

1. True or False. *Justify for full credit.* (25 pts)
- (a) If all the observations in a data set are identical, then the variance for this data set is zero.
 - (b) If $P(A) = 0.4$ and $P(B) = 0.5$, then $P(A \text{ AND } B) = 0.2$.
 - (c) The mean is always equal to the median for a normal distribution.
 - (d) A 95% confidence interval is wider than a 90% confidence interval of the same parameter.
 - (e) In a two-tailed hypothesis testing at significance level α of 0.05, the test statistic is calculated as 2. If $P(X > 2) = 0.03$, then we have sufficient evidence to reject the null hypothesis.

Refer to the following frequency distribution for Questions 2, 3, 4, and 5. *Show all work. Just the answer, without supporting work, will receive no credit.*

A random sample of 100 students was chosen from UMUC STAT 200 classes. The frequency distribution below shows the distribution for study time each week (in hours).

Study Time (in hours)	Frequency	Relative Frequency
0.0 – 4.9	5	
5.0 - 9.9	13	
10.0 - 14.9		0.22
15.0 -19.9	42	
20.0 – 24.9		
Total	100	

- 2. Complete the frequency table with frequency and relative frequency. (5 pts)
- 3. What percentage of the study times was at least 15 hours? (5 pts)
- 4. In what class interval must the median lie? Explain your answer. (5 pts)
- 5. Does this distribution have positive skew or negative skew? Why? (5 pts)

Refer to the following information for Questions 6, 7, and 8. *Show all work. Just the answer, without supporting work, will receive no credit.*

A fair 6-faced die is rolled two times. Let A be the event that the outcome of the first roll is a multiple of 3, and B be the event that the outcome of second roll is greater than 4.

- 6. How many outcomes are there in the sample space? (5 pts)
- 7. What is the probability that the outcome of the second roll is greater than 4, given that the first roll is a multiple of 3? (10 pts)

8. Are A and B independent? Why or why not? (5 pts)
-

Refer to the following situation for Questions 9, 10, and 11.

The five-number summary below shows the grade distribution of two STAT 200 quizzes.

	Minimum	Q1	Median	Q3	Maximum
Quiz 1	12	40	60	95	100
Quiz 2	20	35	50	90	100

For each question, give your answer as one of the following: (a) Quiz 1; (b) Quiz 2; (c) Both quizzes have the same value requested; (d) It is impossible to tell using only the given information. Then explain your answer in each case. (5 pts each)

9. Which quiz has less interquartile range in grade distribution?
 10. Which quiz has the greater percentage of students with grades 90 and over?
 11. Which quiz has a greater percentage of students with grades less than 60?
-

Refer to the following information for Questions 12 and 13. *Show all work. Just the answer, without supporting work, will receive no credit.*

There are 1000 juniors in a college. Among the 1000 juniors, 300 students are taking STAT200, and 150 students are taking PSYC300. There are 50 students taking both courses.

12. What is the probability that a randomly selected junior is taking at least one of these two courses? (10 pts)
 13. What is the probability that a randomly selected junior is taking STAT200, given that he/she is taking PSYC300? (10 pts)
-

14. There are 8 books in the "Statistics is Fun" series. Mimi would like to choose 2 books from the series for her summer reading. How many different ways can the two books be selected? (5 pts)
-

15. Let random variable x represent the number of girls in a family of three children.
 (a) Construct a table describing the probability distribution. (5 pts)
 (b) Determine the mean and standard deviation of x . (Round the answer to two decimal places) (10 pts)
-

16. Rabbits like to eat the cucumbers in Mimi's garden. There are 10 cucumbers in her garden which will be ready to harvest in about 10 days. Based on her experience, the probability of a cucumber being eaten by the rabbits before harvest is 0.60. *Show all work. Just the answer, without supporting work, will receive no credit.*

- (a) Let X be the number of cucumbers that Mimi harvests (that is, the number of cucumbers not eaten by rabbits). 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? (5 pts)
 - (b) Find the probability that Mimi harvests at least 2 of the 10 cucumbers. (round the answer to 3 decimal places) (10 pts)
 - (c) How many cucumbers can she expect to harvest? (5 pts)
-

Refer to the following information for Questions 17, 18, and 19. *Show all work. Just the answer, without supporting work, will receive no credit.*

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

- 17.** What is the probability that a randomly selected pecan tree is between 9 and 12 feet tall? (10 pts)
 - 18.** Find the 80th percentile of the pecan tree height distribution. (5 pts)
 - 19.** If a random sample of 225 pecan trees is selected, what is the standard deviation of the sample mean? (5 pts)
-

20. A random sample of 100 SAT scores has a sample mean of 1500. Assume that SAT scores have a population standard deviation of 300. Construct a 95% confidence interval estimate of the mean SAT scores. *Show all work. Just the answer, without supporting work, will receive no credit.* (10 pts)

21. Consider the hypothesis test given by

$$H_0 : p = 0.5$$

$$H_1 : p < 0.5$$

In a random sample of 100 subjects, the sample proportion is found to be $\hat{p} = 0.47$.

- (a) Determine the test statistic. *Show all work; writing the correct test statistic, without supporting work, will receive no credit.*
- (b) Determine the p -value for this test. *Show all work; writing the correct P -value, without supporting work, will receive no credit.*
- (c) Is there sufficient evidence to justify the rejection of H_0 at the $\alpha = 0.01$ level? Explain. (15 pts)

22. Mimi was curious if regular exercise really helps weight loss, hence she 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?

Subject	Weight (pounds)	
	Before	After
1	159	130
2	170	160
3	185	180
4	165	165
5	200	190

Assume we want to use a 0.01 significance level to test the claim.

- Identify the null hypothesis and the alternative hypothesis.
- Determine the test statistic. *Show all work; writing the correct test statistic, without supporting work, will receive no credit.*
- Determine the p-value. *Show all work; writing the correct critical value, without supporting work, will receive no credit.*
- Is there sufficient evidence to support the claim that regular exercise helps weight loss? Justify your conclusion.

(20 pts)

23. A STAT 200 instructor is interested in whether there is any variation in the final exam grades between her two classes. Data collected from the two classes are as follows:

Section 1: $n_1 = 31$, $\bar{x}_1 = 78$, $s_1 = 10$
 Section 2: $n_2 = 30$, $\bar{x}_2 = 72$, $s_2 = 14$

Her null hypothesis and alternative hypothesis are:

$$H_0: \sigma_1^2 = \sigma_2^2, \quad \text{and} \quad H_a: \sigma_1^2 < \sigma_2^2$$

- Determine the test statistic. *Show all work; writing the correct test statistic, without supporting work, will receive no credit.*
- Determine the p-value for this test. *Show all work; writing the correct P-value, without supporting work, will receive no credit.*
- Is there sufficient evidence to justify the rejection of H_0 at the significance level of 0.05? Explain.

(10 pts)

-
24. A random sample of 4 professional athletes produced the following data where x is the number of endorsements the player has and y is the amount of money made (in millions of dollars).

x	0	1	3	5
y	1	2	3	8

- (a) Find an equation of the least squares regression line. *Show all work; writing the correct equation, without supporting work, will receive no credit.* (15 pts)
- (b) Based on the equation from part (a), what is the predicted value of y if $x = 4$? *Show all work and justify your answer.* (5 pts)
-

25. The UMUC Daily News reported that the color distribution for plain M&M's was: 40% brown, 20% yellow, 20% orange, 10% green, and 10% tan. Each piece of candy in a random sample of 100 plain M&M's was classified according to color, and the results are listed below. Use a 0.05 significance level to test the claim that the published color distribution is correct. *Show all work and justify your answer.*

Color	Brown	Yellow	Orange	Green	Tan
Number	42	21	12	7	18

- (a) Identify the null hypothesis and the alternative hypothesis.
- (b) Determine the test statistic. *Show all work; writing the correct test statistic, without supporting work, will receive no credit.*
- (c) Determine the p-value. *Show all work; writing the correct critical value, without supporting work, will receive no credit.*
- (d) Is there sufficient evidence to support the claim that the published color distribution is correct? Justify your answer.

(15 pts)