Chapter 9: Comparing Two Groups

Section 9.1: Categorical Response: How Can We Compare Two Proportions?

Learning Objectives

1. Bivariate Analyses
2. Independent Samples and Dependent Samples
3. Categorical Response Variable
4. Example
5. Standard Error for Comparing Two Proportions
6. Confidence Interval for the Difference Between Two Population Proportions
7. Interpreting a Confidence Interval for a Difference of Proportions

Learning Objectives

9. Significance Tests Comparing Population Proportions
10. Examples
11. Class Exercises

Learning Objective 1: Bivariate Analyses

Methods for comparing two groups are special cases of bivariate statistical methods: there are two variables

- **The outcome variable on which comparisons are made is the *response variable***
- **The binary variable that specifies the groups is the *explanatory variable***

- Statistical methods analyze how the outcome on the response variable *depends on* or is *explained by* the value of the explanatory variable

Learning Objective 2: Independent Samples

Most comparisons of groups use independent samples from the groups:

- **The observations in one sample are *independent* of those in the other sample**
  - **Example: An observational study that separates subjects into groups according to their value for an explanatory variable**

Learning Objective 2: Dependent Samples

- **Dependent samples result when the data are *matched pairs* – each subject in one sample is matched with a subject in the other sample**
  - **Example: set of married couples, the men being in one sample and the women in the other.**

Learning Objective 3: Categorical Response Variable

For a categorical response variable

- **Inferences compare groups in terms of their population proportions in a particular *category***
- **We can compare the groups by the difference in their population proportions:**

  $(p_1 - p_2)$

7

Learning Objective 4: Example: Aspirin, the Wonder Drug

- **Experiment:**
  - **Subjects were 22,071 male physicians**
  - **Every other day for five years, study participants took either an aspirin or a placebo**
  - **The physicians were randomly assigned to the aspirin or to the placebo group**
  - **The study was double-blind: the physicians did not know which pill they were taking, nor did those who evaluated the results**

8

Learning Objective 4: Example: Aspirin, the Wonder Drug

Results displayed in a contingency table:

TABLE 10.1: Whether Subject Had a Heart Attack, for Placebo and Aspirin Treatment Groups

	Heart Attack		
Group	Yes	No	Total
Placebo	189	10,845	11,034
Aspirin	104	10,933	11,037

9

Learning Objective 4: Example: Aspirin, the Wonder Drug

- **What is the response variable?**
  - **The response variable is whether the subject had a heart attack, with categories 'yes' or 'no'**
- **What are the groups to compare?**
  - **The groups to compare are:**
    - **Group 1: Physicians who took a placebo**
    - **Group 2: Physicians who took aspirin**

10

Learning Objective 4: Example: Aspirin, the Wonder Drug

- **Estimate the difference between the two population parameters of interest**
  - **p_1: the proportion of the population who would have a heart attack if they participated in this experiment and took the *placebo***
  - **p_2: the proportion of the population who would have a heart attack if they participated in this experiment and took the *aspirin***

11

Learning Objective 4: Example: Aspirin, the Wonder Drug

Sample Statistics:

$$\hat{p}_1 = 189/11034 = 0.017$$
$$\hat{p}_2 = 104/11037 = 0.009$$
$$(\hat{p}_1 - \hat{p}_2) = 0.017 - 0.009 = 0.008$$

12

Learning Objective 4:
Example: Aspirin, the Wonder Drug

- **To make an inference about the difference of population proportions, ($p_1 - p_2$), we need to learn about the variability of the sampling distribution of:** $(\hat{p}_1 - \hat{p}_2)$

13

Learning Objective 5:
Standard Error for Comparing Two Proportions

- **The difference, $(\hat{p}_1 - \hat{p}_2)$, is obtained from sample data**
- **It will vary from sample to sample**
- **This variation is the standard error of the sampling distribution of $(\hat{p}_1 - \hat{p}_2)$:**

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

14

Learning Objective 6:
Confidence Interval for the Difference Between Two Population Proportions

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

- **The z-score depends on the confidence level**
- **This method requires:**
  - **Categorical response variable for two groups**
  - **Independent random samples for the two groups**
  - **Large enough sample sizes so that there are at least 10 “successes” and at least 10 “failures” in each group**

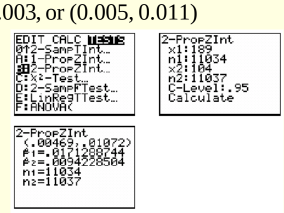
15

Learning Objective 6:
Confidence Interval Comparing Heart Attack Rates for Aspirin and Placebo

- **95% CI**:

$$(.017-.009) \pm 1.96\sqrt{\frac{.017(1-.017)}{11034} + \frac{.009(1-.009)}{11037}} =$$

$$0.008 \pm 0.003, \text{ or } (0.005, 0.011)$$



16

Learning Objective 6:
Confidence Interval Comparing Heart Attack Rates for Aspirin and Placebo

- **Since both endpoints of the confidence interval (0.005, 0.011) for (p_1- p_2) are positive, we infer that (p_1- p_2) is positive**
- **Conclusion: The population proportion of heart attacks is *larger* when subjects take the placebo than when they take aspirin**

17

Learning Objective 6:
Confidence Interval Comparing Heart Attack Rates for Aspirin and Placebo

- **The population difference (0.005, 0.011) is small**
- **Even though it is a small difference, it may be important in public health terms**
- **For example, a decrease of 0.01 over a 5 year period in the proportion of people suffering heart attacks would mean 2 million fewer people having heart attacks**

18

Learning Objective 6: Confidence Interval Comparing Heart Attack Rates for Aspirin and Placebo

- The study used male doctors in the U.S
  - The inference applies to the U.S. population of male doctors
- Before concluding that aspirin benefits a larger population, we'd want to see results of studies with more diverse groups

19

Learning Objective 7: Interpreting a Confidence Interval for a Difference of Proportions

- Check whether 0 falls in the CI
- If so, it is plausible that the population proportions are equal
- If all values in the CI for (p_1- p_2) are positive, you can infer that (p_1- p_2) >0
- If all values in the CI for (p_1- p_2) are negative, you can infer that (p_1- p_2) <0
- Which group is labeled '1' and which is labeled '2' is arbitrary

20

Learning Objective 7: Interpreting a Confidence Interval for a Difference of Proportions

- The magnitude of values in the confidence interval tells you how large any true difference is
- If all values in the confidence interval are near 0, the true difference may be relatively small in practical terms

21

Learning Objective 8: Significance Tests Comparing Population Proportions

1. Assumptions:

- Categorical response variable for two groups
- Independent random samples

22

Learning Objective 8: Significance Tests Comparing Population Proportions

Assumptions (continued):

- Significance tests comparing proportions use the sample size guideline from confidence intervals: Each sample should have at least about 10 "successes" and 10 "failures"
- Two–sided tests are robust against violations of this condition
  - At least 5 "successes" and 5 "failures" is adequate

23

Learning Objective 8: Significance Tests Comparing Population Proportions

2. Hypotheses:

- The null hypothesis is the hypothesis of *no difference* or *no effect*:

H_0: $p_1=p_2$

The alternative hypothesis is the hypothesis of interest to the investigator

H_a: $p_1 \neq p_2$ (two-sided test)
H_a: $p_1<p_2$ (one-sided test)
H_a: $p_1>p_2$ (one-sided test)

24

Learning Objective 8: Significance Tests Comparing Population Proportions

Pooled Estimate

- **Under the presumption that $p_1 = p_2$, we estimate the common value of p_1 and p_2 by the proportion of the *total* sample in the category of interest**
- **This pooled estimate is calculated by combining the number of successes in the two groups and dividing by the combined sample size ($n_1 + n_2$)**

25

Learning Objective 8: Significance Tests Comparing Population Proportions

3. The test statistic is:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $\hat{p}$ is the pooled estimate

26

Learning Objective 8: Significance Tests Comparing Population Proportions

4. P-value: Probability obtained from the standard normal table of values even more extreme than observed *z* test statistic

5. Conclusion: Smaller P-values give stronger evidence against H_0 and supporting H_a

27

Learning Objective 9: Example: Is TV Watching Associated with Aggressive Behavior?

- **Various studies have examined a link between TV violence and aggressive behavior by those who watch a lot of TV**
- **A study sampled 707 families in two counties in New York state and made follow-up observations over 17 years**
- **The data shows levels of TV watching along with incidents of aggressive acts**

28

Learning Objective 9: Example: Is TV Watching Associated with Aggressive Behavior?

TABLE 10.3: TV Watching by Teenagers and Later Aggressive Acts

	Aggressive Act		
TV Watching	Yes	No	Total
Less than 1 hour per day	5	83	88
At least 1 hour per day	154	465	619

29

Learning Objective 9: Example: Is TV Watching Associated with Aggressive Behavior?

- Define Group 1 as those who watched less than 1 hour of TV per day, on the average, as teenagers
- Define Group 2 as those who averaged at least 1 hour of TV per day, as teenagers
- p_1 = population proportion committing aggressive acts for the lower level of TV watching
- p_2 = population proportion committing aggressive acts for the higher level of TV watching

30

Learning Objective 9:
Example: Is TV Watching Associated with Aggressive Behavior?

- Test the Hypotheses:

  H_0: $(p_1 - p_2) = 0$

  H_a: $(p_1 - p_2) \neq 0$

  using a significance level of 0.05
- Test statistic:

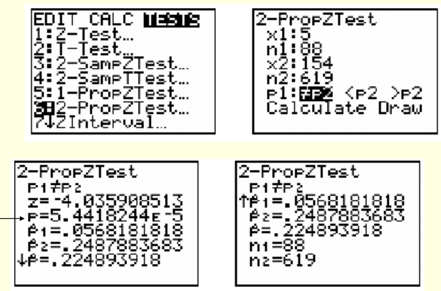
$$\hat{p} = \frac{5+154}{88+619} = 0.225$$

$$se_0 = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{0.225(0.775)\left(\frac{1}{88}+\frac{1}{619}\right)} = 0.0476$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{se_0} = \frac{0.057 - 0.249}{0.0476} = \frac{-0.192}{0.0476} - 4.04$$

31

Learning Objective 9:
Example: Is TV Watching Associated with Aggressive Behavior?



32

Learning Objective 9:
Example: Is TV Watching Associated with Aggressive Behavior?

- **Conclusion: Since the P-value is less than 0.05, we reject H_0**
- **We conclude that the population proportions of aggressive acts differ for the two groups**
- **The sample values suggest that the population proportion is higher for the higher level of TV watching**

33

Learning Objective 9:
Test of Significance: Two Proportions
Summer Jobs Example

- A university financial aid office polled a simple random sample of undergraduate students to study their summer employment.
- Not all students were employed the previous summer. Here are the results:

Summer Status	Men	Women
Employed	718	593
Not Employed	79	139
Total	797	732

- Is there evidence that the proportion of male students who had summer jobs differs from the proportion of female students who had summer jobs?

34

Learning Objective 9:
Test of Significance: Two Proportions
Summer Jobs Example

Hypotheses:

- Null: The proportion of male students who had summer jobs is the same as the proportion of female students who had summer jobs. [H_0: $p_1 = p_2$]
- Alt: The proportion of male students who had summer jobs differs from the proportion of female students who had summer jobs. [H_a: $p_1 \neq p_2$]

35

Learning Objective 9:
Test of Significance: Two Proportions
Summer Jobs Example

Test Statistic:

- n_1 = 797 and n_2 = 732
  (both large, so test statistic follows a Normal distribution)
- Pooled sample proportion:

$$\hat{p} = \frac{718+593}{797+732} = \frac{1311}{1529}$$

- Test statistic:

$$z = \frac{\frac{718}{797} - \frac{593}{732}}{\sqrt{\frac{1311}{1529}\left(1-\frac{1311}{1529}\right)\left(\frac{1}{797}+\frac{1}{732}\right)}} = 5.07$$

36

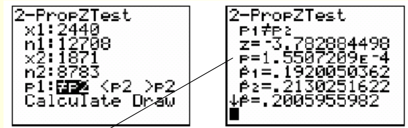
Learning Objective 9:
Test of Significance: Two Proportions
Summer Jobs Example

- ***Hypotheses:*** $H_0: p_1 = p_2$
  $H_a: p_1 \neq p_2$
- ***Test Statistic:***
  $z = 5.07$
- ***P-value:***
  P-value = $2P(Z > 5.07) = 0.000000396$ *(using a computer)*
- ***Conclusion:***
  Since the *P*-value is quite small, there is very strong evidence that the proportion of male students who had summer jobs differs from that of female students.

37

Learning Objective 9:
Test of Significance: Two Proportions
Drinking and unplanned sex

- In a study of binge drinking, the percent who said they had engaged in unplanned sex because of drinking was 19.2% out of 12708 in 1993 and 21.3% out of 8783 in 2001
- Is this change statistically significant at the 0.05 significance level?



The P-value is 0.0002 < .05. The results are statistically significant. But are they practically significant?

38

Learning Objective 10:
Test of Significance: Two Proportions
Class Exercise 1

- A survey of one hundred male and one hundred female high school seniors showed that thirty-five percent of the males and twenty-nine percent of the females had used marijuana previously. Does this survey indicate a difference in proportions for the population of high school seniors? Test at α=5%,

39

Learning Objective 10:
Test of Significance: Two Proportions
Class Exercise 2

- A random sample of 500 persons were questioned regarding political affiliation and attitude toward government sponsored mandatory testing of AIDS. The results were as follows:

	favor	Undecided	Opposed	Total
Dem	135	80	65	200
Rep	95	60	65	220
Total	230	140	130	

Is there a difference in the proportions of Democrats and Republicans who are undecided regarding mandatory testing for AIDS? Test at α=5%

40