Lecture 5 practice

Stats 7 Summer Session II 2022

Identify hypotheses

Write the null and alternative hypotheses in words and then symbols for each of the following situations.

(a) A tutoring company would like to understand if most students tend to improve their grades (or not) after they use their services. They sample 200 of the students who used their service in the past year and ask them if their grades have improved or declined from the previous year.

H_0: $p = 0.5$ (Neither a majority nor minority of students' grades improved)

H_A: $p \neq 0.5$ (Either a majority or a minority of students' grades improved)

Online communication

A study suggests that 60% of college student spend 10 or more hours per week communicating with others online. You believe that this is incorrect and decide to collect your own sample for a hypothesis test. You randomly sample 160 students from your dorm and find that 70% spent 10 or more hours a week communicating with others online. A friend of yours, who offers to help you with the hypothesis test, comes up with the following set of hypotheses. Indicate any errors you see.

$$H_0: \hat{p} < 0.6$$

$$H_A: \hat{p} > 0.7$$

Online communication

A study suggests that 60% of college student spend 10 or more hours per week communicating with others online. You believe that this is incorrect...

$$H_0: \hat{p} < 0.6$$

$$H_A: \hat{p} > 0.7$$

- The hypotheses should be about the population proportion (p), not the sample proportion.
- The null hypothesis should have an equal sign.
- The alternative hypothesis should have a not equals sign
- It should reference the null value, $p_0 = 0.6$, not the observed sample proportion.

Online communication

A study suggests that 60% of college student spend 10 or more hours per week communicating with others online. You believe that this is incorrect...

$$H_0: \hat{p} < 0.6$$

$$H_A: \hat{p} > 0.7$$

The correct way to set up these hypotheses is:

$$H_0: p = 0.6$$

$$H_A: p \neq 0.6$$

Where is the proportion of college students that spend 10 or more hours per week communicating with others online.

Cyberbullying rates

Teens were surveyed about cyberbullying, and 54% to 64% reported experiencing cyberbullying (95% confidence interval). Answer the following questions based on this interval.

(a) A newspaper claims that a majority of teens have experienced cyberbullying. Is this claim supported by the confidence interval? Explain your reasoning.

This claim is reasonable, since the entire interval lies above 50%. I would make it clear we are talking about an estimate, not fact.

Cyberbullying rates

Teens were surveyed about cyberbullying, and 54% to 64% reported experiencing cyberbullying (95% confidence interval). Answer the following questions based on this interval.

(b) A researcher conjectured that 70% of teens have experienced cyberbullying. Is this claim supported by the confidence interval? Explain your reasoning

The value of 70% lies outside of the interval, so we have convincing evidence that the researcher's conjecture is wrong.

Cyberbullying rates

Teens were surveyed about cyberbullying, and 54% to 64% reported experiencing cyberbullying (95% confidence interval). Answer the following questions based on this interval.

(c) Without actually calculating the interval, determine if the claim of the researcher from part (b) would be supported based on a 90% confidence interval?

A 90% confidence interval will be narrower than a 95% confidence interval. Even without calculating the interval, we can tell that 70% would not fall in the interval, and we would reject the researcher's conjecture based on a 90% confidence level as well.

Testing for Fibromyalgia

A patient named Diana was diagnosed with Fibromyalgia, a long-term syndrome of body pain, and was prescribed anti-depressants. Being the skeptic that she is, Diana didn’t initially believe that anti-depressants would help her symptoms. However after a couple months of being on the medication she decides that the anti-depressants are working, because she feels like her symptoms are in fact getting better.

(a) Write the hypotheses in words for Diana’s skeptical position when she started taking the anti-depressants.

H_0: Anti-depressants do not affect the symptoms of Fibromyalgia.

H_A: Anti-depressants do affect the symptoms of Fibromyalgia (either helping or harming).

Testing for Fibromyalgia

A patient named Diana was diagnosed with Fibromyalgia, a long-term syndrome of body pain, and was prescribed anti-depressants. Being the skeptic that she is, Diana didn’t initially believe that anti-depressants would help her symptoms. However after a couple months of being on the medication she decides that the anti-depressants are working, because she feels like her symptoms are in fact getting better.

(b) What is a Type 1 Error in this context?

Concluding that anti-depressants either help or worsen Fibromyalgia symptoms when they actually do neither.

Testing for Fibromyalgia

A patient named Diana was diagnosed with Fibromyalgia, a long-term syndrome of body pain, and was prescribed anti-depressants. Being the skeptic that she is, Diana didn’t initially believe that anti-depressants would help her symptoms. However after a couple months of being on the medication she decides that the anti-depressants are working, because she feels like her symptoms are in fact getting better.

(c) What is a Type 2 Error in this context?

Concluding that antidepressants do not affect Fibromyalgia symptoms when they actually do.

Testing for food safety

A food safety inspector is called upon to investigate a restaurant with a few customer reports of poor sanitation practices. The food safety inspector uses a hypothesis testing framework to evaluate whether regulations are not being met. If he decides the restaurant is in gross violation, its license to serve food will be revoked.

(a) Write the hypotheses in words.

H_0: The restaurant meets food safety and sanitation regulations.

H_A: The restaurant does not meet food safety and sanitation regulations.

Testing for food safety

A food safety inspector is called upon to investigate a restaurant with a few customer reports of poor sanitation practices. The food safety inspector uses a hypothesis testing framework to evaluate whether regulations are not being met. If he decides the restaurant is in gross violation, its license to serve food will be revoked.

(b) What is a Type 1 Error in this context?

The food safety inspector concludes that the restaurant does not meet food safety and sanitation regulations and shuts down the restaurant when the restaurant is actually safe.

Testing for food safety

A food safety inspector is called upon to investigate a restaurant with a few customer reports of poor sanitation practices. The food safety inspector uses a hypothesis testing framework to evaluate whether regulations are not being met. If he decides the restaurant is in gross violation, its license to serve food will be revoked.

(c) What is a Type 2 Error in this context?

The food safety inspector concludes that the restaurant meets food safety and sanitation regulations and the restaurant stays open when the restaurant is actually not safe.

Testing for food safety

A food safety inspector is called upon to investigate a restaurant with a few customer reports of poor sanitation practices. The food safety inspector uses a hypothesis testing framework to evaluate whether regulations are not being met. If he decides the restaurant is in gross violation, its license to serve food will be revoked.

(d) Which error is more problematic for the restaurant owner? Why?

A Type 1 Error may be more problematic for the restaurant owner since his restaurant gets shut down even though it meets the food safety and sanitation regulations.

Testing for food safety

A food safety inspector is called upon to investigate a restaurant with a few customer reports of poor sanitation practices. The food safety inspector uses a hypothesis testing framework to evaluate whether regulations are not being met. If he decides the restaurant is in gross violation, its license to serve food will be revoked.

(e) Which error is more problematic for the diners? Why?

A Type 2 Error may be more problematic for diners since the restaurant deemed safe by the inspector is actually not.

Testing for food safety

A food safety inspector is called upon to investigate a restaurant with a few customer reports of poor sanitation practices. The food safety inspector uses a hypothesis testing framework to evaluate whether regulations are not being met. If he decides the restaurant is in gross violation, its license to serve food will be revoked.

(f) As a diner, would you prefer that the food safety inspector requires strong evidence or very strong evidence of health concerns before revoking a restaurant's license? Explain your reasoning.

Strong evidence. Diners would rather a restaurant that meet the regulations get shut down than a restaurant that doesn't meet the regulations not get shut down.

Nutrition labels

The nutrition label on a bag of potato chips says that a one ounce (28 gram) serving of potato chips has 130 calories and contains ten grams of fat, with three grams of saturated fat. A random sample of 35 bags yielded a confidence interval for the number of calories per bag of 128.2 to 139.8 calories. Is there evidence that the nutrition label does not provide an accurate measure of calories in the bags of potato chips?

Because 130 is inside the confidence interval, we do not have convincing evidence that the true average is any different than what the nutrition label suggests.

Practical vs statistical significance

Determine whether the following statement is true or false, and explain your reasoning: “With large sample sizes, even small differences between the null value and the observed point estimate can be statistically significant.”

True. If the sample size gets ever larger, then the standard error will become ever smaller. Eventually, when the sample size is large enough and the standard error is tiny, we can find **statistically significant (α < p-value)** yet very small, **practically insignificant**, differences between the null value and point estimate **($p_0 \approx \hat{p}$)** (assuming they are not exactly equal).

Same observation, different sample size

Suppose you conduct a hypothesis test based on a sample where the sample size is n = 50, and arrive at a p-value of 0.08. You then refer back to your notes and discover that you made a careless mistake, the sample size should have been n = 500. Will your p-value increase, decrease, or stay the same? Explain.

The p-value will decrease because we will have a larger test statistic. With a larger sample size we have more power to detect differences.

Recap: Hypothesis testing for a population proportion

1. Set the hypotheses
   - H_0: $p = p_0$
   - H_A: $p <$ or $>$ or $\neq p_0$
2. Calculate the point estimate
3. Check assumptions and conditions
   - Independence: random sample/assignment
   - Success failure condition: $np_0 \geq 10$ and $n(1-p_0) \geq 10$
4. Calculate a *test statistic* $Z = \frac{\text{Point Estimate} - \text{Mean}}{\text{SE}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
5. Calculate p-value, area depends on H_A (draw a picture!)
6. Make a decision, and interpret it in context

- If p-value $< \alpha$, reject H_0, data provide evidence for H_A
- If p-value $> \alpha$, do not reject H_0, data do not provide evidence for H_A

Gender pay gap in medicine

A study examined the average pay for men and women entering the workforce as doctors for 21 different positions.

(a) If women are paid less than men, then we would expect more than half of those positions to have men paid more than women. Write appropriate hypotheses to test this scenario.

In effect, we're checking whether men are paid more than women.

We'll use p to represent the fraction of cases where men are paid more than women.

$H_0: p = 0.5$, $H_A: p > 0.5$

Gender pay gap in medicine

A study examined the average pay for men and women entering the workforce as doctors for 21 different positions.

(b) Men were, on average, paid more in 19 of those 21 positions. Supposing these 21 positions represent a simple random sample, complete a hypothesis test with a significance level of 0.05.

Summarize our data: $\hat{p} = 19 / 21$, $n = 21$, $\alpha = 0.05$

Check the conditions:

- Simple random sample so we have independence
- $np_0 = n(1 - p_0) = 21 * 0.5 = 10.5$

Gender pay gap in medicine

(b) ... complete a hypothesis test with a significance level of 0.05.

H_0: $p = 0.5$, H_A: $p > 0.5$, $\hat{p} = 19 / 21$, $n = 21$, $\alpha = 0.05$

Calculate our test statistic:

$$Z = \frac{\text{Point Estimate} - \text{Mean}}{\text{SE}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

$$Z = \frac{19/21 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{21}}} = 3.71$$

Gender pay gap in medicine

(b) ... complete a hypothesis test with a significance level of 0.05.

H_0: $p = 0.5$, H_A: $p > 0.5$, $\hat{p} = 19 / 21$, $n = 21$, $\alpha = 0.05$, test statistic = 3.71

Calculate our p-value:

We have a one-tail right sided alternative hypothesis test so we are calculating area above our test statistic.

```
> 1 - pnorm(3.71, mean = 0, sd = 1)
[1] 0.0001036296
```

Our p-value is 0.0001. This means it is extremely unlikely to observe a test statistic as large or larger as what we observed, if the null were true.

Gender pay gap in medicine

(b) ... complete a hypothesis test with a significance level of 0.05.

H_0: $p = 0.5$, H_A: $p > 0.5$, $\hat{p} = 19 / 21$, $n = 21$, $\alpha = 0.05$, p-value = 0.0001

Finally make a conclusion:

Statisticians thought process:

p-value is less than the significance level so we reject the null hypothesis in support of the alternative.

Conclusion in context:

We found evidence (P-value = 0.0001 $< \alpha = 0.05$) that men are paid more than women as doctors in a higher proportion of cases.

Minimum wage

Do a majority of US adults believe raising the minimum wage will help the economy, or is there a majority who do not believe this? A Rasmussen Reports survey of a random sample of 1,000 US adults found that 42% believe it will help the economy. Conduct an appropriate hypothesis test to help answer the research question. Use a significance level of 0.05.

First, let's set up our hypothesis and identify our information:

H_0: $p = 0.5$, H_A: $p \neq 0.5$, $\hat{p} = 0.42$, $n = 1000$, $\alpha = 0.05$

Now, we need to check the conditions for a 1 proportion hypothesis test.

- Simple random sample gets us independence
- Success-failure conditions are satisfied since $1000 \times 0.5 = 500 \geq 10$

Minimum wage

Do a majority of US adults believe raising the minimum wage will help the economy, or is there a majority who do not believe this?

H_0: p = 0.5, H_A: p ≠ 0.5, $\hat{p}$ = 0.42, n = 1000, α = 0.05

Next we calculate our test statistic:

$$Z = \frac{\text{Point Estimate} - \text{Mean}}{\text{SE}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

$$Z = \frac{0.42 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{1000}}} = -5.06$$

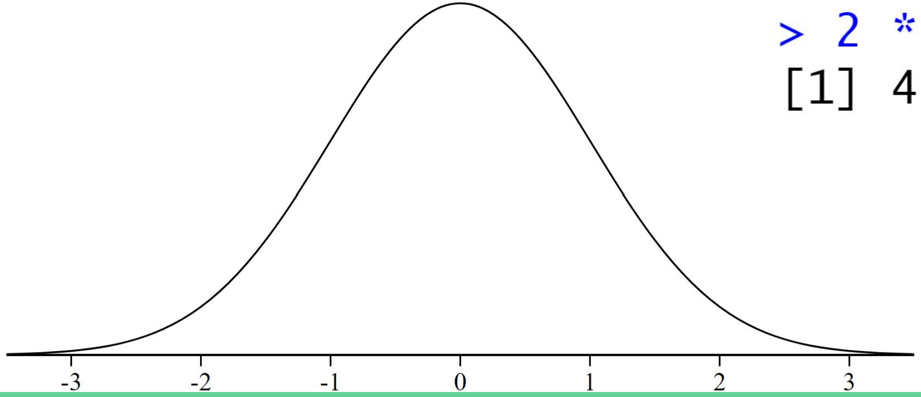
Minimum wage

Do a majority of US adults believe raising the minimum wage will help the economy, or is there a majority who do not believe this?

H_0: $p = 0.5$, H_A: $p \neq 0.5$, $\hat{p} = 0.42$, $n = 1000$, $\alpha = 0.05$, test statistic = -5.06

Now we can calculate the p-value:

Note that the alternative hypothesis is not equal to so we are doing a two sided hypothesis test.




```
> 2 * pnorm(-5.06, mean = 0, sd = 1)
[1] 4.192565e-07
```

P-value < 0.0001

Minimum wage

Do a majority of US adults believe raising the minimum wage will help the economy, or is there a majority who do not believe this?

H_0: $p = 0.5$, H_A: $p \neq 0.5$, $\hat{p} = 0.42$, $n = 1000$, $\alpha = 0.05$, p-value < 0.0001

Finally make a conclusion:

Statisticians thought process:

p-value is less than the significance level so we reject the null hypothesis in support of the alternative.

Conclusion in context:

We found evidence (P-value $< 0.0001 < \alpha = 0.05$) that the percent of US adults who believe raising the minimum wage will help the economy is not 50%.

Fireworks on July 4th

A local news outlet reported that 56% of 600 randomly sampled Kansas residents planned to set off fireworks on July 4th. Determine the margin of error for the 56% point estimate using a 95% confidence level.

Check the conditions:

- We have a random sample so independence is satisfied
- Success failure condition is satisfied since

  n(1 - p) = 600 * (1 - 0.56)= 264 ≥ 10 and np = 600 * 0.56 = 336 ≥ 10

So we can calculate the margin of error:

$$ME = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96 * \sqrt{(\frac{0.56 * (1 - 0.56)}{600})} = 0.0397 \approx 4\%$$

Taste test

Some people claim that they can tell the difference between a diet soda and a regular soda in the first sip. A researcher wanting to test this claim randomly sampled 80 such people. He then filled 80 plain white cups with soda, half diet and half regular through random assignment, and asked each person to take one sip from their cup and identify the soda as diet or regular. 53 participants correctly identified the soda.

Taste test

Some people claim that they can tell the difference between a diet soda and a regular soda in the first sip. A researcher wanting to test this claim randomly sampled 80 such people...

(a) Do these data provide strong evidence that these people are any better or worse than random guessing at telling the difference between diet and regular soda?

H_0: $p = 0.5$, H_A: $p \neq 0.5$, $n = 80$, $\hat{p} = 53 / 80$

Check the conditions:

- Random sample and $np_0 = n(1 - p_0) = 80 * 0.5 = 40 \geq 10$

Calculate the test statistic:

$$Z = \frac{53/80 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{80}}} = 3.07$$

Taste test

Some people claim that they can tell the difference between a diet soda and a regular soda in the first sip. A researcher wanting to test this claim randomly sampled 80 such people...

(a) Do these data provide strong evidence that these people are any better or worse than random guessing at telling the difference between diet and regular soda?

H_0: $p = 0.5$, H_A: $p \neq 0.5$, $n = 80$, $\hat{p} = 53 / 80$, test statistic = 3.07

We have a two sided alternative hypothesis test so the p-value is

```
> 2 * pnorm(-3.07, mean = 0, sd = 1)
[1] 0.002140588
```

Taste test

Some people claim that they can tell the difference between a diet soda and a regular soda in the first sip. A researcher wanting to test this claim randomly sampled 80 such people...

(a) Do these data provide strong evidence that these people are any better or worse than random guessing at telling the difference between diet and regular soda?

H_0: $p = 0.5$, H_A: $p \neq 0.5$, $n = 80$, $\hat{p} = 53 / 80$, p-value = 0.0021

P-value is less than a significance level of 0.05 (assumed) so reject H_0.

We found strong evidence (p-value = $0.0021 < \alpha = 0.05$) that the proportion of people correctly identifying a soda is different than just random guessing.

Taste test

Some people claim that they can tell the difference between a diet soda and a regular soda in the first sip. A researcher wanting to test this claim randomly sampled 80 such people...

(b) Interpret the p-value in this context. (p-value = 0.0021)

If in fact people cannot tell the difference between diet and regular soda and they were randomly guessing, the probability of getting a random sample of 80 people, where 53 or more identify a soda correctly (or 53 or more identify a soda incorrectly), would be 0.0021.

National health plan

Kaiser Family Foundation presented the results of a poll evaluating support for a generic “National Health Plan” in the US in 2019, reporting that 55% of Independents are supportive. If we wanted to estimate this number to within 1% with 90% confidence, what would be an appropriate sample size?

Because a sample proportion ($\hat{p}$ = 0.55) is available, we use this for the sample size calculations.

The margin of error for a 90% confidence interval is $z^\star \times SE$. We want this to be less than 0.01, where we use $\hat{p}$ in place of p:

We have $\hat{p}$ = 0.55 but we need the $z^\star$ for a 90% confidence level:

```
> qnorm((1 - 0.9)/2, mean = 0, sd = 1)
[1] -1.644854
```

National health plan

$$ME \leq 0.01$$

$$1.64 * \sqrt{\frac{0.55(1-0.55)}{n}} \leq 0.01$$

$$1.64^2 * \frac{0.55(1-0.55)}{n} \leq 0.01^2$$

$$1.64^2 * \frac{0.55(1-0.55)}{0.01^2} \leq n$$

$$6696.223 \leq n$$

So the sample size must be at least 6697 to have a margin of error less than 0.01

National health plan continued

In the poll we discussed it was found that 79% of 347 Democrats and 55% of 617 Independents support a National Health Plan.

(a) Calculate a 95% confidence interval for the difference between the proportion of Democrats and Independents who support a National Health Plan ($p_D - p_I$), and interpret it in this context. We have already checked conditions for you.

Recall the conditions:

- Independence within and between groups
- Success failure conditions but with the pooled estimate of the proportion

We are assuming these are met so we can proceed

National health plan continued

In the poll we discussed it was found that 79% of 347 Democrats and 55% of 617 Independents support a National Health Plan.
(a) Calculate a 95% confidence interval for the difference between the proportion of Democrats and Independents who support a National Health Plan ($p_D - p_I$)...

$$\text{Point estimate} \pm z^* \times SE$$

$$(\hat{p}_D - \hat{p}_I) \pm z^* \times \sqrt{\frac{\hat{p}_D(1-\hat{p}_D)}{n_D} + \frac{\hat{p}_I(1-\hat{p}_I)}{n_I}}$$

$$(0.79 - 0.55) \pm 1.96 \times \sqrt{\frac{0.79 * (1 - 0.79)}{347} + \frac{0.55 * (1 - 0.55)}{617}}$$

$$0.24 \pm 0.0581$$

$$(0.1819, 0.2981)$$

National health plan continued

In the poll we discussed it was found that 79% of 347 Democrats and 55% of 617 Independents support a National Health Plan.

(a) Calculate a 95% confidence interval for the difference between the proportion of Democrats and Independents who support a National Health Plan ($p_D - p_I$), and interpret it in this context.

95% CI: (0.1819, 0.2981)

We are 95% confident that the proportion of Democrats who support the plan is 18.1% to 29.9% higher than the proportion of Independents who support the plan.

Offshore drilling

A survey asked 827 randomly sampled registered voters in California “Do you support? Or do you oppose? Drilling for oil and natural gas off the Coast of California? Or do you not know enough to say?” Below is the distribution of responses, separated based on whether or not the respondent graduated from college.

(a) What percent of college graduates and what percent of the non-college graduates in this sample support drilling for oil and natural gas off the Coast of California?

	College Grad	
	Yes	No
Support	154	132
Oppose	180	126
Do not know	104	131
Total	438	389

$\hat{p}_{CG} = 154 / 438$ and $\hat{p}_{NCG} = 132 / 389$

Offshore drilling

(b) Conduct a hypothesis test to determine if the data provide strong evidence that the proportion of college graduates who support off-shore drilling in California is different than that of non-college graduates.

$\hat{p}_{CG}$ = 154 / 438 and $\hat{p}_{NCG}$ = 132 / 389

H_0: $p_{CG} - p_{NCG} = 0$, H_A: $p_{CG} - p_{NCG} \neq 0$

Conditions:

Simple random sample and pooled proportion

$\hat{p}_{pooled}$ = (154 + 132) / (438 + 389) = 0.346 so n*$\hat{p}_{pooled}$ = 286 ≥ 10.

Test statistic:

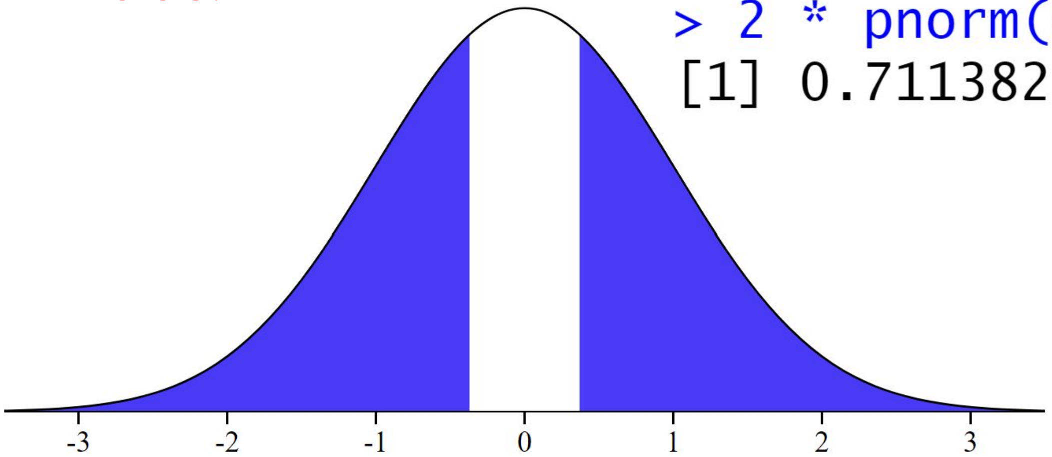
$$Z = \frac{\text{Point estimate} - \text{Mean}}{\text{SE}} = \frac{(\hat{p}_{CG} - \hat{p}_{NCG}) - 0}{\sqrt{\frac{\hat{p}_{pooled}(1-\hat{p}_{pooled})}{n_{CG}} + \frac{\hat{p}_{pooled}(1-\hat{p}_{pooled})}{n_{NCG}}}} = \frac{(154/438 - 132/389) - 0}{\sqrt{\frac{0.346*(1-0.346)}{438} + \frac{0.346*(1-0.346)}{389}}} = 0.37$$

Offshore drilling

(b) Conduct a hypothesis test to determine if the data provide strong evidence that the proportion of college graduates who support off-shore drilling in California is different than that of non-college graduates.

H_0: $p = 0.5$, H_A: $p \neq 0.5$, $\hat{p}_{CG} = 154 / 438$, $\hat{p}_{NCG} = 132 / 389$, test statistic = 0.37

P-value:




```
> 2 * pnorm(-0.37, mean = 0, sd = 1)
[1] 0.7113825
```

Offshore drilling

(b) Conduct a hypothesis test to determine if the data provide strong evidence that the proportion of college graduates who support off-shore drilling in California is different than that of non-college graduates.

H_0: $p = 0.5$, H_A: $p \neq 0.5$, $\hat{p}_{CG} = 154 / 438$, $\hat{p}_{NCG} = 132 / 389$, p-value = 0.7114

Since the p-value is less than the significance level, assumed to be $\alpha = 0.05$, we fail to reject the null hypothesis.

Conclusion in context:

The data do not provide strong evidence (p-value = $0.7114 < \alpha = 0.05$) of a difference between the proportions of college graduates and non-college graduates who support off-shore drilling in California.

Credits

Examples adapted from OpenIntro Statistics (4th edition) by David Diez, Mine Cetinkaya-Rundel, and Christopher D Barr https://www.openintro.org/book/os/ protected under the Creative Commons License