Name: _____ Hour: _____ Date: _____

Lesson 10.1: Day 3: Are some groups underrepresented?

According to phys.org, Black and Hispanic females are underrepresented in STEM programs compared to non-STEM programs. A certain university would like to see if this is true for their student population. They took a random sample of 300 STEM students and found that 12 were Black or Hispanic females. A separate random sample of 500 non-STEM students had 75 Black or Hispanic females.

Do the data provide convincing evidence that Black and Hispanic females are underrepresented in STEM programs? Use a 5% significance level.

STATE: Parameter: $p_1 - p_2$ → true difference in proportions of Black and Hispanic females in Stem and non stem

Statistic: $\hat{p}_1 - \hat{p}_2 = .04 - .15 = -.11$

Hypotheses: $p_1 - p_2 = 0$

$p_1 - p_2 < 0$

Significance level: 5%

PLAN: Name of procedure: Two sample z test for $p_1 - p_2$

Check conditions:

Random:
"Random sample of 300"
"Random sample of 500"

10%:
$300 < \frac{1}{10}$ STEM students
$500 < \frac{1}{10}$ non stem students

Large Counts
$300 \times .04 = 12$
$300 \times .96 = 288$
$500 \times .15 = 75$
$500 \times .85 = 425$
≥ 10

DO: Mean: $\mu_{\hat{p}_1 - \hat{p}_2} = 0$

Standard deviation: $\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{.04 \times .96}{300} + \frac{.15 \times .85}{500}}$

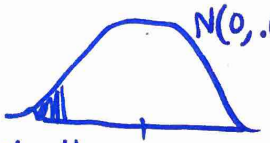
General Formula: Test Stat $= \frac{\text{Stat} - \text{Null}}{SD}$

Since $p_1 - p_2 = 0$ we must use $\hat{p}_c = \frac{12 + 75}{300 + 500} = \frac{87}{800} = .109$

Specific Formula: $z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_c(1-\hat{p}_c)}{n_1} + \frac{\hat{p}_c(1-\hat{p}_c)}{n_2}}}$

Work: $z = \frac{(.04 - .15) - 0}{\sqrt{\frac{.109(1-.109)}{300} + \frac{.109(1-.109)}{500}}} = -4.838$

Picture:



Test statistic: −4.84

P-value: .0000006

CONCLUDE:

Assuming there is no difference in proportions of Black and Hispanic females in Stem & nonstem, there is a .0000006 probability of getting a difference in sample proportions of −.11 or farther purely by chance. We reject the null. This is statistically significant. We conclude the alternative

Name: _____ Hour: _____ Date: _____

Lesson 10.1: Day 3: A Significance Test for $p_1 - p_2$

Important ideas:

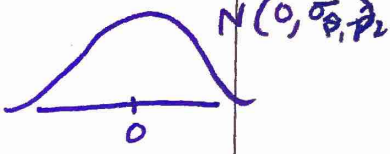
Hypotheses:

H_0: $p_1 - p_2 = 0$

H_a: $p_1 - p_2 \gtrless 0$ $\neq$

$\hat{p}_c$ = Both proportions combined

$\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2 = 0$



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

Check Your Understanding

To study the long-term effects of preschool programs for poor children, researchers designed an experiment. They recruited 123 children who had never attended preschool from low-income families in Michigan. Researchers randomly assigned 62 of the children to attend preschool (paid for by the study budget) and the other 61 to serve as a control group who would not go to preschool. One response variable of interest was the need for social services as adults. Over a 10-year period, 38 children in the preschool group and 49 in the control group have needed social services.

1. Do these data provide convincing evidence that preschool reduces the later need for social services for children like the ones in this study? Justify your answer.

State: $p_1 - p_2$ → the difference in prop. of preschool-control who need S.S.

$\hat{p}_1 - \hat{p}_2 = -.19$

$\hat{p}_c = 87/123 = .71$

$\alpha = .05$

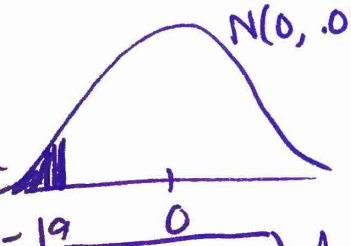
H_0: $p_1 - p_2 = 0$

H_a: $p_1 - p_2 < 0$

Plan: 2 sample z test for $p_1 - p_2$ Conditions met.

Do: Test stat = $\frac{\text{Stat} - \text{Null}}{\text{SD}}$

$$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{.71 \cdot .29}{62} + \frac{.71 \cdot .29}{61}} = .082$$



$$= \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sigma_{\hat{p}_1 - \hat{p}_2}} = \frac{-.19 - 0}{.082} = -2.32 \rightarrow \boxed{.0102}$$

Conclude: Assuming there is no difference in proportions there is a .0102 prob. of getting $\hat{p}_1 - \hat{p}_2 = -.19$ or less purely by chance. Reject the null.

2. Based on your conclusion to Question 1, could you have made a Type I error or a Type II error? Explain your reasoning.

Type I error because we rejected the null.