AP Statistics – **10.1B**	Name: UPDATED KEY 2016
Goal: 2 Sample TOH for proportions	Date:

I. **Formulas You Need to Know**

① TOH

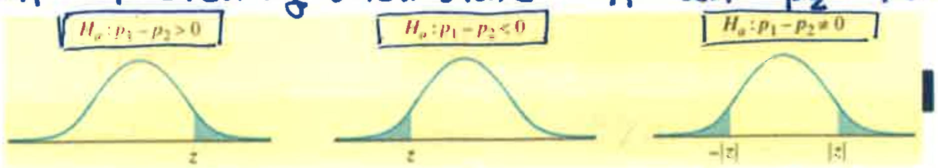
$H_0: p_1 = p_2$

$H_0: p_1 - p_2 = 0$

H_0 = No Difference

Two-Sample z Test for the Difference Between Proportions

② H_A: Tip clearly understand "p_1" an "p_2" To make easy to interpret



$H_A: p_1 > p_2$ $\quad$ $H_A: p_1 < p_2$ $\quad$ $H_A: p_1 \neq p_2$

Recommend Start with H_A.

STEP 1: Calculate Pooled Sample Proportion

③ **Calculate pooled (or combined) sample proportion.**

YOU MUST MEMORIZE THIS!

Pooled Sample Proportion

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

$$\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2}$$

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

④

Formula for Hypothesis test:

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

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

$p_1 = p_2$

So . . .

$p_1 - p_2 = 0$

USE Green Sheet to Set up the Z test statistic

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

Calculator 2-PropZtest

$\hat{p}$ means $\hat{p}_c$

Greensheet For TOH, use 2 sample Special case S.D. because $H_0: p_1 = p_2$

Important!

For TOH, The SE ($\hat{p}_c$)

Helpful Tips for 2-Sample Proportions Formula

MAKE SURE TO UNDERSTAND THESE CONCEPTS

Z statistic Formula

~~Use $\hat{p}_C$ in place of both p_1 and p_2 in the expression for the denominator of the test statistic :~~

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

← USE S.D. FROM GREEN SHEET

GREEN SHEET

Two-Sample

Statistic	Standard Deviation of Stastistic
Difference of Sample Proportions	$\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$ Special case when $p_1 = p_2$ $\sqrt{p(1-p)}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

Use for C.I.

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

Test Tip: write p= $\frac{X1+X2}{n1+n2}$

Important! Know your green sheet

remember for ToH

P is $\hat{P}_C$

Memorize

10-18

BEFORE DEFINING P_1 an P_2 Read Problem To make P_1 and P_2 EASY TO INTERPRET

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?

ToH: HEARING LOSS 2005-06 > 1988-94

Define as P_1 Define as P_2

III. Test of Hypothesis With Treatments

Example #2: ***Cash for quitters***

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 > TRADITIONAL METHOD

Define P_1 Define P_2

IV. 2-Tail Test of Hypothesis

Example #3: ***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.
b) Use a confidence interval to determine if there is convincing evidence and explain your conclusion.

School 1: $\hat{p} = \frac{19}{80} = .2375$

School 2: $\hat{p} = \frac{26}{150} = .1733$

$\boxed{P_1}$

$\boxed{P_2}$

Tip

- Define P_1 with the higher proportion to interpret positive differences.

10.18

EXAMPLE #1 HEARING LOSS

$\boxed{\text{1 TAIL TOH}}$

Given Info

$\hat{p}_1 = .195$
$n_1 = 1800$
$x_1 = .195(1800) = 351$

$\hat{p}_2 = .15$
$n_2 = 300$
$x_2 = .15(300) = 450$

① DEFINE PARAMETERS:

p_1 = proportion of all teens WITH HEARING LOSS 2005-06
p_2 = proportion of all teens with hearing loss 1988-94

$H_0: p_1 - p_2 = 0$ (the proportion remains the same)
$H_A: p_1 - p_2 > 0$ (proportion has improved over time)
(Think: 2005 > 1985)

② NAME TEST AND CHECK CONDITIONS

TEST – 2 SAMPLE Z TEST FOR $p_1 - p_2$ ($\alpha = .05$)

CONDITIONS:

① Random: BOTH SAMPLES WERE RANDOMLY SELECTED

② Independent:

Both must be stated
* The samples were taken independently
* There are more than 10(1800) = 18,000 teens in 2005-06
There are more than 10(3,000) = 30,000 teens in 1988-94.

③ NORMAL: all counts are at least 10

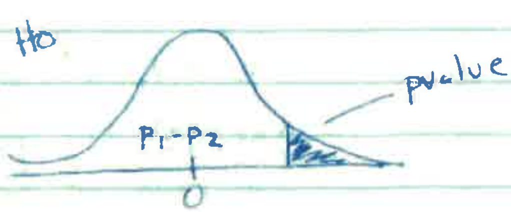
$n_1\hat{p}_1 = (.195)(1800) = 351 \geq 10$
$n_1\hat{q}_1 = 1800 - 351 = 1449 \geq 10$
$n_2\hat{p}_2 = (.15)(3000) = 450 \geq 10$
$n_2\hat{q}_2 = (.85)(3000) = 2550 \geq 10$

IMPORTANT NOTE!
- For 1 Sample tests, we use (p or μ) To check Normal Condition.
- For "2 Sample tests," we must use $\hat{p}$ since we do NOT KNOW p_1 OR p_2.

③ MECHANICS:

MUST MEMORIZE →

$$\text{Pooled } \hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2} = \frac{351 + 450}{1800 + 3000} = \frac{801}{4800} = .167$$



10-1B

EXAMPLE #1 (Continued):

TEST STATISTIC - See GREEN SHEET

$$\text{STANDARDIZED TEST STATISTIC} = \frac{\text{STATISTIC} - \text{PARAMETER}}{\text{SD OF STATISTIC}}$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - \cancel{(p_1 - p_2)}^{\,0}}{\sqrt{p(1-p)} \cdot \sqrt{1/n_1 + 1/n_2}}$$

USE POOLED $\hat{p}_c$

$\hat{p}_1 = .195$ $n_1 = 1800$
$\hat{p}_2 = .15$ $n_2 = 3{,}000$
$\hat{p}_c = .167$

$$z = \frac{.195 - .15}{\sqrt{(.167)(.833)} \cdot \sqrt{\frac{1}{1800} + \frac{1}{3000}}} = 4.05$$

Pvalue

$P(z > 4.05) = .000026 \approx 0$

(4) CONCLUSION

Since the pvalue is so small and less than $\alpha = .05$, we reject H_0 in favor of H_A.
We have convincing evidence that the proportion of all teens with hearing loss has increased from 1988 - 1994 to 2005 - 2006

CALCULATOR Command:

[STAT] [TESTS] 6: 2-PropZTest

$x1 = 351$
$n_1 = 1800$
$x2 = 450$
$n_2 = 3000$
$p1: > p2$

$p_1 > p_2$
$z = 4.05$
$p = .000025$
$\hat{p}_1 = .195$
$\hat{p}_2 = .15$
$\hat{p} = .166875$ pooled $\hat{p}_c$

EXAMPLE #2 CASH FOR QUITTERS

$\boxed{\text{TOH WITH TREATMENTS}}$

① **Hypothesis** TOH CASH INCENTIVE (p_1) > TRADITIONAL (p_2)

p_1 = true QUITTING RATE FOR EMPLOYEES WHO GET A FINANCIAL INCENTIVE

p_2 = TRUE QUITTING RATE FOR EMPLOYEES WHO DON'T GET A FINACIAL INCENTIVE

$H_0: p_1 - p_2 = 0$

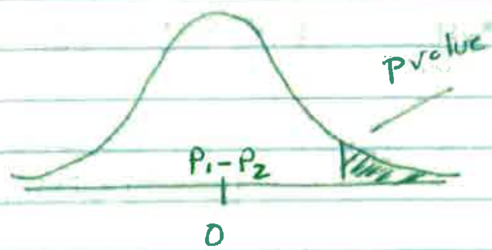
$H_A: p_1 - p_2 > 0$

OR

$H_0: p_1 = p_2$ → "NO DIFFER."

$H_A: p_1 > p_2$ → "CASH > TRAD."

Ho



$p_1 = .15$ $n_1 = 878/2 = 439$

$p_2 = .05$ $n_2 = 439$

$x_1 = (.15)(439)$

$x_1 = 66$

$x_2 = (.05)(439)$

$x_2 = 22$

TEST AND CONDITIONS

② Test: 2 sample ZTest for $p_1 - p_2$ ($\alpha = .05$)

CONDITIONS: "THIS IS AN EXPERIMENT!"

① Random: THE TREATMENTS WERE RANDOMLY ASSIGNED

② Independent: THE RANDOM ASSIGNMENT ALLOWS US TO VIEW THESE 2 GROUPS AS INDEPENDENT. WE MUST ASSUME EACH EMPLOYEE'S DECISION TO QUIT IS INDEPENDENT OF THE OTHER'S DECISIONS

③ Normal: ALL COUNTS ARE AT LEAST 10

$.15(439) = 66 \geq 10$ ✓ $.05(439) = 22 \geq 10$

$.85(439) = 373 \geq 10$ ✓ $.95(439) = 417 \geq 10$

NOTICE: I do not write np or $n\hat{p}$ to avoid making a mistake ($n\hat{p}$ is correct). Showing the numbers is safe and REQUIRED.

10-13

EXAMPLE #2 (CONTINUED):

3 MECHANICS:

MUST GIVE → $\hat{P}_c = \frac{66+22}{439+439} = .100$

$x_1 = 66$
$x_2 = 22$
$n_1 = n_2 = 439$

Recommend – FILL IN FORMULA + USE CALC FOR Z

$$Z = \frac{(.15 - .05) - 0}{\sqrt{(.1)(.9)}\sqrt{\frac{1}{439} + \frac{1}{439}}} = 4.94$$

$\hat{p}_1 = .15$
$\hat{p}_2 = .05$
$\hat{p}_c = .100$

MUST GIVE → $Z = 4.94$

MUST GIVE → $\text{pValue} = P(Z \geq 4.94) = 3.9E^{-7} \approx 0$

CALC Command
[STAT] [TESTS] 6: 2-PropZTest

$x_1 = 66$
$n_1 = 439$
$x_2 = 22$
$n_2 = 439$
$p1: > p2$

$p_1 > p_2$
$z = 4.94$
$p = 3.8E^{-7}$
$\hat{p}_1 = .15$
$\hat{p}_2 = .05$
$\hat{p} = .100$ (pooled $\hat{p}$)

4 Conclusion:

Since the P-Value is less than our .05 significance level, we reject H_0.

We have convincing evidence that the financial incentives help employees like these quit smoking.

10-18

#3A — TOH

Tip: School 1 has a higher $\hat{p}$ (.2375) than School 2 $\hat{p}$ (.1733).

Example: Hungry Children

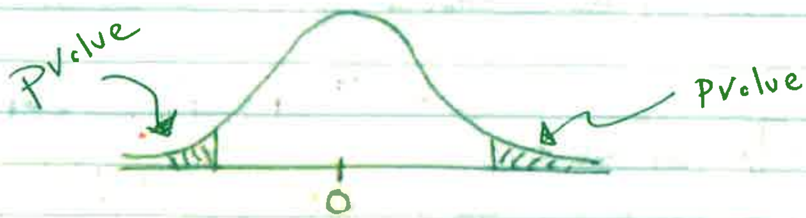
School 1: $n_1 = 80$ $\quad x_1 = 19$ ← No Breakfast $\qquad \alpha = .05$

School 2: $n_2 = 150$ $\quad x_2 = 26$ $\qquad$ Are proportions different?

Parameters: p_1 = true proportion that don't eat breakfast at School 1
p_2 = true proportion that don't eat breakfast at School 2

Hypothesis:
$H_0: p_1 - p_2 = 0$ (NO DIFFERENCE) $\quad$ OR $\quad H_0: p_1 = p_2$
$H_A: p_1 - p_2 \neq 0$ (DIFFERENT) $\quad\quad\quad\quad H_A: p_1 \neq p_2$

Sketch Graph!



NAME TEST: 2 Sample Z Test for $p_1 - p_2$ ($\alpha = .05$)

Conditions

Random: The students from both schools were selected at random

Normal: The successes (19, 26) and failures (61, 124) are both Greater than 10.

STATE EITHER WAY

School 1 $\quad n_1\hat{p}_1 = 19 \geq 10$ ✓ $\quad n_1\hat{q}_1 = \frac{61}{80} \geq 10$ ✓

School 2 $\quad n_2\hat{p}_2 = 26 \geq 10$ ✓ $\quad n_2\hat{q}_2 = \frac{124}{150} \geq 10$ ✓

Independent:

① The 2 schools are independent

② It is reasonable that the samples are less than 10% of the stated population (1,500 students).

School 1 = 80(10) = 800 ≤ 1,500 ✓ $\quad$ School 2 = 150(10) = 1,500 ≤ 1,500 ✓

10-1B

#3A (continued)

CALCULATIONS:

NAME TEST: 2 SAMPLE Z TEST FOR $p_1 - p_2$

Give Sample Stats:

$\hat{p}_1 = .2375 \quad n_1 = 80$

$\hat{p}_2 = .173 \quad n_2 = 150$

MUST SHOW CALC FOR POOLED $\hat{p}_c$!!

$$\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2} = \frac{19+26}{80+150} = .1957$$

USE GREEN SHEET TO HELP CALCULATE TEST STATISTIC:

$$\text{STANDARDIZED TEST STATISTIC} = \frac{\text{STATISTIC} - \text{PARAMETER}}{\text{SD OF STATISTIC}}$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (\overset{0}{\cancel{p_1 - p_2}})}{\sqrt{p(1-p)} \cdot \sqrt{1/n_1 + 1/n_2}}$$

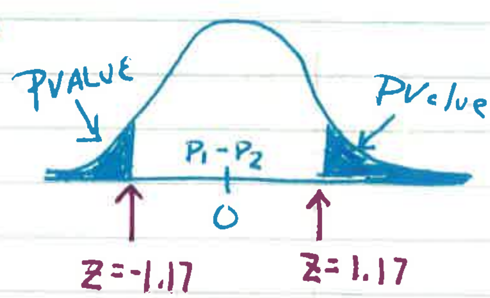
$p_1 = p_2$

TEST STATISTIC FOR PROPORTIONS IS ALWAYS "Z".

Tip: label Green Sheet SD {1st CI, 2nd TOH}

UNDERSTAND HOW TO USE GREEN SHEET BUT YOU DO NOT NEED TO WRITE THIS FORMULA.

DON'T FORGET TO GRAPH:



TEST STATISTIC (fill in):

$$z = \frac{.2375 - .1733}{\sqrt{(.1957)(.8043)} \cdot \sqrt{1/80 + 1/150}}$$

pooled $\hat{p}$

$$\boxed{z = 1.17}$$

check calc

FIRST Take advantage of CALC!

Write info on the side

STAT TESTS

6: 2PropZTest

$x_1 = 19$

$n_1 = 80$

$x_2 = 26$

$n_2 = 150$

$\neq p2$

CALC

$p_1 \neq p_2$

$z = 1.168$ ✓

$p = .2427$

$\hat{p}_1 = .2375$

$\hat{p}_2 = .173$

$\hat{p} = .1957$

10-1B

#3A (Continued) "Hungry Children"

PVALUE:

$P(Z \leq -1.17)$ OR $P(Z \geq 1.17) = \boxed{.242}$

check w/calc

OR $2 \cdot P(Z \geq 1.17) = 2(.121) = \boxed{.242}$

CALC command:
Normalcdf
(1.17, E99, 0, 1)

Conclusion (4 parts required):

Since the pvalue (2.42) is greater than $\alpha = .05$, WE FAIL TO REJECT H_0. THERE IS NOT SUFFICIENT EVIDENCE TO CONCLUDE THAT THERE IS A DIFFERENCE BETWEEN THE 2 SCHOOLS ON THE PROPORTION OF STUDENTS THAT DID NOT EAT BREAKFAST.

10-4B

EXAMPLE #3B   CI for Hungry Children

Parameters are the same:

p_1 = true proportion that don't eat breakfast at School 1

p_2 = " " " don't eat breakfast at School 2

TEST: 2 sample Z-interval for $p_1 - p_2$ (95% CL)

Conditions: same

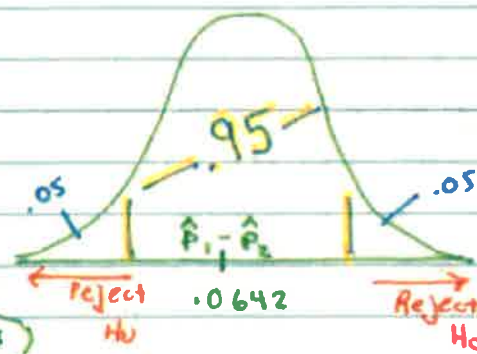
Sampling Distribution:

$\hat{p}_1 = 19/80 = .2375$ $n_1 = 80$

$\hat{p}_2 = 26/150 = .1733$ $n_2 = 150$

$\mu_{\hat{p}_1 - \hat{p}_2} = .2375 - .1733 = .0642$

$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{(.24)(.76)}{80} + \frac{(.17)(.83)}{150}} = .057$



95% CI

$$\hat{p}_1 - \hat{p}_2 \pm z^* \sqrt{\frac{(.24)(.76)}{80} + \frac{(.17)(.83)}{150}}$$

CALC:
STAT
TESTS
B: 2PropZInt
x1 = 19
n1 = 80
x2 = 26
n2 = 150
C-level = .95
(-.047, .1754)

$.0642 \pm 1.96\,(.057)$

$.0642 \pm .11172$ ← ME

$(-.04752,\ .17592)$

Conclusion:

Since our interval -.05 to .18, DOES INCLUDE "0" which is our population proportion ($p_1 - p_2 = 0$). we would fail to reject our Ho. Our interval captures the true population parameter (0) and conclude there is NOT a difference who do not eat breakfast at the 2 schools.