STA 1020 – Spring/Summer 2024 – Final Exam

Total: 150 points

Directions:

Work each problem on a separate sheet of paper. In order to receive full credit, be sure to show all work including any relevant steps. Use units where appropriate. Calculators are permitted. Make sure to write clearly and label each problem number and part. If I cannot read something, it will be marked wrong.

It is expected that you will work this exam on your own, which means that you will not consult with current classmates, including those in other sections of STA 1020, or with former students, or anyone in general. Any questions should be directed to your instructor. Further, the use of outside sources (including but not limited to: Chegg, Wolfram Alpha, Bartleby, any artificial intelligence (AI) system, etc.) is strictly prohibited.

You may use your lecture notes, Hawkes, your instructors lecture videos/notes, cheat sheets, the eBook, and any relevant materials posted on our Canvas site. You may also review the lectures as needed.

Time limit:

The Final Exam is available from 12 am until 11:59 pm on Monday, July 29th. So, if you begin the exam at 10:59 pm, you will only have 1 hour to complete it.

Make sure you allow enough time to work all problems AND scan and upload your work to Canvas. No email submissions, please use Canvas. And there is no need to write out the problem.

If you run into any technical problems, email YOUR instructor right away.

At the top of your exam paper, please print the following statement and sign your name:

I affirm that the work I am turning in is mine, and mine alone and that I did not, in any way, breach the academic integrity expected of me. (Sign your name)

Failure to include this statement could result in a deduction of points or refusal to grade your final exam.

1. [10] A high school counselor is interested in the proportion of U.S. graduating seniors who plan to attend a 4-year university after high school. She decides to sample the high school seniors at her school. From this year's graduating class, 77% of them plan to attend a 4-year university.
   a. [2] What is the population of interest?
   b. [2] What is the sample?
   c. [2] What is the parameter?
   d. [2] What is the statistic?
   e. [2] Is this an observational study or an experiment?

2. [12] The distribution below shows the time (in hours) students spent playing video games on weekends.
   a. [1] Classify this data as quantitative or qualitative.
   b. [1] Classify this data as continuous, discrete, or neither.
   c. [2] What type of graph is this?
   d. [2] Describe the shape of the distribution.
   e. [2] How many students play less than 15 hours of video games on weekends?
   f. [2] What time range contains the mode for this data?
   g. [2] For this graph, which is larger, the mean or the median?

Hours Spent Playing Video Games on Weekends

3. [12] The following stem-and-leaf plot represents the test scores for 22 students in a class on their most recent test.

Test Scores by Student

Stem	Leaves							
6	1	4	4	5	8	8		
7	4	5	6	9				
8	3	4	6	7	9			
9	4	5	6	6	8	9	9	

Key: $6 \mid 1 = 61$

a) [6] Which test score is the 14th percentile?
b) [6] What percentile is represented by a score of 84? **Round to the nearest whole percentage.**

4. [8] A box contains 18 large marbles and 10 small marbles. Each marble is either green or white. 12 of the large marbles are green, and 5 of the small marbles are white. If a marble is randomly selected from the box, what is the probability that it is small or green? **Express your answer as a fraction or a decimal number rounded to four decimal places.**

5. [4] Suppose you like to keep a jar of change on your desk. Currently, the jar contains the following: 18 pennies, 24 dimes, 25 quarters, 7 nickels. What is the probability that you reach into the jar and randomly grab a dime and then, *without replacement*, a nickel? **Express your answer as a fraction or a decimal number rounded to four decimal places.**

6. Mrs. Griese's literature class has 86 students, classified by academic year and gender, as illustrated in the table. Mrs. Griese randomly chooses one student to collect yesterday's work.

Mrs. Griese's Literature Class		
	Males	**Females**
Freshmen	18	6
Sophomores	18	6
Juniors	8	7
Seniors	17	6

[6] What is the probability that she selects a female, given that she chooses randomly from only the sophomores? **Express your answer as a fraction or a decimal rounded to four decimal places**.

7. [10] Below is a valid discrete probability distribution, answer the questions that follow.

X	1	2	3	4	5	6	7
$P(X = x)$	0.12	0.1	??	0.34	0.22	0.03	0.11

    a. [4] Find $P(X = 3)$
    b. [2] Find $P(X \geq 5)$
    c. [2] Find $P(X < 3)$
    d. [2] Find $P(X > 7)$

8. [20] Suppose that the mean commute time of off-campus students is 26 minutes with a standard deviation of 8 minutes. Assume that commute times are approximately normal.
    a. [6] What is the probability that it takes a student more than 20 minutes? **Round answers to 4 decimal places.**
    b. [6] How long would a student's commute time need to be for the student to be in the 90th percentile? **Round answers to 1 decimal place.**
    c. [8] If the sample size is 12 students, what is the probability that sample mean differs from the population mean by less than 5 minutes? **Round answers to 4 decimal places.**

9. [14] Suppose a sample of 310 tenth graders is drawn. Of the students sampled, 55 read at or below the eighth grade level.
    a. [2] Using the data, estimate the proportion of 10th graders reading at or below the 8th grade level. **Give your answer as a fraction or a decimal number rounded to 4 decimal places.**
    b. [8] Using the data, construct the 85% confidence interval for the population proportion of 10th graders who read at or below the 8th grade level. **Round your answers to 4 decimal places.**
    c. [4] Interpret the confidence level you found in part (b)

10. [4] Lisa is confused. She inadvertently deleted the confidence levels for two confidence intervals that she computed. One is supposed to be 85% whereas the other is supposed to be 99%. Which is which, and how do you know? (this means give a brief explanation)
CI #1: $(43.5, 49.7)$ CI #2: $(40.9, 52.3)$

11. [6] A radio station states that the mean age of its listeners is 23. One radio station worker claims that the mean age is higher than 23. The radio station worker conducts a hypothesis test and rejects the null hypothesis. In reality, the mean age of its listeners is 23. Was an error made? If so, what type?

12. [12] A hospital director is told that 33% of the treated patients are uninsured. The director wants to test the claim that the percentage of uninsured patients is under the expected percentage. A sample of 260 patients found that 65 were uninsured. At the 0.01 level, is there enough evidence to support the director's claim? **Show all steps to receive credit.**

13. [10] A political researcher wishes to know if political affiliation and age are related. He has collected data on 291 people in three age categories as shown below. Assume an $\alpha = 0.05$ level of significance is used to test whether age and political affiliation are related.

Age	Democrat	Republican	Independent	Total
18-31	37	27	13	77
32-51	44	33	18	95
52-68	31	59	29	119
Total	112	119	60	291

   a. [2] State the hypotheses for this test.
   b. [4] What is the expected count for the age group 32-51 and Independent? **Round to 4 decimal places.**
   c. [2] State the degrees of freedom for this test.
   d. [2] If the decision was to reject the null hypothesis, what would be the conclusion for this test?

14. [4] Suppose that you are told that the correlation between GPA and ACT scores is $r = 0.62$. What percent of the variation (change) in ACT scores is due to the regression model? **Express your answer as a percent and round the percent to the nearest tenth.**

15. [18] Consider the following data and answer the questions that follow.

Caffeine Consumption (1 unit = 1 cup of coffee)	1	2	5	2	1	13	21
Avg. Hrs of Sleep	9	9	8	6	7	6	4

   a. [2] Which variable is the response variable?
   b. [10] Create a scatterplot for this data.
   c. [2] For this data, estimate the coefficient correlation (r) in words (i.e., positive, negative, no correlation).
   d. [4] Suppose it is known the slope of the LSR equation is -0.311. Interpret what this value means in terms of the context of these variables.