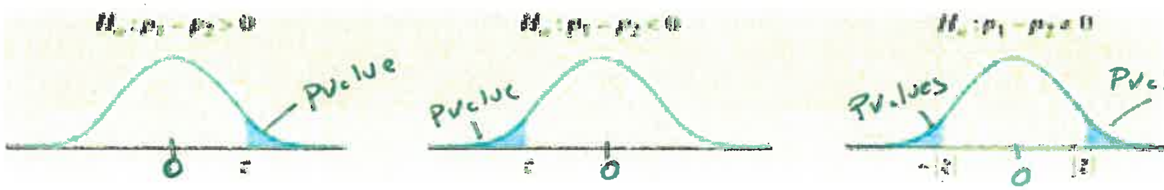
AP Statistics – **10.1B**	Name: 2021 KEY
Goal: 2 Sample TOH for proportions	Date:

I. Formulas You Need to Know

TOH must Define p_1 + p_2

$H_0: p_1 = p_2$

Two-Sample z Test for the Difference Between Proportions



$H_a: p_1 > p_2$ $H_a: p_1 < p_2$ $H_a: p_1 \neq p_2$

Tip: define p_1 with the largest %.

STEP 1: Calculate Pooled Sample Proportion

Since we do not know p_1 or p_2, we must estimate it

$$\hat{p}_c = \frac{\text{count of successes in both samples combined}}{\text{count of individuals in both samples combined}} = \frac{X_1 + X_2}{n_1 + n_2}$$

STEP 2: Calculate Z-Test Statistic for $p_1 - p_2$

Formula for Hypothesis test:

$$\text{Test statistic} = \frac{\text{statistic} - \text{parameter}}{\text{SD of statistic}}$$

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - \cancel{(p_1 - p_2)}}{\sqrt{\hat{p}(1-\hat{p}) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$p_1 = p_2$

So

$p_1 - p_2 = 0$

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

only use TOH

$\hat{p}_c$ – pooled sample proportion

GREEN SHEET

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

Use for C.I.

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

Use for H.T. and p and q are the pooled Phat

Resource for more help: Watch Ben Bishop 10.1 video #3 https://youtu.be/0wfdlHCNXdE

II. 1-Tail Test of Hypothesis

Example #1: ***Hearing loss***

Are teenagers going deaf? In a study of 3000 randomly selected teenagers in 1988-1994, 15% showed some hearing loss. In a similar study of 1800 teenagers in 2005-2006, 19.5% showed some hearing loss. (Source: Arizona Daily Star, 8-18-2010). Does these data give convincing evidence that the proportion of all teens with hearing loss has increased?

HEARING LOSS

p_1 → 2005-06 vs 1988-94 ← p_2

19.5% 15%

p_1 = true proportion of all teens with hearing loss on 2005-06

p_2 = " " " " " " " " 1988-94

H_0: $p_1 = p_2$ OR $p_1 - p_2 = 0$

H_A: $p_1 > p_2$ OR $p_1 - p_2 > 0$

[has hearing loss increased?]

2005 > 1988

$\alpha = 0.05$

Name the test:

2 sample Z-test for difference of proportions

CONDITIONS

Random: 2 RANDOM samples were taken

Independent: The years do not over lap, therefore the 2 groups are independent

$n_1 = 1{,}800 \le \frac{1}{10}$ (all teens in 2005-06)

$n_2 = 3{,}000 \le \frac{1}{10}$ (all teens in '88-94)

Normal: Check both samples w/ $\hat{p}_c$ →

2005-06

$.167(1800) = 301 \ge 10$

$.833(1800) = 1499 \ge 10$

1988-94

$.167(3000) = 501 \ge 10$

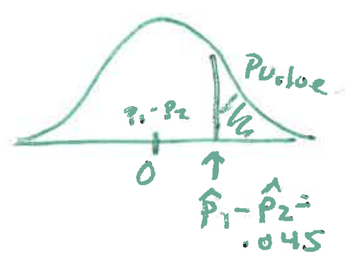
$.833(3000) = 2499 \ge 10$

$\hat{p}_1 = .195$

$\hat{p}_2 = .15$

$$\hat{p}_c = \frac{351+450}{1800+3000} = \frac{801}{4800} = 0.167$$

$\hat{p}_1 - \hat{p}_2 = .195 - 0.15 = 0.045$



Use calc

$$Z = \frac{(.195 - 0.15) - \cancel{(p_1 - p_2)}^{\,0}}{\sqrt{(.167)(.833)\cdot\left(\frac{1}{1800} + \frac{1}{3000}\right)}} = 4.05$$

Pvalue = $P(Z \ge 4.05) \approx 0$

Conclusion

Since the pvalue (≈ 0) is less than $\alpha = 0.05$, we reject H_0. We have convincing evidence that the proportion of all teens with hearing loss has increased from 1988-94 to 2005-06.

III. Test of Hypothesis With Treatments **Example #2:** *Cash for quitters*

Resource for more help: Watch Ben Bishop 10.1 video #4 https://youtu.be/bRoQwpOC5J4

In an effort to reduce health care costs, General Motors sponsored a study to help employees stop smoking. In the study, half of the subjects were randomly assigned to receive up to $750 for quitting smoking for a year while the other half were simply encouraged to use traditional methods to stop smoking. None of the 878 volunteers knew that there was a financial incentive when they signed up. At the end of one year, 15% of those in the financial rewards group had quit smoking while only 5% in the traditional group had quit smoking. Do the results of this study give convincing evidence that a financial incentive helps people quit smoking? (Source: Arizona Daily Star, 2-11-09).

TOH: CASH incentive (p_1) > Traditional (p_2)

p_1 = true proportion of employees QUIT smoking given cash incentive

p_2 = true proportion of employees QUIT smoking with NO CASH

$H_0: p_1 = p_2$ $\alpha = .05$

$H_A: p_1 > p_2$

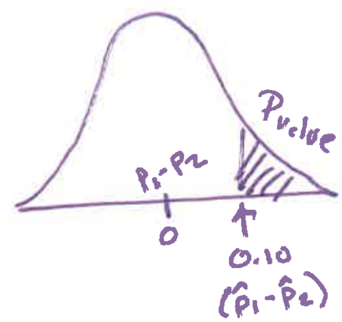
Name: 2 sample z test for difference of proportions

$\hat{p}_1 = 0.15$ $n_1 = 439$

$\hat{p}_2 = 0.05$ $n_2 = 439$

$\hat{p}_1 - \hat{p}_2 = 0.10$

$\hat{p}_c = \dfrac{66+22}{439+439} = 0.100$



$Z = 4.94$

$\text{Pvalue} = P(Z \geq 4.94) \cong 0$

Conditions

Random: The treatments were randomly assigned

Independent: Since the treatments were randomly assigned, we can say the groups are independent

Random: (use $\hat{p}_c$)
w/CASH → $0.1(439) = 44 \geq 10$ ✓ $0.9(439) = 395 \geq 10$
No CASH → $0.1(439) = 44 \geq 10$ ✓ $0.9(439) = 395 \geq 10$

Conclusion: Since the pvalue (~0) is less than $\alpha = 0.05$ we reject H_0.

We have convincing evidence that the financial incentive for GM employees helped them quit smoking compared to the group that did not get the financial incentive.

IV. 2-Tail Test of Hypothesis

Example #3a: ***Hungry Children***

Researchers designed a survey to compare the proportions of children who come to school without eating breakfast in two low-income elementary schools. An SRS of 80 students from School 1 found that 19 had not eaten breakfast. At School 2, an SRS of 150 students included 26 who had not had breakfast. More than 1500 students attend each school. Do these data give convincing evidence of a difference in the population proportions?

a) Use a significance test to determine if there is convincing evidence and explain your conclusion.

Parameters

p_1 = true proportion of students that don't eat lunch at School 1.

p_2 = true proportion of students that don't eat lunch at school 2.

$H_0: p_1 = p_2$

$H_a: p_1 \neq p_2$

$\alpha = 0.05$

Test: 2 sample Z test for the difference of proportions

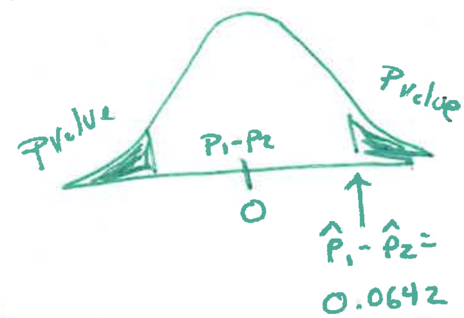
School 1

$n_1 = 80 \quad \hat{p}_1 = 19/80 = 0.2375$

School 2

$n_2 = 150 \quad \hat{p}_2 = 26/150 = 0.1733$

$\hat{p}_1 - \hat{p}_2 = 0.0642$



Conditions

Random: Random samples taken at both schools

Independent: The 2 schools are independent

$n_1 = 80 \leq \frac{1}{10}$ (School 1 total population)

$n_2 = 150 \leq \frac{1}{10}$ (School 2 total population)

$\hat{p}_c = \frac{19+26}{80+150} = 0.196$

Random: School 1 = $0.196(80) = 16 \geq 10$ ✓ $\quad 0.804(80) = 64 \geq 10$ ✓

School 2 = $0.196(150) = 29 \geq 10$ ✓ $\quad 0.804(150) = 121 \geq 10$ ✓

Calculations

$$Z = \frac{(0.2375 - 0.1733) - 0}{\sqrt{(0.196)(0.804)\left(\frac{1}{80} + \frac{1}{150}\right)}}$$

$\boxed{Z = 1.17}$

Pvalue = $2 \cdot P(Z \geq 1.17) = 0.242$

TI 84

[2PropZtest]

$x_1 = 19$

$n_1 = 80$

$x_2 = 26$

$n_2 = 150$

$\neq p_2$

IV (CONT) 2 TAIL TEST OF HYPOTHESIS FOR

Hungry Children

Conclusion

Since the pvalue (0.242) is greater than $\alpha = 0.05$, we fail to reject H_0. We do NOT have sufficient evidence to conclude there is a difference between the 2 schools for the proportion of their students that did NOT EAT BREAKFAST.

V. **2-Tail CI for Inference** **Example #3b:** ***Hungry Children***

b) Use a confidence interval to determine if there is convincing evidence and explain your conclusion.

Parameters

p_1 = true proportion of kids that come to school that don't eat breakfast at school 1

p_2 = " " do not eat breakfast at school 2

TEST: 2 sample z interval for difference of proportions

Conditions

Random: 2 random samples taken

Independent: the 2 schools are independent

school 1 $n_1 = 80 \le \frac{1}{10}$(all students @ school 1)

school 2 $n_2 = 150 \le \frac{1}{10}$(all students @ school 2)

Random: school 1 $(\hat{p} + \hat{q}) \rightarrow n\hat{p} = 19 \ge 10$ ✓

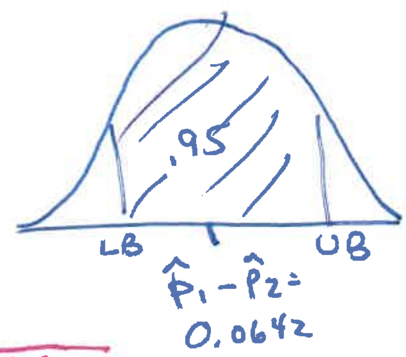
school 2 $(\hat{p} + \hat{q}) \rightarrow n\hat{q} = 26 \ge 10$ ✓

Sample data

$n_1 = 80$ $\hat{p}_1 = 19/80 = 0.2375$

$n_2 = 150$ $\hat{p}_2 = 26/150 = 0.1733$

$\hat{p}_1 - \hat{p}_2 = 0.2375 - 0.1733 = 0.0642$



Calculate 95% CI:

$$0.2375 - 0.1733 \pm 1.96\sqrt{\frac{(.24)(.76)}{80} + \frac{(.17)(.83)}{150}}$$

$$0.0642 \pm 1.96(0.057)$$ SE

$$0.0642 \pm 0.1172$$ ME

$[-0.048, 0.176]$

Conclusion:

Since our 95% confidence interval (−0.05, 0.18) does include "0", we fail to reject Ho at $\alpha = 0.05$.

Therefore, we do not have convincing evidence to conclude there is a difference between School 1 and School 2 true proportion of children who do not eat breakfast before coming to school.