The Big Picture:
Where Chapter 10 Fits

Ch. 8
Confidence
Intervals

Ch. 9
Significance
Tests

The Big Picture:
Where Chapter 10 Fits

Ch. 8 Confidence Intervals

The Big Picture: Where Chapter 10 Fits

Ch. 9 Significance Tests

for a difference in proportions or diff. in means

Ch. 8 Confidence Intervals

The Big Picture: Where Chapter 10 Fits

Ch. 9 Significance Tests

for a difference in proportions or diff. in means

We use these procedures when comparing
- two independent rand. samples
- tw groups in a randomized experiment.

Chapter 10: The Big Ideas

1. Most studies are comparative, which requires that we investigate the difference between two samples, two groups in an experiment, or paired data.

2. Inference for the difference in two proportions or the difference in two means is based on the sampling distributions of these differences. Inference for a mean difference uses paired *t* procedures.

3. The logic of inference is the same as it was in Chapter 8 (confidence intervals) and Chapter 9 (significance tests), although the details differ somewhat.

mostly

4. The calculations we perform when doing inference for experiments are the same as when doing inference for random samples.

PACING ~~9~~ 8 **days**

Chapter 10: Comparing Two Populations or Groups

10.1 Comparing Two Proportions	~~2~~ 3 Days
10.2 Comparing Two Means	2 Days
10.3 Comparing Two Means: Paired Data	2 Days
Review, FRAPPY, and Test	2 Days
	8 days

Next Test – Wed. Feb. 12th

Targets today

Already did in Ch. 7

DETERMINE whether the conditions are met for doing inference about a difference between two proportions.

CONSTRUCT and INTERPRET a confidence interval for a difference between two proportions.

As usual ...

- Start by diving in
- We'll formalize later

Lesson 10.1 Day 2: Which grade is more likely to go to Prom?

At many high schools, Prom is an annual dance that only Juniors and Seniors can purchase tickets for. The student council at a large high school is wondering if Juniors or Seniors are more likely to attend Prom. They take a random sample of 50 Juniors and find that 28 are planning on attending Prom. They select a random sample of 45 Seniors and 29 are planning on attending. Construct and interpret a 95% confidence interval for the difference in proportions of Juniors and Seniors who are planning on attending Prom.

1. What is the **point estimate** for…

   the proportion of Juniors planning on attending prom? $\widehat{p_1}$ = ________

   the proportion of Seniors planning on attending prom? $\widehat{p_2}$ = ________

   the difference in the proportion of Jrs and Srs planning on attending prom? $\widehat{p_1} - \widehat{p_2}$ = ________

CLASS OF 2020 **Lesson 10.1 Day 2: Which grade is more likely to go to Prom?** CLASS OF 2021

At many high schools, Prom is an annual dance that only Juniors and Seniors can purchase tickets for. The student council at a large high school is wondering if Juniors or Seniors are more likely to attend Prom. They take a random sample of 50 Juniors and find that 28 are planning on attending Prom. They select a random sample of 45 Seniors and 29 are planning on attending. Construct and interpret a 95% confidence interval for the difference in proportions of Juniors and Seniors who are planning on attending Prom.

1. What is the **point estimate** for… $28/50 = .56$

   the proportion of Juniors planning on attending prom? $\widehat{p_1}$ = .56

   the proportion of Seniors planning on attending prom? $\widehat{p_2}$ = .64 $29/45 = .64$

   the difference in the proportion of Jrs and Srs planning on attending prom? $\widehat{p_1} - \widehat{p_2}$ = −0.08

2. Check the conditions needed in order to construct a confidence interval.

   Random:

   10%:

   Large Counts:

2. Check the conditions needed in order to construct a confidence interval.

Random: RS of 50 JRS ✓
RS of 45 SRS

10%: assume $50 < \frac{1}{10}$(all JRS at school) ✓
$45 < \frac{1}{10}$(all SRS at school

Large Counts:

JRS $50(.56) = 28$
$50(.44) = 22$ ≥ 10

SRS $45(.64) = \cancel{29}$ 28.8
$45(.36) = \cancel{21}$ 16.2 ≥ 10 ✓

Let's analyze the conditions

2. Check the conditions needed in order to construct a confidence interval.

Random: RS of 50 JRS
RS of 45 SRS

Independ. rand. samp. so we can generalize to the population

10%: assume $50 < \frac{1}{10}$(all JRS at school)
$45 < \frac{1}{10}$(all SRS at school

Large Counts:

JRS $50(.56) = 28$ ≥ 10
$50(.44) = 22$

SRS $45(.64) = 29$ ≥ 10
$45(.36) = 21$

2. Check the conditions needed in order to construct a confidence interval.

Random: RS of 50 JRS
RS of 45 SRS

Independ. rand. samp. so we can generalize to the population

10%: assume $50 < \frac{1}{10}$(all JRS at school)
$45 < \frac{1}{10}$(all SRS at school

So we can sample without replacement

Large Counts:

JRS $50(.56) = 28$ ≥ 10
$50(.44) = 22$

SRS $45(.64) = 29$ ≥ 10
$45(.36) = 21$

2. Check the conditions needed in order to construct a confidence interval.

Random: RS of 50 JRS
RS of 45 SRS

Independ. rand. samp.
so we can generalize
to the population

10%: assume $50 < \frac{1}{10}$(all JRS at school)
$45 < \frac{1}{10}$(all SRS at school

so we can sample
without replacement

Large Counts:

JRS
$50(.56) = 28$
$50(.44) = 22$
≥ 10

SRS
$45(.64) = 29$
$45(.36) = 21$
≥ 10

So the sampling distrib.
of $\hat{p}_1 - \hat{p}_2$ is
approx. normal

3. Construct and interpret a 95% confidence interval for the difference in proportions of Juniors and Seniors who are planning on attending prom.

General Formula:

Specific Formula:

Work:

Conclude:

3. Construct and interpret a 95% confidence interval for the difference in proportions of Juniors and Seniors who are planning on attending prom.

General Formula: Point Estimate $\pm$ margin of error

Specific Formula: $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Work:

Conclude:

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

in significance tests $\longrightarrow$ When $p_1 = p_2$ is assumed:

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

where $\hat{p}_c = \frac{X_1 + X_2}{n_1 + n_2}$

3. Construct and interpret a 95% confidence interval for the difference in proportions of Juniors and Seniors who are planning on attending prom.

General Formula: Point Estimate ± margin of error

Specific Formula: $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Work: $-.08 \pm 1.96 \sqrt{\frac{.56(.44)}{50} + \frac{.64(.36)}{45}}$

Conclude:

3. Construct and interpret a 95% confidence interval for the difference in proportions of Juniors and Seniors who are planning on attending prom.

General Formula: Point Estimate ± margin of error

Specific Formula: $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Work: $-.08 \pm 1.96 \sqrt{\frac{.56(.44)}{50} + \frac{.64(.36)}{45}} \rightarrow -.08 \pm .20 \rightarrow (-.28, .12)$

Conclude:

3. Construct and interpret a 95% confidence interval for the difference in proportions of Juniors and Seniors who are planning on attending prom.

General Formula: Point Estimate $\pm$ margin of error

Specific Formula: $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p}_1)}{n_1} + \frac{p_2(1-\hat{p}_2)}{n_2}}$

Work: $-.08 \pm 1.96\sqrt{\frac{.56(.44)}{50} + \frac{.64(.36)}{45}} \rightarrow -.08 \pm .20 \rightarrow (-.28, .12)$

Conclude:

We are

3. Construct and interpret a 95% confidence interval for the difference in proportions of Juniors and Seniors who are planning on attending prom.

General Formula: Point Estimate $\pm$ margin of error

Specific Formula: $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p}_1)}{n_1} + \frac{p_2(1-\hat{p}_2)}{n_2}}$

Work: $-.08 \pm 1.96\sqrt{\frac{.56(.44)}{50} + \frac{.64(.36)}{45}} \rightarrow -.08 \pm .20 \rightarrow (-.28, .12)$

Conclude:

We are 95% confident that the interval from -.28 to .12 (JRS - SENIORS) captures the true difference in proportion of juniors and seniors going to the prom

4. Does the interval provide convincing evidence that Juniors have a lower proportion planning on going to prom or is it plausible that there is no difference between the two classes? Explain.

4. Does the interval provide convincing evidence that Juniors have a lower proportion planning on going to prom or is it plausible that there is no difference between the two classes? Explain.

Because the interval contains 0
it is plausible that there is
no difference.

4. Does the interval provide convincing evidence that Juniors have a lower proportion planning on going to prom or is it plausible that there is no difference between the two classes? Explain.

Because the interval contains 0
it is plausible that there is
no difference.

We do not have convincing evidence
there is a difference.

4. Does the interval provide convincing evidence that Juniors have a lower proportion planning on going to prom or is it plausible that there is no difference between the two classes? Explain.

Because the interval contains 0
it is plausible that there is
no difference.

We do not have convincing evidence
there is a difference.

(-.28, .12)

if (+, +) →
if (−, −) →
if (−, +) →

4. Does the interval provide convincing evidence that Juniors have a lower proportion planning on going to prom or is it plausible that there is no difference between the two classes? Explain.

Because the interval contains 0 it is plausible that there is no difference.

We do not have convincing evidence there is a difference.

(-.28, .12)

if (+, +) → higher prop. of Juniors

if (−, −) →

if (−, +) →

4. Does the interval provide convincing evidence that Juniors have a lower proportion planning on going to prom or is it plausible that there is no difference between the two classes? Explain.

Because the interval contains 0 it is plausible that there is no difference.

We do not have convincing evidence there is a difference.

(-.28, .12)

if (+, +) → higher prop. of Juniors

if (−, −) → " " " SRS

if (−, +) →

4. Does the interval provide convincing evidence that Juniors have a lower proportion planning on going to prom or is it plausible that there is no difference between the two classes? Explain.

Because the interval contains 0 it is plausible that there is no difference.

We do not have convincing evidence there is a difference.

(-.28, .12)

if (+, +) → higher prop. of Juniors
if (-, -) → " " " SRS
if (-, +) → could be no difference

Constructing a Confidence Interval for $p_1 - p_2$

for a difference of two proportions

Important ideas:

State:

Plan

Do

Conclude:

Constructing a Confidence Interval for $p_1 - p_2$

for a difference of two proportions

Important ideas:

State: $p_1 - p_2$ true diff in proportions

Plan: Two-sample z interval for $p_1 - p_2$

1. independ. random samples
2. 10% condition
3. Large Counts

Do

Conclude:

Constructing a Confidence Interval for $p_1 - p_2$

for a difference of two proportions

Important ideas:

State: $p_1 - p_2$ true diff in proportions

Plan: Two-sample z interval for $p_1 - p_2$

1. independ. random samples
2. 10% condition
3. Large Counts

Do

Pt Est ± MOE → $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Conclude: We are ...

Constructing a Confidence Interval for $p_1 - p_2$

for a difference of two proportions

Important ideas:

State: $p_1 - p_2$ true diff in proportions

Plan: Two-Sample z interval for $p_1 - p_2$
1. independ. random samples
2. 10% condition
3. Large Counts

Do:
Pt Est ± MOE → $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Conclude: We are ...

Convincing Evidence?

$(+, +)$ p_1 greater

$(-, -)$ p_2 greater

$(-, +)$ Not convincing evidence of diff (interval contains 0)

CAUTION:
Never suggest that you believe the difference between the true proportions is 0 just because 0 is in the interval!

Check Your Understanding

A Pew Research Center poll asked independent random samples of working women and men how much they value job security. Of the 806 women, 709 said job security was very or extremely important, compared with 802 of the 944 men surveyed. Construct and interpret a 95% confidence interval for the difference in the proportion of all working women and men who consider job security very or extremely important.

skip the plan step

STATE: Parameter: Confidence level:

PLAN:

Name of procedure:

Check conditions:

DO:

General Formula:

Specific Formula:

Work:

Answer:

CONCLUDE:

Check Your Understand: A Pew Research Center poll asked independent random samples of working women and men how much they value job security. Of the 806 women, 709 said job security was very or extremely important, compared with 802 of the 944 men surveyed. Construct and interpret a 95% confidence interval for the difference in the proportion of all working women and men who consider job security very or extremely important.

STATE

PLAN

DO

CONCLUDE

Because of short day

Skip

STATE $p_1 - p_2 \rightarrow$ true difference in the proportion of working women and men who consider job security very important.

95% confidence

$\hat{p}_1 = \frac{709}{806} = .88$

$\hat{p}_2 = \frac{802}{944} = .85$

PLAN Two sample z interval for $p_1 - p_2$

Random : "R.S. of women and men".

10% : $806 < \frac{1}{10}$(all working women)
$944 < \frac{1}{10}$(all working men)

Large Counts: $806(.88) = 709$
$806(.88) = 97$
$944(.85) = 802$
$944(.15) = 142$
≥ 10

DO Pt. Estim ± MOE

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

$$0.88 - .85 \pm 1.96 \sqrt{\frac{.88(.12)}{806} + \frac{.85(.15)}{944}}$$

$(-.0019, .0621)$

DO Pt. Estim ± MOE

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

$$0.88 - .85 \pm 1.96 \sqrt{\frac{.88(.12)}{806} + \frac{.85(.15)}{944}}$$

$\rightarrow (-.0019, .0621)$

CONCLUDE

We are 95% confident that the interval from $-.0019$ to $.062$ captures the true diff (women-men) in proportions of working men and women.

$\boxed{\text{DO}}$ Pt. Estim ± MOE

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

$$0.88 - .85 \pm 1.96 \sqrt{\frac{.88(.12)}{806} + \frac{.85(.15)}{444}}$$

$$\rightarrow (-.0019, .0621)$$

2-PropZInt

Daren Starnes recommends you start using Technology on FR QUESTIONS Now that you have STRUCTURE SOLIDIFIED but encourages you to still write general specific, etc

$\boxed{\text{CONCLUDE}}$

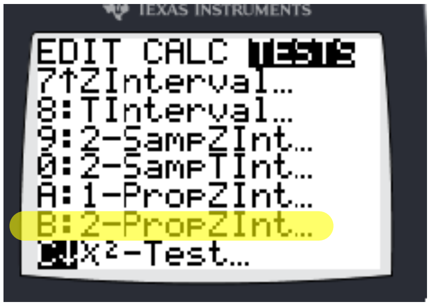
We are 95% confident that the interval from −.0019 to .062 captures the true diff (women − men) in proportions of working men and women.

The TI-83/84 can be used to construct a confidence interval for $p_1 - p_2$.

```
NORMAL FLOAT AUTO REAL RADIAN MP
         2-PropZInt
 x1:39
 n1:57
 x2:43
 n2:55
 C-Level:0.99
 Calculate
```

```
NORMAL FLOAT AUTO REAL RADIAN MP
         2-PropZInt
 (-0.3114,0.11623)
 p̂1=0.6842105263
 p̂2=0.7818181818
 n1=57
 n2=55
```

STAT
↓
TESTS



Heads Up!

2-Samp Z Int is for comparing means when the population standard deviation is known (not in this case)

So far, we have focused on doing inference using data that were produced by *random sampling*.

The following example shows how to construct and interpret a confidence interval for a difference in proportions from a ***randomized comparative experiment.***

Construct and interpret a confidence interval for a difference in proportions from a *randomized comparative experiment.*

BACK PAIN: Patients with lower back pain are often given nonsteroidal anti-inflammatory drugs (NSAIDs) like naproxen to help ease their pain. Researchers wondered if taking Valium along with the naproxen would affect pain relief. To find out, they recruited 112 patients with severe lower back pain and randomly assigned them to one of two treatments: naproxen and Valium or naproxen and placebo. After 1 week, 39 of the 57 subjects who took naproxen and Valium reported reduced lower back pain, compared with 43 of the 55 subjects in the naproxen and placebo group.

(a) Construct and interpret a 99% confidence interval for the difference in the proportion of patients like these who would report reduced lower back pain after taking naproxen and Valium versus after taking naproxen and placebo for a week.

State

plan ← Is the 10% condition even required? Don't worry about other conditions

For now

$\boxed{\text{State}}$

You can use the wording of the question to def your parameters and to write the conclusion

(called "parroting" the stem of the question)

STATE 99% C.I. for $p_1 - p_2$

where p_1 = true prop. of patients who would report reduced lower back pain taking naxproxen & Valium

p_2 = true propor. of patients who would report reduced lower back pain after taking naxproxen and a placebo

PLAN TWO-sample Z Interval for $p_1 - p_2$

✓ Random : Randomly assigned patients to take naproxen and Valium or Naproxen and a placebo.

✓ Large Counts : 39 and 57−39 = 18 43 and 55−43 = 12 all ≥ 10

10% condition not needed since researchers did not sample patients without replacement from a larger population

Construct and interpret a confidence interval for a difference in proportions from a *randomized comparative experiment.*

BACK PAIN: Patients with lower back pain are often given nonsteroidal anti-inflammatory drugs (NSAIDs) like naproxen to help ease their pain. Researchers wondered if taking Valium along with the naproxen would affect pain relief. To find out, they recruited 112 patients with severe lower back pain and randomly assigned them to one of two treatments: naproxen and Valium or naproxen and placebo. After 1 week, 39 of the 57 subjects who took naproxen and Valium reported reduced lower back pain, compared with 43 of the 55 subjects in the naproxen and placebo group.

(a) Construct and interpret a 99% confidence interval for the difference in the proportion of patients like these who would report reduced lower back pain after taking naproxen and Valium versus after taking naproxen and placebo for a week.

State

plan ← Is the 10% condition even required? Don't worry about other conditions

Do ← do quickly w/technology

Conclude

$\boxed{\text{DO}}$

$\hat{p}_1 = \frac{39}{57} = .684 \quad \hat{p}_2 = \frac{43}{55} = .782$

Point Estim $\pm$ MOE

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

$$(.684 - .782) \pm 2.576\sqrt{\frac{.684(.316)}{57} + \frac{.782(.218)}{55}}$$

$$-0.098 \pm .214$$

$$(-.312, 0.116)$$

$\boxed{\text{CONCLUDE}}$

We are 99% confident that the interval from $-.312$ to $.116$ captures $p_1 - p_2$ = true difference in the proportions of patients who would report reduce pain after taking naproxen and valium versus after taking naproxen and a placebo.

(b) Based on the confidence interval in part (a), what conclusion would you make about whether taking Valium along with naproxen affects pain relief? Justify your answer.

Because the interval includes 0 as a possible value for $p_1 - p_2$, we don't have convincing evidence that taking Valium along with naproxen affects pain relief for patients like these.

10.1....4, 7, 9, 11, 13, and 35, 36

study pp. 625-630