TENNESSEE MATHEMATICS TEACHER' ASSOCIATION

SIXTY-FOURTH ANNUAL MATHEMATICS CONTEST

2022

Statistics

Prepared by:

Tamah Fridman
University of Tennessee
Knoxville, TN

Reviewed by:

Mathematics Faculty
Austin Peay State University
Clarksville, TN
University of TN at Martin
Martin, TN

Scoring formula: 4 x (Number Right) – (Number Wrong) + 40

Directions:

Do not open this booklet until you are told to do so

This is a test of your competence in high school mathematics. For each problem, determine the best answer and indicate your choice by making a heavy black mark in the proper place on the separate answer sheet provided. You must use a pencil with a soft lead (No. 2 lead or softer).

This test has been constructed so that most of you are not expected to answer all the questions. Do your best on the questions you feel you know how to work. You will be penalized for incorrect answers, so wild guesses are not advisable.

If you change your mind about an answer, be sure to erase completely. Do not mark more than one answer for any problem. Make no stray marks of any kind on the answer sheet. The answer sheets will not be returned to you; if you wish a record of your performance, mark your answers in this booklet also. You will keep the booklet after the test is completed.

When told to do so, open your test booklet and begin. You will have exactly eighty minutes to work.

TMTA **2022** Statistics Exam

1. To compile a frequency table, a student collected the following data on the location of Carolina wren nests: 45 tree and building cavities, 48 low tree branches and bushes, 62 building eaves, 51 vines. Which words best describes this data?

   a. continuous, interval

   b. ratio, nominal

   c. discrete, ratio

   d. interval, ordinal

   e. ordinal, ratio

2. According to U.S. Census Bureau data from 2017, the average wage was \$48,251.57. To be in the top half (50%) of all earners, you needed to earn more than \$29,999.99 a year. To be in the top X%, a person had to earn roughly

X	1%	10%	20%	30%	40%	50%	60%	70%
earn $\geq$	250,000	95,000	65,000	50,000	40,000	30,000	20,000	15,000

   What can you say about the distribution of the data?

   a. The distribution is symmetric.

   b. The distribution is skewed to the left.

   c. The distribution is skewed to the right.

   d. The distribution is bimodal.

   e. The distribution is unimodal.

3. Clementine studied the distribution of blue crocodile egg masses by measuring eggs' mass in the nests that she could find, and she concluded that the distribution of egg mass is lognormal with a mean value of 55 g and a standard deviation of 4 g. After obtaining all data and results, she discovered that her scale was miscalibrated and showed a reading that was 1.2 grams lower than the true value every time. Therefore, her results must be adjusted to:

   a. $\mu = 56.2$ g and $\sigma = 4.0$ g

   b. $\mu = 56.2$ g and $\sigma = 5.2$ g

   c. $\mu = 54.8$ g and $\sigma = 5.2$ g

   d. $\mu = 54.8$ g and σ must be recomputed

   e. depends on the values of the data, everything should be recomputed

4. A normally distributed error of a Pop timer has a mean $\mu = 0$ s and a standard deviation $\sigma = 0.02$ seconds. What is the probability that your next timer measurement will have an absolute value of error of more than 0.03 s?

   a. 0.5000

   b. 0.6667

   c. 0.0668

   d. 0.1336

   e. 0.1102

5. A study is investigating whether a certain costly innovation in architecture will improve the seismic stability of high rise buildings in San Francisco. The null hypothesis states that it there is does not improve the stability (therefore, it should not be implemented). Which statement is true:

   a. the error of Type 1 occurs when the correct solution that the innovation does not improve seismic stability is rejected

   b. the error of Type 2 occurs when the correct solution that the innovation does not improve seismic stability is rejected

   c. the error of Type 1 can cause economical damage

   d. the error of Type 2 can cause loss of lives

   e. the error of Type 2 occurs when incorrect Null hypothesis is rejected

6. A simple linear regression model was created to fit the data resulting in the formula $y = 2.35x - 27.5$ with $R^2 = 0.6126$. Which of the following is the correct interpretation of the data:

   a. There is no correlation between x and y.

   b. Most of the variance is explained by the model.

   c. R^2 is too low therefore the coefficient estimates are biased.

   d. R^2 is sufficiently high to prove that your regression model provides an adequate fit to your data.

   e. R^2 value is low, and it is always a problem when R^2 is low.

2

7. You ran an experiment and found that wild brown mice that eat more chocolate perform maze tasks faster. Which of the following statements is NOT true?

   a. It is impossible to prove whether chocolate in the diet is the cause of the mice's success because correlation does not mean causation.

   b. It is possible that mice that eat chocolate lose weight and therefore can run faster. So eating more chocolate and performing faster just correlate, but the weight has the cause-correlation relationship with both variables.

   c. Chocolate possibly kills mice parasites, so they become healthier and perform better.

   d. It is possible to investigate causation by redesigning the experiment in a way that allows you to control or randomize other variables that can affect the performance rate, and therefore see if the chocolate dose is still the factor that affects the performance.

   e. Causation cannot be proven statistically.

8. You sampled the weight distribution of all your relatives and determined that it is skewed to the right. What can we state with certainty?

   a. the sample size is too small

   b. the distribution skews to the right because you did not separate men and women; once you do that, both distributions will be symmetric with zero skew

   c. the range is too small compared to the mean value

   d. biometric data tend to have a lognormal distribution since they cannot be negative; therefore, the distribution is expected to skew to the right. Your data reflects this tendency.

   e. Everybody needs to lose weight.

9. Astronomers define star brightness in terms of apparent magnitude – how bright the star appears from Earth, and absolute magnitude – how bright the star appears at a standard distance of 32.6 light-years, or 10 parsecs, from itself. Two stars that appear equally faint to you (low apparent magnitude) may be in fact a nearby faint star and a very bright star that is located far far away. To compute absolute magnitude from apparent magnitude, you have to know the distance to the star. In the figure, you see The Color - Absolute Magnitude Diagram of the Hyades Star Cluster (*Johnson, H. L.; Mitchell, R. I.; Iriarte, B., Astrophysical Journal, vol. 136, p.75*). The lower the color index B-V, the bluer the star. The higher Absolute Magnitude M, the brighter the star. Estimate the value of Pearson's correlation coefficient r for the data in the interval $[0.2, 1.2]$.
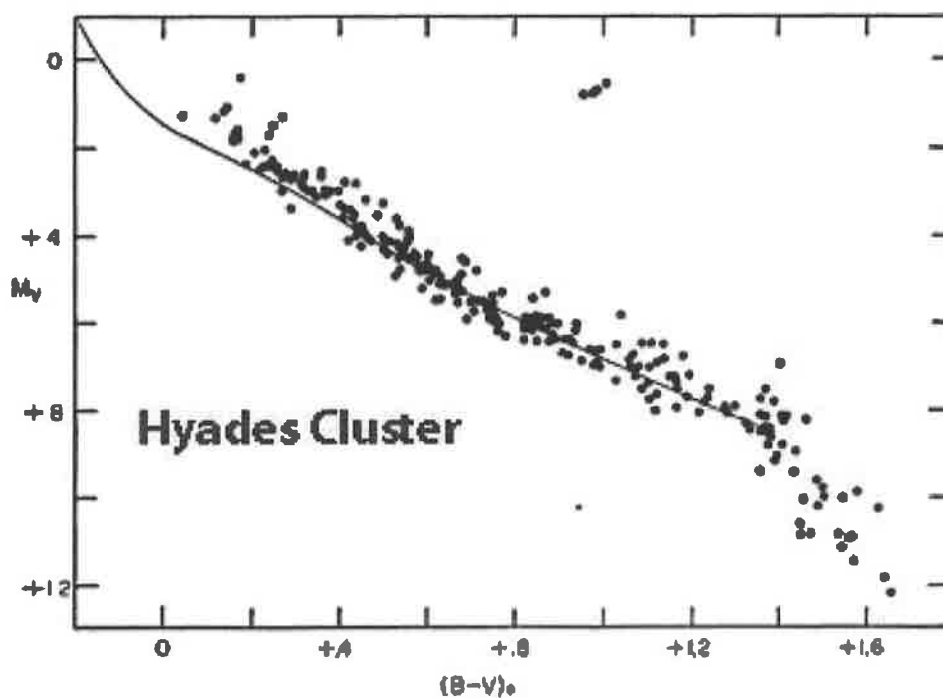
a. 0.9

b. 0.5

c. 0

d. -0.5

e. -0.9

10. As we discussed in problem 9, stars vary in brightness and color. Brightness and color correlate. The brighter the star, the farther it can be seen. You decided to study star color distribution, wrote to NASA, and obtained a perfect picture of the night sky. You selected 10,000 stars at random uniformly distributed in the picture, and accessed a NASA program that could tell you the color of the selected stars. You got 1,000 values for the color index (B-V), and your data comprises rational numbers in the interval $[0.2, 1.2]$. Which statement is true:

    a. Since you selected stars at random, your sample is representative of the star population without bias.

    b. Your method of sampling has self-selection bias.

    c. Your method of sampling has undercoverage bias.

    d. Your method of sampling has survivorship bias.

    e. Since you selected stars at random, your color index distribution is normal.

11. For the following stem and leaf plot, what is the mode?

Stem	Leaf
0	2 2 4 7
1	0 2 2 5 6 6 7 8
2	1 2 3 3 3 7 8
3	2 3 6 6 7 7

    a. 21

    b. 23

    c. 20.36

    d. 16

    e. 35

12. Which statement is NOT correct:

    a. The width of the confidence interval determines the precision of the estimate.

    b. The smaller σ, the smaller the confidence interval.

    c. The larger the size of the sample, the smaller the confidence interval.

    d. To compute a confidence interval, we use two-tailed t-value.

    e. A 95% confidence interval is wider than a 99% confidence interval.

13. While scuba diving in Australia, you collected five pearls of the following sizes in millimeters: 8.5, 8.8, 9, 9.2, and 9.5. Compute your sample's standard deviation.

    a. 0.500

    b. 0.381

    c. 0.341

    d. 0.145

    e. 0.116

14. Five volumes of the Feynman Lectures on Physics are placed on a bookshelf at random. What is the probability that volumes 1 and 2 are standing next to each other?

    a. $\dfrac{1}{5!}$

    b. $\dfrac{1}{4!}$

    c. $\dfrac{1}{4}$

    d. $\dfrac{1}{5}$

    e. $\dfrac{2}{5}$

15. What is the probability to roll either an odd number or a number that is less than 4, if the dice is fair?

    a. $\dfrac{1}{6}$

    b. $\dfrac{1}{3}$

    c. $\dfrac{1}{2}$

    d. $\dfrac{2}{3}$

    e. $\dfrac{5}{6}$

16. For the probability distribution in the table, which of the following statements is correct?

x	0.2	0.3	0.5	0.6	0.9	1.0
P(x)	0.22	0.20	0.30	0.03	0.25	0.00

    a. the mode is 0.30 and the range is 0.8

    b. $P(x \geq 0.5) = 0.58$ and the median is 0.55

    c. the median is 0.5 and the range is $\{0.2, 0.3, 0.5, 0.6, 0.9\}$

    d. the mean is 0.497 and the range is 0.7

    e. $P(x \leq 0.3) = 0.42$ and the probability distribution is continuous.

17. The mean of a sample is 64 and the standard deviation is 15. There are 144 data in the sample. What is the 90% confidence interval for this sample?

    a. [61.9, 66.1]

    b. [59.9, 68.1]

    c. [49.0, 79.0]

    d. [61.5, 66.5]

    e. [61.1, 66.9]

18. Glitter Washing Machines company published a normalized data histogram proving that most of their machines do not require repairs for 10 years. What descriptive statistic did they use to provide this information?

    a. Mean

    b. Mode

    c. Median

    d. Standard deviation

    e. Range

19. A professor gave the same test in two large sections of his class, and noticed that the means for the exam were different with $\mu_1 < \mu_2$, and both grade distributions were not Gaussian. The professor thinks about analyzing the disparity using a t-test. Which of the following statements is correct:

    a. He should not use a t-test because the grade distributions are not Normal.

    b. He should use a t-test and his Null Hypothesis must be: the means of the distributions are equal.

    c. He can use a t-test only if the number of students in the sections is the same.

    d. He should use a t-test and his Null Hypothesis must be: $\mu_1 < \mu_2$.

    e. He should use a paired t-test.

20. A Reddit source claims that Brad Pitt movies featuring him consuming 200 or more calories of food average more money and receive higher critic scores than his movies where he eats less or nothing. To arrive at this conclusion scientifically, one should perform a _____.

    a. t-test on 2 pairs of data sets with resulting $p \leq 0.05$ in each test

    b. F-test on 2 pairs of data sets with resulting $p \geq 0.95$ in each test

    c. χ^2 test on 2 data sets with resulting $p \leq 0.05$

    d. Pearson's correlation test on 2 data sets with resulting $r \geq 0.95$

    e. ANOVA on 4 data sets with resulting $p \leq 0.05$

21. To estimate the percentage of people who are planning to vote for a particular independent candidate in a presidential election at a 95% confidence level and the margin of error 2%, how many people do you need to interview? (Assume random sample, no bias).

    a. $n \geq 1022$

    b. $n \geq 1335$

    c. $n \geq 1850$

    d. $n \geq 2031$

    e. $n \geq 2401$

22. A 2018 study by the NIH/National Institute on Deafness and Other Communication Disorders found that 6.66% Americans over the age of 40 experience phantom odors, that is, a person would smell an odor that is not there. We might use this to infer that in a sample of 300 people over the age of 40,

    a. 1 person in 15 experiences phantom odors

    b. 1 person in 15 does not experience phantom odors

    c. 1 person in 20 experiences phantom odors

    d. 15 persons in 100 do experience phantom odors

    e. 15 persons in 100 do not experience phantom odors

23. There are 10 white and 5 red balls in an urn. You pull out one ball, then return it to the urn, and pull out a ball again. What is the probability that both balls were the same color?

    a. 0.22

    b. 0.44

    c. 0.50

    d. 0.55

    e. 0.78

24. You are working with a dataset of $(x, y) =$(body mass, brain mass) of various species, some of which do not even have a brain. You plan to test a linear regression hypothesis. Which of the following data transformations can be used:

    a. $x' = \sqrt{x}$, $y' = \frac{1}{y}$

    b. $x' = \log x$, $y' = \log(1 + y)$

    c. $x' = x$, $y' = \log y$

    d. $x' = \frac{1}{x}$, $y' = \frac{1}{y}$

    e. we cannot use data transformation for testing regression hypothesis

9

25. Jay spent the summer in a Russian Hacker camp, and for his hacking thesis made a scatter plot of starting salaries of his 100 fellow major graduates as a function of their GPAs. His goal was to investigate whether higher GPA corresponded to a higher starting salary, i.e., the value of GPA in the job market. He checked for the outliers and analyzed them. Seven of his peers went to graduate school, so their salaries came out as 0. One person had a low GPA and yet his starting salary was five times higher than the next highest salary; it turned out that this "outlier" came from a very affluent family. For the purpose of the study, Jay considers the following options: 1. eliminate the 7 graduate students from the data 2. eliminate the affluent peer's data 3. perform a t-test 4. perform a linear regression analysis. What would be the correct way to proceed?

    a. 1, 2, 3, 4

    b. 3, 4

    c. 1, 4

    d. 1, 2, 3

    e. 1, 2, 4

26. Dr.Yulie conducts a human subject study funded by the NIH on the benefits of a new cancer drug B with respect to an established drug A. The benefits include longer survival and better quality of life represented by severity and nature of side effects. Dr. Yulie recruited 58 patients with 4th stage cancer. What statement is NOT true:

    a. Dr. Yulie needs to separate the patients into two groups, drug A and Drug B recipients.

    b. Gender and age differences should be accounted for.

    c. A t-test must be conducted to accept or reject the Null Hypothesis that the means of the survival time under both treatments are equal.

    d. The study must be double blind.

    e. Dr. Yulie needs to recruit healthy people for the control group.

10

27. In a particular country, 1% of a population is sick with a certain virus and needs treatment. The test that is available to determine whether a person is infected, has a 4% false positive rate, and a false negative rate of 0. If a patient tested positive for the virus, what is the probability that he really has it?

    a. 0.67

    b. 0.25

    c. 0.20

    d. 0.17

    e. 0.13

28. A fair die is thrown 120 times. What is the probability that a "6" rolled up 26 times or more?

    a. 0.092

    b. 0.217

    c. 0.392

    d. 0.463

    e. 0.655

29. A manufacturer bought 10,000 warped geometry microchips from three suppliers: 3800 from supplier A, 3500 from B, and 2700 from supplier C. 4% of supplier A chips are defective, so are 6% of supplier B's chips, and 3% of supplier C's chips. What is the probability that a randomly selected chip is defective?

    a. 0.0130

    b. 0.0451

    c. 0.1300

    d. 0.0443

    e. 0.0513

30. You are taking a multiple choice exam, and do not know how to solve 10 problems. Each problem has 5 possible answers. You fill the answers to these 10 problems at random. What is the probability that you will get exactly 3 correct answers?

    a. 0.200

    b. 0.210

    c. 0.008

    d. 0.790

    e. 0.300

31. You are planning to go camping tomorrow and would hate to get caught in the rain. Weather.com predicts an 80% chance of clear sky (that is, no rain). Given that that "no rain" prognoses of weather.com come true in 95% cases, and "rain" prognoses come true in 99% cases, what is the probability that there will be no rain tomorrow?

    a. 0.760

    b. 0.762

    c. 0.800

    d. 0.950

    e. 0.958

32. The probability of having a baby boy is 0.515. For a family with 3 children, what is the probability that at least one of them is a boy? Assume that gender is binary (boy/girl).

    a. 0.886

    b. 0.863

    c. 0.777

    d. 0.667

    e. 0.62

33. An FBI officer analyzed his target shooting practice results, and noticed that when he makes 3 shots, his probability of having hit the target with at least one shot is 0.973. What is the probability that he hits the target on his single shot?

    a. 0.85

    b. 0.80

    c. 0.70

    d. 0.65

    e. 0.60

34. Dr. Fisher noticed that all of his patients who reported suddenly diminished response to antidepressant drug A get their drug from supplier T. He decided to check whether the active ingredient in the drug by supplier T is below the standard dose which is 40 microgram. The mass of the active ingredient must be normally distributed with $\sigma = 2$ microgram. Dr. Fisher obtained 100 pills from different batches of drug A from supplier T and found that the mean for his sample was $\bar{X} = 38.3$ microgram. Find the 95% confidence interval for the mass of the active ingredient in pills produced by T.

    a. $[38.0, 42.0]$

    b. $[36.3, 40.3]$

    c. $[37.9, 38.7]$

    d. $[37.4, 39.1]$

    e. $[37.3, 39.2]$

35. You study how children growing up in rural farm communities versus urban areas are affected by asthma and pollen allergies. You compiled the following table:

	asthma	allergies	neither	total			
rural farm	10	50	40	100			
urban	22	60	18	100			

Compute the value of χ^2 test statistic.

   a. $\chi^2 = 17.57$

   b. $\chi^2 = 13.75$

   c. $\chi^2 = 2.50$

   d. $\chi^2 = 2.29$

   e. $\chi^2 = 1.38$

36. You are working to develop drugs to cure a new lethal infection, zombi rabies. You developed 3 drugs: the first cures 60% of patients, the second cures 50%, and the third cures 70%. The drugs cannot be combined. You made 6 pills of the first drug, 4 of the second, and 5 of the third. Your sloppy graduate student puts all the pills in the same box. Suddenly, your friend falls sick with zombi rabies. You give him a random pill from the box, and your friend gets cured. Which drug did you most likely give him, drug 1, or 2, or 3?

   a. drug 1

   b. drug 2

   c. drug 3

   d. drug 1 or 3 at equal probabilities

   e. not enough data

14

37. You want to bring 6 souvenirs home from a trip, and you zeroed in on four worthy items. In how many ways can you select 6 souvenirs out of four kinds of goods? (The store is well stocked.)

    a. 4096

    b. 1296

    c. 576

    d. 84

    e. 20

38. On average, six cars pass the Henley-Main street cross section per minute. What is the probability that there will be no cars for one minute?

    a. 0.0024

    b. 0.0018

    c. 0.1000

    d. 0.1667

    e. 0.8333

39. In an Organic Chemistry class, 4 students prepared very well for a Pass/Fail midterm, 18 prepared reasonably well, and 8 prepared moderately well. A very well-prepared student passes the midterm with probability 0.95, a reasonably well-prepared student passes with probability 0.75, and a moderately well-prepared student passes the midterm with a probability 0.5. Bob passed the midterm. What is the probability that he was prepared moderately well?

    a. 0.667

    b. 0.5

    c. 0.267

    d. 0.188

    e. 0.133

15

40. Zoe participated in the Young Engineer Competition Part 1, where the top 5% were awarded a trip to Vienna, and in Part 2, where the top 5% were awarded a trip to Tokyo, and the next 20% (in the top 25% but not in the top 5%) were awarded a trip to Santa Fe. The scores of the Part 1 participants were distributed uniformly in the interval [50,100], and the scores of Part 2 participants were distributed normally with $\mu = 50$, $\sigma = 16$. Zoe scored 97 in Part 1 and 66 in part 2. What award(s) did she get? Hint: draw the distribution for Part 1 and find its y coordinate.

    a. a trip to Vienna

    b. a trip to Tokyo

    c. a trip to Santa Fe

    d. trips to Vienna and Tokyo

    e. trips to Vienna and Santa Fe