent Samples

The estimated minimum equal sample sizes $n_1 = n_2$ needed to estimate the difference $\mu_1 - \mu_2$ in two population means to within E units at $100(1 - \alpha)\%$ confidence is

$$n_1 = n_2 = \frac{(z_{\alpha/2})^2(\sigma_1^2 + \sigma_2^2)}{E^2} \text{ (rounded up)}$$

In all the examples and exercises the population standard deviations σ_1 and σ_2 will be given.

EXAMPLE 13

A law firm wishes to estimate the difference in the mean delivery time of documents sent between two of its offices by two different courier companies, to within half an hour and with 99.5% confidence. From their records it will randomly sample the same number n of documents as delivered by each courier company. Determine how large n must be if the estimated standard deviations of the delivery times are 0.75 hour for one company and 1.15 hours for the other.

Solution:

Confidence level 99.5% means that $\alpha = 1 - 0.995 = 0.005$ so $\alpha / 2 = 0.0025$. From the last line of [Figure 12.3 "Critical Values of "](#) we obtain $z_{0.0025} = 2.807$.

To say that the estimate is to be "to within half an hour" means that $E = 0.5$. Thus

$$n = \frac{(z_{\alpha/2})^2 (\sigma_1^2 + \sigma_2^2)}{E^2} = \frac{(2.807)^2 (0.75^2 + 1.15^2)}{0.5^2} = 59.40953746$$

which we round up to 60, since it is impossible to take a fractional observation. The law firm must sample 60 document deliveries by each company.

Estimating $\mu_1 - \mu_2$ with Paired Samples

As we mentioned at the end of [Section 9.3 "Comparison of Two Population Means: Paired Samples"](#), if the sample is large (meaning that $n \geq 30$) then in the formula for the confidence interval we may replace $t_{\alpha/2}$ by $z_{\alpha/2}$, so that the confidence interval formula becomes $\bar{d} \pm E$ for

$$E = z_{\alpha/2} \frac{s_d}{\sqrt{n}}$$

The number s_d is an estimate of the standard deviations σ_d of the population of differences. We must assume that we either know or can reasonably approximate σ_d . Thus, assuming that large samples will be required to meet the criteria given,

we can solve the displayed equation for n to obtain an estimate of the number of pairs needed in the sample.

Minimum Sample Size for Estimating the Difference in the Means of Two Populations Using Paired Difference Samples

The estimated minimum number of pairs n needed to estimate the difference $\mu_d = \mu_1 - \mu_2$ in two population means to within E units at $100(1 - \alpha)\%$ confidence using paired difference samples is

$$n = \frac{(z_{\alpha/2})^2 \sigma_d^2}{E^2} \text{ (rounded up)}$$

In all the examples and exercises the population standard deviation of the differences σ_d will be given.

EXAMPLE 14

A automotive tire manufacturer wishes to compare the mean lifetime of two tread designs under actual driving conditions. They will mount one of each type of tire on n vehicles (both on the front or both on the back) and measure the difference in remaining tread after 20,000 miles of driving. If the standard deviation of the differences is assumed to be 0.025 inch, find the minimum samples size needed to estimate the difference in mean depth (at 20,000 miles use) to within 0.01 inch at 99.9% confidence.

Solution:

Confidence level 99.9% means that $\alpha = 1 - 0.999 = 0.001$ so $\alpha / 2 = 0.0005$. From the last line of [Figure 12.3 "Critical Values of "](#) we obtain $z_{0.0005} = 3.291$.

To say that the estimate is to be “to within 0.01 inch” means that $E = 0.01$. Thus

$$n = \frac{(z_{\alpha/2})^2 \sigma_d^2}{E^2} = \frac{(3.291)^2 (0.025)^2}{(0.01)^2} = 67.69175625$$

which we round up to 68. The manufacturer must test 68 pairs of tires.

Estimating $p_1 - p_2$

The confidence interval formula for estimating the difference $p_1 - p_2$ between two population proportions is $\hat{p}_1 - \hat{p}_2 \pm E$, where

$$E = z_{\alpha/2} \sqrt{\frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2 (1 - \hat{p}_2)}{n_2}}$$

To say that we wish to estimate the mean to within a certain number of units means that we want the margin of error E to be no larger than that number. The number $z_{\alpha/2}$ is determined by the desired level of confidence.

We cannot solve for both n_1 and n_2 , so we have to make an assumption about their relative sizes. We will specify that they be equal. With these assumptions we obtain the minimum sample sizes needed by solving the displayed equation for $n_1 = n_2$.

Minimum Equal Sample Sizes for Estimating the Difference in Two Population Proportions

The estimated minimum equal sample sizes $n_1 = n_2$ needed to estimate the difference $p_1 - p_2$ in two population proportions to within E percentage points at $100(1 - \alpha)\%$ confidence is

$$n_1 = n_2 = \frac{(z_{\alpha/2})^2 (\hat{p}_1 (1 - \hat{p}_1) + \hat{p}_2 (1 - \hat{p}_2))}{E^2} \text{ (rounded up)}$$

Here we face the same dilemma that we encountered in the case of a single population proportion: the formula for estimating how large a sample to take contains the numbers \hat{p}_1 and \hat{p}_2 , which we know only after we have taken the sample. There are two ways out of this dilemma. Typically the researcher will have some idea as to the values of the population proportions p_1 and p_2 , hence of what the sample proportions \hat{p}_1 and \hat{p}_2 are likely to be. If so, those estimates can be used in the formula.

The second approach to resolving the dilemma is simply to replace each of \hat{p}_1 and \hat{p}_2 in the formula by 0.5. As in the one-population case, this is the **most conservative estimate**, since it gives the largest possible estimate of n . If we have an estimate of only one of p_1 and p_2 we can use that estimate for it, and use the conservative estimate 0.5 for the other.

EXAMPLE 15

Find the minimum equal sample sizes necessary to construct a 98% confidence interval for the difference $p_1 - p_2$ with a margin of error $E = 0.05$,

- assuming that no prior knowledge about p_1 or p_2 is available; and
- assuming that prior studies suggest that $p_1 \approx 0.2$ and $p_2 \approx 0.3$.

Solution:

Confidence level 98% means that $\alpha = 1 - 0.98 = 0.02$ so $\alpha / 2 = 0.01$. From the last line of [Figure 12.3 "Critical Values of "](#) we obtain $z_{0.01} = 2.326$.

- Since there is no prior knowledge of p_1 or p_2 we make the most conservative estimate that $\hat{p}_1 = 0.5$ and $\hat{p}_2 = 0.5$. Then

$$\begin{aligned} n_1 = n_2 &= \frac{(z_{\alpha/2})^2 (\hat{p}_1 (1 - \hat{p}_1) + \hat{p}_2 (1 - \hat{p}_2))}{E^2} \\ &= \frac{(2.326)^2 ((0.5)(0.5) + (0.5)(0.5))}{0.05^2} \\ &= 1082.0552 \end{aligned}$$

which we round up to 1,083. We must take a sample of size 1,083 from each population.

- Since $p_1 \approx 0.2$ we estimate \hat{p}_1 by 0.2, and since $p_2 \approx 0.3$ we estimate \hat{p}_2 by 0.3. Thus we obtain

$$\begin{aligned}
 n_1 = n_2 &= \frac{(z_{\alpha/2})^2 (\hat{p}_1 (1 - \hat{p}_1) + \hat{p}_2 (1 - \hat{p}_2))}{E^2} \\
 &= \frac{(2.326)^2 ((0.2)(0.8) + (0.3)(0.7))}{0.05^2} \\
 &= 800.720848
 \end{aligned}$$

which we round up to 801. We must take a sample of size 801 from each population.

KEY TAKEAWAYS

- If the population standard deviations σ_1 and σ_2 are known or can be estimated, then the minimum equal sizes of independent samples needed to obtain a confidence interval for the difference $\mu_1 - \mu_2$ in two population means with a given maximum error of the estimate E and a given level of confidence can be estimated.
- If the standard deviation σ_d of the population of differences in pairs drawn from two populations is known or can be estimated, then the minimum number of sample pairs needed under paired difference sampling to obtain a confidence interval for the difference $\mu_d = \mu_1 - \mu_2$ in two population means with a given maximum error of the estimate E and a given level of confidence can be estimated.
- The minimum equal sample sizes needed to obtain a confidence interval for the difference in two population proportions with a given maximum error of the estimate and a given level of confidence can always be estimated. If there is prior knowledge of the population proportions p_1 and p_2 then the estimate can be sharpened.

EXERCISES

BASIC

1. Estimate the common sample size n of equally sized independent samples needed to estimate $\mu_1 - \mu_2$ as specified when the population standard deviations are as shown.
 - a. 90% confidence, to within 3 units, $\sigma_1 = 10$ and $\sigma_2 = 7$
 - b. 99% confidence, to within 4 units, $\sigma_1 = 6.8$ and $\sigma_2 = 9.3$
 - c. 95% confidence, to within 5 units, $\sigma_1 = 22.6$ and $\sigma_2 = 31.8$
2. Estimate the common sample size n of equally sized independent samples needed to estimate $\mu_1 - \mu_2$ as specified when the population standard deviations are as shown.
 - a. 80% confidence, to within 2 units, $\sigma_1 = 14$ and $\sigma_2 = 23$
 - b. 90% confidence, to within 0.3 units, $\sigma_1 = 1.3$ and $\sigma_2 = 0.8$
 - c. 99% confidence, to within 11 units, $\sigma_1 = 42$ and $\sigma_2 = 37$
3. Estimate the number n of pairs that must be sampled in order to estimate $\mu_d = \mu_1 - \mu_2$ as specified when the standard deviation s_d of the population of differences is as shown.
 - a. 80% confidence, to within 6 units, $\sigma_d = 26.5$
 - b. 95% confidence, to within 4 units, $\sigma_d = 12$
 - c. 90% confidence, to within 5.2 units, $\sigma_d = 11.3$
4. Estimate the number n of pairs that must be sampled in order to estimate $\mu_d = \mu_1 - \mu_2$ as specified when the standard deviation s_d of the population of differences is as shown.
 - a. 90% confidence, to within 20 units, $\sigma_d = 75.5$
 - b. 95% confidence, to within 11 units, $\sigma_d = 31.4$
 - c. 99% confidence, to within 1.8 units, $\sigma_d = 4$
5. Estimate the minimum equal sample sizes $n_1 = n_2$ necessary in order to estimate $p_1 - p_2$ as specified.
 - a. 80% confidence, to within 0.05 (five percentage points)
 - a. when no prior knowledge of p_1 or p_2 is available
 - b. when prior studies indicate that $p_1 \approx 0.20$ and $p_2 \approx 0.65$
 - b. 90% confidence, to within 0.02 (two percentage points)

- a. when no prior knowledge of p_1 or p_2 is available
 - b. when prior studies indicate that $p_1 \approx 0.75$ and $p_2 \approx 0.63$
 - c. 95% confidence, to within 0.10 (ten percentage points)
 - a. when no prior knowledge of p_1 or p_2 is available
 - b. when prior studies indicate that $p_1 \approx 0.11$ and $p_2 \approx 0.37$
6. Estimate the minimum equal sample sizes $n_1 = n_2$ necessary in order to estimate $p_1 - p_2$ as specified.
- a. 80% confidence, to within 0.02 (two percentage points)
 - a. when no prior knowledge of p_1 or p_2 is available
 - b. when prior studies indicate that $p_1 \approx 0.78$ and $p_2 \approx 0.65$
 - b. 90% confidence, to within 0.05 (two percentage points)
 - a. when no prior knowledge of p_1 or p_2 is available
 - b. when prior studies indicate that $p_1 \approx 0.12$ and $p_2 \approx 0.24$
 - c. 95% confidence, to within 0.10 (ten percentage points)
 - a. when no prior knowledge of p_1 or p_2 is available
 - b. when prior studies indicate that $p_1 \approx 0.14$ and $p_2 \approx 0.21$

APPLICATIONS

- 7. An educational researcher wishes to estimate the difference in average scores of elementary school children on two versions of a 100-point standardized test, at 99% confidence and to within two points. Estimate the minimum equal sample sizes necessary if it is known that the standard deviation of scores on different versions of such tests is 4.9.
- 8. A university administrator wishes to estimate the difference in mean grade point averages among all men affiliated with fraternities and all unaffiliated men, with 95% confidence and to within 0.15. It is known from prior studies that the standard deviations of grade point averages in the two groups have common value 0.4. Estimate the minimum equal sample sizes necessary to meet these criteria.
- 9. An automotive tire manufacturer wishes to estimate the difference in mean wear of tires manufactured with an experimental material and ordinary production tire, with 90% confidence and to within 0.5 mm. To eliminate

extraneous factors arising from different driving conditions the tires will be tested in pairs on the same vehicles. It is known from prior studies that the standard deviations of the differences of wear of tires constructed with the two kinds of materials is 1.75 mm. Estimate the minimum number of pairs in the sample necessary to meet these criteria.

10. To assess to the relative happiness of men and women in their marriages, a marriage counselor plans to administer a test measuring happiness in marriage to n randomly selected married couples, record the their test scores, find the differences, and then draw inferences on the possible difference. Let μ_1 and μ_2 be the true average levels of happiness in marriage for men and women respectively as measured by this test. Suppose it is desired to find a 90% confidence interval for estimating $\mu_d = \mu_1 - \mu_2$ to within two test points. Suppose further that, from prior studies, it is known that the standard deviation of the differences in test scores is $\sigma_d \approx 10$. What is the minimum number of married couples that must be included in this study?
11. A journalist plans to interview an equal number of members of two political parties to compare the proportions in each party who favor a proposal to allow citizens with a proper license to carry a concealed handgun in public parks. Let p_1 and p_2 be the true proportions of members of the two parties who are in favor of the proposal. Suppose it is desired to find a 95% confidence interval for estimating $p_1 - p_2$ to within 0.05. Estimate the minimum equal number of members of each party that must be sampled to meet these criteria.
12. A member of the state board of education wants to compare the proportions of National Board Certified (NBC) teachers in private high schools and in public high schools in the state. His study plan calls for an equal number of private school teachers and public school teachers to be included in the study. Let p_1 and p_2 be these proportions. Suppose it is desired to find a 99% confidence interval that estimates $p_1 - p_2$ to within 0.05.
 - a. Supposing that both proportions are known, from a prior study, to be approximately 0.15, compute the minimum common sample size needed.
 - b. Compute the minimum common sample size needed on the supposition that nothing is known about the values of p_1 and p_2 .

ANSWERS

1.
 - a. $n_1 = n_2 = 45$,
 - b. $n_1 = n_2 = 56$.
 - c. $n_1 = n_2 = 234$

3.
 - a. $n_1 = n_2 = 33$.
 - b. $n_1 = n_2 = 35$.
 - c. $n_1 = n_2 = 13$

5.
 - a.
 - a. $n_1 = n_2 = 329$,
 - b. $n_1 = n_2 = 255$.

 - b.
 - a. $n_1 = n_2 = 3383$,
 - b. $n_1 = n_2 = 2846$.

 - c.
 - a. $n_1 = n_2 = 193$,
 - b. $n_1 = n_2 = 128$

7. $n_1 = n_2 \approx 80$

9. $n_1 = n_2 \approx 34$

11. $n_1 = n_2 \approx 769$