1. True or False. Justify for full credit.

(a) If all the observations in a data set are identical, then the variance for this data set is zero.

(b) If A and B are disjoint, P(A) = 0.4 and P(B) = 0.5, then P(A 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 98% confidence interval of the same parameter.

(e) It’s easier to reject the null hypothesis in a hypothesis test at 0.05 significance level than at 0.01 significance level.

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

(a) A study was conducted at a local college to analyze the average GPA of students graduated from UMUC in 2016. 100 students graduated from UMUC in 2016 were randomly selected, and the average GPA for the group is 3.5. The value 3.5 is a

(i) statistic (ii) parameter (iii) cannot be determined

(b) 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

(c) On the day of the Virginia Primary Election, UMUC News Club organized an exit poll at three polling stations were randomly selected and all voters were surveyed as they left those polling stations. This type of sampling is called:

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

3. 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). (Show all work. Just the answer, without supporting work, will receive no credit.)

|  |  |  |
| --- | --- | --- |
| Study Time (in hours) | Frequency | Relative Frequency |
| 0.0 – 4.9 | 2 |  |
| 5.0 - 9.9 | 13 |  |
| 10.0 – 14.9 | 20 |  |
| 15.0 – 19.9 |  | 0.45 |
| 20.0 – 24.9 |  |  |
| total | 100 |  |

(a) Complete the frequency table with frequency and relative frequency. Express the relative frequency to two decimal places.

(b) What percentage of the study times was at least 15 hours?

(c) In what class interval must the median lie? 5.0 – 9.9, 10.0 -14.9, 15.0 – 19.9, or 20.0 – 24.9? Why?

4. The five-number summary below shows the grade distribution of a STAT 200 quiz for a sample of 60 students.

Answer each question based on the given information, and explain your answer in each case.

(a) What is the range in the grade distribution?

(b) Which of the following score bands has the most students?

(i) 30 – 50

(ii) 50 - 70

(iii) 85 – 100

(Iv) Cannot be determined

(c) How many students in the sample are in the score band between 65 and 100?

5. A basket contains 3 white balls, 2 yellow balls, and 5 red balls. Consider selecting one ball at a time from the basket. (Show all work. Just the answer, without supporting work, will receive no credit.)

(a) Assuming the ball selection is with replacement. What is the probability that the first ball is white and the second ball is also white?

(b) Assuming the ball selection is without replacement. What is the probability that the first ball is yellow and the second ball is red?

6. There are 1000 juniors in a college. Among the 1000 juniors, 300 students are taking STAT200, and 150 students are taking PSYC300. There are 100 students taking both courses. Let S be the event that a randomly selected student takes STAT200, and P be the event that a randomly selected student takes PSYC300. (Show all work. Just the answer, without supporting work, will receive no credit.)

(a) Provide a written description of the complement event of (S OR P).

(b) What is the probability of complement event of (S OR P)?

7. Consider rolling a fair 6-faced die twice. Let A be the event that the product of the two rolls is at most 5, and B be the event that the first one is a multiple of 3.

(a) What is the probability that the product of the two rolls is at most 5 given that the first one is a multiple of 3? Show all work. Just the answer, without supporting work, will receive no credit.

(b) Are event A and event B independent? Explain.

8. Answer the following two questions. (Show all work. Just the answer, without supporting work, will receive no credit).

(a) A bike courier needs to make deliveries at 6 different locations. How many different routes can he take?

(b) Mimi has eight books from the Statistics is Fun series. She plans on bringing three of the eight books with her in a road trip. How many different ways can the three books be selected?

9. Let random variable x represent the number of heads when a fair coin is tossed three times.

(a) Construct a table describing the probability distribution.

(b) Determine the mean and standard deviation of x. (Round the answer to two decimal places)

10. Mimi just started her tennis class three weeks ago. On average, she is able to return 20% of her opponent’s serves. Assume her opponent serves 8 times.

(a) Let X be the number of the serves that Mimi returns. 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 that she returns at least 1 of the 8 serves from her opponent. (round the answer to 3 decimal places) Show all work. Just the answer, without supporting work, will receive no credit.

11. The heights of pecan trees are normally distributed with a mean of 10 feet and a standard deviation of 2 feet. 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 7 and 11 feet tall? (round the answer to 4 decimal places) (b) Find the 40th percentile of the pecan tree height distribution. (round the answer to 2 decimal places)

12. Based on the performance of all individuals who tested between July 1, 2012 and June 30, 2015, the GRE Quantitative Reasoning scores are normally distributed with a mean of 152.47 and a standard deviation of 8.93. (https://www.ets.org/s/gre/pdf/gre\_guide\_table1a.pdf). Show all work. Just the answer, without supporting work, will receive no credit. (a) Consider all random samples of 49 test scores. What is the standard deviation of the sample means? (Round your answer to three decimal places) (b) What is the probability that 49 randomly selected test scores will have a mean test score that is greater than 150? (Round your answer to four decimal places)

13. An insurance company checks police records on 600 randomly selected auto accidents and notes that teenagers were at the wheel in 90 of them. Construct a 95% confidence interval estimate of the proportion of auto accidents that involve teenage drivers. Show all work. Just the answer, without supporting work, will receive no credit.

14. In a study designed to test the effectiveness of acupuncture for treating migraine, 100 patients were randomly selected and treated with acupuncture. After one-month treatment, the number of migraine attacks for the group had a mean of 2 and standard deviation of 1.5. Construct a 95% confidence interval estimate of the mean number of migraine attacks for people treated with acupuncture. Show all work. Just the answer, without supporting work, will receive no credit.

15. Mimi is interested in testing the claim that banana is the favorite fruit for more than 50% of the adults. She conducted a survey on a random sample of 100 adults. 58 adults in the sample chose banana as his / her favorite fruit.

Assume Mimi wants to use a 0.10 significance level to test the claim.

(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 for this test. Show all work; writing the correct P-value, without supporting work, will receive no credit.

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

16. In a study of memory recall, 5 people were given 10 minutes to memorize a list of 20 words. Each was asked to list as many of the words as he or she could remember both 1 hour and 24 hours later. The result is shown in the following table.

|  |  |  |
| --- | --- | --- |
|  | Numbers of words recalled | |
| Subject | 1 hour later | 24 hours later |
| 1 | 14 | 12 |
| 2 | 18 | 15 |
| 3 | 11 | 9 |
| 4 | 13 | 12 |
| 5 | 12 | 12 |

Is there evidence to suggest that the mean number of words recalled after 1 hour exceeds the mean recall after 24 hours? Assume we want to use a 0.05 significance level to test the claim.

(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 P-value, without supporting work, will receive no credit.

(d) Is there sufficient evidence to support the claim that the mean number of words recalled after 1 hour exceeds the mean recall after 24 hours? Justify your conclusion.

17. In a pulse rate research, a simple random sample of 600 men results in a mean of 80 beats per minute, and a standard deviation of 11.3 beats per minute. Based on the sample results, the researcher concludes that the pulse rates of men have a standard deviation less than 12 beats per minutes. Use a 0.05 significance level to test the researcher’s claim.

(a) Identify the null hypothesis and 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 for this test. Show all work; writing the correct P-value, without supporting work, will receive no credit.

(d) Is there sufficient evidence to support the researcher’s claim? Explain.

18. The UMUC MiniMart sells four different types of teddy bears. The manager reports that the four types are equally popular. Suppose that a sample of 500 purchases yields observed counts of 150, 125, 105, and 120 for types 1, 2, 3, and 4, respectively.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| type | 1 | 2 | 3 | 4 |
| number | 150 | 125 | 105 | 120 |

Assume we want to use a 0.05 significance level to test the claim that the four types are equally popular.

(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 P- value, without supporting work, will receive no credit.

(d) Is there sufficient evidence to support the manager’s claim that the four types are equally popular? Justify your answer.

19. A STAT 200 instructor believes that the average quiz score is a good predictor of final exam score. A random sample of 10 students produced the following data where x is the average quiz score and y is the final exam score.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | 80 | 93 | 50 | 60 | 100 | 40 | 85 | 70 | 75 | 85 |
| Y | 70 | 96 | 50 | 70 | 96 | 38 | 83 | 65 | 77 | 87 |

(a) Find an equation of the least squares regression line. Show all work; writing the correct equation, without supporting work, will receive no credit.

(b) Based on the equation from part (a), what is the predicted final exam score if the average quiz score is 90? Show all work and justify your answer.

20. A study of 10 different weight loss programs involved 200 subjects. Each of the 10 programs had 20 subjects in it. The subjects were followed for 12 months. Weight change for each subject was recorded. We want to test the claim that the mean weight loss is the same for the 10 programs.

|  |  |  |  |
| --- | --- | --- | --- |
| Source of variation | Sum of squares (ss) | Degrees of freedom (df) | Mean square (ms) |
| Factor (between) | 65.4 |  |  |
| Error (within) |  |  |  |
| total | 653.05 | 199 | n/a |

(a) Complete the following ANOVA table with sum of squares, degrees of freedom, and mean square (Show all work):

(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 P-value, without supporting work, will receive no credit.

(d) Is there sufficient evidence to support the claim that the mean weight loss is the same for the 10 programs at the significance level of 0.05? Explain.