Math 247: Test 3
(Wright, Fall 2018)

Name:_____

Class Time: _____

Part I: (50 points) Do this part individually. You will turn in your own copy of this work.

- *This exam is due <u>at the beginning of class</u> on Tuesday, November 4, 2018. <u>Be sure that all answers are written in your own words</u>; i.e., do not write verbatim the same answer as another student.*
- *You are welcome to work with other students in the class (and I encourage you to do so!) but please do not ask tutors or other instructors to answer the questions on the exam for you.*

Scoring will be based on <u>neatness and organization</u> of your work, <u>accuracy</u>, and thoughtful, well-written answers using complete sentences!

1. (3 pts) If you wanted to obtain a Simple Random Sample (SRS) of 30 Business majors at Cuesta College, what would you have to do? Assume there are about 400 students who are Business majors at Cuesta and use that number to describe the process you would have to go through.

   If instead you just used 30 students in your Accounting class as the sample, what would this sample be called?

   Sample of _____

   If you wandered around campus asking people whether they were Business majors, then using those people as your sample, would this be a raondom sample? Why or why not?

2. (3 pts) Deepak hosts a podcast and he is curious how much his listeners like his show. He decides to conduct an online poll. He asks his listeners to visit his website and participate in the poll. He finds 89% of 200 people who responded say they love his show.

   What type of bias does this survey have? (Specific name of type of bias)

   Is the bias positive (overestimate) or negative (underestimate)? Explain your answer.

3. (6 pts) A company with a fleet of 150 cars found that the emissions systems of 7 out of the 22 they randomly chose and tested failed to meet pollution control guidelines. They want to use this sample to test whether more than 20% of the fleet is out of compliance.

(a) What would the hypotheses be? (Write them in <u>words</u> and using appropriate <u>symbols</u>).

(b) If you wanted to use the z-Test for One Proportion, would this scenario satisfy the conditions for the test? Write down each of the conditions that must be checked and explain why or why not it is satisfied.

(c) Is it appropriate to use the z-Test for 1 Proportion in this study? Why or why not?

4. (3 pts) Donated blood is tested for infectious diseases and other contaminants. Since most donated blood is safe, it saves time and money to test batches of donated blood rather than test individual samples. A certain test is performed to see if a certain toxin is present, and the entire batch is discarded if the toxin is detected. This is like using a null and an alternative hypothesis to determine whether to discard or keep the batch. The hypotheses being tested could be stated as:

$H0$: The batch does not contain the toxin.
Ha: The batch contains the toxin.
What would be the consequence of a Type I error in this context?Choose 1 answer:

(a) The batch is discarded when it actually contains the toxin.
(b) The batch is discarded when it actually doesn't contain the toxin.
(c) The batch is kept when it actually contains the toxin.
(d) The batch is kept when it actually doesn't contain the toxin.

5. (3 pts) Regulations from the Environmental Protection Agency say that soil used in play areas should not have lead levels that exceed 400parts per million (ppm). Before beginning construction at a new site, an agent will take a sample of soil and run a significance test on the mean lead level in the soil. If the mean lead level in the sample is significantly higher than 400 ppm then the soil is deemed unsafe and construction cannot continue. Here are the hypotheses for this test:

$H0$: $\mu = 400$ ppm (soil is safe):

Ha: : $\mu > 400$ ppm (soil is unsafe)

(where μ is the mean lead level in the soil at the new site).

What would be the consequence of a Type II error in this setting? Choose 1 answer.

(a) Construction continues when the soil is actually safe.
(b) Construction stops when the soil is actually safe.
(c) Construction continues when the soil is unsafe.
(d) Construction stops when the soil is actually unsafe.

6. (10 pts) Nanda saw a report that claimed 57% of US adults primarily get their news from television. She was curious how faculty at Cal Poly get the news, so she surveyed a random sample of 100 Cal Poly professors and made a 95% confidence interval to estimate the proportion of professors who get their news from TV.

Her resulting interval was (0.44, 0.60).

(a) Interpret the confidence interval in the context of the problem.

(b) Graph the confidence interval on a number line.

(c) How many people in Nanda's sample get their news from television? (Hint: what is the center of the confidence interval?)

(d) What is the margin of error? (Again, reason this out from the confidence interval)

(e) What conclusion can Nanda make based on this confidence interval? Choose 1 answer.
   a. The proportion of Cal Poly faculty who get their news from television is significantly higher than the proportion of overall US adults.
   b. The proportion of Cal Poly faculty who get their news from television is significantly lower than the proportion of overall US adults.
   c. The proportion of Cal Poly faculty who get their news from television is not significantly different from the proportion of overall US adults.

7. (14 pts) A survey conducted five years ago by the health center at a college showed that 15% of the students smoked at the time. After implementing a smoking ban, a new survey was conducted to determine whether the percentage of smokers percentage has changed. A random sample of 200 students from the college was taken, and it was found that 21 of them smoke. Do these data provide evidence to suggest that the percentage of students who smoke now has <u>changed</u> after the implementation of the smoking ban? Use a significance level of .10.

Hypotheses (in words and using correct symbols):

Sample proportion (use correct notation):

<u>Assume the conditions are met</u> for performing a z-Test for 1 Proportion. Perform the hypothesis test using StatCrunch. Write the results below:

Hypothesis test results: (round to 3 decimal places)

Proportion	Count	Total	Sample Prop.	Std. Err.	Z-Stat	P-value

Find the Standard Error by hand (show work!) and confirm that it matches the results from StatCrunch.

The SE value is the _____ of the Sampling Distribution of p-hat.

Find the Test statistic, z, by hand and confirm that it matches the results from StatCrunch.

Sketch the sampling distribution of $\hat{p}$, with the <u>null hypothesis</u>, <u>sample proportion</u>, and <u>P-value</u> clearly illustrated and labeled. Put the <u>z-axis</u> under the $\hat{p}$ axis, and mark where the Test Statistic is.

Conclusion of Hypothesis Test (give a thorough answer!):

You should have found that the result was statistically significant at the .10 significance level. Can we then say that the smoking ban *caused* a decrease in smoking for the students at this university? Why or why not (I'm looking for a very specific answer here!)

8. (8 points) (a) Using the same data as #6, find a 90% confidence interval by hand for the proportion of students who smoke. Confirm your results using StatCrunch.

"A survey conducted five years ago by the health center at a college showed that 15% of the students smoked at the time. After implementing a smoking ban, a new survey was conducted to determine whether the percentage of smokers percentage has changed. A random sample of 200 students from the college was taken, and it was found that 21 of them smoke."

(b) Interpret the confidence interval in words in the context of the problem.

(c) Explain how the confidence interval supports the conclusion of the hypothesis test, again, in the context of the problem.

Part II: (50 points) **Making Inferences about Two Populations**

You may do this part of the exam with <u>one other student</u> in the class.

Please turn in ONE COPY with BOTH NAMES on it. Staple this cover sheet to your report.

- *This exam is due <u>at the beginning of class</u> on Thursday, 11/8/18*
- *Your work should be typed except for calculations and graphs, which can be hand-written.*
- *Your work should incorporate all of the StatCrunch work mentioned below; i.e., copy and paste the StatCrunch results into your write-up.*

Scoring will be based on <u>organization</u> of your work, <u>accuracy</u>, and thoughtful, well-written answers using complete sentences!

You are going to do a study comparing two populations of interest with regard to proportions. You will be comparing to see whether there is a significant difference in these two groups with regard to some issue (see "Examples" under #1).

o Determine a research question around two populations of interest and some question of interest (a qualitative variable).   You will ask your subjects a political/sociological/business/etc. question with a YES or NO answer. You can ask a tame question, or you can ask something edgier and more controversial but in either case **make your actual question as neutral as possible**.

> **Examples** of <u>possible</u> questions with a qualitative variable: Choose ONE or come up with your own. (I'd rather you would come up with your own question.) Voter Participation: Will/did you vote in the current election? Animal Welfare: Do you think the U.S. should have stricter animal welfare laws? Energy: Do you think the U.S government should invest more in developing alternative energy sources? Environment: Do you believe that human activity is a cause of global warming? Death Penalty: Do you support having a death penalty for certain crimes? Education Priorities: Do you think art and music should be taught in elementary school?

o Choose two different populations of interest.
> **Examples**: college athletes vs. college non-athletes; your major vs. not your major; college students vs. non-students; Cuesta students vs. Cal Poly students; people who attend church vs. people who do not attend church; men vs. women; gender neutral vs. gender identified.

o   Choose a qualitative (categorical) and quantitative variable of interest* (see below*) and then <u>write your research question.</u> Your research question is going to be whether there is a significant difference between the populations regarding each of these variables.

o Gather data from a sample of 25 people from <u>each</u> population. <u>Describe how you collected your data.</u>

o Use StatCrunch to conduct a hypothesis test for Two Proportions on your data.

o Use StatCrunch to find the 95% confidence interval for your data.

o Put all of this work together in a report, following the requirements listed below.

Your typed report, written at the college level (complete sentences, correct spelling, grammar, punctuation) should include the following: *Please number your responses as shown below.*

1. Introduction: State your research question and explain why you find this question interesting.

    Example (not to be used in your study…make up a different research question than this!): "Do athletes vs. non-athletes have significantly different voter participation rates? I find this question interesting because I think that even though student athletes are busy, they may be more community-minded and responsible, so may be more inclined to vote."

2. Sample Description: Describe how you chose the participants and how you gathered your data from them. Include the question you asked the participants, and any issues you had in getting responses. Describe any bias that your sampling method may have created. Finally, <u>organize your data in a table (or list it) and include this with the exam</u>.

3. Conduct a hypothesis test using your data, with all 4 steps clearly labeled.

    Write the hypotheses using math symbols and also in words related to the problem.

    State the test you are using, list the conditions that should be met, and check the conditions carefully. Write <u>complete sentences</u> and thoroughly explain how the problem does or does not satisfy each condition and/or any assumptions you have to make.

    For example, for the condition of "random sample" if you write just "yes" or "no" or "assume" then you won't get credit for checking that condition.

    For the conditions of "independence within samples" and "independence between samples", explain why the answers your subjects gave are or are not independent. This will tie in with how you collected your data.

    Proceed with the hypothesis test even though the conditions won't be fully satisfied by your samples.

    Computation will be performed using StatCrunch. Cut and paste the StatCrunch results into your write-up.

    Interpretation should be very thorough and should address the original research question.

4. Use StatCrunch to get a confidence interval for the <u>difference</u> in the Two Proportions. <u>Interpret</u> the confidence interval and specifically say whether or not it indicates that there is a significant difference between the two groups and <u>how you can tell</u>.

5. Write a summary of your findings including your reaction. Did your findings confirm your beliefs or were you surprised by the results? What did you learn?