SECTION 11.1C Exercises

Do work on Notebook Paper. Follow the format of "Test of Significant Template." EACH PROBLEM (3) should take at least 1 full page.

15. **Who owns iPods?** As part of the Pew Internet and American Life Project, researchers surveyed a random sample of 800 teens and a separate random sample of 400 young adults. For the teens, 79% said that they own an iPod or MP3 player. For the young adults, this figure was 67%. Is there a significant difference between the population proportions? State appropriate hypotheses for a significance test to answer this question. Define any parameters you use.

DO 15+17 TOGETHER

17. **Who owns iPods?** Refer to Exercise 15.

(a) Carry out a significance test at the $\alpha = 0.05$ level.

(b) Construct and interpret a 95% confidence interval for the difference between the population proportions. Explain how the confidence interval is consistent with the results of the test in part (a).

21. **Did the random assignment work?** A large clinical trial of the effect of diet on breast cancer assigned women at random to either a normal diet or a low-fat diet. To check that the random assignment did produce comparable groups, we can compare the two groups at the start of the study. Ask if there is a family history of breast cancer: 3396 of the 19,541 women in the low-fat group and 4929 of the 29,294 women in the control group said "Yes."[15] If the random assignment worked well, there should *not* be a significant difference in the proportions with a family history of breast cancer.

(a) How significant is the observed difference? Carry out an appropriate test to help answer this question.

(b) Describe a Type I and a Type II error in this setting. Which is more serious? Explain.

Exercises 23 through 26 involve the following setting. Some women would like to have children but cannot do so for medical reasons. One option for these women is a procedure called in vitro fertilization (IVF), which involves injecting a fertilized egg into the woman's uterus.

23. **Prayer and pregnancy** Two hundred women who were about to undergo IVF served as subjects in an experiment. Each subject was randomly assigned to either a treatment group or a control group. Women in the treatment group were intentionally prayed for by several people (called *intercessors*) who did not know them, a process known as intercessory prayer. The praying continued for three weeks following IVF. The intercessors did not pray for the women in the control group. Here are the results: 44 of the 88 women in the treatment group got pregnant, compared to 21 out of 81 in the control group.[17]

Is the pregnancy rate significantly higher for women who received intercessory prayer? To find out, researchers perform a test of $H_0: p_1 = p_2$ versus $H_a: p_1 > p_2$, where p_1 and p_2 are the actual pregnancy rates for women like those in the study who do and don't receive intercessory prayer, respectively.

(a) Name the appropriate test and check that the conditions for carrying out this test are met.

(b) The appropriate test from part (a) yields a P-value of 0.0007. Interpret this P-value in context.

(c) What conclusion should researchers draw at the $\alpha = 0.05$ significance level? Explain.

(d) The women in the study did not know if they were being prayed for. Explain why this is important.

15 + 17

PARAMETERS: p_1 = actual proportion of teens with IPOD/MP3
p_2 = actual proportion of young adults with IPOD or MP3

HYPOTHESIS: $H_0: p_1 = p_2$ OR $H_0: p_1 - p_2 = 0$
$H_A: p_1 \neq p_2$ $H_A: p_1 - p_2 \neq 0$

TEST: 2 SAMPLE Z TEST FOR p
SIGNIFICANCE LEVEL: $\alpha = .05$

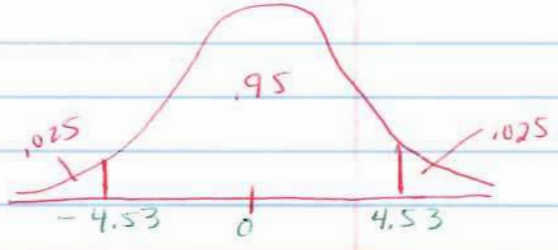
CONDITIONS:
① Random - Both samples were randomly selected
② Independent - There are more than 800(10) = 8,000 teens and 400(10) = 4,000 young adults that live in the US.
③ Normal Condition was met. Success + Failure are at least 10
teens = 632, 168 young adults = 268, 132

Sampling Distribution:

STAT TEST 2 PROP Z TEST

$x_1 = 632$
$n_1 = 800$ ✓
$\hat{p}_1 = .79$ ✓

$x_2 = 268$
$n_2 = 400$ ✓
$\hat{p}_2 = .67$ ✓



$$\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2} = \frac{632 + 268}{800 + 400} = .75$$ ✓

$Z = 4.53$

$p \approx 0$
$\hat{p} = .75$

17 CONT TEST STATISTIC: $z = \dfrac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}_c \hat{q}_c} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

$\hat{p}_1 = .79$ $\hat{p}_c = .75$
$\hat{p}_2 = .67$ $\hat{q}_c = .25$
$n_1 = 800$
$n_2 = 400$

$$z = \frac{.79 - .67}{\sqrt{.75(.25)} \cdot \sqrt{\frac{1}{800} + \frac{1}{400}}}$$

$$z = \frac{.12}{(.4330)(.0612)} = \frac{.12}{.02651} = 4.53$$

$z = 4.53$ ✓

PVALUE $P(Z \leq -4.53)$ OR $P(Z \geq 4.53) \approx 0$
normalcdf$(-E99, -4.53, 0, 1) \approx 0$ ↗

Pvalue $\approx 0 < .05$ (alpha) Reject H_0

(a) Conclude: Since the pvalue $< .05$, we reject H_0 and conclude that the actual proportions of teens and young adults who own IPods/MP3 players are different.

(b) TEST 2 SAMPLE ZINTERVAL FOR $p_1 - p_2$ (95% CI)

$$\hat{p}_1 - \hat{p}_2 \pm z^* \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = .79 - .67 \pm 1.96 \sqrt{\frac{(.79)(.21)}{800} + \frac{(.67)(.33)}{400}}$$

$.12 \pm 1.96(.0276)$
$.12 \pm .054$ $(.066, .174)$ ✓

(CALC; STAT TESTS; 2-PROP ZINT (.0661, .174))

CONCLUDE: we are 95% confident that the interval .066 to .174 captures the difference in proportions of teens and young adults who own IPods or MP players. This is consistent with our decision to Reject H_0. In both cases we ruled out the difference of proportion being 0 as a plausible value.

11.1C HW

21

PARAMETERS:

p_1 = the actual proportion of women with a family history of breast cancer who are assigned to the low fat diet

p_2 = the actual proportion of women ... assigned normal diet

HYPOTHESIS: $H_0: p_1 = p_2$ $H_A: p_1 \neq p_2$

SIGNIFICANCE: CALCULATE

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

CONDITIONS: Random – This was a randomized study

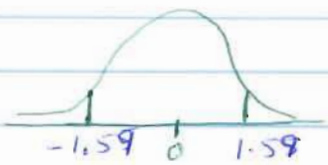
Independent – Due to the random assignment, these 2 groups of women can be viewed as independent.

Normal – The samples are large and success and failures are all above 10 – Low fat: 3396 + 16,145
Normal diet: 4,929 + 24,365

Sampling Distribution

2-PROP Z TEST

$x_1 = 3396$
$n_1 = 19,541$
$\hat{p}_1 = .1737$
$x_2 = 4929$
$n_2 = 29,294$
$\hat{p}_2 = .1683$



$\hat{p}_c = \frac{3396 + 4929}{19,541 + 29,294} = .1705$

$z = 1.59$
$p = .111$

LET me know if you GOT $Z = 1.55$ $P\text{value} = .0836$

TEST STATISTICS

$$z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}_c \hat{q}_c} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$z = \frac{.1737 - .1683}{\sqrt{.171(.829)}\sqrt{\frac{1}{19541} + \frac{1}{29294}}} \quad .0054$$

$z = 1.55$

PVALUE

$P(Z \leq -1.59)$ OR $P(Z \geq 1.59) =$

normal cdf(1.59, E99, 0, 1) =

$.056 * 2 = .111$

P VALUE = .111 SINCE P IS large "FAIL TO REJECT"

11.1C HW

21 CONT

(a) Conclude: Since the pvalue (.11) is large > .05, we fail to reject H_0. We do not have enough evidence to conclude there is a statistically significant difference in the proportion of women assigned to the 2 groups (low-fat vs regular diet) who have a family history of breast cancer

21b A TYPE I EROR would be to say that the groups are significantly different when they are not.

A Type II ERROR would be to say that the 2 GROUPS ARE NOT SIGNIFICANTLY DIFFERENT WHEN THEY ARE. A TYPE 2 ERROR WOULD WOULD BE MORE SERIOUS BECAUSE THE EXPERIMENT WOULD PROCEED ASSUMING THAT THE 2 GROUPS WERE SIMILAR TO BEGIN WITH.. ANY CONCLUSIONS ABOUT THE DIFFERENCE BETWEEN THE 2 GROUPS AT THE END OF THE STUDY WOULD THEN BE SUSPECT

11.1C HW

23

PARAMETERS:

p_1 = actual proportion of women getting pregnant who do receive prayers

p_2 = actual proportion of women getting pregnant who did NOT receive prayers

HYPOTHESIS: $H_0: p_1 = p_2$

$H_A: p_1 > p_2$ "Do prayers help?"

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

CONDITIONS: Random - this was a randomized experiment

Independent - due to the random assignment, these 2 groups of women can be viewed as independent.

Normal - Condition met since success & failures at least 10.

Prayer: 44, 44 non-prayer: 21, 60

Sampling Distribution

TEST STATISTIC:

2PROP Z TEST

$x_1 = 44$ $x_2 = 21$

$n_1 = 88$ $n_2 = 81$

$\hat{p}_1 = .5$ $\hat{p}_2 = .259$

$\hat{p}_C = \frac{44+21}{88+81} = .385$ ✓

$z = 3.21$

$p \approx 0$

$$z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}_C \hat{q}_C} \cdot \sqrt{1/n_1 + 1/n_2}}$$

$$z = \frac{.5 - .259 = .241}{\sqrt{(.385)(.615)} \cdot \sqrt{1/88 + 1/81}}$$

$z = 3.21$ ✓

PVALUE $P(z > 3.21) = \text{normalcdf}(3.21, E99, 0, 1) \approx 0.0007$

pvalue = .0007 < .05 (alpha) Reject H_0

11.1C HW

23 Cont

(a) answered on prior page

(b) $P_{value} = .0007$ ⟶ IF THERE IS NO DIFFERENCE IN PREGNANCY RATES OF WOMEN WHO ARE BEING PRAYED FOR AND THOSE WHO ARE NOT, THERE IS A 0.07% CHANCE OF SEEING AS MANY OR MORE PREGNANCIES WHILE BEING PRAYED FOR AS WE DID.

(c) SINCE PVALUE $< .05$, WE REJECT H_0. WE HAVE ENOUGH EVIDENCE TO CONCLUDE THAT THE PROPORTION OF PREGNANCIES AMONG WOMEN LIKE THESE WHO ARE PRAYED FOR IS HIGHER THAN THAT AMONG WOMEN WHO ARE NOT PRAYED FOR.

(d) IF THE WOMEN HAD KNOWN WHETHER THEY WERE BEING PRAYED FOR, THIS MIGHT HAVE AFFECTED THEIR BEHAVIOR IN SOME WAY (EVEN UNCONSCIOUSLY) THAT WOULD HAVE AFFECTED WHETHER THEY BECAME PREGNANT OR NOT.