## **2024 Regional Statistics Competition** – Page 1 Arkansas Council of Teachers of Mathematics

This competition includes 25 multiple-choice questions and three open-response questions that might be used as tie breakers. For questions 1 through 25 (the multiple-choice questions), mark your answer choice in the appropriate location on the sheet provided. After completing questions 1 through 25, answer each tie breaker question in sequential order (i.e., complete Tie Breaker #1 first, then Tie Breaker #2, and then Tie Breaker #3 last). Be sure that your name is printed on each of the tie breaker questions. When time is called, you will be asked to turn in your multiple-choice question answer sheet and your written responses to the tie breaker questions.

- 1. The values of  $Q_1$ ,  $Q_2$ , and  $Q_3$  in a positively skewed distribution are calculated. What is the expected relationship between  $(Q_2 Q_1)$  and  $(Q_3 Q_1)$ ?
  - a.  $(Q_2 Q_1)$  is greater than  $(Q_3 Q_1)$ .
  - b.  $(Q_2 Q_1)$  is equal to  $(Q_3 Q_1)$ .
  - c.  $(Q_2 Q_1)$  is less than  $(Q_3 Q_1)$ .
  - d. Cannot be determined without examining the data.
- A group of 200 6<sup>th</sup> grade students were given a standardized test and obtained scores ranging from 42 to 88.
  If the scores tended to "bunch up" in the low 80s, the shape of the distribution would be which of the following?
  - a. Symmetrical
  - b. Negatively skewed
  - c. Positively skewed
  - d. Cannot be determined.
- 3. After examining data collected over the past 10 years, researchers at a theme park found the following for N first-time guests: 2250 visited during the summer months, 675 visited during the fall, 1300 visited during the winter, and the rest visited during the spring. If the relative frequency for first-time guests who visited during the winter is 0.26, what is the relative frequency of first-time guests who visited in the spring?
  - a. 0.135
  - b. 0.45
  - c. 0.26
  - d. 0.155
  - e. Not enough information given.

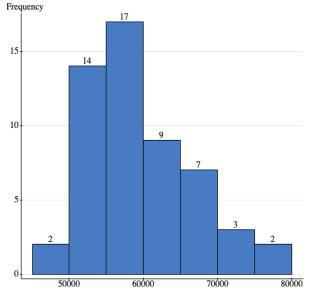
#### <u>2024 Regional Statistics Competition</u> – Page 2 Arkansas Council of Teachers of Mathematics

- 4. A 20-item statistics test was graded using the following procedure: A correct response is scored +1, a blank response is scored 0, and an incorrect response is scored −1. The highest possible score is +20, the lowest score is −20. Suppose you calculated the variance of the test scores and got a result of −3. From this result, you conclude which of the following?
  - a. The class did very poorly on the test.
  - b. The test was easy for the class.
  - c. Some students received negative scores.
  - d. A computational error certainly was made.

## For questions 5-7, refer to the histogram, which shows the

## frequency of salaries for a sample of teachers.

- 5. How many teachers were included in this sample?
  - a. 100
  - b. 51
  - c. 54
  - d. 7
  - e. It cannot be determined from the graph.
- 6. What proportion of teachers have a salary less than \$55,000? Round your answer to two decimal places.
  - a. 0.30
  - b. 0.33
  - c. 0.35
  - d. 0.26
  - e. It cannot be determined from the graph.
- Is it unlikely to find a teacher who has a salary of at least \$75,000? (Consider an event to be unlikely if it occurs less than 5% of the time)
  - a. No, the appropriate probability is greater than 0.05.
  - b. Yes, the appropriate probability is greater than 0.05.
  - c. No, the appropriate probability is less than 0.05.
  - d. Yes, the appropriate probability is less than 0.05.
  - e. Unable to determine.



#### <u>2024 Regional Statistics Competition</u> – Page 3 Arkansas Council of Teachers of Mathematics

## 8. Given a distribution of 200 IQ scores, you find the mean is considerably greater than the median.

This means the distribution is which one of the following?

- a. Negatively skewed
- b. Positively skewed
- c. Symmetrical
- d. Cannot be determined.
- 9. The standard error of the mean is which one of the following?
  - a. Standard deviation of a sample distribution.
  - b. Standard deviation of the population distribution.
  - c. Mean of the sampling distribution of the standard deviation.
  - d. Standard deviation of the sampling distribution of the mean.
- 10. Adding just one or two extreme scores to the high end of a large distribution of scores will have a greater effect on which one of the following?
  - a. Median
  - b. Mean
  - c. Mode
  - d. Interquartile Range
  - e. It will have no effect.
- 11. A quiz consists of 10 multiple choice questions. Each question has five possible answers, one of which is correct. A student must get 60% or better on the quiz to pass. If a student randomly guesses, what is the probability that the student will pass the quiz?
  - a. 0.377
  - b. 0.205
  - c. 0.006
  - d. 0.060
  - e. None of the above.

- 12. A quiz consists of 100 multiple choice questions. Each question has five possible answers, only one of which is correct. Find the mean and the standard deviation of the number of correct answers.
  - a. mean: 50; standard deviation: 7.07106781
  - b. mean: 20; standard deviation: 4
  - c. mean: 20; standard deviation: 4.47213595
  - d. mean: 50; standard deviation: 4
  - e. None of the above.
- 13. A highly selective boarding school will only admit students who place at least 1.5 z-scores above the mean on a standardized test. This test has a mean of 110 and a standard deviation of 12. What is the minimum score that an applicant must make on the test to be accepted?
  - a. 92
  - b. 98
  - c. 128
  - d. 122
  - e. None of the above.
- 14. In interpreting a boxplot of a data set we note that 1) the median is to the left of the center of the box and 2) the right whisker line is longer than the left whisker line. We can conclude that
  - a. The data is positively skewed.
  - b. The data is negatively skewed.
  - c. The data is symmetric.
  - d. Skewness or symmetry cannot be determined from this information.
- 15. According to government data, the probability that an adult never visited a museum is 15%. In a random survey of 10 adults, what is the probability that at least eight adults visited a museum?
  - a. 0.002
  - b. 0.200
  - c. 0.820
  - d. 0.800
  - e. None of the above.

- 16. A physical fitness association includes the mile run in its secondary-school fitness test. The time for this event for boys in secondary school is known to possess a normal distribution with a mean of 450 seconds and a standard deviation of 50 seconds. The fitness association wants to recognize the fastest 10% of the boys with certificates of recognition. What time would the boys need to beat to earn a certificate of recognition from the fitness association?
  - a. 514 sec
  - b. 532.25 sec
  - c. 367.75 sec
  - d. 386 sec
  - e. None of the above.
- 17. Determine the point estimate of the population mean and margin of error for the confidence interval with lower bound= 7 and upper bound= 25.
  - a.  $\bar{x} = 16, E = 18$
  - b.  $\bar{x} = 16, E = 9$
  - c.  $\bar{x} = 25, E = 9$
  - d.  $\bar{x} = 7, E = 18$
  - e. None of the above.
- 18. Jessica's percentile score on a math test was 40 while Alex's percentile score on the same test was 80. Which one of the following is true?
  - a. Jessica correctly answered half as many items as Alex did.
  - b. Alex's math achievement is twice as good as Jessica's.
  - c. Alex correctly answered more items than Jessica.
  - d. All of the above.
- 19. Suppose that the correlation coefficient between the two variables *X* and *Y* is 0.81. What can you say about the proportion of variance in *Y* that is explained by *X*?
  - a. It will be the same as correlation coefficient.
  - b. It will be smaller than the correlation coefficient.
  - c. It will be higher than the correlation coefficient.
  - d. Cannot be determined.

## **<u>2024 Regional Statistics Competition</u> – Page 6** Arkansas Council of Teachers of Mathematics

- 20. A scatter plot shows a linear relationship between *X* and *Y*. On the scatter plot, below average *X* values tend to associate with above average *Y* values and above average *X* values tend to associate with below average *Y* values. Which one of the following is true about the correlation coefficient?
  - a. The correlation coefficient is negative.
  - b. The correlation coefficient is positive.
  - c. The correlation coefficient is zero.
  - d. Cannot be determined.

Given events J and K: P(J) = 0.18, P(K) = 0.37, P(J or K) = 0.45. Use these to answer Questions 21 & 22.

- 21. Find *P*(*J* and *K*)
  - a. 0.1
  - b. 0.45
  - c. 0.55
  - d. 0.9
  - e. Cannot be determined.
- 22. Find the probability of the complement to event (*J* and *K*).
  - a. 0.1
  - b. 0.45
  - c. 0.55
  - d. 0.9
  - e. Cannot be determined.
- 23. Suppose that you play a game in which you roll a six-sided die. If you roll a 6, you win \$10. If you roll a 4 or 5, you win \$5. Otherwise, you pay \$7. How much will you win/lose if you play this game 100 times?
  - a. You will win \$16.67.
  - b. You will lose \$16.67.
  - c. You will win \$0.17.
  - d. You will lose \$0.17.
  - e. Cannot be determined.

## Suppose a study of speeding violations and drivers who use cell phones produced the following data: Refer to the table below for questions 24 - 25.

	Speeding violation in the last year	No speeding violation in the last year
Uses cell phone while driving	25	280
Does not use cell phone while driving	45	405

24. Find the probability that a driver had no violation in the last year and used a cell phone while driving.

- a. 280/685
- b. 280/755
- c. 685/755
- d. 710/755
- e. Cannot be determined.

25. Find the probability that a driver was a cell phone user given the driver had a violation in the last year.

- a. 25/70
- b. 25/45
- c. 25/755
- d. 70/755
- e. Cannot be determined.

## **<u>2024 Regional Statistics Competition</u> – Page 8** Arkansas Council of Teachers of Mathematics

#### **<u>2024 Regional Statistics Competition</u> – Page 9** Arkansas Council of Teachers of Mathematics

Name
------

Tie	Breaker	#1

Name: \_\_\_\_\_

School: \_\_\_\_\_

A university dean is interested in determining the proportion of students who receive some sort of financial aid. Rather than examine the records for all students, the dean randomly selects 200 students and finds that 118 of them are receiving financial aid. Use a 95% confidence interval to estimate the true proportion of students on financial aid. Express the answer in the form  $\hat{P} \pm E$  and round to the nearest thousandth.

## **<u>2024 Regional Statistics Competition</u> – Page 10** Arkansas Council of Teachers of Mathematics

#### **2024 Regional Statistics Competition – Page 11** Arkansas Council of Teachers of Mathematics

Name
------

#### **Tie Breaker #2**

Name: \_\_\_\_\_

School: \_\_\_\_\_

A lumber manufacturing plant produces standard "2x4" wooden beams used in house construction. Despite the name, the width of the beam is actually 3.5 inches. Because of the variability in the manufacturing process, the width of the beam is approximately normally distributed with a standard deviation of 0.02 inches. Use this information to answer the following questions.

a. Beams with a width less than 3.45 inches or greater than 3.55 inches are discarded. What proportion of beams are discarded?

b. If 2,000 beams are manufactured in a day, how many beams would you expect to discard?

c. If an order of 10,000 beams is received, how many beams should the plant manager expect to manufacturer so that all beams have a width between 3.45 inches and 3.55 inches?

## **<u>2024 Regional Statistics Competition</u> – Page 12** Arkansas Council of Teachers of Mathematics

#### **2024 Regional Statistics Competition** – Page 13 Arkansas Council of Teachers of Mathematics

Name \_\_\_\_\_

Tie Breaker #3

Name: \_\_\_\_\_

School: \_\_\_\_\_

A pharmaceutical company wishes to test a new drug with the expectation of lowering cholesterol levels. Ten subjects are randomly selected and pretested. The results are listed below. The subjects were placed on the drug for a period of 6 months, after which their cholesterol levels were tested again. The results are listed below. (All units are milligrams per deciliter.) Test the company's claim that the drug lowers cholesterol levels. Use  $\alpha = 0.05$ . Assume that the distribution is normally distributed.

Subject	1	2	3	4	5	6	7	8	9	10
Before	195	225	202	195	175	250	235	268	190	240
After	180	220	210	175	170	250	205	250	190	225

a. What are the null and alternative hypotheses?

b. What is the test statistic?

c. What is the *p*-value?

d. What is your conclusion? Justify your answer.

Name \_\_\_\_\_

## **<u>2024 Regional Statistics Competition</u> – Page 14** Arkansas Council of Teachers of Mathematics

Name \_\_\_\_\_

## **2024 Regional Statistics Competition** – Page 15 Arkansas Council of Teachers of Mathematics

#### **Multiple Choice Answers**

1. C	11. C	21. A
2. B	12. B	22. D
3. D	13. C	23. B
4. D	14. A	24. B
5. C	15. C	25. A
6. A	16. D	
7. D	17. B	
8. B	18. C	
9. D	19. B	
10. B	20. A	

## <u>TB1:</u>

We calculate  $\hat{P} = \frac{118}{200} = 0.59$ .

# Our standard error is $E = 1.96 \cdot \sqrt{\frac{0.59 \cdot (1-0.59)}{200}} = 0.068$

The interval is  $\hat{P} \pm E = 0.59 \pm 0.068 = (0.522, 0.658)$ 

NORMAL FLOAT AUTO REAL RADIAN MP	NORMAL FLOAT AUTO REAL RADIAN MP
<b>1-PropZInt</b>	<mark>1-PropZInt</mark>
x:118	(0.52184,0.65816)
n:200	<sup>\$=0.59</sup>
C-Level:0.95	n=200
Calculate	■

#### <u>TB2:</u>

a. Our mean is 3.5 and standard deviation is 0.02. We can calculate the area between 3.45 and 3.55.

NORMAL FLOAT AUTO REAL RADIAN MP	NORMAL FLOAT AUTO REAL RADIAN MP
normalcdf lower:3.45 upper:3.55 μ:3.5 σ:0.02 Paste	normalcdf(3.45,3.55,3.5,0. 0.9875806403 1-0.9875806403 0.0124193597

We keep 0.988=98.8% of beams. We discard 0.012=1.2% of beams.

b. If 2000 beams are produced, we would expect to discard 25 beams.

$$2000\cdot 0.012\approx 25$$

c. If an order of 10,000 beams is received, we need to "overproduce" by 1.2%. We need the 98.8% we keep to result in the 10,000 beam order.

$$0.988 \cdot x = 10,000$$
$$x = \frac{10,000}{0.988} \approx 10,126$$

We need to produce 10,126 beams.

NORMAL FLOAT	AUTO	REAL	RADIAN	MP	Ī
normalcdf	(3.	45,3	3.55,	3.5,	ø.
1-0.98758	2064		9875	8064	03
		0.	0124	1935	97.
2000*0.01				3871	94
10000/0.9	875	8064	403 1012	5.75	54
				= - <i>(</i> . <del>-</del>	π. <b>!</b> .

## <u>TB3:</u>

## This uses a paired one tail t-test. We compare the difference between the two given lists of data.

## Approaches may vary, but here are two options for the problem:

Hypotheses:	$H_0: \mu_d = 0$ $H_1: \mu_d < 0$ Where $d = After - Before$	$H_0: \mu_d = 0$ $H_1: \mu_d > 0$ Where $d = Before - After$		
Data Entry:	NORMAL FLOAT AUTO REAL RADIAN MP        L1      L2      L3      L4      L5      H        195      180      -15            225      220      -5            202      210      8            250      210      8            195      175      -20            195      175      -20            195      176      -5            250      250      0            268      250      -18           240      225      -15	NORMAL      FLOAT      AUTO      REAL      RADIAN      MP        L1      L2      L3      L4      L5      4        195      180      15          225      220      5          202      210      -8          195      175      20          250      250      0          255      205      30          255      205      18          190      190      0          L4(1)=      L4(1)=		
Calculate Results:	NORMAL FLOAT AUTO REAL RADIAN MP    I      Inpt:Data Stats    I      µ0:0    List:L3      Freq:1    I      µ:≠µ0    Yµ0      Color:    BLUE      Calculate Draw      NORMAL FLOAT AUTO REAL RADIAN MP      Impt: 2.752409413      p=0.011194199      x= -10      Sx=11.48912529      n=10	NORMAL FLOAT AUTO REAL RADIAN MP      Inpt:Data      Stats      µ0:0      List:L3      Freq:1      µ:≠µ0 <µ0 >µ0      Color:      BLUE      Calculate      Calculate      Draw      NORMAL FLOAT AUTO REAL RADIAN MP      Impo      Impo      Impo      Color:      BLUE      Calculate      Draw      Impo      Impo		
Test Statistic:	t = -2.75	t = 2.75		
<i>p</i> -value:	<i>p</i> -value = 0.011	<i>p</i> -value = 0.011		
Conclusion:	(Same conclusion for both options.) We were given $\alpha = 0.05$ . Since <i>p</i> -value< $\alpha$ , or 0.011 < 0.05, we reject the null hypothesis. There is a statistically significant difference between the "before" and "after" data.			