Math 247: Test 3 (Wright, Spring 2019) **Name:** KEY

_____/ 100 points

1. (6 pts) Suppose you wanted to take a poll to find out whether Cuesta students would support a new fee to fund a new Tutoring Center. You decide to gather a sample of 30 students for your poll

   (a) If you walked around campus during the day and asked 30 students whether they support the fee, you have a sample that is (circle the answer)

   Random and Representative of the Population

   Not Random but Represenatative of the Population

   Random but not Representative of the Population

   (Neither Random nor Representative of the Population)

   (b) If you asked a Random Sample of 30 students the question "Considering the fact that students already pay a lot of money to go to college, would you support having a new fee for a new Tutoring Center?", what type of bias would your sample have and would it be positive or negative bias?

   Type of Bias: Leading Question Bias

   Positive or (Negative) Bias (circle your answer)

2. (6 pts) Assume there are about 300 students who are Nursing majors at Cuesta.
   If you wanted to obtain a Simple Random Sample (SRS) of 20 Nursing majors at Cuesta College, describe the process you would go through to get the sample.

   Assign a number, 1-300 to each student
   Use a Random Number Generator to pick
   20 numbers from 1 to 300.
   These 20 people would be your random Sample.

   If instead you just used 20 students in a pre-nursing class, what would this sample be called?

   A Sample of Convenience.

3. (3 pts) Statistical inference includes which of the following (circle the correct answer):

   (a) Using a sample to prove that something is true about a population with 100% certainty

   (b) Using a sample to prove that something is false about a population with 100% certainty

   (c) Using a sample statistic to estimate a population parameter with a level of confidence that is always less than 100%.

   (d) Using a sample to prove something about the sample.

We are never 100% certain in statistical analysis (inferential)

4. (6 pts) Suppose in conducting a study, you've done everything correctly in gathering data, in doing the analysis via hypothesis testing, then in forming a conclusion based on the P-value.

There is still the possibility, due to Sampling variability, that the evidence led you to a conclusion that is incorrect. This would be an invisible error.

If the evidence led you to reject the null hypothesis, you could have made a Type I (False positive) error.

If the evidence led you to not reject the null hypothesis, you could have made a Type II (False negative) error.

5. (6 pts) Suppose a hypothesis test was done correctly, with perfect random, independent sampling and all conditions for the test satisfied. If the P-value for the test was .25, which of the following would be correct (circle **all** correct statements…there may be more than one!):

P-value = .25 > .05 = α

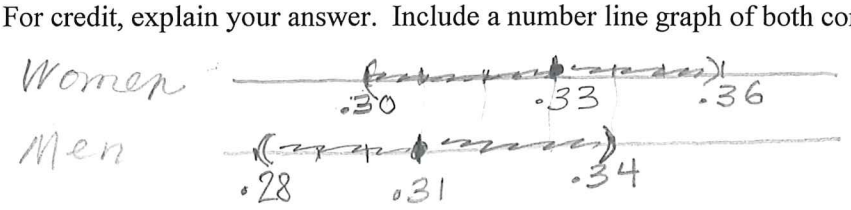
a. We would reject the null hypothesis.
b. We would accept the null hypothesis. — We never accept the null (default value)
c. We would fail to reject the null hypothesis.
d. There is a 25% chance we would get these data values and test statistic due to sampling variability, if the null is true.
e. The result of the test is statistically significant.
f. We would know with 100% certainty that the null hypothesis is true.

we are never 100% certain

6. (6 pts) A magazine article cited a poll which found that 33% of the women pet owners have dogs and 31% of the men pet owners have dogs. The margin of error was +/- 3%, with a 95% level of confidence. The magazine article had the title "Dogs Are Actually Woman's Best Friend".

Is this title correct or misleading, based on the results of the poll? I.e., can we conclude from this poll that a higher percentage of ALL women (pet owners) have dogs as compared to men pet owners?

For credit, explain your answer. Include a number line graph of both confidence intervals.

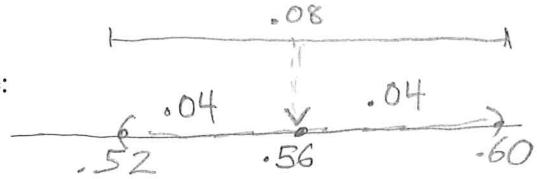


POPULATION = Proportion who own dogs

The title is misleading! Since the proportion of ALL women pet owners is between 30% and 36% and the proportion of men is between 28% and 34% (at 95% confidence level) there is NO statistically significant difference between these two groups!

7. (5 pts) A poll on a proposition showed that the proportion of all voters supporting it is between .52 and .60, using a 95% level of confidence.

(a) Graph the interval on a number line:



(b) What proportion of the sample supported the proposition? .56 = 56%

(c) What is the margin of error? .04 = 4%

8. (18 pts) An e-commerce research company claims that 32% of people who click on a particular ad eventually buy the item. Suppose a random sample of 200 people who clicked on the ad is taken and 58 of them went on to buy the item.

$n = 200$ $\quad x = 58$

8 (a) Based on the sample, find the 95% confidence interval for the proportion of ALL ad-clickers who will go on to buy the product. z* values are provided.
(StatCrunch results are given below for reference, but you MUST show work by hand for credit!)

Confidence Level	z*
80%	1.282
90%	1.645
95%	1.960
99%	2.576

$\hat{p} = \frac{x}{n} = \frac{58}{200} = .29$

$x = 58$ ad-clickers who bought the item

$SE_{est} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$= \sqrt{\frac{.29(.71)}{200}}$

$= .032$ 3 pts

CI: $\hat{p} \pm z^* SE_{est}$ 4 pts

$.29 \pm 1.96(.032)$

$.29 \pm .063$

$(.29 - .063, .29 + .063)$

$= (.227, .353)$

For reference:

95% confidence interval results:

Proportion	Count	Total	Sample Prop.	Std. Err.	L. Limit	U. Limit
p	58	200	0.29	0.032085822	0.22711294	0.35288706

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

We are 95% confident that the proportion of Ad-Clickers who went on to buy the product was between 22.7% and 35.3%

5 (c) Does the confidence interval support or not support the claim of the e-commerce research company? Briefly explain.

This confidence interval supports the company's claim since 32% is in the CI.
(the CI captures 32%)

9. (2 pts) What is the relationship between the P-value for a one-tailed test and the P-value for a two-tailed test, assuming you are using the same hypotheses and data?

this → is ½ of

Bad wording

$P\text{-value}_{\text{one-tail}} = \frac{1}{2} \cdot P\text{-value}_{\text{two tails}}$ OR $P\text{-value}_{\text{two-tails}} = 2 \cdot P\text{-value}_{\text{one-tail}}$

null

P-value for two tails = double / twice P-value for one tail

10. (2 pts) What is the "default" significance level that is most often used in hypothesis tests? $\alpha = .05$

11. (2 pts)If you want to conduct a hypothesis test at the .01 significance level, what confidence level should you use if you want to include a confidence interval in your analysis?

Confidence Level = 99%

12. (2 pts) What are the 4 steps of hypothesis testing? You can use one word for each step (no explanation required).

Step 1: Hypothesize

Step 2: Prepare

Step 3: Compute

Step 4: Interpret

13. (6 pts) What 3 conditions have to be checked for a z-Test for One Proportion hypothesis test or One Proportion Confidence Interval?

1. Random sample with independent observations
2. Large Sample: Expected successes $E = np \geq 10$ Expected failures $E = n(1-p) \geq 10$
3. Large Pop: Pop size $\geq 10 \cdot n$

We have to check conditions to make sure that the Central Limit Theorem applies.

This Theorem tells us that the Sampling Distribution of $\hat{p}$ is approximately Normal as long as the conditions are met.

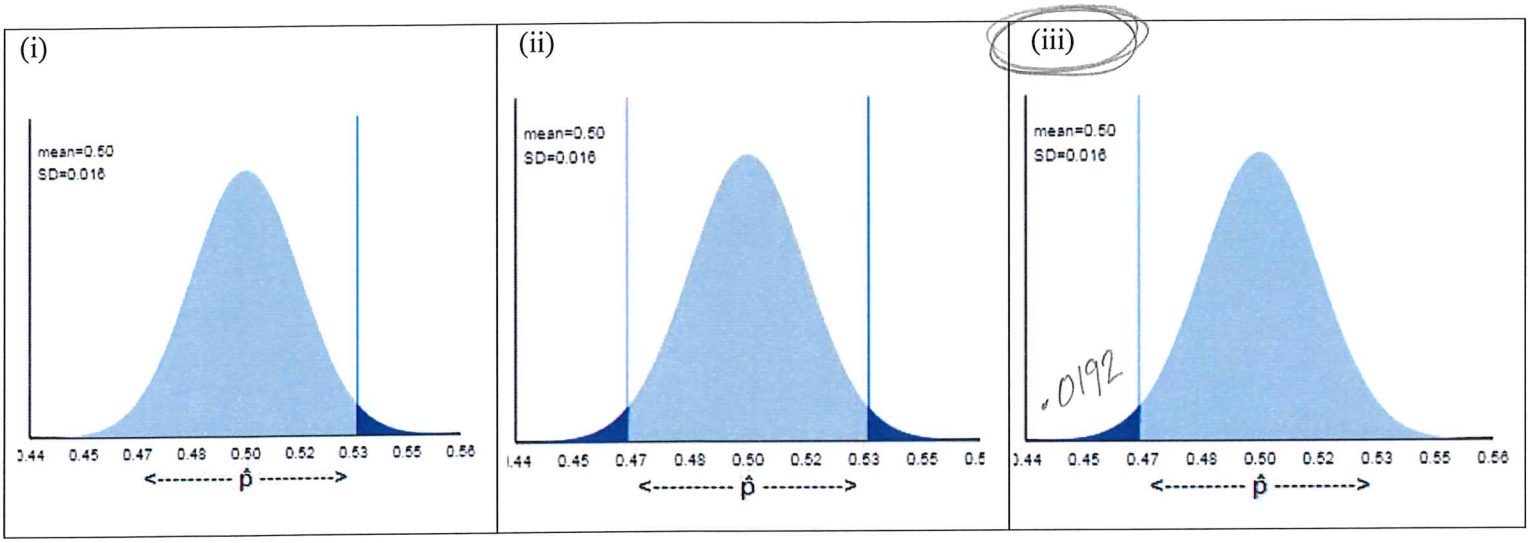
14. (12 pts) In the past, an average of 50% of employed people said they were completely or very satisfied with their jobs. Recently, the General Social Survey sampled 1016 employed people and asked them how satisfied they were with their jobs. Of the 1016 people sampled, 475 said they were completely satisfied or very satisfied with their jobs. $n = 1016$ $X = 475$ $\hat{p} = \frac{475}{1016} = .468$

3 (a) If you were to test to see whether the proportion of completely or very satisfied workers has decreased from the previous proportion of 50%, what would the hypotheses be? Write in math symbols and in words. <

	Math symbols	Words
H_0 :	$p = .50$	The proportion of satisfied workers is still 50%
H_a :	$p < .50$	The proportion has decreased.

#14: continued…

2 (b) Choose the graph below that correctly represents the P-value for the test. (No calculations necessary!)



3 (c) The P-value is .0192. What is the conclusion to the hypothesis test?

P-value = .0192 < .05 = α (default)

1 (Reject H_0, accept H_a

There is a slight chance (.0192) we would have seen a $\hat{p}$ as small as .468 if the proportion of ALL workers who are satisfied was still 50%.

2 (It's more likely (we have convincing evidence) that this proportion has decreased, i.e. there is a statistically significant decrease in the proportion of satisfied workers.

2 (d) If the alternative hypothesis was testing for whether happiness with work had *changed* instead of *decreased*, what would the P-value be?

P-value = 2·(.0192) = .0384

2 (e) What (invisible) error could we make with this conclusion? (Circle your answer)

(Type I (False Positive)) – It could be the proportion is still 50% even though the evidence made us conclude there was a decrease.

Type II (False Negative)

No error, we know that the sample proportion tells us exactly what the population proportion is.

15. (16 pts) In 2012, the *Ventura County Star* (newspaper) reported that 63% of employers allow employees to work from home sometimes. Suppose a random sample of 400 employers shows that 235 allow some work from home. Test the hypothesis that the percentage of all employers who allow some work from home is not 63%. Use a significance level of .05.

$n = 400$
$X = 235$
success = work at home
$\hat{p} = \frac{235}{400} = .588$

2 Step 1: (Write hypotheses with math symbols and in words.)

$H_0: p = .63$ — The proportion of ALL employers who allow some work at home is 63%

$H_a: p \neq .63$ — The proportion is not 63% (may be larger or smaller)

6 Step 2: Use 1 proportion Z-Test.

1. Random and independent? Random - stated; Independent - assume
2. Large sample? Success $E = np_0 = 400(.63) = 252 \geq 10$? Yes
   Failure $E = n(1-p_0) = 400(.37) = 148 \geq 10$? Yes
3. Large Pop? Pop of ALL employers $\geq 10(400) = 4{,}000$? Reasonable to assume.

4 Step 3: Find the Test statistic, z, by hand. Your work must include a sketch the sampling distribution of $\hat{p}$, with the null hypothesis, sample proportion, and P-value clearly illustrated and **labeled**. Put the z-axis under the $\hat{p}$ axis, and mark where the Test Statistic is.

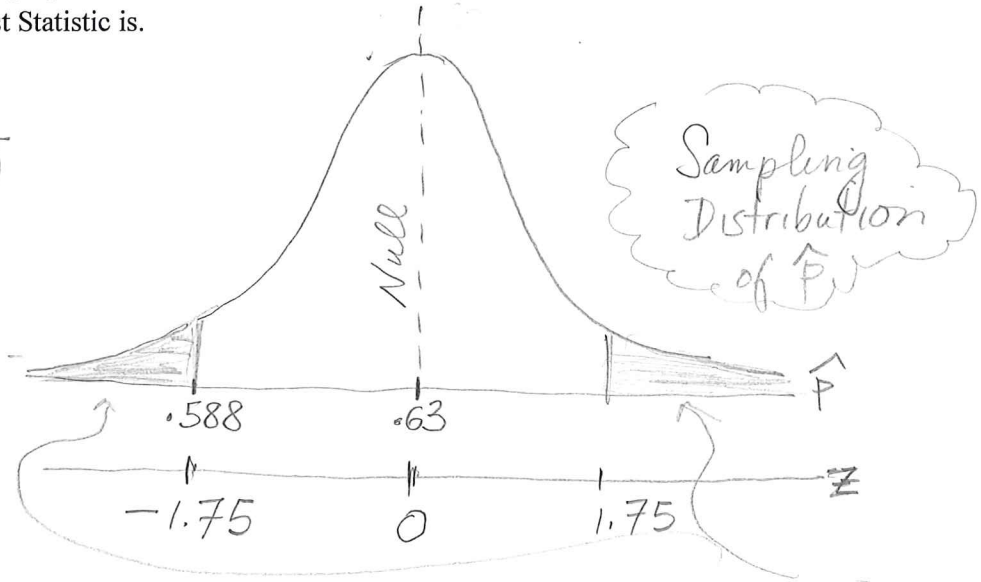
$\hat{p} = .588$

$SE = \sqrt{\frac{.63(1-.63)}{400}}$

$SE = .024$

$Z = \frac{.588 - .63}{.024}$

$= -1.75$



For reference:

Hypothesis test results:

Proportion	Count	Total	Sample Prop.	Std. Err.	Z-Stat	P-value
p	235	400	0.5875	0.024140215	-1.7605477	0.0783

4 Step 4: P-value = .0783 > .05 = α

Fail to reject H_0

There is insufficient evidence to conclude that the proportion has changed.

There has not been a significant change in the proportion of employers who allow employees to work at home.