1. **(1 point)** `carroll_problib/statistics/InterStats/CI2.pg`
In 2013, the Pew Research Foundation reported that "50% of U.S. adults report that they live with one or more chronic conditions." However, this value was based on a sample, so it may not be a perfect estimate for the population parameter of interest on its own.

The research foundation created three confidence intervals at three different confidence levels. Match the confidence interval to the confidence level.

___1. A 99% confidence interval.
___2. A 95% confidence interval.
___3. A 99.5% confidence interval.

A. The researchers estimate that between 38.51% and 61.49% of U.S. adults live with one or more chronic conditions
B. The researchers estimate that between 30.61% and 69.39% of U.S. adults live with one or more chronic conditions
C. The researchers estimate that between 34.86% and 65.14% of U.S. adults live with one or more chronic conditions

Answer(s) submitted:

-
-
-

(incorrect)

2. **(1 point)** `carroll_problib/statistics/InterStats/CI3.pg`
A recent poll found that 55% of U.S. adult Twitter users get at least some news on Twitter.

(a) Which confidence interval has the smallest confidence level?

- A. The fraction of U.S. adult Twitter users who get some news from Twitter is between 39.10% and 70.90%
- B. The fraction of U.S. adult Twitter users who get some news from Twitter is between 42.58% and 67.42%
- C. The fraction of U.S. adult Twitter users who get some news from Twitter is between 45.58% and 64.42%

(b) Which confidence interval has the largest confidence level?

- A. The fraction of U.S. adult Twitter users who get some news from Twitter is between 39.10% and 70.90%
- B. The fraction of U.S. adult Twitter users who get some news from Twitter is between 42.58% and 67.42%
- C. The fraction of U.S. adult Twitter users who get some news from Twitter is between 45.58% and 64.42%

Answer(s) submitted:

-
-

(incorrect)

3. **(1 point)** `carroll_problib/statistics/InterStats/CI4.pg`
For each problem, select the best response.

(a) A 95% confidence interval for the mean μ of a population is computed from a random sample and found to be 10 ± 4. We may conclude

- A. that there is a 95% probability that the true mean is 10 and a 95% chance the true margin of error is 4.
- B. that there is a 95% probability that μ is between 6 and 14.
- C. 95% of the population is between 6 and 14.
- D. that if we took many, many additional samples and from each computed a 95% confidence interval for μ, approximately 95% of these intervals would contain μ.
- E. All of the above.

(b) Suppose you collect a simple random sample of size n from a population and from the data collected you computed a 95% confidence interval for the mean of the population. Which of the following would produce a new confidence interval with larger width (larger margin of error) based on these same data?

- A. Use a smaller confidence level.
- B. Use a larger confidence level.
- C. Use the same confidence level, but compute the interval n times. Approximately 5% of these intervals will be larger.
- D. Nothing can guarantee absolutely that you will get a larger interval. One can only say the chance of obtaining a larger interval is 0.05.
- E. None of the above.

Answer(s) submitted:

-
-

(incorrect)

4. **(1 point)** `carroll_problib/statistics/InterStats/CI5.pg`

When calculating a confidence interval for a proportion we needs a few ingredients.

1) We need a sample proportion, $\hat{p}$ and a sample size, n.
2) We need to make sure that a normal model is appropriate for the underlying sampling distribution. To do so, we check that $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$. These conditions state that we are expecting at least 10 successes and failures given the proportion and the sample size that we're dealing with.
3) We need a standard error (based on the sample proportion and the sample size), $SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
4) Next, we need a critical z-score, z^*
5) The margin of error is calculated as $MOE = z^* \times SE$.

The confidence interval is calculated as

$$\text{Lower Bound } = \hat{p} - z^* \times SE \qquad \text{Upper Bound } = \hat{p} + z^* \times SE$$

This is the same as

$$\text{Lower Bound } = \hat{p} - MOE \qquad \text{Upper Bound } = \hat{p} + MOE$$

(A) In a chili cookoff, 52 people sample all of the chili and 20% vote to give the "Fire Hot Mouth Burner" the best chili in the contest. Since we haven't sampled everyone at the entire cook off we have to use a confidence interval to estimate the proportion of chili eating contest-goers that might choose this as the best chili of the contest.

(i) $\hat{p} =$ ____
(ii) $n =$ ____
(iii) $n\hat{p} =$ ____ (this is the number that voted for the "Fire Hot Mouth Burner")
(iv) $n(1-\hat{p}) =$ ____ (this is the number that didn't vote for the "Fire Hot Mouth Burner")
(v) Since both of the previous answers are greater than or equal to 10 we can proceed with a normal model for the confidence interval. We now need to get the critical z-score. We will use a 92% confidence level. In MS Excel we need to use the "=norm.s.inv( )" command to determine the appropriate critical z-score.
=norm.s.inv(____)
(vi) This gives a critical z-score of
$z^* =$ ____
(vii) The standard error is:
$SE =$ ____
(viii) The margin of error is:
$MOE =$ ____
(ix) Hence, we are 92% confident that the true proportion of people that will vote for "Fire Hot Mouth Burner" is between ____ and ____.

(B) In a hospital, an orthopedic doctor collects a random sample of 67 previous patients and checks to see if pre-surgical physical therapy strength training helps to speed their recovery after a total hip replacement. She finds that in 61.19% of the 67 patients that indeed it does. Since the doctor didn't sample every possible surgical patient she has to use a confidence interval to estimate the proportion of total hip surgical patients that will benefit from pre-surgical PT.

(i) $\hat{p} =$ ____
(ii) $n =$ ____
(iii) $n\hat{p} =$ ____ (this is the number showed improved recovery time.)
(iv) $n(1-\hat{p}) =$ ____ (this is the number that didn't show improved recovery time.)
(v) Since both of the previous answers are greater than or equal to 10 we can proceed with a normal model for the confidence interval. We now need to get the critical z-score. We will use a 99% confidence level. In MS Excel we need to use the "=norm.s.inv( )" command to determine the appropriate critical z-score.
=norm.s.inv(____)
(vi) This gives a critical z-score of
$z^* =$ ____
(vii) The standard error is:
$SE =$ ____
(viii) The margin of error is:
$MOE =$ ____
(ix) Hence, we are 92% confident that the true proportion of surgical total hip patients that will have improved recovery time with pre-surgical PT is between ____ and ____.

Answer(s) submitted:

-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-

(incorrect)

5. **(1 point)** `carroll_problib/statistics/InterStats/CI6.pg`

When calculating a confidence interval for a proportion we needs a few ingredients.

1) We need a sample proportion, $\hat{p}$ and a sample size, n.
2) We need to make sure that a normal model is appropriate for the underlying sampling distribution. To do so, we check that $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$. These conditions state that we are expecting at least 10 successes and failures given the proportion and the sample size that we're dealing with.
3) We need a standard error (based on the sample proportion and

the sample size), $SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
4) Next, we need a critical z-score, z^*
5) The margin of error is calculated as $MOE = z^* \times SE$.
The confidence interval is calculated as

Lower Bound $= \hat{p} - z^* \times SE$ Upper Bound $= \hat{p} + z^* \times SE$

This is the same as

Lower Bound $= \hat{p} - MOE$ Upper Bound $= \hat{p} + MOE$

(a) Consider a situation where $\hat{p} = 0.3$ and $N = 106$. Calculate the standard error, the critical z-score for 95% confidence, and the margin of error.
$SE =$ _____
$z - critical$ = norm.s.inv(_____) = _____
$MOE =$ _____

(b) Consider a situation where $\hat{p} = 0.69$ and $N = 573$. Calculate the standard error, the critical z-score for 94% confidence, and the margin of error.
$SE =$ _____
$z - critical$ = norm.s.inv(_____) = _____
$MOE =$ _____

Answer(s) submitted:

-
-
-
-
-
-
-
-

(incorrect)

---

6. (1 point) `carroll_problib/statistics/InterStats/CI7.pg`
In a 2010 Survey USA poll, 66% of the 122 respondents between the ages of 18 and 34 said they would vote in the 2010 general election for Proposition 19, which would change California law to legalize marijuana and allow it to be regulated and taxed. At a 95% confidence level, this sample as a 8.4% margin of error. Based on this information, determine if the following statements are true or false.

(a) We are 95% confident that between 57.600% and 74.400% of the California voters in this sample support Proposition 19. [?/True/False]

(b) We are 95

(c) IF we consider many random samples of 122 California voters between the ages of 18 and 34, and we calculated 95% confidence intervals for each, 95% of them will include the true population proportion of Californians who support Proposition 19. [?/True/False]

(d) In order to decrease the margin of error to 0.042% we would need to quadruple (multiply by 4) the sample size. [?/True/False]

(e) Based on this confidence interval, there is sufficient evidence to conclude that a majority of California voters between the ages of 18 and 34 support Proposition 19. [?/True/False]

Answer(s) submitted:

-
-
-
-
-

(incorrect)

---

7. (1 point) `carroll_problib/statistics/InterStats/CI8.pg`
We are interested in estimating the proportion of graduates at a mid-sized university who found a job within one year of completing their undergraduate degree. Suppose we conduct a survey and find out that 340 of the 449 randomly sampled graduates found jobs. The graduating class under consideration included 4700 students.
For each of the following use appropriate computational software and be sure that your answer has at least 4 decimal digits of accuracy.

(a) What is the value of the point estimate for the proportion of graduates who found a job within one year.
$\hat{p} =$ _____
(b) We need to check if the conditions for constructing a confidence interval based on these data are met. Recall that the conditions are that we expect at least 10 successes, at least 10 failures, and we are sampling independent individuals. To check the last one we need to make sure that we are gathering no more than about 10% of the total population. This is especially important since the survey is done without replacement.
How many successes do we expect? $n\hat{p} =$ _____ ≥ 10
How many failures do we expect? $n(1-\hat{p}) =$ _____ ≥ 10
Based on the 10% rule stated above, can we assume that the individuals are independent? [?/Yes/No]

(c) Calculate a 95% confidence interval for the proportion of graduates who found a job within one year of completing their undergraduate degree at this university. (round your critical z-score to two decimal places)
Lower Bound: $\hat{p} - z^* \times SE =$ _____
Upper Bound: $\hat{p} + z^* \times SE =$ _____

(d) Select all of the following that are appropriate interpretations of the confidence interval.

- A. If we were to create many 95% confidence intervals in the same way with the same sample sizes, then all of them would contain the true proportion of the graduates who found a job within one year of graduation.
- B. We are 95% confident that the true proportion of people who found a job within one year of graduation is between the lower bound and the upper bound.
- C. If we were to create many 95% confidence intervals in the same way with the same sample sizes, then 95% of them would contain the true proportion of the graduates who found a job within one year of graduation.
- D. We are 95% confident that the proportion of people in our sample that found a job within one year of graduation is $\hat{p}$.

(e) If we were to create a 99% confidence interval instead then the resulting interval would be

- ?
- wider than
- narrower than
- the same size as

the 95% confidence interval.

Answer(s) submitted:

-
-
-
-
-
-
-
-

(incorrect)

---

8. (1 point) `carroll_problib/statistics/InterStats/HT_1p_1.pg`

A 2012 survey of 400 American adults indicates that 18% of cell phone owners do their browsing on their phone rather than a computer or other device.

According to an online article, a report from a mobile research company indicates that 38% of Chinese mobile web users only access the internet through their cell phones. We wish to conduct a hypothesis test to determine if these data provide strong evidence that the proportion of Americans who only use their cell phones to access the internet **is less than** the Chinese proportion of 38%.

(a) The Null Hypothesis, in mathematical symbols, is: H_0 : $p =$ _____

(b) The Alternative Hypothesis, in mathematical symbols, is: H_A : p [?/=/¡/¿/not equal] _____

In both the null and alternative hypotheses, the symbol p is the proportion of American adult cell phone owners who do their browsing on their phone.

(c) The standard error based on the null hypothesis is: _____

(d) Our data indicates a proportion of Americans who use their phones to browse is $\hat{p} = 0.18$. We need to use this to find a test statistics (in this case called a z-score). Recall that a test statistics is a measure for how many standard errors our data point is away from our assumed mean.

test statistic = _____

Find the p-value for this hypothesis test using Excel. You should find that it is quite small!

(e) What is this p-value the probably of?

- A. The p-value is the probability of the null hypothesis being true assuming that the American adult cell phone owners and Chinese mobile users are the same in their browsing habits.
- B. The p-value is the probability of the alternative hypothesis being true assuming that the American adult cell phone owners and Chinese mobile users are the same in their browsing habits.
- C. The p-value is the probability of finding the 18% of American adult cell phone owners who use their phones for browsing.
- D. The p-value is the probability of finding our sample assuming that American adult cell phone owners and Chinese mobile users are the same in their browsing habits.

(f) Select all of the following that are appropriate conclusions from this test.

- A. We fail to reject the null hypothesis.
- B. We have evidence to suggest American adult cell phone users do less browsing than Chinese mobile users.
- C. We reject the null hypothesis.
- D. We do not have evidence to suggest American adult cell phone users do less browsing than Chinese mobile users

Answer(s) submitted:

-
-
-
-
-
-
-

(incorrect)

---

9. (1 point) `carroll_problib/statistics/InterStats/One_Sample_p/baby_smoke_1.pg`

To start this problem, download the data file below:

<u>THE LINK TO THE DATA IS HERE</u>

In the data file you will find 1,000 entries with information about new mothers in North Carolina. For this problem we are going to build a confidence interval to estimate the proportion of babies born in NC that are considered full term.

(a) Use a pivot table to find the number of full term babies in the data set.

Number of full term babies = _____

(b) There might be a few entries in the data that don't have the premie status recorded and must be removed from our forthcoming statistical analysis. Taking the NA's into account, how many total babies are included in the sample?

Number of babies = _____

(c) Now find the proportion of babies in the data set that are full term. This serves as a point estimate for the population of North Carolina mothers.

$\hat{p} =$ _____ (use at least 4 decimal digits of accuracy)

(d) Recall that the standard error for a confidence interval on one proportion is given by

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

Calculate the standard error for the proportion of babies that are full term. _____ (use at least 4 decimal digits of accuracy)

(e) We would like to build a 91% confidence interval for the proportion of full term babies in NC. To do so we first need to find the critical z-score associated with this confidence level:
critical z-score = norm.s.inv(_____) = _____

(f) Now recall that the lower and upper bounds on the confidence interval are given by:

Lower Bound $= \hat{p} - z^* \times SE$ Upper Bound $= \hat{p} + z^* \times SE$

Calculate the lower bound: _____
Calculate the upper bound: _____

(g) Now that we have the lower and the upper bound for the interval we need to interpret what we've done. Select all of the following that are correct interpretations of the confidence interval.

- A. If we were to create many many 91% confidence intervals in the same way with the same sample size then 100% of them would contain the true proportion of babies carried to full term in NC.
- B. We are 91% confident that the true proportion babies that are carried to full term in NC is between the lower and the upper bound.
- C. If we were to create many many 91% confidence intervals in the same way with the same sample size then 91% of them would contain the true proportion of babies carried to full term in NC.
- D. We are 91% confident that the proportion of babies found in the data set is between the lower bound and the upper bound.

Answer(s) submitted:

-
-
-
-
-
-
-
-
-

(incorrect)

---

10. **(1 point)** `carroll_problib/statistics/InterStats/One_Sample_p/baby_smoke_2.pg`

To start this problem, download the data file below:
THE LINK TO THE DATA IS HERE

In the data file you will find 1,000 entries with information about new mothers in North Carolina. For this problem we are going to build a confidence interval to estimate the proportion new mothers who smoke in NC.

(a) Now find the proportion of mothers in the data set that are smokers. This serves as a point estimate for the population of North Carolina mothers. Be sure to watch out for missing data!
$\hat{p} =$ _____
(b) You should now go find the standard error for the sampling distribution behind the point estimate ... I'll wait while you do that ...
(c) We would like to build a 99% confidence interval for the proportion of mothers who are smokers in NC. To do so we first need to find the critical z-score associated with this confidence level:
critical z-score = norm.s.inv(_____) = _____
(e) Now find the lower and upper bounds that make up the confidence interval.
Calculate the lower bound: _____
Calculate the upper bound: _____

Now that we have the lower and the upper bound for the interval we need to interpret what we've done. Select all of the following that are correct interpretations of the confidence interval.

- A. If we were to create many many 99% confidence intervals in the same way with the same sample size then 99% of them would contain the true proportion of mothers who are smokers in NC.
- B. We are 99% confident that the true proportion mothers that are smokers in NC is between the lower and the upper bound.
- C. If we were to create many many 99% confidence intervals in the same way with the same sample size then 100% of them would contain the true proportion of mothers who are smokers in NC.
- D. We are 99% confident that the proportion of mothers who are smokers in the data set is between the lower bound and the upper bound.

Answer(s) submitted:

-
-
-
-
-
-

(incorrect)

---

11. **(1 point)** `carroll_problib/statistics/InterStats/One_Sample_p/baby_smoke_3HT.pg`

To start this problem, download the data file below:
THE LINK TO THE DATA IS HERE

In the data file you will find 1,000 entries with information about new mothers in North Carolina. It is common knowledge that each baby has about a 50% chance of being female. IS this true for "premie" babies in North Carolina?

(a) Start by downloading the data set and filtering and sorting so you can see the "premie" babies. Then determine the proportion of premies that are female.

$\hat{p} =$ _____

(b) Let's set up a hypothesis test to see if our data indicates some significant difference between our common knowledge. In particular, we are interested in seeing if the mothers of premies are less likely to have female babies. Let p be the proportion of premie babies that are female.

Null Hypothesis: $p =$ _____

Alternative Hypothesis: p [?/=/¡/¿/not equal to] _____

(c) We are sampling far less than 10% of the population and we certainly found more than 10 successes and failures in our sample. Hence, a normal model for the sampling distribution is reasonable. What is the standard error for the underlying sampling distribution?

$SE_0 =$ _____

(d) Now let's get a test statistic. Recall that the words "test statistic" for a normal model is also called a "z score". That is, we are interested in finding the number of standard errors that our sample statistics, $\hat{p}$, falls above or below the mean of the sampling distribution.

test statistic = _____

(e) Now we will get a p-value. Since we are dealing with a normal model and we have a z-score we can use the "=NORM.S.DIST( )" command in Excel.

p-value = NORM.S.DIST(_____ , 1) = _____

(f) Based on this p-value, what is our decision using $\alpha = 0.05$?

- A. Reject the null hypothesis
- B. Fail to reject the null hypothesis

(g) Now that you've made the decision in statistical terms, what does your decision mean in practical terms?

- A. We have evidence to say that mothers of premies are equally likely to have males as females.
- B. We do not have evidence to say that mothers of premies are less likely to have females.
- C. We have evidence to say that mothers of premies are less likely to have females.
- D. We do not have evidence to say that mothers of premies are less likely to have males.
- E. We do not have evidence to say that mothers of premies are equally likely to have males as females.

Answer(s) submitted:

-
-
-
-
-
-
-
-
-
-

(incorrect)

---

12. **(1 point)** `carroll_problib/statistics/InterStats/One_Sample_p/HealthcareLaw2012.pg`

On June 28, 2012 the U.S. Supreme Court upheld the much debated 2010 healthcare law, declaring it constitutional. A Gallup poll released the day after this decision indicates that 46% of 1,012 Americans agree with this decision. At a 95% confidence level, this sample has a 3% margin of error. Based on this information, determine if the following statements are true or false.

(a) We are 95

(b) We are 95

(c) If we considered many random samples of 1,012 Americans, and we calculated the sample proportions of those who support the decision of the U.S. Supreme Court, 95

(d) The margin of error at a 90

Answer(s) submitted:

-
-
-
-

(incorrect)

---

13. **(1 point)** `carroll_problib/statistics/InterStats/One_Sample_p/VegetarianStudents.pg`

Suppose that 8% of college students are vegetarians. Determine if the following statements are true or false.

(a) The distribution of the sample proportions of vegetarians in random samples of size 126 is approximately normal since we expect at least 10 successes and 10 failures in the sample. [?/true/false]

(b) A random sample of 140 college students where 10.4% are vegetarians would be considered unusual. [?/true/false]

(c) A random sample of 560 college students where 10.4

(d) The standard error of the sampling distribution for samples of size 140 is

- ?
- two times
- three times
- four times
- eight times
- ten times

larger than the standard error of the sampling distribution for samples of size 560.

Answer(s) submitted:

-
-
-
-

(incorrect)

---

14. **(1 point)** `carroll_problib/statistics/InterStats/One_Sample_p/Fireworks.pg`

In late June 2012, Survey USA published results of a survey stating that 59% of the 575 randomly sampled Kansas residents planned to set off fireworks on July 4th. Determine the margin of error and the confidence interval for the 59% point estimate using a 95
Margin of Error = ____
Lower Estimate = ____
Upper Estimate = ____

Answer(s) submitted:

-
-
-

(incorrect)

---

15. **(1 point)** `carroll_problib/statistics/InterStats/One_Sample_p/SampleSize_Given_MOE.pg`

Before conducting a study it is often a good idea to decide on the necessary sample size. To do so, statisticians often first decide on an acceptable margin of error and then work backward to the sample size. Since they haven't run the study yet it is common practice to approximate the sample proportion as 50% since this is the worst case scenario.
Recall that the standard error for a single proportion is

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}},$$

Note: The calculation to find the sample size will likely give a decimal answer so be sure to round appropriately so that the actual margin of error will be less than what the researchers want.

(a) If researchers decide on a 4.8% margin of error for a 95% confidence interval. What sample size should they pursue? ____

(b) If researchers decide on a 2.1% margin of error for a 99% confidence interval. What sample size should they pursue? ____

Answer(s) submitted:

-
-

(incorrect)

---

16. **(1 point)** `carroll_problib/statistics/InterStats/Two_Sample_p/baby_smoke_4HT.pg`

To start this problem, download the data file below:
THE LINK TO THE DATA IS HERE

In the data file you will find 1,000 entries with information about new mothers in North Carolina. We would like to answer the question: "Do mothers who are smokers give birth to a higher proportion of low-weight babies?" This is naturally a hypothesis test, but since we will be comparing smokers and non-smokers this will be a test of the difference of two proportions.

(a) Start by downloading the data set and determine the two proportions in our data set.
Proportion of low-weight babies in the group of mothers who are smokers = $\hat{p}_1 =$ ____
Proportion of low-weight babies in the group of mothers who are non-smokers = $\hat{p}_2 =$ ____
Now we need to get one value in order to run our hypothesis test. We are interested in determining if the proportion of low-weight babies among the smoker group is larger than that in the non-smoker group. Hence, we will calculate $\hat{p}_1 - \hat{p}_2$ and expect a positive number.
$\hat{p}_1 - \hat{p}_2 =$ ____

(b) Let's set up a hypothesis test to see if our data indicates some significant difference between our common knowledge. In particular, is smoking linked to an increase in the prevalence of low-weight babies.
Null Hypothesis: $p_1 - p_2 =$ ____
Alternative Hypothesis: $p_1 - p_2$ [?/=/¡/¿/not equal to] ____
(c) We are sampling far less than 10% of the population and we certainly found more than 10 successes and failures in our sample. Hence, a normal model for the sampling distribution is reasonable. To get a standard error for this problem we need to take a bit of care. We are assuming that the two populations (smokers and non-smokers) have the same proportion of low-weight babies, and the standard error needs to be created based on this assumption.
* We first need to get a pooled proportion of low-weight babies in the data set (total number of low weight babies divided by the total number of babies)
$\hat{p}_{pooled} =$ ____
* Now use the pooled proportion to find a standard error for the underlying sampling distribution. T The formula for this is

$$SE_{pooled} = \sqrt{\hat{p}_{pooled}(1-\hat{p}_{pooled})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$SE_{pooled} =$ ____
(d) Now let's get a test statistic. Recall that the words "test statistic" for a normal model is also called a "z score". That is, we are interested in finding the number of standard errors that our sample statistic, $\hat{p}_1 - \hat{p}_2$, falls above or below the mean of the sampling distribution.
test statistic = ____
(e) Now we will get a p-value using the normal model that we have built.

p-value = _____

(f) Based on this p-value, what is our decision using $\alpha = 0.05$?

- A. Fail to reject the null hypothesis
- B. Reject the null hypothesis

(g) Now that you've made the decision in statistical terms, what does your decision mean in practical terms?

- A. We do not have evidence to say that mothers who smoke have proportionally more low-weight babies than non-smoking mothers.
- B. We have evidence to say that mothers who smoke have proportionally more low-weight babies than non-smoking mothers.
- C. We do not have evidence to say that mothers who smoke have the same proportion of low-weight babies as non-smoking mothers.
- D. We have evidence to say that mothers who smoke have the same proportion of low-weight babies as non-smoking mothers.

Answer(s) submitted:

•
•
•
•
•
•
•
•
•
•
•
•

(incorrect)

---

17. **(1 point)** `carroll_problib/statistics/InterStats/Two_Sample_p/FluShot_HT.pg`

To start this problem, download the data file below:

THE LINK TO THE DATA IS HERE

This year, the Carroll College Health Sciences Department studied a new flu vaccine in hopes that this year's flu season will be more under control than in past years. The company that developed the vaccine states that it will be equally effective for men and women.

To test the claim of the drug company the Carroll College Health Science researchers chose a simple random sample of 100 women and 200 men from the Carroll population. At the end of the semester, each person was given a binary score (1 or 0) where 1 =caught a cold and 0 =did not catch a cold.

Based on the data in the Excel spreadsheet, is there a statistical difference between the proportion of men and women who become sick at the $\alpha = 0.01$ level?

Test Statistic = _______

p-value = _______

Decision =

- ?
- We have evidence that the two groups are different
- We do not have evidence that the two groups are different
- We can make no conclusion

Answer(s) submitted:

•
•
•

(incorrect)

---

18. **(1 point)** `carroll_problib/statistics/InterStats/2SampleProp_SocialExp.pg`

A "social experiment" conducted by a TV program questioned what people do when they see a very obviously bruised woman getting picked on by her boyfriend. On two different occasions at the same restaurant, the same couple was depicted. In one scenario the woman was dressed "provocatively" and in the other scenario the woman was dressed "conservatively". The table below shows how many restaurant diners were present under each scenario, and whether or not they intervened.

	Provocatively Dressed	Conservatively Dressed	To
Intervene	9	35	4
Did Not Intervene	35	21	5
Total	44	56	1

Is there a statistical difference between the proportion of people who intervened when the woman was dressed conservatively versus when the woman was dressed provocatively? Supply a p-value and your decision below.

p-value = _________

Decision:

- A. The way the woman dresses appears to have an effect on whether or not people will intervene
- B. The way the woman dresses does not appear to have an effect on whether or not people will intervene
- C. We cannot make a decision based on this data

Answer(s) submitted:

•
•

(incorrect)

---

19. **(1 point)** `carroll_problib/statistics/InterStats/2SamplePropHTCI.pg`

1. In a study of red/green color blindness, 1000 men and 2750 women are randomly selected and tested. Among the men, 94 have red/green color blindness. Among the women, 6 have red/green color blindness. We wish to test the claim that men have a higher rate of red/green color blindness. Using the following test:

$$H_0 : p_{men} - p_{women} = 0 \quad \text{against} \quad H_A : p_{men} - p_{women} > 0$$

First find the following proportions:
$\hat{p}_{men} =$ ________
$\hat{p}_{women} =$ ________
$\hat{p}_{men} - \hat{p}_{women} =$ ________

Now we will conduct the hypothesis test:
The test statistic is ________
The p-value is ________

Is there sufficient evidence to support the claim that men have a higher rate of red/green color blindness than women using the $\alpha = 0.05$ significance level?

- A. No
- B. Yes

Next, construct the 95% confidence interval for the difference between the color blindness rates of men and women.

________ $< (p_1 - p_2) <$ ________

Which of the following is the correct interpretation for your answer in part 2?

- A. We can be 95% confident that that the difference between the rates of red/green color blindness for men and women in the sample lies in the interval
- B. There is a 95% chance that that the difference between the rates of red/green color blindness for men and women lies in the interval
- C. We can be 95% confident that the difference between the rates of red/green color blindness for men and women lies in the interval
- D. None of the above

Answer(s) submitted:

-
-
-
-
-
-
-
-
-

(incorrect)

---

20. **(1 point)** `carroll_problib/statistics/InterStats/2SampleProp HT2.pg`

Independent random samples, each containing 50 observations, were selected from two populations. The samples from populations 1 and 2 produced 23 and 18 successes, respectively.
Test $H_0 : (p_1 - p_2) = 0$ against $H_A : (p_1 - p_2) > 0$. Use $\alpha = 0.09$
(a) The test statistic is ________
(b) The P-value is ________
(c) The final conclusion is

- A. There is not sufficient evidence to reject the null hypothesis that $(p_1 - p_2) = 0$.
- B. We can reject the null hypothesis that $(p_1 - p_2) = 0$ and accept that $(p_1 - p_2) > 0$.

Answer(s) submitted:

-
-
-

(incorrect)

---

21. **(1 point)** `carroll_problib/statistics/InterStats/2SampleProp HT3.pg`

Suppose a group of 700 smokers (who all wanted to give up smoking) were randomly assigned to receive an antidepressant drug or a placebo for six weeks. Of the 336 patients who received the antidepressant drug, 122 were not smoking one year later. Of the 364 patients who received the placebo, 48 were not smoking one year later. Given the null hypothesis $H_0 : (p_{drug} - p_{placebo}) = 0$ and the alternative hypothesis $H_A : (p_{drug} - p_{placebo}) \neq 0$, conduct a test to see if taking an antidepressant drug can help smokers stop smoking. Use $\alpha = 0.01$,
(a) The test statistic is ________
(b) The P-value is ________
(c) The final conclusion is

- A. There is not sufficient evidence to determine whether the antidepressant drug had an effect on changing smoking habits after one year.
- B. There seems to be evidence that the patients taking the antidepressant drug have a different success rate of not smoking after one year than the placebo group.

Answer(s) submitted:

-
-
-

(incorrect)