Math 247: Test 4 Take Home **Name:** KEY

Class day and time: ____________________

MW class: Due at the beginning of class on Monday, 5/8/19.
TR class: Due at the beginning of class on Tuesday, 5/9/19.

I encourage you to work with other students in the class but the final work you hand in must be your own. Your answers must match every step of your work; otherwise, you may lose most or all of the points for the problem. You may consult with tutors but do not ask them to solve the problems for you!

Make sure your work is clear, legible and well organized. ***Include units!***

Test points: 50 pts organization and clarity_______/5 solutions_______/45

StatCrunch Project: 50 pts _______/50

Please read all of these instructions!
For full credit, please include the following in your solutions:

- a sketch of a normal curve (provided)
- the **axis** labeled with both z, and x or $\overline{x}$, as appropriate
- the **center** of the curve labeled with the appropriate value
- appropriate **shading** under the curve, based on the problem.

Also, be sure to LABEL all values (don't just write down a number without labeling it with a variable or a name). Include units on answers.

(4) 1. (10 pts) The owner of a rental car agency wants to estimate the mean gas mileage, in miles per gallon, of the cars in the fleet of 2000 cars. A random sample of 20 cars yields a sample mean gas mileage of 27.5 mpg with sample standard deviation of 3.5 mpg.

(a) By hand (not using StatCrunch), construct a 95% confidence interval for the mean gas mileage of the entire fleet of rental cars.

$n = 20$
$\overline{x} = 27.5$ mpg
$s = 3.5$ mpg
$df = 19$
$t^* = 2.093$
(for 95% Confidence)
$SE = \frac{s}{\sqrt{n}} = \frac{3.5}{\sqrt{20}} = .7826\ldots$

CI: point estimate ± Margin of Error
$\overline{x} \pm t^* SE$
$27.5 \pm 2.093(.7826)$
27.5 ± 1.638
$(27.5 - 1.638, 27.5 + 1.638)$
$(25.862, 29.138)$ mpg

(1) (b) Now use StatCrunch to construct the 95% confidence interval. (You don't have to include the printout for this problem.)

Report the StatCrunch Confidence Interval here (fill the values from StatCrunch, with all decimal points; i.e., don't round!

95% confidence interval results:

Mean	Sample Mean	Std. Err.	DF	L. Limit	U. Limit
μ	27.5	.78262379	19	25.86195	29.13805

(3) (c) ~~(d)~~ **Interpret the confidence intervals in words, with units, in the context of the problem.**

Based on the evidence in the sample, we are 95% confident that the mean gas mileage for the entire fleet of cars is between 25.9 and 29.1 mpg.

required: "all" or "entire" or "population" 'not ideal

(2) (d) ~~(e)~~ **The owner had thought the average gas mileage was 30 mpg. Does the work you did above support this belief? Explain.**

No, this work does not support this belief since 30 mpg was not in the confidence interval.

2. One of the controversies around nuclear power and power plants is the release of radiation into the environment, which can get into the water supply. There are strict regulations for the maximum allowable radioactivity from alpha particles in supplies of drinking water, namely an average of 15 picocuries per litre (pCi/L) or less.

Suppose a random sample of 20 specimens of water from SLO city's water supply gave a mean of 15.8 pCi/L in measured radiation with a standard deviation of 2.5 pCi/L.

(2 pts) **(a) Just looking at the sample, does it appear that the city has too much radioactivity in the water? Write a complete sentence to explain.**

Yes, it looks bad since the sample mean exceeds the acceptable limit of radiation ($\bar{x} = 15.8 > 15$ pCi/L) However, there is quite a bit of variability in the sample results ($s = 2.5$ pCi/L) so it's difficult to say whether this data will provide evidence that the water supply's radiation level is significantly higher than 15 pCi/L.

(16 pts) **(b) What hypotheses would we use to determine whether this sample actually indicates the entire water supply is out of compliance; i.e., that the mean amount of radiation in all the water supply exceeds 15 pCi/L?**

(2)

$H_0: \mu = 15$ pCi/L

The mean radiation level in the entire water supply is 15 pCi/L or less (best answer!) (Safe water)

$H_a: \mu > 15$ pCi/L

The mean radiation level is higher than 15 pCi/L (unsafe water!)

Positive (reject H_0) $\Rightarrow$ something's going on (in this case, water is unsafe!)

Negative (fail to reject H_0) $\Rightarrow$ nothing's going on (the water is fine!)

Based on the hypotheses, but before actually performing the test, describe what a False Positive (Type I) error would be <u>specifically</u> for this situation.

(3) False positive $\Rightarrow$ we conclude something's going when there really isn't.

A false positive would mean the evidence led us to conclude the water was unsafe, when in reality it was actually safe; i.e, we conclude the radiation levels <u>exceed</u> 15 pCi/L when in fact they dont.

What would the real world consequences be of making this type of error?

(2) People would be warned to not drink the water when actually the water is safe. Diablo might be fined, efforts to clean up the water would be made, etc.

Again, just based on the hypotheses, but before actually performing the test, describe what a False Negative (Type II) error would be <u>specifically</u> for this situation.

False negative $\Rightarrow$ we conclude nothing's going on when actually there is!

(3) A false negative would mean we concluded the radiation levels were not significantly above safe levels when actually the water IS above safe levels of radiation (on average)

What would the real world consequences be of making this type of error

(2) People would not be alerted to the truth that the water <u>has</u> unsafe levels of radiation; they would end up drinking unsafe water.

(2) **Which would be more serious in this situation, a Type I or Type II error? Explain your answer.**

Type II is much more serious – better that we be too cautious than not cautious enough when it comes to health and public safety.

(2) **Choose a level of significance based on your answer above.** $\alpha =$ .10

(Choose from .01, .05, .10)

Note: A smaller α makes it harder to reject the null. We want it to be <u>easier</u> to reject the null, so make α larger.

(22 pts) (c) Now perform the entire test to see whether this data provides evidence that the mean level of radiation in all the water supply exceeds 15 pCi/L. (Do your work on the next page.)

- Make sure your hypotheses are written in math symbols and in words.
- In the "Prepare" step, use the level of significance you chose in part (b). Make sure you state which test you're using and carefully check the conditions and clearly state what assumptions you have to make.
- In "Compute", you may use StatCrunch to do the calculations, but include a carefully labeled sampling distribution curve with the null, observed value, test-statistic, and P-value all clearly marked and labeled.
- In "Interpret" be sure to state what the conclusion is in the context of the problem; make sure your answer includes the word "significant" at some point.

(2) 1.

$H_0: \mu = 15$ pCi/L — The entire water supply is within the limits for safe water; i.e. the mean radiation level does not exceed 15 pCi/L

$H_a: \mu > 15$ pCi/L — The water supply is not within safe limits; the mean radiation level exceeds 15 pCi/L.

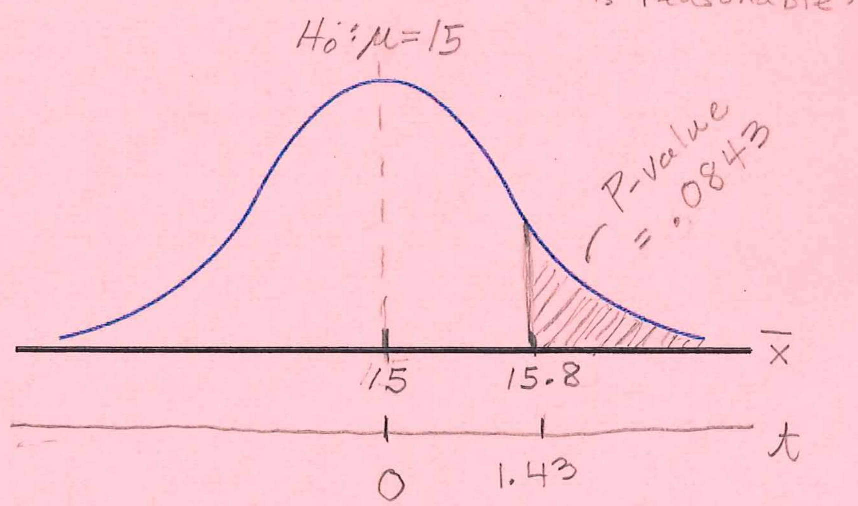
(9) 2.

Choose $\alpha = .10$
Use the 1 sample t-test
Conditions:
1. Random sample? Yes, stated.
   Independent observations? Assume.
2. Sample size is small ($n < 25$) so assume distribution of radiation measurement is normal.
3. Pop of all possible water samples $\geq 10 \cdot (20) = 200$ samples? Yes, this is reasonable!

(6) 3.

Compute using StatCrunch:

$n = 20$
$\bar{x} = 15.8$ pCi/L
$S = 2.5$ pCi/L



Hypothesis test results: (Fill in from StatCrunch results…you do not have to include a printout.)

Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
μ	15.8	.55901699	19	1.4310835	.0843

Answers will vary here!

4. (5) P-value = .0843 < .10 = α
Reject Ho, accept Ha. (POSITIVE!)
We have enough evidence to conclude the mean radiation in the entire water supply exceeds 15pCi/L so is above safe limit standard!
In short, the mean radiation level is significantly higher than the legal limit!

Note: If you chose $\alpha = .05$ or $\alpha = .01$, notice that you would fail to reject the null, leading to a conclusion that you don't have enough evidence to say the water is unsafe.
Thus, you wouldn't caution people to NOT drink the water!

StatCrunch Project: Comparing Groups by an Observational Study (50 points)

ASSIGNMENT: Staple your work (described below) to this exam.

First, pick something measurable that you want to compare between two groups.

Examples of possible comparisons between two groups:

- Do female students work more hours than male students?
- Do engineering majors sleep less than psychology majors?
- Who exercises more hours per week, students under the age of 25 or over the age of 25?
- Who studies more hours per week based on major/gender/political affiliation/ethnicity, etc.?
- Who does more volunteer work, based on major/gender/political affiliation/ ethnicity, etc.?
- Think of your own question you'd like to answer about the differences between two groups!

Next, select a sample of 10 people for each group (so two samples).

<u>**Use StatCrunch**</u> to do the following: (<u>**Include a printout of the StatCrunch work**</u>).

- Type the data into two columns and label the columns
- Construct a boxplot of each data set to check the assumption of normality.
- Use StatCrunch to find the Summary Statistics for your two samples
- Use Statcrunch for the compute step for a Two-Tailed, Two-sample t-test comparing the means from each group. Be sure include all four steps in your work! Use a significance level of .05.
- Use StatCrunch to construct a 95% confidence interval for the difference in means.

<u>Type</u> a paper that includes well-written and organized answers to the following questions:

1. What is your research question and why are you interested in this topic?
2. How did you gather your data?
3. What is the data? List the raw data, not just the summary.
4. Perform all 4 steps of the hypothesis test. Do the work as described in problem #2(c).

For step 2, be sure to thoroughly describe how your data does or does not meet the <u>conditions for the test</u>;

- o For instance, for the condition of random sample, don't write just "Random sample? No." Instead explain why your sample is not random and include positive or negative bias that may be in your sample values based on how you gathered the data.
- o For the condition of normality, include the boxplot in your discussion of whether it's reasonable, or not, to assume the population data (if you had that census data) would be normally distributed.

5. Critique your study. What did you learn that you didn't know before?