9.3 - Confidence Intervals for the Difference Between Two Population Proportions or Means

When a sample survey produces a proportion or a mean as a response, we can use the methods in section 9.1 and section 9.2 to find a confidence interval for the true population values. In this section, we discuss confidence intervals for comparative studies. How do we assess the difference between two proportions or means when they come from a comparative observational study or experiment? To address this question, we first need a new rule.

Standard Error of a Difference

When two samples are independent of each other,

Standard Error for a Difference between two sample summaries =

$$\sqrt{(\text{standard error in first sample})^2 + (\text{standard error in second sample})^2}$$

Example 9.6

A medical researcher conjectures that smoking can result in the wrinkled skin around the eyes. The researcher recruited **150 smokers** and **250 nonsmokers** to take part in an observational study and found that **95** of the **smokers** and **105** of the **nonsmokers** were seen to have prominent wrinkles around the eyes (based on a standardized wrinkle score administered by a person who did not know if the subject smoked or not). Some results from the study are found in **Table 9.2**.

Table 9.2. *Results of the Smoking and wrinkles study (example 9.6)*

Smokers	**Nonsmokers**	
Sample Size	150	250
Sample Proportion with Prominent Wrinkles	95/150 = 0.63	105/250 = 0.42
Standard Error for Proportion	$\sqrt{\frac{0.63(0.37)}{150}} = 0.0394$	$\sqrt{\frac{0.42(0.58)}{250}} = 0.0312$

How do the smokers compare to the non-smokers? The difference between the two sample proportions is 0.63 - 0.42 = 0.21. We would like to make a CI for the true difference that would exist between these two groups in the population. So we compute

$$\text{Standard Error for Difference} = \sqrt{0.0394^2 + 0.0312^2} \approx 0.05$$

If we think about all possible ways to draw a sample of 150 smokers and 250 non-smokers then the differences we'd see between sample proportions would approximately follow the normal curve. Thus, a 95% Confidence Interval for the differences between these two proportions in the population is given by:

$$\text{Difference Between the Sample Proportions} \pm z^*(\text{Standard Error for Difference})$$

or

$$0.21 \pm 2(0.05) \;\; \text{or} \;\; 0.21 \pm 0.1$$

Notice that this 95% confidence interval goes from 0.11 to 0.31. Since the interval does not contain 0, we see that the difference seen in this study was "significant."

Another way to think about whether the smokers and non-smokers have significantly different proportions with wrinkles is to calculate a 95% Confidence Interval for each group separately. For the smokers, we have a confidence interval of 0.63 ± 2(0.0394) or 0.63 ± 0.0788. The interval for smokers goes from about 0.55 up to 0.71. For the non-smokers, we have a confidence interval of 0.42 ± 2(0.0312) or 0.42 ± 0.0624. The interval for non-smokers goes from about 0.36 up to 0.48. The interval for the smokers (which starts at 0.55) and the interval for the non-smokers (which ends at 0.48) do not overlap - that is another sign that the differences seen in this study were "significant."

Statistical Significance and Confidence Intervals

- **If the two confidence intervals do not overlap**, we can conclude that there is a statistically significant difference in the two population values at the given level of confidence; or alternatively
- **If the confidence interval for the difference does not contain zero**, we can conclude that there is a statistically significant difference in the two population values at the given level of confidence.

The first rule is the "more conservative" one since there are some circumstances when the interval for the difference does not contain zero but there is some overlap in the individual confidence intervals.

Importantly, the formula for the standard deviation of a difference is for two *independent* samples. It would not apply to dependent samples like those gathered in a matched pairs study.

Lesson

1: Statistics: Benefits, Risks, and Measurements

2: Characteristics of Good Sample Surveys and Comparative Studies

3: Getting the Big Picture and Summaries

4: Bell-Shaped Curves and Statistical Pictures

5: Relationships Between Measurement Variables

6: Relationships Between Categorical Variables

7: Understanding Uncertainty

Example 9.7

A general rule used clinically to judge normal levels of strength is that a person's dominant hand should have about 10% higher grip strength than their non-dominant hand. The idea is that the preferential use of your dominant hand in everyday activities might act as a form of endurance training for the muscles of the hand resulting in the strength differential. If this theory about the underlying reason for the strength differential is true then there should be less of a difference in young children than in adults. Data from a study of 60 right-handed boys under 10 years old and 60 right-handed men aged 30-39 are shown in **Table 9.3**.

Table 9.3 *Grip Strength (kilograms) Average and Standard Deviation by Hand and Age*

	Boys < 10 years old (n=60)	**Men 30-39 years old (n=60)**
Right Hand	$\bar{x}$ = 6.2 kg s = 2.1 kg	$\bar{x}$ =40.3 kg s = 9.3 kg
Left Hand	$\bar{x}$ =5.9 kg s = 2.2 kg	$\bar{x}$ =35.6 kg s = 8.8 kg
Difference	$\bar{x}$ =0.3 kg s = 0.8 kg	$\bar{x}$ =4.7 kg s = 3.6 kg

8: The Diversity of Samples
9: Confidence Intervals
9.1 - Confidence Intervals for a Population Proportion
9.2 - Confidence Intervals for a Population Mean
9.3 - Confidence Intervals for the Difference Between Two Population Proportions or Means
9.4 - Test Yourself!
9.5 - Have Fun With It!
10: Hypothesis Testing
11: Significance Testing Caveats & Ethics of Experiments

Resources

References
Reviews