STAT 200
OL4 Sections
Final Exam
Fall 2019

The final exam will be posted at 12:01 am on December 13, 2019, and it is due at 11:59 pm on December 15, 2019 Eastern Time.

This is an open-book exam. You may refer to your text and other course materials for the current course as you work on the exam, and you may use a calculator, applets, or Excel. You must complete the exam individually. Neither collaboration nor consultation with others is allowed. It is a violation of the UMUC Academic Dishonesty and Plagiarism policy to use unauthorized materials or work from others.

Answer all 20 questions. Make sure your answers are as complete as possible, particularly when it asks for you to show your work. Answers that come straight from calculators, programs or software packages without any explanation will not be accepted. If you need to use technology (for example, Excel, online or hand-held calculators, statistical packages) to aid in your calculation, you must cite the sources and explain how you get the results. For example, state the Excel function along with the required parameters when using Excel; describe the detailed steps when using a hand-held calculator; or provide the URL and detailed steps when using an online calculator, and so on.

Record your answers and work on the separate answer sheet provided.
This exam has 20 problems; 5% for each problems.

You must include the Honor Pledge on the title page of your submitted final exam. Exams submitted without the Honor Pledge will not be accepted.

1. The World Health Organization wishes to estimate the mean density of people per square kilometer, they collect data on 56 countries. *Justify for full credit.*

(a) Which of the following is the variable?

(i) All countries in the world
(ii) Density of people per square kilometer
(iii) Set of densities of people per square kilometer of all countries
(iv) Set of densities of people per square kilometer of 56 countries

(b) Which of the following is the population?

(i) All countries in the world
(ii) Density of people per square kilometer
(iii) Set of densities of people per square kilometer of all countries
(iv) Set of densities of people per square kilometer of 56 countries

2. Choose the best answer. *Justify for full credit.*

(a) The hotel ratings are usually on a scale from 0 star to 5 stars. The level of this measurement is

(i) interval
(ii) nominal
(iii) ordinal
(iv) ratio

(b) In a career readiness research, 100 students were randomly selected from the psychology program, 150 students were randomly selected from the communications program, and 120 students were randomly selected from cyber security program. This type of sampling is called:

(i) cluster
(ii) convenience
(iii) systematic
(iv) stratified

3. True or False. Justify for full credit.

(a) If the variance from a data set is zero, then all the observations in this data set must be identical.

(b) The median of a normal distribution curve is always zero.

4. A school district wanted to assess the effectiveness of a new math readiness program for fifth graders. The school district is divided into the individual fifth grade classrooms and 10 classrooms are randomly selected. All of the children in each of the 10 selected classrooms are assessed.

(a) What type of sampling method is being used?

(b) Please explain your answer.

5. A study was conducted to determine whether the mean braking distance of four-cylinder cars is greater than the mean braking distance of six-cylinder cars. A random sample of 20 four-cylinder cars and a random sample of 20 six-cylinder cars were obtained, and the braking distances were measured.

(a) What would be the appropriate hypothesis test for this analysis?
- (i) t-test for two independent samples
- (ii) t-test for dependent samples
- (iii) z-test for population mean
- (iv) correlation

(b) Explain the rationale for your selection in (a). Specifically, why would this be the appropriate statistical approach?

6. A study of 10 different weight loss programs involved 500 subjects. Each of the 10 programs had 50 subjects in it. The subjects were followed for 12 months. Weight change for each subject was recorded. The researcher wants to test the claim that all ten programs are equally effective in weight loss.

(a) Which statistical approach should be used?
- (i) confidence interval
- (ii) t-test
- (iii) ANOVA

(iv) Chi square

(b) Explain the rationale for your selection in (a). Specifically, why would this be the appropriate statistical approach?

7. A STAT 200 professor took a sample of 10 midterm exam scores from a class of 30 students. The 10 scores are shown in the table below:

95	67	76	47	85	70	87	80	67	72

(a) What is the sample mean?

(b) What is the sample standard deviation? (Round your answer to two decimal places)

(c) If you leveraged technology to get the answers for part (a) and/or part (b), what technology did you use? If an online applet was used, please list the URL, and describe the steps. If a calculator or Excel was used, please write out the function.

8. UMUC Stat Club must appoint a president, a vice president, and a treasurer. There are 10 qualified candidates.

(a) How many different ways can the officers be appointed?

(b) Please describe the method used and the reason why it is appropriate for answering the question. *Just the answer, without the description and reason, will receive no credit.*

9. Sara has eight new summer outfits. She plans to pack three of the new summer outfits in her trip to Tokyo.

(a) How many different ways can the three summer outfits be selected?

(b) Please describe the method used and the reason why it is appropriate for answering the question. Just the answer, without the description and reason, will receive no credit.

10. There are 4 suits (heart, diamond, clover, and spade) in a 52-card deck, and each suit has 13 cards. Suppose your experiment is to draw one card from a deck and observe what suit it is. Express the probability in fraction format. (*Show all work. Just the answer, without supporting work, will receive no credit.)*

(a) Find the probability of drawing a heart or diamond.

(b) Find the probability that the card is not a spade.

11. Let random variable *x* represent the number of heads when a fair coin is tossed two times.

(a) Construct a table describing the probability distribution.

x	P(x)
0	
1	
2	

(b) Determine the mean and standard deviation of *x*. *Show all work. Just the answer, without supporting work, will receive no credit.*

12. Mimi plans make a random guess at 10 true-or-false questions. Answer the following questions:

(a) Let X be the number of correct answers Mimi gets. As we know, the distribution of X is a binomial probability distribution. What is the number of trials (n), probability of successes (p) and probability of failures (q), respectively?

(b) Find the probability that she gets at most 5 correct answers. (Round the answer to 3 decimal places.

Refer to the following information for Questions 13 and 14.

The heights of pecan trees are normally distributed with a mean of 10 feet and a standard deviation of 2 feet.

13. Show all work. Just the answer, without supporting work, will receive no credit.

(a) What is the probability that a randomly selected pecan tree is between 9 and 12 feet tall? (Round the answer to 4 decimal places)

(b) Find the 75th percentile of the pecan tree height distribution. (Round the answer to 2 decimal places)

14. Show all work. Just the answer, without supporting work, will receive no credit.

(a) For a sample of 64 pecan trees, state the standard deviation of the sample mean (the "standard error of the mean"). (Round your answer to three decimal places)

(b) Suppose a sample of 64 pecan trees is taken. Find the probability that the sample mean heights is between 9.5 and 10 feet. (Round your answer to four decimal places)

15. A survey showed that 720 of the 1000 adult respondents believe in global warming.

(a) Construct a 90% confidence interval estimate of the proportion of adults believing in global warming. (Round the lower bound and upper bound of the confidence interval to three decimal places) *Include description of how confidence interval was constructed.*

(b) Describe the results of the survey in everyday language.

16. A city built a new parking garage in a business district. For a random sample of 64 days, daily fees collected averaged $2,000, with a standard deviation of $400.

(a) Construct a 90% confidence interval estimate of the mean daily parking fees collected. (Round the lower bound and upper bound of the confidence interval to two decimal places) *Include description of how confidence interval was constructed.*

(b) Describe the confidence interval in everyday language.

17. The UMUC MiniMart sells five different types of teddy bears. The manager reports that the five types are equally popular. Suppose that a sample of 100 purchases yields observed counts 25, 19, 15, 17, and 24 for types 1, 2, 3, 4, and 5, respectively**.**

Type	1	2	3	4	5
Number	25	19	15	17	24

Use a 0.10 significance level to test the claim that the five types are equally popular.

(a) Identify the appropriate hypothesis test and explain the reasons why it is appropriate for analyzing this data.

(b) Identify the null hypothesis and the alternative hypothesis.

(c) Determine the test statistic. (Round your answer to two decimal places)

(d) Determine the p-value. (Round your answer to two decimal places)

(e) Compare p-value and significance level α. What decision should be made regarding the null hypothesis (e.g., reject or fail to reject) and why?

(f) Is there sufficient evidence to support the claim that the five types are equally popular? Justify your answer.

18. David was curious if regular excise really helps weight loss, hence he decided to perform a hypothesis test. A random sample of 5 UMUC students was chosen. The students took a 30-minute exercise every day for 6 months. The weight was recorded for each individual before and after the exercise regimen. Does the data below suggest that the regular exercise helps weight loss? Assume David wants to use a 0.05 significance level to test the claim.

	Weight (pounds)	
Subject	Before	After
1	190	180
2	170	160
3	185	190
4	160	160
5	200	190

(a) What is the appropriate hypothesis test to use for this analysis: z-test for two proportions, t-test for two proportions, t-test for two dependent samples (matched pairs), or t-test for two independent samples? Please identify and explain why it is appropriate.

(b) Let $\mu 1$ = mean weight before the exercise regime. Let $\mu 2$ = mean weight after the exercise regime. Which of the following statements correctly defines the null hypothesis?

(i) $\mu 1 - \mu 2 > 0$ ($\mu d > 0$)

(ii) $\mu 1 - \mu 2 = 0$ ($\mu d = 0$)

(iii) $\mu 1 - \mu 2 < 0$ ($\mu d < 0$)

(c) Let $\mu 1$ = mean weight before the exercise regime. Let $\mu 2$ = mean weight after the exercise regime. Which of the following statements correctly defines the alternative hypothesis?

(a) $\mu 1 - \mu 2 > 0$ ($\mu d > 0$)

(b) $\mu 1 - \mu 2 = 0$ ($\mu d = 0$)

(c) $\mu 1 - \mu 2 < 0$ ($\mu d < 0$)

(d) Determine the test statistic. Round your answer to three decimal places. *Show all work; writing the correct test statistic, without supporting work, will receive no credit.*

(e) Determine the p-value. Round your answer to three decimal places. *Show all work; writing the correct critical value, without supporting work, will receive no credit.*

(f) Compare p-value and significance level α. What decision should be made regarding the null hypothesis (e.g., reject or fail to reject) and why?

(g) Is there sufficient evidence to support the claim that regular exercise helps weight loss? Justify your conclusion.

19. A grocery store manager is interested in testing the claim that banana is the favorite fruit for more than 50% of the adults. The manager conducted a survey on a random sample of 100 adults. The survey showed that 56 adults in the sample chose banana as his/her favorite fruit. Assume the manager wants to use a 0.05 significance level to test the claim.

(a) What is the appropriate hypothesis test to use for this analysis? Please identify and explain why it is appropriate.

(b) Identify the null hypothesis and the alternative hypothesis.

(c) Determine the test statistic. Round your answer to two decimal places. *Describe method used for obtaining the test statistic.*

(d) Determine the p-value. Round your answer to three decimal places. *Describe method used for obtaining the p-value.*

(e) Compare p-value and significance level α. What decision should be made regarding the null hypothesis (e.g., reject or fail to reject) and why?

(f) Is there sufficient evidence to support the claim that banana is the favorite fruit for more than 50% of the adults? Explain your conclusion.

20. A business analyst believes that December holiday sales in 2016 are a good predictor of December holiday sales in 2017. A random sample of 8 toys stores produced the following data

where *x* is the amount of December holiday sales in 2016 and *y* is the amount of December sales in 2017, in dollars.

x	*y*
10257	11689
6556	6438
7224	8662
9987	9454
11568	12004
8453	8021
4235	6048
5576	4850

(a) Find an equation of the least squares regression line. Round the slope and y-intercept value to two decimal places. *Describe method for obtaining results.*

(b) Based on the equation from part (a), what is the predicted 2017 December holiday sales if the 2016 December holiday sales is 6,000 dollars? *Show all work and justify your answer.*

(c) Based on the equation from part (a), what is the predicted 2017 December holiday sales if the 2016 December holiday sales is 20,000 dollars? *Show all work and justify your answer.*

(d) Which predicted 2017 holiday sales that you calculated for (b) and (c) do you think is closer to the true predicted 2017 holiday sales and why?