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?

0 / 1 point

- (●) n = 104. Out of these 104 there are only a few successes (15), but relatively many failures (89).
- ( ) n = 104. Out of these 104 there are an equal number of successes and failures (52 each).
- ( ) None of these options.
- ( ) n = 104. These observations are a simple random sample and make up less than 10% of the population.

⊗ **Incorrect**

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 met since both are greater than 10.

2. When checking conditions for calculating a confidence interval for a proportion, you should use which number of successes and failures?

- (●) Expected (based on the null value)
- ( ) Depends on the context
- ( ) Observed
- ( ) Not applicable. The number of successes and failures (observed or otherwise) is not part of the conditions required for calculating a confidence interval for a proportion.

⊗ **Incorrect**

For confidence intervals use $\hat{p}$ (observed sample proportion) when calculating the standard error and checking the success/failure condition. For hypothesis tests use p_0 (null value) when calculating the standard error and checking the success/failure condition.

Use the observed number of successes and failures when calculating a confidence interval for a proportion, but not when doing a hypothesis test. In a hypothesis test for a proportion, you should use $np0$ and $n(1-p_0)$ successes and failures; that is, the expected number based on the **null proportion**.

3. In May 2011, Gallup asked 1,721 students in grades five through twelve if their school teaches them about money and banking. Researchers are interested in finding out if a majority of students receive such education. Which of the following is the correct set of hypotheses?

- ○ H_0: $\hat{p} = 0.5$; H_A : $\hat{p} \neq 0.5$
- ○ H_0 : $\mu = 0.5$; H_A : $\mu > 0.5$
- ○ H_0 : $p < 0.5$; H_A : $p > 0.5$
- ◉ H_0 : $p = 0.5$; H_A : $p > 0.5$

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 students receiving this education is greater than 50% (i.e. makes them "a majority").

4. The campaign manager for a congressional candidate claims that the candidate has **more than** 50% support from the district's electorate. A newspaper collects a simple random sample of 500 likely voters in this district and estimates the support for this candidate to be 52%. The p-value for the hypothesis test evaluating the campaign manager's claim is 0.19. Which of the below is **correct**?

- ○ The success-failure condition is not met, so this p-value is not reliable.
- ◉ If in fact 50% of likely voters support this candidate, the probability of obtaining a random sample of 500 likely voters where 52% or more support the candidate is 0.19.
- ○ 95% of random samples of size 500 will estimate the support for this candidate to be 52%.
- ○ The data provide convincing evidence for the campaign manager's claim.

Correct

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

5. Gallup conducts an annual poll of U.S. residents. Approximately 1,000 residents across all 50 states and Washington D.C. are asked "Do you believe the use of marijuana should be made legal?" The distribution of responses by date of survey is shown in the table below. Imagine a hypothesis test evaluating whether there is a difference from 2012 to 2013 between proportions of "yes" responses. Using the information in the table below, calculate the standard error for this hypothesis test. Choose the closest answer.

		time of survey	
		Nov 2012	October 2013
response	yes	493	596
	no	514	401
	undecided	30	31
	total	1037	1028

- (●) 0.022
- ( ) 0.00048
- ( ) 0.5798
- ( ) 0.4754
- ( ) 0.5274

Correct

First calculate

$$\hat{p}_{\text{pool}} = \frac{493 + 596}{1037+1028} \approx 0.53$$

Then $SE =$

$$\sqrt{\frac{0.53 \times (1-0.53)}{1037} + \frac{0.53 \times (1-0.53)}{1028}}$$

$$= 0.022$$

6. "In statistical inference for proportions, standard error (SE) is calculated differently for hypothesis tests and confidence intervals." Which of the following is the **best** justification for this statement?

- ○ Because in hypothesis testing we're interested in the variability of the true population distribution, and in confidence intervals we're interested in the variability of the sampling distribution.
- ○ Because statistics is full of arbitrary formulas.
- ○ Because if we used the same method for hypothesis tests as we did for confidence intervals, the calculation would be impossible.
- ◉ Because in hypothesis testing, we assume the null hypothesis is true, hence we calculate SE using the null value of the parameter. In confidence intervals, there is no null value, hence we use the sample proportion(s).

Correct

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

Note that the reason for the difference in calculations of standard error is the same as in the case of the single proportion: when the null hypothesis claims that the two population proportions are equal, we need to take that into consideration when calculating the standard error for the hypothesis test, and use a common proportion for both samples.

7. At the beginning of a semester an anonymous survey was conducted on students in a statistics class. Two of the questions on the survey were about gender and whether or not students have equal, more, or less energy in the afternoon compared to the morning. Below are the results.

	Equal	Less	More
Female	18	37	24
Male	9	15	24

What test should we perform to see if gender and energy level are associated?

- ○ F test
- ○ ANOVA
- ◉ Chi-square test of independence

8. A report on your local TV news station claims that 60% of the city's residents support using limited city funds to hire and train more police officers. A second local news station has picked up this story, and they claim that certainly less than 60% of residents support the additional hiring. In order to test this claim the second news station takes a random sample of 100 residents and finds that 56 of them (56%) support the use of limited funds to hire additional police officers. Which of the following is the correct set-up for calculating the p-value for this test?

- ○ In a bag place 100 chips, 56 red and 44 blue. Randomly sample 100 chips, with replacement, and record the proportion of red chips in the sample. Repeat this many times, and calculate the proportion of samples where at least 60% of the chips are red.
- ◉ Randomly sample 100 residents of a nearby city, and record the number of residents in the sample who support the hiring of additional police officers. Repeat this many times and calculate the proportion of samples where at least 56% of the residents support the hiring.
- ○ Roll a 10-sided die 100 times and record the proportion of times you get a 6 or lower. Repeat this many times, and calculate the number of simulations where the sample proportion is 56% or less.
- ○ Roll a 10-sided die 100 times and record the proportion of times you get a 6 or lower. Repeat this many times, and calculate the number of simulations where the sample proportion is 60% or less.

⊗ **Incorrect**

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.

It is incorrect to sample residents of a nearby city because our hypothesis is about residents of a particular city.

9. True or false: The χ^2 statistic is always non-negative.

- (●) True
- ( ) False

✓ **Correct**

10. 80% of Americans start the day with a cereal breakfast. Based on this information, determine if the following statement is true or false.

"The sampling distribution of the proportions of Americans who start the day with a cereal breakfast in random samples of size 40 is right skewed."

- ( ) True
- (●) False

✓ **Correct**

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 not met, and the true population is closer to 1 than 0, so the sampling distribution will be left skewed.

11. At a stop sign, some drivers come to a full stop, some come to a 'rolling stop' (not a full stop, but slow down), and some do not stop at all. We would like to test if there is an association between gender and type of stop (full, rolling, or no stop). We collect data by standing a few feet from a stop sign and taking note of type of stop and the gender of the driver. Below is a contingency table summarizing the data we collected. If gender is not associated with type of stop, how many males would we expect to not stop at all? Choose the closest answer.

		gender	
		female	male
stop	full stop	6	6
	rolling stop	16	15
	no stop	4	3

- ○ 3.64
- ○ 6.24
- ○ 5.76
- ◉ 3.36

Correct

Calculate expected counts in two-way tables as

$$E = \frac{rowtotal \times columntotal}{grandtotal}$$

12. True / False: A randomization test for $H_0 : p_1 = p_2$ uses a randomization distribution centered at 0.

- ◉ 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

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?

- ○ $n = 104$. 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 successes (15), but relatively many failures (89).
- ◉ $n = 104$. Out of these 104 there are an equal number of successes and failures (52 each).
- ○ None of these options.

Incorrect

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 met since both are greater than 10.

2. When performing a hypothesis test on proportions (either where $H_0 : p = p_0$ or where $H_0 : p_1 = p_2$) you should use the **observed** number of successes and failures when checking conditions.

- ○ True
- ◉ False
- ○ Depends on the context
- ○ Observed proportion for one sample, expected proportion for two samples.

Correct

For confidence intervals use $\hat{p}$ (observed sample proportion) when calculating the standard error and checking the success/failure condition. For hypothesis tests use p_0 (null value) when calculating the standard error and checking the success/failure condition.

Use the observed number of successes and failures when calculating a confidence interval for a proportion, but not when doing a hypothesis test. In a hypothesis test for a proportion, you should use np_0 and $n(1-p_0)$ successes and failures; that is, the expected number based on the **null proportion**.

3. Scientists researching the effect of dietary supplements such as multivitamins and supplements of folic acid, magnesium or zinc on the health of elderly women followed 38,772 women, whose average age was 62 at the start of the study, over a period of 19 years and recorded the death rates of those who did and did not use dietary supplements. The researchers found that women using supplements were more likely than women who did not to have died during that period. Write the hypotheses for testing if the death rate of elderly women who use dietary supplements is different than the death rate of elderly women who do not use dietary supplements.

- ○ $H_0 : \mu_{used} = \mu_{not}$; $H_A : \mu_{used} \neq \mu_{not}$
- ○ $H_0 : \hat{p}_{used} = \hat{p}_{not}$; $H_A : \hat{p}_{used} \neq \hat{p}_{not}$
- ○ $H_0 : p_{used} = p_{not}$; $H_A : p_{used} > p_{not}$
- ◉ $H_0 : p_{used} = p_{not}$; $H_A : p_{used} \neq p_{not}$

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

4. The campaign manager for a congressional candidate claims that the candidate has **more than** 50% support from the district's electorate. A newspaper collects a simple random sample of 500 likely voters in this district and estimates the support for this candidate to be 52%. The p-value for the hypothesis test evaluating the campaign manager's claim is 0.19. Which of the below is **correct**?

- ○ The data provide convincing evidence for the campaign manager's claim.
- ○ 95% of random samples of size 500 will estimate the support for this candidate to be 52%.
- ○ The success-failure condition is not met, so this p-value is not reliable.
- ◉ If in fact 50% of likely voters support this candidate, the probability of obtaining a random sample of 500 likely voters where 52% or more support the candidate is 0.19.

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

5. Voters in the U.S. state of New Jersey voted on a measure that would legalize sports betting. Before the vote, SurveyUSA asked a random sample of registered voters the following question: "'If betting on sports were to become legal in New Jersey, would you yourself place bets?" The distribution of responses by age group is shown in the table below. Consider a hypothesis test evaluating whether there is a difference between proportions of 18 to 34 year olds and 35+ year olds who would never place bets. Using the information from the table, calculate the standard error for this hypothesis test. Choose the closest answer.

		age group		
		18 to 34	35+	total
response	regularly	3	30	33
	occasionally	69	128	197
	never	71	224	295
	not sure	1	7	8
	total	144	389	533

- ○ 0.0024
- ◉ 0.0485
- ○ 0.5535
- ○ 0.4931
- ○ 0.5758

Correct

First calculate

$$\hat{p}_{\text{pool}}$$

$$= \frac{71 + 224}{144+389} \approx 0.55$$

Then $SE =$

$$\sqrt{\frac{0.55 \times (1-0.55)}{144} + \frac{0.55 \times (1-0.55)}{389}}$$

$$= 0.0421$$

6. To evaluate the following hypotheses

$H_0: p = 0.3$

$H_A: p \neq 0.3$

we use a random sample of 50 observations where $\hat{p} = 0.36$. Which of the following is the correct standard error? Choose the closest answer.

- ○ 0.0096
- ○ 0.0092
- ○ 0.0042
- ◉ 0.0648
- ○ 0.0297
- ○ 0.0679

Correct

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

Note that the reason for the difference in calculations of standard error is the same as in the case of the single proportion: when the null hypothesis claims that the two population proportions are equal, we need to take that into consideration when calculating the standard error for the hypothesis test, and use a common proportion for both samples.

For a hypothesis test,

$$SE = \sqrt{\frac{p_0(1-p_0)}{n}} = \sqrt{\frac{0.3 \times 0.7}{50}} \approx 0.0648$$

7. At the beginning of a semester an anonymous survey was conducted on students in a statistics class. Two of the questions on the survey were about gender and whether or not students have equal, more, or less energy in the afternoon compared to the morning. Below are the results.

	Equal	Less	More
Female	18	37	24
Male	9	15	24

What test should we perform to see if gender and energy level are associated?

- ○ F test
- ○ ANOVA
- ◉ Chi-square test of goodness of fit
- ○ hypothesis test for a single mean
- ○ Comparing two means
- ○ Comparing two proportions
- ○ Chi-square test of independence
- ○ Z test

Incorrect

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're evaluating the dependence of two categorical variables, not evaluating whether a single categorical variable follows a hypothesized distribution.

8. A report on your local TV news station claims that 60% of the city's residents support using limited city funds to hire and train more police officers. A second local news station has picked up this story, and they claim that certainly less than 60% of residents support the additional hiring. In order to test this claim the second news station takes a random sample of 100 residents and finds that 56 of them (56%) support the use of limited funds to hire additional police officers. Which of the following is the correct set-up for calculating the p-value for this test?

- ○ Roll a 10-sided die 100 times and record the proportion of times you get a 6 or lower. Repeat this many times, and calculate the number of simulations where the sample proportion is 56% or less.
- ○ Randomly sample 100 residents of a nearby city, and record the number of residents in the sample who support the hiring of additional police officers. Repeat this many times and calculate the proportion of samples where at least 56% of the residents support the hiring.
- ○ Roll a 10-sided die 100 times and record the proportion of times you get a 6 or lower. Repeat this many times, and calculate the number of simulations where the sample proportion is 60% or less.
- ◉ In a bag place 100 chips, 56 red and 44 blue. Randomly sample 100 chips, with replacement, and record the proportion of red chips in the sample. Repeat this many times, and calculate the proportion of samples where at least 60% of the chips are red.

⊗ **Incorrect**

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.

The chips in the bag should represent the null hypothesis (60% red chips) rather than the observed data (56% red chips).

9. True or false: The χ^2 statistic is always non-negative.

- ○ False
- ◉ True

Correct

10. 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 30 people from this population?

- ○ uniform
- ◉ nearly normal
- ○ right-skewed
- ○ left-skewed

Incorrect

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 not met.

11. At a stop sign, some drivers come to a full stop, some come to a 'rolling stop' (not a full stop, but slow down), and some do not stop at all. We would like to test if there is an association between gender and type of stop (full, rolling, or no stop). We collect data by standing a few feet from a stop sign and taking note of type of stop and the gender of the driver. What are the hypotheses for testing for an association between gender and type of stop?

- ○ H_0: Males and females are equally likely to come to a full stop.

  H_A: Males and females are not equally likely to come to a full stop.

- ○ H_0: Males and females are equally likely to come to a rolling stop.

  H_A: Males are more likely than females to come to a rolling stop.

- ○ H_0: Gender and type of stop are associated.

  H_A: Gender and type of stop are independent.

- ◉ H_0: Gender and type of stop are independent.

  H_A: Gender and type of stop are associated.

Correct

When evaluating the independence of two categorical variables where at least one has more than two levels, use a chi-square test of independence.

H_0 : The two variables are independent.

H_A : The two variables are dependent.

12. True / False: A randomization test for $H_0 : p_1 = p_2$ uses a randomization distribution centered at 0.

- ○ True
- ◉ False

Incorrect

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

Congratulations! You passed!

Grade received 91.66%

Latest Submission Grade 91.67%

To pass 80% or higher

Go to next item

1. Which of the following is **not** required for the distribution of the sample **proportion** to be nearly normal?

1 / 1 point

- ○ There should be at least 10 successes.
- ○ Observations should be independent.
- ◉ Sample size should be at least 30 and the population distribution should not be extremely skewed.
- ○ There should be at least 10 failures.

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 $\sqrt{\frac{p(1-p)}{n}}$.

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

When considering the distribution of the sample proportion, we don't have a requirement of $n \geq 30$. To determine if the sample size of categorical data is high enough, we instead check the success-failure condition.

2. When performing a hypothesis test on proportions (either where $H_0 : p = p_0$ or where $H_0 : p_1 = p_2$) you should use the **observed** number of successes and failures when checking conditions.

- ○ True
- ◉ False
- ○ Depends on the context
- ○ Observed proportion for one sample, expected proportion for two samples.

Correct

For confidence intervals use $\hat{p}$ (observed sample proportion) when calculating the standard error and checking the success/failure condition. For hypothesis tests use p_0 (null value) when calculating the standard error and checking the success/failure condition.

Use the observed number of successes and failures when calculating a confidence interval for a proportion, but not when doing a hypothesis test. In a hypothesis test for a proportion, you should use np_0 and $n(1-p_0)$ successes and failures; that is, the expected number based on the **null proportion**.

3. Scientists researching the effect of dietary supplements such as multivitamins and supplements of folic acid, magnesium or zinc on the health of elderly women followed 38,772 women, whose average age was 62 at the start of the study, over a period of 19 years and recorded the death rates of those who did and did not use dietary supplements. The researchers found that women using supplements were more likely than women who did not to have died during that period. Write the hypotheses for testing if the death rate of elderly women who use dietary supplements is different than the death rate of elderly women who do not use dietary supplements.

- ○ $H_0 : \hat{p}_{used} = \hat{p}_{not}$; $H_A : \hat{p}_{used} \neq \hat{p}_{not}$
- ◉ $H_0 : p_{used} = p_{not}$; $H_A : p_{used} \neq p_{not}$
- ○ $H_0 : \mu_{used} = \mu_{not}$; $H_A : \mu_{used} \neq \mu_{not}$
- ○ $H_0 : p_{used} = p_{not}$; $H_A : p_{used} > p_{not}$

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

4. You and a friend are about to visit the aviary at the local zoo for the first time. A trustworthy zookeeper says the aviary holds about 3,000 birds. Your friend read somewhere that 10% of those birds are cardinals, but he thinks there are really more cardinals than that. You're both great at identifying cardinals so you decide to test this claim with a hypothesis test on the true proportion p of cardinals in the aviary. You walk around the aviary together and get a simple random sample by spotting 250 birds. Of these, 35 were cardinals and 215 were not cardinals. The p-value is 0.0175. Which of the following is **false**?

- ○ The success-failure condition is met.
- ◉ $\hat{p} = 0.14$
- ○ $H_0 : p = 0.10$
- ○ If in fact 10% of the birds in the aviary are cardinals, the probability of obtaining a random sample of 250 birds where exactly 14% are cardinals is 0.0175.

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

5. Gallup conducts an annual poll of U.S. residents. Approximately 1,000 residents across all 50 states and Washington D.C. are asked "Do you believe the use of marijuana should be made legal?" The distribution of responses by date of survey is shown in the table below. Imagine a hypothesis test evaluating whether there is a difference from 2012 to 2013 between proportions of "yes" responses. Using the information in the table below, calculate the standard error for this hypothesis test. Choose the closest answer.

		time of survey	
		Nov 2012	October 2013
response	yes	493	596
	no	514	401
	undecided	30	31
	total	1037	1028

- ○ 0.4754
- ○ 0.5274
- ○ 0.5798
- ○ 0.00048
- ◉ 0.022

Correct

First calculate

$$\hat{p}_{\text{pool}}$$

$$= \frac{493 + 596}{1037+1028} \approx 0.53$$

Then $SE =$

$$\sqrt{\frac{0.53 \times (1-0.53)}{1037} + \frac{0.53 \times (1-0.53)}{1028}}$$

$$= 0.022$$

6. "In statistical inference for proportions, standard error (SE) is calculated differently for hypothesis tests and confidence intervals." Which of the following is the **best** justification for this statement?

- (●) Because in hypothesis testing, we assume the null hypothesis is true, hence we calculate SE using the null value of the parameter. In confidence intervals, there is no null value, hence we use the sample proportion(s).
- ( ) Because in hypothesis testing we're interested in the variability of the true population distribution, and in confidence intervals we're interested in the variability of the sampling distribution.
- ( ) Because if we used the same method for hypothesis tests as we did for confidence intervals, the calculation would be impossible.
- ( ) Because statistics is full of arbitrary formulas.

Correct

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

Note that the reason for the difference in calculations of standard error is the same as in the case of the single proportion: when the null hypothesis claims that the two population proportions are equal, we need to take that into consideration when calculating the standard error for the hypothesis test, and use a common proportion for both samples.

7. At the beginning of a semester an anonymous survey was conducted on students in a statistics class. Two of the questions on the survey were about gender and whether or not students have equal, more, or less energy in the afternoon compared to the morning. Below are the results.

	Equal	Less	More
Female	18	37	24
Male	9	15	24

What test should we perform to see if gender and energy level are associated?

- ○ Comparing two means
- ○ Z test
- ○ Chi-square test of goodness of fit
- ○ ANOVA
- ○ hypothesis test for a single mean
- ○ Comparing two proportions
- ○ F test
- ◉ Chi-square test of independence

Correct

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.

A variety of studies suggest that 10% of the world population is left-handed. It is also claimed that artists are more likely to be left-handed. In order to test this claim we take a random sample of 40 art students at a college and find that 6 of them (15%) are left handed. Which of the following is the correct set-up for calculating the p-value for this test?

- ○ Roll a 10-sided die 40 times and record the proportion of times you get a 1. Repeat this many times, and calculate the proportion of simulations where the sample proportion is 10% or more.
- ○ In a bag place 40 chips, 6 red and 34 blue. Randomly sample 40 chips, with replacement, and record the proportion of red chips in the sample. Repeat this many times, and calculate the proportion of samples where at least 10% of the chips are red.
- ◉ Roll a 10-sided die 40 times and record the proportion of times you get a 1. Repeat this many times, and calculate the proportion of simulations where the sample proportion is 15% or more.
- ○ Randomly sample 40 non-art students, and record the number of left-handed students in the sample. Repeat this many times and calculate the proportion of samples where at least 15% of the students are left-handed.

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 this problem, generating simulated samples based on the null hypothesis corresponds to rolling the 10-sided die 40 times. Note that we're treating a rolled 1 as "left-handed" where a rolled 2 through rolled 9 are considered "right-handed".

9. One of the early studies linking smoking and lung cancer compared patients hospitalized with lung cancer to similar patients without lung cancer (hospitalized for other reasons), and recorded whether each patient smoked. For a hypothesis test testing whether the proportion of smokers is higher for the patients with lung cancer than for patients without lung cancer, the p-value is less than 0.000001. Does this provide significant evidence that smoking causes lung cancer?

		smoking status		
		smoker	non-smoker	total
cancer	lung cancer	647	2	649
	not sure	622	27	649
	total	1269	29	1298

Study reference: Doll, R. & Hill, A.B. (1950) "Smoking and carcinoma of the lung: preliminary report", *British Medical Journal.*

- ○ Whether or not we can conclude that smoking causes lung cancer depends on the statistical method the researchers used to obtain the p-value.
- ○ No, with the given p-value we would fail to reject H0 in favor of HA.
- ◉ Based on this study we cannot conclude that smoking causes lung cancer,

  regardless of the p-value.
- ○ Yes, with the given p-value we would reject H0 in favor of HA, and conclude that smoking causes lung cancer.

Correct

These data was not (and could not have ethically been!) collected through a controlled experiment; therefore no causal statements can be made using the study. Recall that a controlled experiment (as opposed to an observational study) is needed in order to determine causality.

10. 80% of Americans start the day with a cereal breakfast. Based on this information, determine if the following statement is true or false.

"The sampling distribution of the proportions of Americans who start the day with a cereal breakfast in random samples of size 40 is right skewed."

- ○ True
- ◉ False

Correct

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 not met, and the true population is closer to 1 than 0, so the sampling distribution will be left skewed.

11. At a stop sign, some drivers come to a full stop, some come to a 'rolling stop' (not a full stop, but slow down), and some do not stop at all. We would like to test if there is an association between gender and type of stop (full, rolling, or no stop). We collect data by standing a few feet from a stop sign and taking note of type of stop and the gender of the driver. What are the hypotheses for testing for an association between gender and type of stop?

- ◉ H_0: Gender and type of stop are independent.

  H_A: Gender and type of stop are associated.
- ○ H_0: Males and females are equally likely to come to a rolling stop.

  H_A: Males are more likely than females to come to a rolling stop.
- ○ H_0: Males and females are equally likely to come to a full stop.

  H_A: Males and females are not equally likely to come to a full stop.
- ○ H_0: Gender and type of stop are associated.

  H_A: Gender and type of stop are independent.

Correct

When evaluating the independence of two categorical variables where at least one has more than two levels, use a chi-square test of independence.

H_0 : The two variables are independent.

H_A : The two variables are dependent.

12. Does Weight Watchers work? Researchers randomly divided 500 people into two equal-sized groups. One group spent 6 months on the Weight Watchers program. The other group received a pamphlet about controlling portion sizes. At the end of the study 35% of the subjects in the pamphlet group and 55% of the subjects in the Weight Watchers group had lost at least 10 pounds. To test whether Weight Watchers is more effective for weight loss than pamphlets, a statistician used an index card to represent each subject in the study and wrote whether or not the subject lost at least 10 pounds on the index card. He then shuffled these cards together, and dealt them into two equal-sized groups. Which of the following best describes the expected result?

- (●) The difference between the proportions of cards indicating whether or not the subject lost at least 10 pounds will be about 0.
- ( ) The difference between the proportions of cards indicating whether or not the subject lost at least 10 pounds will be about 20%.
- ( ) If Weight Watchers was effective, the difference between the proportions of cards indicating whether or not the subject lost at least 10 pounds will be more than 20%.

Correct

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

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.

The randomization distribution will be centered at the null value, which is $p_1 - p_2 = 0$.