STAT 216 Exam 2 – Fall 2025 – PRACTICE EXAM

You'll be given a copy of the official cheat sheet (also posted on D2L), which includes the "Golden Ticket" from the last page of your coursepack (page 388), as well as additional formulas during Individual Exam 1 on Wednesday, October 22, in class.

You may also bring one sheet of handwritten notes (8.5" × 11"). Feel free to include anything you'd like on your note sheet and use both sides, but please remember—it must be handwritten (no typed or printed notes).

This practice exam reflects the types of questions you might encounter on Individual Exam 2, though it does not cover every possible topic. Many of the questions are in written response format to aid your study. The actual exam will feature a question format more similar to that of assignments. The practice exam is longer than the real exam to provide additional study material.

STAT 216 Exam 2 – Fall 2025 – PRACTICE EXAM

1. Carpal Tunnel Syndrome (CTS) can be treated both surgically and with the use of splints. In a study comparing the effectiveness of each treatment, 176 volunteer CTS patients were randomly assigned to two groups of 88 subjects each. One group of subjects had surgery to alleviate symptoms associated with CTS; the other group was treated with splints. In the surgery group, 71 of the 88 patients showed an improvement in their symptoms, whereas in the splint group, only 47 of the 88 patients showed an improvement. The results are summarized in the table below. Is there evidence that there is a difference in the rates of improvement in symptoms between CTS patients that undergo surgery and those that use a splint? Use Surgery – Splint for the order of subtraction.

	Surgery	Splint	Total
Improvement	71	47	118
No Improvement	17	41	58
Total	88	88	176

a) Choose one option in each bracket below to correctly determine the study design.

This is a/an { ○ *observational study* / ○ *randomized experiment* } because { ○ *surgery or splint* / ○ *improvement or not* }

{ ○ *was* / ○ *was not* } randomly { ○ *assigned* / ○ *selected* } .

b) Write the null hypothesis, ***in words***, in the context of the problem.

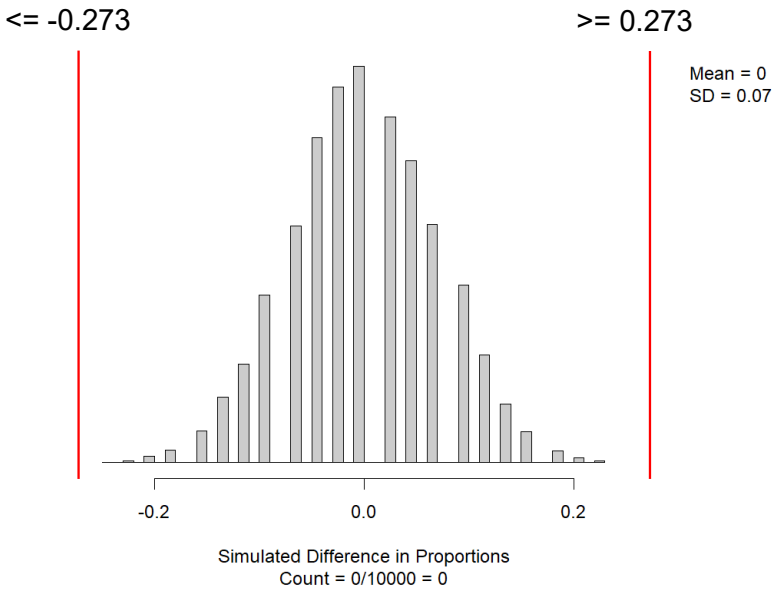
H_0:

c) Is the alternative hypothesis one- or two-sided? Explain your reasoning.

d) What is the appropriate scope of inference for this study, assuming a statistically significant difference was found?

- ○ Treatment (surgery or splint) caused the differences seen in improvement rates for all CTS patients.
- ○ Treatment (surgery or splint) caused the differences seen in improvement rates for CTS patients similar to those in the study.
- ○ There is an association between treatment (surgery or splint) and whether a patient's CTS symptoms improved for all CTS patients.
- ○ There is an association between treatment (surgery or splint) and whether a patient's CTS symptoms improved for CTS patients similar to those in the study.

e) Below is the plot of the simulated null distribution of 10000 simulations from R. Explain how one sample on the null distribution would be created.



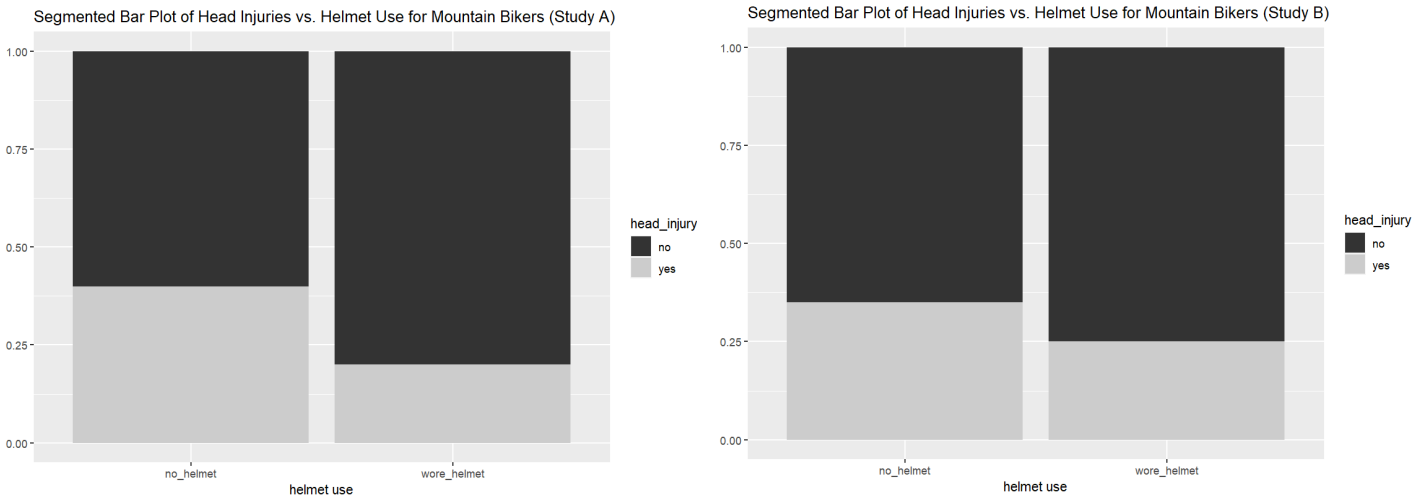
f) Interpret the p-value in context of the problem. Select one.

- ○ In less than 1 out of 1000 simulated samples, we would observe a sample difference in proportions of 0.273 or further from 0, if there is no difference in true proportion of CTS patients who show improvement in their symptoms between those who have surgery and those who use a splint.
- ○ If there is a difference in true proportion of CTS patients who show improvement in their symptoms between those who have surgery and those who use a splint, we would observe a sample difference in proportions of 0.273 or further from 0 with a probability of less than 1 out of 1000.
- ○ The probability of seeing a sample difference in proportion of CTS patients who show improvement in their symptoms between those who have surgery and those who use a splint of 0.273 or further from 0 is less than 0.1%.
- ○ The probability that there is no difference in true proportion of CTS patients who show improvement in their symptoms between those who have surgery and those who use a splint, is less than 0.1%.

g) Using the p-value provided in the R output above, write a **conclusion** in the context of the problem.

2. Researchers want to know if there is a difference in the probability of head injuries for mountain bikers involved in bike wrecks who wear helmets compared to mountain bikers involved in bike wrecks who do not wear helmets.

   a) [2 pts] Suppose two different studies (A and B) are conducted on mountain bikers involved in bike wrecks to address this research. Both study A and study B had 100 mountain bikers within each helmet use group. The data for each study are plotted below.



   Which study (A or B) provides stronger evidence that there is a difference in probability of head injuries for mountain bikers involved in bike wrecks that wear helmets compared to mountain bikers involved in bike wrecks that do not wear helmets? ***Select one.***

   - ○ Study A
   - ○ Study B
   - ○ The strength of evidence would be similar for these two studies.

   b) [2 pts] Suppose that two more studies (C and D) are conducted on this issue.
   - Study C finds that 8 of 40 (0.20) mountain bikers who wore helmets had head injuries and 12 of 40 (0.30) mountain bikers who did not wear helmets had head injuries.
   - Study D finds that 20 of 100 (0.20) mountain bikers who wore helmets had head injuries and 30 of 100 (0.30) mountain bikers who did not wear helmets had head injuries.

   Which study (C or D) provides stronger evidence that there is a difference in probability of head injuries for mountain bikers involved in bike wrecks that wear helmets compared to mountain bikers involved in bike wrecks that do not wear helmets? ***Select one.***

   - ○ Study C
   - ○ Study D
   - ○ The strength of evidence would be similar for these two studies.

   c) [2 pts] Which study (C or D) would have larger power? ***Select one.***

   - ○ Study C
   - ○ Study D
   - ○ Power will stay the same

d) [2 pts] Calculate the relative risk of head injuries for mountain bikers who wore helmets compared to those that did not wear helmets in Study C. ***Show all work!***

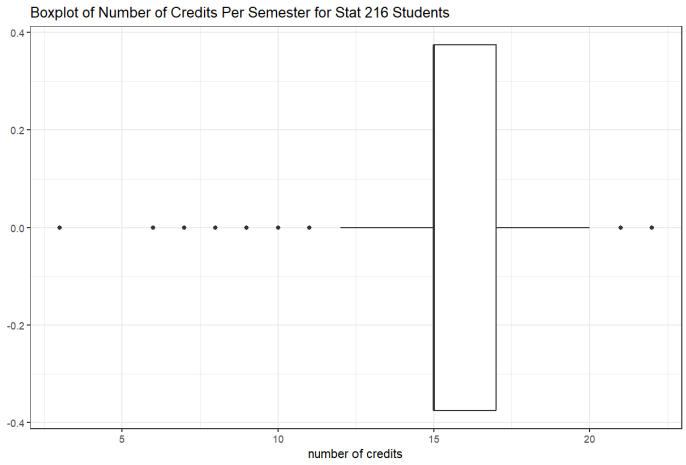
Work:

Answer:_____

e) [3 pts] Interpret the relative risk found in part d) as a percent increase/decrease ***in context of the problem.***

3. On the first day of class, we collected data on students present in Stat class that day. One of the variables measured was the number of credits students were taking this semester. In order to keep scholarships, students must take at least 15 credits each semester. Is there evidence that the number of credits Stat 216 students take per semester differs from 15 credits, on average?

min	Q1	median	Q3	max	mean	sd	n	missing
3	15	15	17	22	15.572	2.340	579	0



a) Give the appropriate alternative hypothesis in proper notation.

H_A:

b) Calculate the standardized sample mean from these data.

Work:

Answer: _____

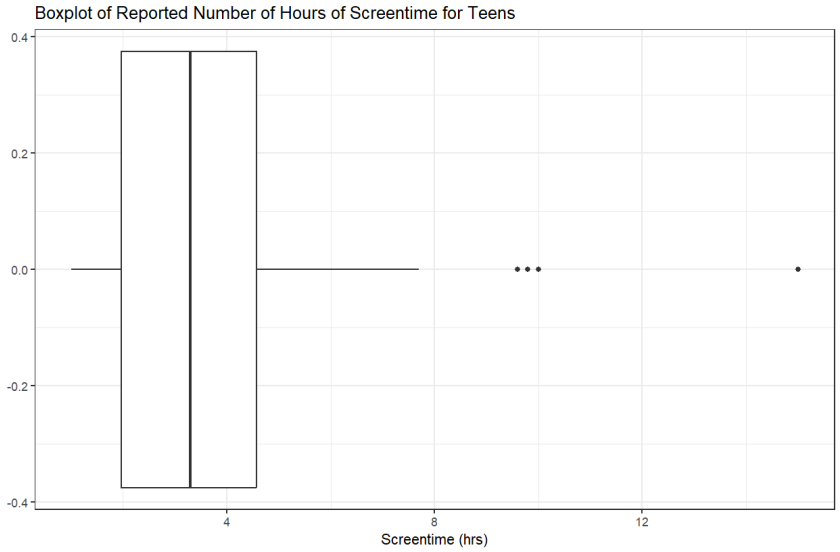
c) Interpret the standardized statistic in the context of the problem.

d) Which theoretical distribution would we use to find the p-value of this test?

e) Can these results be generalized to all MSU students? Justify your answer.

4. The CDC recommends that teens (age 13 – 17) should limit their screentime per day to at most 2 hours per day. A survey sent to public high schools in Gallatin County asked about students' screen use. Results were reported on 88 students. Is there evidence that Gallatin County public high school students report a higher mean number of hours of screentime per day than 2 hours?

min	Q1	median	Q3	max	mean	sd	n	missing
1	1.975	3.3	4.575	15	3.8	2.642	88	0



a) Describe the distribution of screentime using the four characteristics of boxplots.

b) Interpret the value of 4.575 from the provided R output above in context of the problem.

c) Show mathematically why the maximum value of 15 is an outlier.

d) Interpret the sample standard deviation in context of the problem.

e) What does μ represent in context of the study?

f) Is the sample size large enough to use theory-based methods to analyze these data? Explain how you know the condition is or is not met in the context of the study.

g) Regardless of your answer to f), using a multiplier of 2.634, calculate the margin of error for a 99% confidence interval for the parameter of interest.

Work:

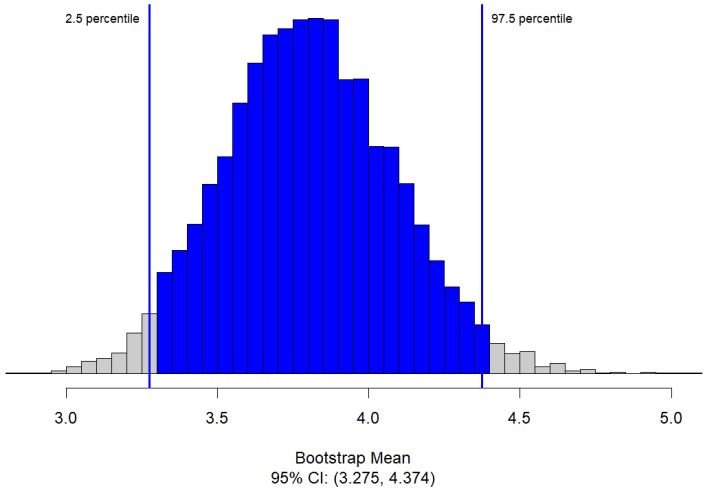
Answer: _____

h) Use your answer to part g) to calculate a theory-based 99% confidence interval for the parameter of interest.

Work:

Answer (written as an interval):

A simulated bootstrap distribution of 10,000 sample means is shown below.



i) Interpret the confidence interval from the bootstrap distribution shown above in context of the problem.

5. In 2018, FiveThirtyEight surveyed a random sample of 1615 American adults (18+ years old) who identified as men in an effort to understand how male gender identity is formed, and how it has changed over the years. Among the questions asked was the following: "Do you think that society puts pressure on men in a way that is unhealthy or bad for them?" FiveThirtyEight noted that there were generational differences in the responses to this question, with 70% of respondents younger than 35 answering yes, compared to only 55% of participants older than 35 saying the same. Researchers would like to know if these data provide evidence of a difference in perceptions about masculinity and society between the two age groups? Use order of subtraction 18 – 35 years old – 35+ years old.

	18 – 35 years old	35+ years old	Totals
Yes (or agree)	328	631	959
No (or disagree)	140	516	656
Column Totals	468	1147	1615

a) What are the observational units (cases) for this study?

b) Choose one answer in each set of brackets to correctly identify the type and role of each variable.

Whether or not the respondent agrees or disagrees is the $\{\bigcirc\ explanatory \;\; \bigcirc\ response\}$ variable and it is $\{\bigcirc\ categorical \;\; \bigcirc\ quantitative\}$.

Age group is the $\{\bigcirc\ explanatory \;\; \bigcirc\ response\}$ variable and it is $\{\bigcirc\ categorical \;\; \bigcirc\ quantitative\}$.

c) Calculate the difference in sample proportions. Show your work, round your answer to three decimal places, and give the appropriate notation (including informative subscripts if required).

Work:

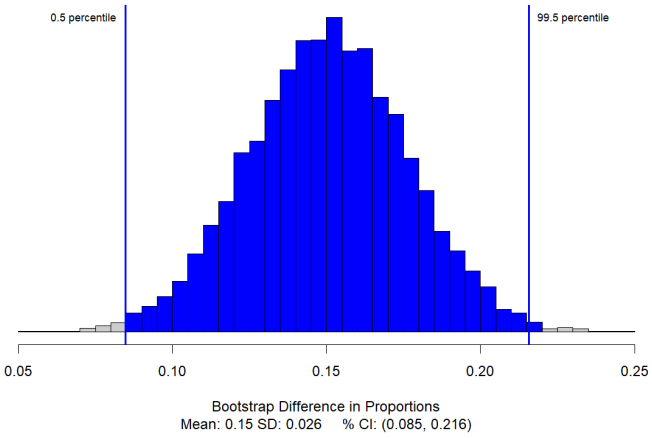
Value of statistic: _____

Notation: _____

d) Below is the bootstrap distribution of 10000 simulations for these data. Where is the distribution centered? Explain why that makes sense.

Center:

Explanation:



e) How much confidence should you have that the interval shown above, (0.085, 0.216), contains the parameter?

f) Based on this confidence interval, can we conclude that age **causes** the differences seen in perceptions of masculinity?

g) Based on the confidence interval, what type of error may have occurred? Explain why.

h) Write the error in context of the study.