CHAPTER 20

COMPARING TWO PROPORTIONS

OVERVIEW

Confidence intervals and tests designed to compare two population proportions are based on the **difference in the sample proportions** $\hat{p}_1 - \hat{p}_2$. The formula for the level C confidence interval is

$$\hat{p}_1 - \hat{p}_2 \pm z^* \, SE$$

where z^* is the critical value for the standard Normal density with area C between $-z^*$ and z^*, and SE is the standard error for the difference in the two proportions computed as

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

In practice, use this confidence interval when the populations are at least 10 successes and at least 10 failures in both samples, both of which are simple random samples from large populations.

To get a more accurate confidence interval, especially for smaller samples, add four imaginary observations - one success and one failure - in each sample. Then, with these new values for the number of failures and successes, use the previous formula for the approximate level C confidence interval. This is the **plus four confidence interval.** You can use it whenever both samples have five or more observations.

Significance tests for the equality of the two proportions, H_0: $p_1 = p_2$, use a different standard error for the difference in the sample proportions, which is based on a **pooled estimate** of the common (under H_0) value of p_1 and p_2,

$$\hat{p} = \frac{\text{count of successes in both samples combined}}{\text{count of observations in both samples combined}}$$

The test uses the z statistic

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}},$$

and P-values are computed using Table A of the standard normal distribution. In practice, use this test when the populations are at least 10 times as large as the samples and the counts of successes and failures are five or more in both samples.

GUIDED SOLUTIONS

Exercise 20.21

KEY CONCEPTS: Testing equality of two population proportions

First verify that it is safe to use the z test for equality of two proportions.

Let p_1 represent the proportion of papers *without* statistical assistance that were rejected without being reviewed in detail, and p_2 the proportion of papers *with* statistical help that were rejected without being reviewed in detail. Recall that a test of the hypothesis H_0: $p_1 = p_2$ uses the z statistic

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where n_1 and n_2 are the sizes of the samples, $\hat{p}_1$ and $\hat{p}_2$ are the estimates of p_1 and p_2, and

$$\hat{p} = \frac{\text{count of successes in both samples combined}}{\text{count of observations in both samples combined}}$$

First state the hypotheses to be tested. Is the alternative hypothesis one-sided or two-sided?

The two sample sizes are

$n_1 =$

$n_2 =$

From the data, the estimates of these two proportions are

$\hat{p}_1 =$

$\hat{p}_2 =$

Compute the pooled estimate of the value common to p_1 and p_2 under H_0:

$$\hat{p} = \frac{\text{count of successes in both samples combined}}{\text{count of observations in both samples combined}} =$$

Compute the test statistic:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} =$$

Compute the P-value:

P-value =

What do you conclude?

Exercise 20.23

KEY CONCEPTS: Large sample confidence interval for the difference of two population proportions

First determine whether the conditions for the large sample confidence interval are met or whether the plus four confidence interval needs to be used.

The two populations are proportions of papers rejected without review when a statistician *is* and *is not* involved in the research. The two sample sizes are

n_1 = number of papers rejected without review *without* a statistician involved =

n_2 = number of papers rejected without review *with* a statistician involved =

and the number of "successes" are

Number of papers in sample rejected without review *without* a statistician involved =

Number of papers in sample rejected without review *with* a statistician involved =

From the data, the estimates of the two proportions are

$\hat{p}_1 =$

$\hat{p}_2 =$

Let p_1 represent the proportion of all papers rejected without review *without* a statistician involved, and p_2 represent the proportion of all papers rejected without review *with* a statistician involved. Recall that a level C confidence interval for $p_1 - p_2$ is

$$(\hat{p}_1 - \hat{p}_2) \pm z^*SE$$

where z^* is the appropriate critical value for the standard Normal density, and SE is the standard error for the difference in the two proportions computed as

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

Use the values of $\hat{p}_1$ and $\hat{p}_2$ you computed to obtain the standard error:

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

For a 95% confidence interval,

$z^* =$

Compute the interval:

$(\hat{p}_1 - \hat{p}_2) \pm z^*SE$:

Exercise 20.27

KEY CONCEPTS: Testing equality of two population proportions; the four-step process

The four step process for testing hypotheses follows.

State. What is the practical question that requires a statistical test?

Plan. Identify the parameters, state null and alternative hypotheses, and choose the type of test that fits your situation.

Solve. Carry out the test in three phases:
(1) Check the conditions for the test you plan to use,
(2) Calculate the test statistic,
(3) Find the *P*-value.

Conclude. Return to the practical question to describe your results in this setting.

To apply the steps to this problem, here are some suggestions. You may want to use Examples 20.4 and 20.5 of the text as a guide.

State. Describe the problem of interest and the data obtained.

Plan. Are there two populations being compared in this problem? What are they?

Define the two proportions of interest.

Is the alternative hypothesis one-sided or two-sided?

Write the null and alternative hypotheses.

What kind of test will you use?

Solve. First check the conditions for using the test.

Write the two sample sizes.

$n_1 =$

$n_2 =$

From the data, the estimates of the two proportions are

$\hat{p}_1 =$

$\hat{p}_2 =$

Compute the pooled estimate of the value common to p_1 and p_2 under H_0:

$$\hat{p} = \frac{\text{count of successes in both samples combined}}{\text{count of observations in both samples combined}} =$$

Now compute the test statistic:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} =$$

Compute the *P*-value:

P-value =

Conclude. What do you conclude?

Exercise 20.33

KEY CONCEPTS: Plus four confidence intervals for the difference between two population proportions

The four-step process for a confidence interval follows.

State. What is the practical question that requires estimating a parameter?

Plan. Identify the parameters, choose a level of confidence, and select the type of confidence interval that fits your situation.

Solve. Carry out the work in two phases:
(1) Check the conditions for the interval you plan to use.
(2) Calculate the confidence interval.

Conclude. Return to the practical question to describe our results in this setting.

To apply the steps to this problem, here are some suggestions. You may want to use Example 20.3 of the text as a guide.

State. What are the populations of interest? What is it we want to estimate?

Plan. What are the parameters of interest in this problem?

What are we going to estimate using a confidence interval? Write this in terms of the parameters.

What confidence interval will you use? Remember that the Plus Four confidence interval is recommended over the traditional large-sample confidence interval. Write the formula for this interval:

Solve.

Are conditions required for use of the confidence interval method you selected in "Plan" (above) satisfied?

Compute the two proportion estimates.

$\hat{p}_1 =$

$\hat{p}_2 =$

Compute the standard error.

$SE =$

For a 90% confidence interval, $z^* = 1.645$, so our 90% confidence interval is

$(\hat{p}_1 - \hat{p}_2) \pm z^*SE$:

Conclude. What can you conclude about the difference in proportions between the two areas?

COMPLETE SOLUTIONS

Exercise 20.21

The count of successes and failures are each five or more in both samples, so the z test for equality of two population proportions can be used.

We are interested in determining whether there is good evidence that the proportion of papers rejected without review is *different* for papers with and without statistical help. Thus, the hypotheses to be tested are

$$H_0: p_1 = p_2$$

$$H_a: p_1 \neq p_2$$

Letting Population 1 be papers *without* statistical help and Population 2 be papers *with* statistical help, the two sample sizes and estimates of the proportions are

$n_1 = 190 \qquad \hat{p}_1 = 135/190 = 0.7105$

$n_2 = 514 \qquad \hat{p}_2 = 293/514 = 0.5700$

The pooled sample proportion is

$$\hat{p} = \frac{\text{count of successes in both samples combined}}{\text{count of observations in both samples combined}} = \frac{135+293}{190+514} = \frac{428}{704} = 0.6080$$

and the z statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.7105 - 0.5700}{\sqrt{0.6080(1-0.6080)\left(\frac{1}{190}+\frac{1}{514}\right)}} = \frac{0.1405}{0.0414} = 3.39$$

Finally, using Table A, we have

P-value = $2 \times P(z \geq 3.39) = (2)(1 - 0.9997) = 0.0006$

which is very strong evidence that a higher proportion of papers submitted without statistical help will be sent back without review. However, because this is an observational study, the evidence does not establish causation - getting statistical help for your paper may not make it less likely to be sent back without review.

Exercise 20.23

The counts of successes and failures are each 10 or more in both samples, so we may use the large-sample confidence interval.

Let Population 1 be papers *without* statistical help, and Population 2 be papers *with* statistical help. Let a success denote the event that a paper is sent back without review. Then for each population, sample sizes, number of successes and estimates of the proportions follow.

Population 1 $n_1 = 190$ number of successes = 135 $\hat{p}_1 = 135/190 = 0.7105$

Population 2 $n_2 = 514$ number of successes = 293 $\hat{p}_2 = 293/514 = 0.5700$

Using these values, the standard error is

$$\text{SE} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{0.7105(1-0.7105)}{190} + \frac{0.5700(1-0.5700)}{514}} = \sqrt{0.001559} = 0.0395$$

For a 95% confidence interval, $z^* = 1.96$. The 95% confidence interval is

$$(\hat{p}_1 - \hat{p}_2) \pm z^*SE = (0.7105 - 0.5700) \pm (1.96)(0.0395) = 0.1405 \pm 0.0774 = 0.0631 \text{ to } 0.2179$$

We are 95% confident that the percentage of papers without statistical help that are not reviewed is between 6.3 and 21.8 percentage points higher than the percentage of papers with statistical help.

Exercise 20.27

State. North Carolina University looked at factors that affected the success of students in a required chemical engineering course. Students must receive a C or better in the course to continue as chemical engineering majors, so we consider a grade of C or better as a success. Is there a difference in the proportions of male and female students who succeeded in the course? The data showed that 23 of the 34 women and 60 of the 89 men succeeded. We view these as SRSs of men and women who would take this course.

Plan. Let p_1 denote the proportion of female students who will succeed, and p_2 the proportion of males who will succeed. We are interested in determining whether there is a difference in these two proportions; hence we test the hypotheses

$$H_0: p_1 = p_2$$
$$H_a: p_1 \neq p_2$$

Solve. We can use the z test when the count of successes and failures are each 5 or more in both samples. For the females there were 23 successes and 34 – 23 = 11 failures, and for the males there were 60 successes and 89 – 60 = 29 failures. The conditions for safely using the test are met.

The two sample sizes are

n_1 = number of female students in the course = 34

n_2 = number of male students in the course = 89

From the data, the estimates of these two proportions are

$\hat{p}_1 = 23/34 = 0.6765$

$\hat{p}_2 = 60/89 = 0.6742$

We compute

$$\hat{p} = \frac{\text{count of successes in both samples combined}}{\text{count of observations in both samples combined}} = \frac{23+60}{34+89} = 83/123 = 0.6748$$

The value of the z-test statistic is thus

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.6765 - 0.6742}{\sqrt{0.6748(1-0.6748)\left(\frac{1}{34}+\frac{1}{89}\right)}} = \frac{0.0023}{\sqrt{0.00892}} = 0.02$$

We compute the P-value using Table A (we need to double the tail area because this is a two-sided test):

P-value = 2 × (0.4920) = 0.9840

Conclude. The data provide no evidence of a difference between the proportion of men and women who succeed.

Exercise 20.33

State. What is the size of the difference in the proportion of mice ready to breed in good acorn years and bad acorn years?

Plan. In a low-acorn year, experimenters added hundreds of thousands of acorns to Area 1 to simulate a good acorn year, while Area 2 was left untouched. They then trapped mice in both areas and counted the number of mice in breeding condition. The data follow.

Population	Population description	Sample size	Number of successes	Sample proportion
1	Area 1 (acorn rich)	$n_1 = 72$	54	$\hat{p}_1 = 54/72 = 0.75$
2	Area 2 (acorn poor)	$n_2 = 17$	10	$\hat{p}_2 = 10/17 = 0.59$

We want a 90% confidence interval for the difference of population proportions $p_1 - p_2$, where p_1 is the proportion of mice ready to breed in good acorn years, and p_2 is the proportion of mice ready to breed in bad acorn years. We'll use the plus four confidence interval procedure.

Solve. We can use the large-sample confidence interval when the populations are much larger than the samples and the counts of successes and failures are 10 or more in both samples. That isn't the case here. Anyway, the plus four confidence interval procedure is more accurate, and it is appropriate when each sample has size at least 5. That procedure is appropriate here. To apply the plus four confidence interval procedure, we add four imaginary observations. The new data summary follows.

Population	Population description	Sample size	Number of successes	Plus four sample proportion
1	Area 1 (acorn rich)	$n_1 + 2 = 74$	$54 + 1 = 55$	$\hat{p}_1 = 55/74 = 0.74$
2	Area 2 (acorn poor)	$n_2 + 2 = 19$	$10 + 1 = 11$	$\hat{p}_2 = 11/19 = 0.58$

The standard error is

$$SE = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{0.74(1-0.74)}{74} + \frac{0.58(1-0.58)}{19}} = \sqrt{0.0026 + 0.0128} = 0.12$$

and for a 90% confidence interval $z^* = 1.645$, so the 90% confidence interval is

$$(\hat{p}_1 - \hat{p}_2) \pm z^*\text{SE} = (0.74 - 0.58) \pm 1.645(0.12) = 0.16 \pm 0.20 \text{ or } -0.04 \text{ to } 0.36.$$

Conclude. We are 90% confident that mice in Area 1 (with abundant crop) are between –4% and 36% more likely to be in breeding condition than those in Area 2, which was left untouched. The confidence interval is quite wide and there isn't much evidence of a difference in the breeding condition of mice in the two areas in either direction.