Statistic and data handling

This exam contains 30 multiple choice questions. A question and the possible answer choices appear. Select the choice you find to be the correct answer of the question. Clicking on a choice saves it as your answer for the question.

Circle the correct response for each question. Make sure that your answer is clearly marked. You will not receive partial credit for any work done.

This is a closed book, closed notes examination. You may use a calculator if you wish. However, cell phones are not permitted for use in any way.

Any discussion or otherwise inappropriate communication between examinees, as well as the appearance of any unnecessary material or cell-phone usage, will be dealt with severely. Violations may result in an “F" for this exam, “F" for the class, suspension, or expulsion.

Print your name at the top of this page in the upper right hand corner.

*Good Luck!!*

HONOR PLEDGE FOR THIS EXAM:

After you have finished the exam, please read the following statement and sign your name below it.

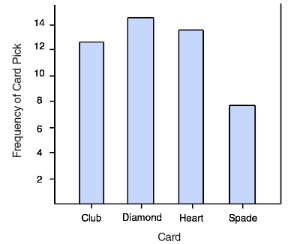
*I promise that I did not discuss any aspect of this exam with anyone other than the instructor, that I neither gave nor received any unauthorized assistance on this exam, and that the work presented herein is entirely my own.*

1. What is the probability of rolling a die and not getting a 1?
2. 5/6
3. 1/6
4. 1/3
5. 2/6
6. if I have a deck of cards, a coin and a die. What is the probability drawing a black card, flipping tails, and rolling an odd number?
7. 1/12
8. 1/24
9. 3/24
10. 3/12
11. Find the mean, median, mode and range of this data set:  19, 18, 21, 16, 15, 17, 20, 18 (∑x =144)
12. all three, i.e. mean, median, mode, equal 18 and the range is 3
13. mean 18, median 16, mode 18 and the range is 6
14. all three, i.e. mean, median, mode, equal 18 and the range is 6
15. mean 18, median 16, mode 18 and the range is 3
16. mean 18, median 18, mode 16 and the range is 6
17. Suppose a simple random sample of 150 students is drawn from a population of 3000 college students. Among sampled students, the average IQ (Intelligence Quotient) score is 115 with a standard deviation of 10. What is the 99% confidence interval for the students' IQ score?
18. 115 + 0.01
19. 115 + 0.82
20. 115 + 2.1
21. 115 + 2.6
22. Which of the following statements are true?

I. Categorical variables are the same as qualitative variables.   
II. Categorical variables are the same as quantitative variables.   
III. Quantitative variables can be continuous variables.

1. I only
2. II only
3. III only
4. I and III
5. I and II
6. One card was picked from a standard deck of cards and the suit was recorded in a bar graph, then it was placed back into the deck and the process was repeated 50 times. Here are the results:

What is the theoretical probability of drawing each suit?



1. 1/2
2. 1/8
3. 1/3
4. 1/4
5. 1/5
6. Which suit has the largest frequency, and what is it?
7. 13/59 Heart
8. 15/50 Diamond
9. 7/50 Spade
10. 15/50 Club
11. 12/59 Club
12. Which of the following statements is true?

I. The center of a confidence interval is a population parameter.   
II. The bigger the margin of error, the smaller the confidence interval.   
III. The confidence interval is a type of point estimate.   
IV. A mean is an example of a point estimate.

1. I only
2. II only
3. III only
4. IV only
5. All the above
6. A sample consists of four observations: {1, 3, 5, 7}. What is the standard deviation?
7. 2
8. 2.58
9. 6
10. 6.67
11. None of the above
12. Which of the following statements is true.

I. The standard error is computed solely from sample attributes.   
II. The standard deviation of a population is computed solely from sample attributes.   
III. The standard error is a measure of central tendency.

1. I only
2. II only
3. III only
4. I and II
5. I and III
6. Nine hundred (900) high school freshmen were randomly selected for a national survey. Among survey participants, the mean grade-point average (GPA) was 2.7, and the standard deviation was 0.4. What is the confidence interval, assuming a 95% confidence level?
7. 2.7 ± 0.013
8. 2.7 ± 0.026
9. 2.7 ± 0.4
10. 2.7 ± 0.8
11. None of the above
12. The amount of time spent studying and your grade point average.
13. time spent studying increase and GPA positively correlated
14. GPA is not correlate with time spent studying
15. as time spent studying increase, GPA decreases
16. GPA inversely depend by time
17. GPA directly depend by time
18. Frankie and Lucy are planning on selling a new iPhone app. This is a scatter plot estimation of how many apps they can sell at different prices. A line of best fit is drawn in for you.

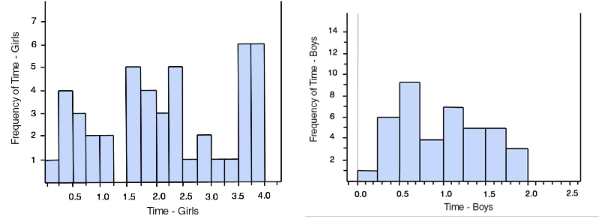


What kind of correlation is shown in this graph?

1. Positive
2. Negative
3. no correlation
4. nonlinear correlation
5. none of above
6. If they price the application at $3.00, how many can they expect to sell, and how much money would they get from the sales?
7. 30 and $120
8. 50 and $200
9. 30 and $90
10. 40 and $120
11. none of the above
12. These histograms graph the amount of time (hours per day) that 46 middle school girls and 40 middle school boys in Rome spend on the website Face-Book. 50 boys and 50 girls took the survey but 4 girls and 10 boys did not use these sites at all. Each is graphed with a bin width of 0.25 hours.

Compare the percentage of boys and girls that spend more than zero but less than 1 hour/day on these sites.

1. about 14% of boys and 4% of girls spend less than 1 hour/day
2. about 50% of boys and about 22% of girls spend less than 1 hour/day
3. About 20% of boys and 50% of girls spend less than 1 hour/day
4. About 52%of boys and 26% of girls spend less than 1 hour/day
5. About 26%of boys and 52% of girls spend less than 1 hour/day



1. What can you conclude based on these graphs?
2. Girls spend much more time on these sites than boys. In addition, their times vary more widely, ranging from zero to 4 hours/day, while the boys' times top out at 2 hours/day
3. Boy and girl spent about the same time on this site
4. Boys spend much more time on these sites than girls. In addition, their times vary more widely, ranging from zero to 4 hours/day, while the boys' times top out at 2 hours/day
5. The times variations are about the same for boys and girls
6. Girls spend much more time on these sites than boys. In addition, boys’ times vary more widely, ranging from zero to 4 hours/day, while the girls' times top out at 2 hours/day
7. A national achievement test is administered annually to 3rd graders. The test has a mean score of 100 and a standard deviation of 15. If Jane's z-score is 1.20 (score in a normalized distribution with mean zero and standard deviation 1), what was her score on the test?
8. 82
9. 88
10. 100
11. 112
12. 118
13. Which of the following is a discrete random variable?
14. The average height of a randomly selected group of boys.
15. The number of Red Blood Cells.
16. The number of presidential elections in the 20th century.
17. I only
18. II only
19. III only
20. I and II
21. II and III
22. Suppose we want to estimate the average weight of an adult male in Naples, Italy. We draw a random sample of 1,000 men from a population of 1,000,000 men and weigh them. We find that the average man in our sample weighs 60 kg, and the standard deviation of the sample is 10 kg. What is the 95% confidence interval.
23. 60 + 0.62
24. 60 + 1.0
25. 60 + 1.96
26. 80 + 10
27. None of the above
28. Twenty-two students were randomly selected from a population of 1000 students. The sampling method was simple random sampling. All of the students were given a standardized English test and a standardized math test. Test results are summarized below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | tudent | English | Math | Difference, d | (d - d)2 | | 1 | 95 | 90 | 5 | 16 | | 2 | 89 | 85 | 4 | 9 | | 3 | 76 | 73 | 3 | 4 | | 4 | 92 | 90 | 2 | 1 | | 5 | 91 | 90 | 1 | 0 | | 6 | 53 | 53 | 0 | 1 | | 7 | 67 | 68 | -1 | 4 | | 8 | 88 | 90 | -2 | 9 | | 9 | 75 | 78 | -3 | 16 | | 10 | 85 | 89 | -4 | 25 | | 11 | 90 | 95 | -5 | 36 | | |  |  |  |  |  | | --- | --- | --- | --- | --- | | Student | English | Math | Difference, d | (d - d)2 | | 12 | 85 | 83 | 2 | 1 | | 13 | 87 | 83 | 4 | 9 | | 14 | 85 | 83 | 2 | 1 | | 15 | 85 | 82 | 3 | 4 | | 16 | 68 | 65 | 3 | 4 | | 17 | 81 | 79 | 2 | 1 | | 18 | 84 | 83 | 1 | 0 | | 19 | 71 | 60 | 11 | 100 | | 20 | 46 | 47 | -1 | 4 | | 21 | 75 | 77 | -2 | 9 | | 22 | 80 | 83 | -3 | 16 | |

=1

What is the 90% confidence interval for the mean difference between student scores on the math and English tests? Assume that the mean differences are approximately normally distributed.

1. 1 + 0.8
2. 1 + 1.0
3. 1 + 1.3
4. 1 + 2.0
5. 1 + 3.6
6. The F statistic is a ratio of two types of variance: variance\_\_\_\_\_\_\_\_\_\_ and within groups.
7. random
8. individual
9. between
10. systematic
11. A Type I error occurs when the null hypothesis is
12. rejected and the research hypothesis is actually false.
13. accepted but and research hypothesis is actually true.
14. rejected and null hypothesis is actually true.
15. accepted and null hypothesis is actually true.
16. In term of set theory, the null and alternative hypotheses divide all possibilities into:
17. two sets that overlap
18. two non-overlapping sets
19. two sets that may or may not overlap
20. as many sets as necessary to cover all possibilities
21. probability space cannot be subdivided
22. An inventor has developed a new, energy-efficient lawn mower engine. He claims that the engine will run continuously for 5 hours (300 minutes) on a single gallon of regular gasoline. From his stock of 2000 engines, the inventor selects a simple random sample of 50 engines for testing. The engines run for an average of 295 minutes, with a standard deviation of 20 minutes. Test the null hypothesis that the mean run time is 300 minutes against the alternative hypothesis that the mean run time is not 300 minutes. Use a 0.05 level of significance. (Assume that run times for the population of engines are normally distributed.)

P(t < -1.77) = 0.04 with 49 degrees of freedom

1. the Null hypothesis cannot be rejected
2. the alternative (μ = 300) is rejected
3. the Null hypothesis cannot be completely rejected
4. the alternative (μ = 300) cannot be completely rejected
5. A sample survey is a study that obtains data from a subset of a population, in order to estimate population attributes. Which of the following statements are true?
6. A sample survey is a type of experiment.
7. An observational study requires fewer resources than an experiment.
8. The best method for investigating causal relationships is an observational study.
9. I only
10. II only
11. III only
12. All of the above.
13. None of the above
14. In the context of regression analysis, which of the following statements are true?

I. When the sum of the residuals is greater than zero, the data set is nonlinear.

II. A random pattern of residuals supports a linear model.

III. A random pattern of residuals supports a non-linear model.

1. I only
2. II only
3. III only
4. I and II
5. I and III
6. In the context of regression analysis, which of the following statements are true? An influential point is an outlier that greatly affects the slope of the regression line. One way to test the influence of an outlier is to compute the regression equation with and without the outlier.
7. When the data set includes an influential point, the data set is nonlinear.
8. Influential points always reduce the coefficient of determination.
9. All outliers are influential data points.
10. I only
11. II only
12. III only
13. All of the above
14. None of the above
15. Suppose a researcher conducts an experiment to test a hypothesis. If she doubles her sample size, which of the following will increase?
16. The power of the hypothesis test.
17. The effect size of the hypothesis test.
18. The probability of making a Type II error.
19. I only
20. II only
21. III only
22. All of the above
23. None of the above
24. You conduct a hypothesis test and you observe values for the sample mean and sample standard deviation when n = 25 that do not lead to the rejection of . You calculate a *p*-value of 0.0667. What will happen to the *p*-value if you observe the same sample mean and standard deviation for a sample > 25?
25. Increase
26. Decrease
27. Stay the same
28. May either increase or decrease
29. Remain constant
30. What is the standard deviation for the data given: 5, 10, 7, 12, 0, 20, 15, 22, 8, 2 (mean 10.1)
31. 6.89
32. 10.1
33. 7.26
34. 9
35. 5.26

Rome, November 24 2016

**Answer corrector table**: put a cross in correspondence to the selected answer

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N° | a | b | c | d | e | N° | a | b | c | d | e |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  | 30 |  |  |  |  |  |

**Answer corrector table**: put a cross in correspondence to the selected answer

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N° | A | b | c | d | e | N° | a | b | c | d | e |
| 1. 1 |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  | 1. 30 |  |  |  |  |  |