Chapter 22 Summary
Comparing Two Proportions

What have we learned?

- We've now looked at the difference in two proportions.
- Perhaps the most important thing to remember is that the concepts and interpretations are essentially the same—only the mechanics have changed slightly.
- Hypothesis tests and confidence intervals for the difference in two proportions are based on Normal models.
    - ○ Both require us to find the standard error of the difference in two proportions.
        - ▪ We do that by adding the variances of the two sample proportions, assuming our two groups are independent.
        - ▪ When we test a hypothesis that the two proportions are equal, we pool the sample data; for confidence intervals we don't pool.

Comparing Two Proportions

- Comparisons between two percentages are much more common than questions about isolated percentages. And they are more interesting.
- We often want to know how two groups differ, whether a treatment is better than a placebo control, or whether this year's results are better than last year's.

Another Ruler

- In order to examine the difference between two proportions, we need another ruler—the standard deviation of the sampling distribution model for the difference between two proportions.
- Recall that standard deviations don't add, but variances do. In fact, the variance of the sum or difference of two independent random quantities is the sum of their individual variances.

The Standard Deviation of the Difference Between Two Proportions

- Proportions observed in independent random samples *are* independent. Thus, we can add their variances. So…
- The standard deviation of the difference between two sample proportions is $SD(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$
- Thus, the standard error is $SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$

Assumptions and Conditions

- Independence Assumptions:
    - ○ Randomization Condition: The data in each group should be drawn independently and at random from a homogeneous population or generated by a randomized comparative experiment.
    - ○ The 10% Condition: If the data are sampled without replacement, the sample should not exceed 10% of the population.
    - ○ Independent Groups Assumption: The two groups we're comparing must be independent *of each other*.

Assumptions and Conditions (cont.)

- Sample Size Condition:
  - *Each* of the groups must be big enough…
  - Success/Failure Condition: Both groups are big enough that at least 10 successes and at least 10 failures have been observed in each.

The Sampling Distribution

- We already know that for large enough samples, each of our proportions has an approximately Normal sampling distribution.
- The same is true of their difference.
- Provided that the sampled values are independent, the samples are independent, and the samples sizes are large enough, the sampling distribution of $\hat{p}_1 - \hat{p}_2$ is modeled by a Normal model with

  Mean: $\mu = p_1 - p_2$ Standard deviation: $SD\left(\hat{p}_1 - \hat{p}_2\right) = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$

Two-Proportion *z*-Interval

- When the conditions are met, we are ready to find the confidence interval for the difference of two proportions:
- The confidence interval is $\left(\hat{p}_1 - \hat{p}_2\right) \pm z^* \times SE\left(\hat{p}_1 - \hat{p}_2\right)$

  where

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

- The critical value *z** depends on the particular confidence level, *C*, that you specify.

Everyone into the Pool

- The typical hypothesis test for the difference in two proportions is the one of no difference. In symbols, H_0: $p_1 - p_2 = 0$.
- Since we are hypothesizing that there is no difference between the two proportions, that means that the standard deviations for each proportion are the same.
- Since this is the case, we combine (pool) the counts to get one overall proportion.
- The pooled proportion is $\hat{p}_{pooled} = \dfrac{Success_1 + Success_2}{n_1 + n_2}$

  where $Success_1 = n_1 \hat{p}_1$ and $Success_2 = n_2 \hat{p}_2$
  - If the numbers of successes are not whole numbers, round them first. (This is the *only* time you should round values in the middle of a calculation.)
- We then put this pooled value into the formula, substituting it for *both* sample proportions in the standard error formula:

  $$SE_{pooled}\left(\hat{p}_1 - \hat{p}_2\right) = \sqrt{\frac{\hat{p}_{pooled} \hat{q}_{pooled}}{n_1} + \frac{\hat{p}_{pooled} \hat{q}_{pooled}}{n_2}}$$

Compared to What?

- We'll reject our null hypothesis if we see a large enough difference in the two proportions.
- How can we decide whether the difference we see is large?
  - Just compare it with its standard deviation.
- Unlike previous hypothesis testing situations, the null hypothesis doesn't provide a standard deviation, so we'll use a standard error (here, pooled).

Two-Proportion *z*-Test

- The conditions for the two-proportion *z*-test are the same as for the two-proportion *z*-interval.
- We are testing the hypothesis H_0: $p_1 = p_2$.
- Because we hypothesize that the proportions are equal, we pool them to find

$$\hat{p}_{pooled} = \frac{Success_1 + Success_2}{n_1 + n_2}$$

- We use the pooled value to estimate the standard error:

$$SE_{pooled}\left(\hat{p}_1 - \hat{p}_2\right) = \sqrt{\frac{\hat{p}_{pooled}\hat{q}_{pooled}}{n_1} + \frac{\hat{p}_{pooled}\hat{q}_{pooled}}{n_2}}$$

- Now we find the test statistic: $z = \dfrac{\hat{p}_1 - \hat{p}_2}{SE_{pooled}\left(\hat{p}_1 - \hat{p}_2\right)}$
- When the conditions are met and the null hypothesis is true, this statistic follows the standard Normal model, so we can use that model to obtain a P-value.

What Can Go Wrong?

- Don’t use two-sample proportion methods when the samples aren’t independent.
    - These methods give wrong answers when the independence assumption is violated.
- Don’t apply inference methods when there was no randomization.
    - Our data must come from representative random samples or from a properly randomized experiment.
- Don’t interpret a significant difference in proportions causally.