Week 4 Practice Quiz

Practice Quiz, 9 questions

✖ **Try again once you are ready.**

Required to pass: 80% or higher

You can retake this as many times as you'd like.

Back to Week 4

Retake

Week 4 Practice Quiz

1 / 1 points

Practice Quiz, 9 questions

1.

Suppose you want to construct a confidence interval for a population proportion. Which of the following, if it were true, would **prevent** you from being able to assume that the distribution of the sample **proportion** is nearly normal?

- *n* = 104. Out of these 104 there are only a few successes (15), but relatively many failures (89).
- *n* = 104. These observations are a simple random sample and make up less than 10% of the population.
- *n* = 104. Out of these 104 there are only a few failures (7), but relatively many successes (97).

Correct

Recognize that the Central Limit Theorem (CLT) is about the distribution of point estimates, and that given certain conditions, this distribution will be nearly normal.

- In the case of the proportion the CLT tells us that if

(1) the observations in the sample are independent,

(2) the sample size is sufficiently large (checked using the success/failure condition: $np \geq 10$ and $n(1-p) \geq 10$),

then the distribution of the sample proportion will be nearly normal, centered at the true population proportion and with a standard error of $\sqrt{\frac{p(1-p)}{n}}$.

$$\hat{p} \sim N\left(mean = p, SE = \sqrt{\frac{p(1-p)}{n}}\right)$$

S-F condition not met since 7 < 10.

- None of these options.

Week 4 Practice Quiz

1 / 1 points

Practice Quiz, 9 questions

2.

In 2013, Edward Snowden leaked details of top-secret NSA spying activities to the media. A poll conducted by USA TODAY / Pew Research Center asked 1,504 people in U.S. whether Snowden's leaks have helped or harmed the public interest. 53% of respondents answered "helped the public interest". You want to test whether a majority of people in the U.S. believe he helped the public interest. Which of the following is the correct set of hypotheses?

- $H_0 : \rho = 0.53; H_A : \rho > 0.53$
- $H_0 : \rho = 0.5; H_A : \rho > 0.5$

Correct

This question revisits the setup of hypothesis testing within the categorical data / proportions.

The wording of the question tells us we're interested in whether the true proportion of people in the US believe he helped the public interest is greater than 50% (i.e. makes them "a majority").

- $H_0 : \rho < 0.5; H_A : \rho > 0.5$
- $H_0 : \rho = 0.53; H_A : \rho < 0.53$

Week 4 Practice Quiz

1 / 1 points

Practice Quiz, 9 questions

3.

In response to complaints from residents about too many (about 15%) of the cars passing by the local school speeding, the police started closely monitoring traffic. You want to check if the police's efforts had an effect on the prevalence of speeding in this area. One day you observe 560 different cars pass by the school, and find that 70 of them were speeding. You calculate a p-value of 0.0976. Assuming the cars are representative of all cars that drive by the school, which of the following is **true**?

- If in fact the police's efforts had an effect, the probability of getting a random sample of 560 cars where 70 or less cars are speeding is 0.0976.
- If in fact the police's efforts didn't have an effect, the probability of getting a random sample of 560 cars where 70 cars are speeding is 0.0976.
- If in fact the police's efforts didn't have an effect, the probability of getting a random sample of 560 cars where 70 or less cars are speeding is 0.0976.
- If in fact the police's efforts didn't have an effect, the probability of getting a random sample of 560 cars where 70 or less or 98 or more cars are speeding is 0.0976.

Correct

p-value = P(observed or more extreme test statistic | H_0 true)

Week 4 Practice Quiz

0/1 points

Practice Quiz, 9 questions

4.

When do we use the **pooled** proportion in calculation of the standard error of the difference of two proportions ($SE_{(\hat{p}1 - \hat{p}2)}$)?

- when using a randomization test to compare $p_1 - p_2$
- when comparing p_1 and p_2 using a theoretical approach, and the null hypothesis is $H_0: p_1 - p_2 = 0$
- when comparing p_1 and p_2 using a theoretical approach, and the null hypothesis is $H_0: p_1 - p_2 =$ (some value other than 0)

This should not be selected

- Note that the standard error calculation for the confidence interval and the hypothesis test are different when dealing with proportions, since in the hypothesis test we need to assume that the null hypothesis is true.
- Note that the calculation of the standard error of the distribution of the difference in two independent sample proportions is different for a confidence interval and a hypothesis test.

Review the associated learning objective.

- when constructing a confidence interval for $p_1 - p_2$

Week 4 Practice Quiz

0/1 points

Practice Quiz, 9 questions

5.

Rock-paper-scissors is a hand game played by two or more people where players choose to sign either 'rock', 'paper', or 'scissors' with their hands. We would like to test if players choose between these three options randomly, or if certain options are favored above others. What hypothesis test should we conduct to answer this research question?

- Chi square test of goodness of fit
- Chi square test of independence

This should not be selected

The question refers to the following learning objective(s):

- Use a chi-square test of goodness of fit to evaluate if the distribution of levels of a single categorical variable follows a hypothesized distribution.
- When evaluating the independence of two categorical variables where at least one has more than two levels, use a chi-square test of independence.

We are testing whether the distribution of levels of a **single** categorical variable follows a hypothesized distribution, not the independence between two categorical variables.

- Compare two means
- Compare two proportions

Week 4 Practice Quiz

1 / 1 points

Practice Quiz, 9 questions

6.

When doing a hypothesis test on a single proportion (i.e. for one categorical variable), we have studied how to calculate the p-value for the hypothesis test, beginning with generating simulated samples. Which of the following is the best description for how you should generate the simulated samples, and why?

- Generate simulated samples based on the null hypothesis because that is the hypothesis we're trying to prove when doing the hypothesis test.
- Generate simulated samples based on the null hypothesis because we need to see how extreme our observed data looks if the null hypothesis were really true.

Correct

The question refers to the following learning objective(s):

In hypothesis testing for one categorical variable, generate simulated samples based on the null hypothesis, and then calculate the number of samples that are at least as extreme as the observed data.

In hypothesis testing for one categorical variable, generate simulated samples based on the null hypothesis, and then calculate the number of samples that are at least as extreme as the observed data.

- Generate simulated samples based on the alternative hypothesis because that is the hypothesis we're trying to prove when doing the hypothesis test.
- Generate simulated samples based on the alternative hypothesis because we need to see how extreme our observed data looks if the alternative hypothesis were really true.

Week 4 Practice Quiz

1 / 1 points

Practice Quiz, 9 questions

7.

True or false: In calculation of the required sample size for a given margin of error of the confidence interval for a population proportion, we should use $\hat{p}$ = 0.5 if we don't have any knowledge about the characteristics of the population.

- True

Correct

- False

Week 4 Practice Quiz

0 / 1 points

Practice Quiz, 9 questions

8.

Suppose in a population 20% of people wear contact lenses. What is the expected shape of the sampling distribution of proportion of contact lens wearers in random samples of 1000 people from this population?

- nearly normal
- left-skewed
- right-skewed

This should not be selected

The question refers to the following learning objective(s):

Note that if the CLT doesn't apply and the sample proportion is low (close to 0) the sampling distribution will likely be right skewed, if the sample proportion is high (close to 1) the sampling distribution will likely be left skewed.

S-F condition met.

- uniform

Week 4 Practice Quiz

1 / 1 points

Practice Quiz, 9 questions

9.

True/False: When the success-failure condition is not met, we should use a T test to compare two proportions.

- True
- False

Correct

Use simulation methods when sample size conditions aren't met for inference for categorical variables.

- Note that the t-distribution is only appropriate to use for means. When sample size isn't sufficiently large, and the parameter of interest is a proportion or a difference between two proportions, we need to use simulation.

In hypothesis testing

- for one categorical variable, generate simulated samples based on the null hypothesis, and then calculate the number of samples that are at least as extreme as the observed data.

- for two categorical variables, use a randomization test.