

## HOMEWORK 17

- Suppose we select a random sample of 100 students and find that 43% said they believe in love at first sight. Which statement is NOT necessarily true?
  - there were 43 students in the sample who said they believe in love at first sight.
  - based on the information provided by the sample, we cannot determine exactly what proportion of the population would say they believe in love at first sight.
  - $\hat{p} = 0.43$
  - $p = 0.43$
- In a February Gallup Poll, about 50% of the respondents said they believed in love at first sight. Let  $p$  be the proportion of *college students* who believe in love at first sight.
  - What null and alternative hypotheses do you think would be appropriate?

$$H_0: p = 0.50$$

$$H_a: p \neq 0.50$$

- Using the class data from the questionnaire, test your hypothesis.

12pm class: 25 out of 41 said they believe in love at first sight.

That is the sample proportion is  $\hat{p} = 25/41 = 0.61$ .

$$\text{Conditions: } np_0 = 41(0.5) = 20.5 > 10 \quad \text{and} \quad n(1 - p_0) = 41(1 - 0.5) = 20.5 > 10$$

$$\text{Using the formula, the test statistic is: } z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} = \frac{0.61 - 0.5}{\sqrt{\frac{0.5(1 - 0.5)}{41}}} = 1.41$$

Using the calculator: STAT  $\rightarrow$  TESTS  $\rightarrow$  5: 1-PropZTest with  $p_0 = 0.5$ ,  $x = 25$ ,  $n = 41$ .

The p-value is 0.16.

2pm class: 17 out of 42 said they believe in love at first sight.

That is the sample proportion is  $\hat{p} = 17/42 = 0.4$ .

$$\text{Conditions: } np_0 = 42(0.5) = 21 > 10 \quad \text{and} \quad n(1 - p_0) = 42(1 - 0.5) = 21 > 10$$

$$\text{Using the formula, the test statistic is: } z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} = \frac{0.4 - 0.5}{\sqrt{\frac{0.5(1 - 0.5)}{42}}} = -1.23$$

Using the calculator: STAT  $\rightarrow$  TESTS  $\rightarrow$  5: 1-PropZTest with  $p_0 = 0.5$ ,  $x = 17$ ,  $n = 42$ .

The p-value is 0.22

c. State your conclusion in context.

For both classes: Since the p-value is greater than 5%, or even than 10%, we don't have enough evidence to reject the null hypothesis. That is, we don't have enough evidence to reject the claim that 50% of college students believe in love at first sight.

3. Currently, about 10% of marriages in the United States end in divorce during the first five years of marriage. A sociologist is studying the effect of having children within the first two years of marriage on the divorce rate. Using hospital birth records, she selected a random sample of 100 couples who had a child within the first two years. Following up on these couples, she finds that 17 of these couples had divorced within the first five years.

a. Set up appropriate null and alternative hypotheses to test whether having children within the first two years of marriage affects the divorce rate. Assume having children could either increase *or* decrease the divorce rate.

$$H_0: p = 0.1$$

$$H_a: p \neq 0.1$$

b. Calculate the value of the test statistic.

The sample statistic is  $\hat{p} = \frac{17}{100} = 0.17$

Using the formula, the test statistic is: 
$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.17 - 0.10}{\sqrt{\frac{0.10(1-0.10)}{100}}} = 2.33$$

Using the calculator: STAT → TESTS → 5: 1-PropZTest with  $p_0 = 0.1$ ,  $x = 17$ ,  $n = 100$ .

c. Determine the p-value of the test.

p-value: 0.0196

d. Suppose a test is performed using a significance level of  $\alpha = .05$ . Would we reject the null hypothesis? What could we conclude, if anything, about the effect of having children within the first two years of marriage on the divorce rate?

Since  $p\text{-value} < 0.05$ , we have enough evidence to reject the null hypothesis. That is, we have enough evidence to reject the claim that 10% of marriages in the United States end in divorce during the first five years of marriage, and we can conclude that the this rate is not 10%.

4. A university administrator guesses that about 20% of students at his college had sent a text message during class at least once. The head of the Faculty Senate believes that the percentage is higher than that. In order to test the accuracy of the administrator's figure, a survey is conducted, and 28% of the students admitted to having sent a text message during class at least once.

a. Write the null and alternative hypotheses in symbol notation.

$H_0: p = 0.20$

$H_a: p > 0.20$

b. Give the meaning of the parameter  $p$ .

$p$  is the proportion of ALL students at that college who had sent a text message during class at least once.

c. What has to be true about the student responses in order for the test to have validity?

Their responses must be honest, and unbiased.

d. What has to be true about the design of the survey in order for the test to have validity?

The sample must be a random sample.

e. What additional information would have to be provided in order for *you* to conduct the hypothesis test?

We need to know the sample size.

f. The test was carried out, and the resulting p-value was 0.03. Fill in the blanks: “If the true proportion of students who have texted during class is 20%, then the probability that a survey like this would result in 28% or more of the students admitting to having texted during class is 0.03 .

g. Fill in the blanks: “The p-value given in (f) provides moderately strong evidence that the proportion of students who have texted during class is higher than 20%.

h. At a significance level of 5%, would we reject the null hypothesis? Justify your answer.

Yes. Since the p-value is less than 5%, we have enough evidence to reject the null hypothesis. That is, at the 5% level we can reject the claim that the proportion of ALL students at that college who had sent a text message during class at least once is 20%.

i. At a significance level of 1%, would we reject the null hypothesis? Justify your answer.

No. Since the p-value is greater than 1%, we don't have enough evidence to reject the null hypothesis. That is, at the 1% level we cannot reject the claim that the proportion of ALL students at that college who had sent a text message during class at least once is 20%.

5. An argument broke out in a fraternity over who got to call the flip of a coin in a friendly bet. One of the frat members demanded to be allowed to make the call—he wanted to call “heads” because he believed that heads comes up more often than tails. As a result of this argument, another member of the fraternity (who was afflicted with obsessive-compulsive disorder) tossed a coin 25,000 times to determine whether one side tended to come up more often than the other. He obtained 12,671 heads.
- a. Determine a 95% confidence interval for the probability that the coin lands heads on a single toss. (This is the same as the long run proportion of tosses that the coin would come up heads if it were tossed indefinitely.)

The sample statistic is  $\hat{p} = \frac{12671}{25000} = 0.50684$

Conditions:  $n\hat{p} = 25000(0.50684) = 12671 > 10$  and  $n(1 - \hat{p}) = 25000(1 - 0.50684) = 12329 > 10$

The 95% confidence interval is:

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = 0.50684 \pm 1.96 \sqrt{\frac{0.50684(1 - 0.50684)}{25000}} = (0.50064, 0.51304)$$

- b. Use the confidence interval from (a) to test whether the coin comes up heads 50% of the time against the alternative that it does not come up heads 50% of the time.

That is,  $H_0: p = 0.50$        $H_a: p \neq 0.50$

Since 0.5 is not in the CI, we would need to reject the null hypothesis. That is, we reject the claim that the proportion of heads is 0.5.

- c. What is the significance level of the test you carried out in part (b)?

Since we used a 95% CI, the significance level must be 5%.

- d. Perform the test in the usual way, by computing the test statistic, determining the p-value, and comparing it to the significance level. Use a significance level of 5%. Is your result consistent with what you obtained for part b?

$$\text{The test statistic is: } z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} = \frac{0.50684 - 0.50}{\sqrt{\frac{0.50(1 - 0.50)}{25000}}} = 2.163$$

Using the calculator STAT → TESTS → 5: 1-PropZTest with  $p_0 = 0.5$ ,  $x = 12671$ ,  $n = 25000$ .

The p-value is 0.03

At the 5% level, we can reject the null hypothesis because the p-value is less than 5%. That is, we can reject the claim that the proportion of heads is 0.5. We can conclude that the coin is not fair. This is the same conclusion as we had in part b.

- e. What would the p-value have been if we were testing our null hypothesis against the alternative that the coin comes up heads *more than* 50% of the time?

Since “more than” is just one tail, the p-value would be half of the two-tailed test’s p-value, that is it would be 0.015.

- f. Based on your answers to (b) and/or (d), are the results *statistically significant*?

Yes, based on the CI, and the p-value, the results are statistically significant.

- g. Do the results have much *practical* significance? Explain. (Hint: The result of part (a) is particularly pertinent to this question.)

No, the results have not much of a practical significance. Since the plausible values are between 0.50064 and 0.51304, 0.5 is PRACTICALLY right on the edge of the CI.