1. An English teacher wants to create a linear regression model that will allow her to predict grades based on the length of papers written. Below is a printout of a computer regression analysis done by the Buchholz statistics team. Which of the following is true?

Predictor	Coeff	StDev	T	P
Constant	47.34	37.71	1.255	.093
Length	1.122	.237	?	.008

    S= 38.42   R-sq = 92.16%

    a. The slope of the regression line is too close to 0 to conduct a test of significance
    b. The model contains 92.16% of the data points.
    c. The slope of the regression line is significant.
    d. 96% of the variability in grade can be explained by the length.
    e. NOTA

2. Which of the following is **false** about standard deviation?
    a. Standard deviation is the square root of the variance
    b. Standard deviation cannot be computed for skewed distributions
    c. For mound-shaped, symmetric distributions, the interval one standard deviation below to one standard deviation above the mean is wider than the interquartile range
    d. For mound-shaped, symmetric distributions, the points one standard deviation below and above the mean correspond to inflection points along the curve.
    e. NOTA

3. In the regression equation $\hat{y} = a + bx$, which of the following statements about a is **not** always true?
    a. It depends on the values of $b, \bar{x}, \bar{y}$.
    b. It is the y-intercept of the line
    c. It is the predicted value of the response variable when the explanatory variable has a value of 0.
    d. It provides important contextual information about the relationship between the variables.
    e. All of these are true

4. Based on the Central Limit Theorem, how many of the statements below are true?
    I. The mean of the sampling distribution gets closer to the mean of the population as the size of the samples in the sampling distribution get larger
    II. The standard deviation of the sampling distribution cannot be calculated if the population is heavily skewed
    III. The standard deviation of the sampling distribution gets closer and closer to the standard deviation of the population as the sample size grows
    IV. The shape of the sampling distribution is normal regardless of the shape of the population distribution
    a. 0   b. 1   c. 2   d. 3   e. 4

5. What is the probability that if you roll a pair of fair six-sided dice that you get a sum of eight or you do not get doubles?
    a. $\dfrac{35}{36}$ b. $\dfrac{5}{18}$ c. $\dfrac{31}{36}$ d. $\dfrac{11}{36}$ e. NOTA

6. A random variable X has a probability distribution that when graphed forms a triangle with vertices at the origin, $(8,0)$ and $\left(4,\dfrac{1}{4}\right)$. What is the probability that X will be less than 2?
   a. This is not a valid probability distribution
   b. .5
   c. .25
   d. .125
   e. NOTA

7. A sample of 200 Mu sponsors indicated that the mean amount of time they spend each week outside the classroom doing math is 18.6 hours. Three of Middleton's crack statisticians independently calculated different confidence intervals of the true mean amount of time spent each week outside the classroom doing math for all Mu sponsors. The confidence intervals created were A: (12.83, 24.37); B: (13.654, 23.346); and C: (15.281, 21.919). Which conclusion is valid?
   a. All are calculated correctly with different levels of confidence.
   b. A and C have reasonable intervals but B does not.
   c. B and C have reasonable intervals but A does not.
   d. A and B have reasonable intervals but C does not.
   e. NOTA

8. The statement " The median of a distribution is approximately equal to the mean of the distribution" can be made true by adding which of the following:
   a. For all distributions.
   b. Only for roughly symmetric, mound-shaped distributions.
   c. For skewed distributions.
   d. For roughly symmetric distributions.
   e. NOTA

9. A talk radio host runs a survey on a controversial issue by asking viewers to call one of two 800 numbers: one signifying yes, the other signifying no. The results of this survey
   a. Represents a valid sample since conflicting biases will cancel out.
   b. May exhibit non-response bias.
   c. Represents a valid sample of the viewing population.
   d. May exhibit substantial variation.
   e. NOTA

10. Spujitsu Incorporated has decided to hire people in certain proportions to demonstrate that it is an equal opportunity employer. Management decides that 65% of their new employees will be female, and 30% of the employees will be non-whites. 25% of the new employees will be white males. What percentage of the employees will be non-white women?
    a. Impossible to answer without more information
    b. 20%
    c. $66\dfrac{2}{3}\%$
    d. 45%
    e. NOTA

11. The attendance officer at Rickards believes that kids are more likely to skip school as the week progresses. To investigate this claim, the officer collects the number of absences over a one month period and organizes by day of the week as shown below.

Monday	Tuesday	Wednesday	Thursday	Friday
310	271	253	318	348

A chi-square analysis was performed to test the claim that there is a relationship between the day of the week and the number of absences. What is the P-value of the test?
a. $P < 0.0005$
b. $.0005 < P < .001$
c. $.001 < P < .01$
d. $.01 < P < .025$
e. $P > .025$

12. The gator football team reports a mean time of 22 minutes for its players to do the three mile run. A transfer from Miami runs it in 30 minutes and is excited to hear his time was at the third quartile. What standard deviation did the gator team use if the times were normally distributed? (round to nearest minute)
a. 16 minutes
b. 12 minutes
c. 7 minutes
d. 3 minutes
e. NOTA

13. A Vestavia statistician computed a least squares regression line and found that the correlation coefficient was .49. A Florida student pointed out that the variables had been switched. The Vestavia statistician corrected the mistake and then computed the correct correlation. The correct correlation is?
a. $\dfrac{1}{49}$
b. $-.49$
c. $.49$
d. $\dfrac{-1}{49}$
e. Impossible to know

14. An oil company plans to drill wells at a particular site until it finds one that will produce oil. Assuming each well has a probability of .1 of producing oil, what is the probability that it will take at least 5 tries to get the first oil-producing well?
a. .6561   b. .5905   c. .4095   d. .3439   e. NOTA

15. The University of Florida police department adds up the number of parking tickets written by the departments 248 officers during the month of October. The department wants to determine the 95% confidence interval for the true mean number of tickets in October written by each officer. Which of the following is the reason why the department cannot compute the confidence interval?
a. October is not a typical month because of all the gator football home games
b. It is not known if the data are normally distributed
c. A confidence interval cannot be constructed when population data are known
d. If a few officers wrote an inordinate number of tickets, they could skew the data
e. The department is fine in computing the interval in this situation

16. Researchers want to know if a new breathing technique will help golfers who get the yips while putting. To investigate this idea, the researchers teach the breathing technique to 43 golfers who have the yips. After one month of breathing with the new technique 30 of the volunteers report a significant reduction in getting the yips. The Lincoln statistics team criticizes this experiment. Which of the following is likely to be the main criticism of the design?
a. No P-value was given for people to decide how important the results were.
b. The volunteers were not divided into blocks that had different levels of the yips
c. No control group was used to see if there was a placebo effect.
d. Golfers without the yips were not included in this experiment.
e. None of these are valid criticisms of the design

17. If you were to select 7 digit phone numbers at random what is the probability that you would select your number if the following restrictions apply. You cannot begin with a 0 or a 1 and the number 911 cannot be used for the first three digits?

a. $\dfrac{1}{8000000}$ b. $\dfrac{1}{5670000}$ c. $\dfrac{1}{7999999}$ d. $\dfrac{1}{7990000}$ e. NOTA

18. Causation cannot be inferred for
    a. Valid matched pairs designs.
    b. Valid randomized block designs.
    c. Valid observational studies where n is large.
    d. Valid completely randomized designs.
    e. Causation can be inferred for all of these.

19. Which of the following best describes the effect on a binomial distribution if the number of trials is held constant but the probability of success increases?
a. The mean increases but the standard deviation decreases.
b. The mean and standard deviation both decrease.
c. The mean and standard deviation both increase
d. The mean decreases but the standard deviation increases.
e. The mean increases but the standard deviation can increase or decrease.

20. A study was done to compare the weight of bald eagle eggs in Alaska versus those in Florida. 20 eggs from Alaska and 20 from Florida were taken and weighed with the result shown below. For the purpose of our analysis we will consider both sets to be SRS of eggs from their respective populations. A 95% confidence interval for the difference of means and a two sample t-test were carried out with the results as:
95% PCT FOR MU(ALASKA)-MU(FLORIDA):(-54,68)
TTEST MU(ALASKA)-MU(FLORIDA):T= .78
Based on the information, which of the following statements is true?
a. A matched pair t-test should have been used.
b. The sample sizes are too small to draw any conclusions
c. The population mean of Alaska bald eagle egg weight is higher than Florida bald eagle egg weight.
d. Statistically, there is no difference between the population means of the two groups
e. NOTA

21. Which of the following plots is least useful in determining the shape of a distribution?
a. Stem plot
b. Box plot
c. Histogram
d. Dot plot
e. All of these are very useful in determining shape.

22. The Buchholz math team recently completed a study to compare the length of time to graduate college between males and females. For each person, three variables were measured: gender, major, and length of time to graduate. Which of the following statements is correct?
   a. Gender and major are explanatory variables; length of time is a response variable
   b. Gender and length of time are explanatory variables; major is a response variable
   c. Gender is an explanatory variable; length of time is a response variable; major is a possible confounding variable
   d. Length of time is the explanatory variable; major and gender are response variables.
   e. Length of time is an explanatory variable; gender is a response variable; major is a possible confounding variable.

23. Which of the following is not possible?
   a. The IQR is equal to the range
   b. The mean is negative and the standard deviation is positive
   c. The five number summary has 4 identical values
   d. The standard deviation is greater than the mean
   e. All of these are possible

24. Some former Mu competitor created two drugs to treat students who choke under the intense Mu pressure. Drug A will cure 75% of the cases but has unwanted side effects (you don't want to know what they are) in 40% of the cases. Drug B will cure 50% of the cases and has unwanted side effects in 30%. 60 chokers are given Drug A and 80 chokers receive Drug B. If each choker receives only one drug, what is the expected number of patients who will experience unwanted side effects?
   a. 30   b. 48   c. 49   d. 85   e. 88

25. The Vero math team created a new game while waiting around for the awards to begin at a recent invitational. The game uses just 4's, 6's, and 8's from many different decks of cards. 40% of the cards are 4's, 40% of the cards are 8's the rest are 6's. In each round, each person is given a turn choosing four cards with replacement and adding the results. The person with the highest result wins the game. If you were to advise a newcomer to the game what sum she should expect, your advise would be which of the following?
   a. 6
   b. 12
   c. 24
   d. There is no way to predict the expect sum based on the information given
   e. NOTA

26. The Middleton statisticians were asked by their sponsor to survey Mu participants to find out if they approve of the job the FAMAT board is doing. They created a 90% confidence interval that had a margin of error of 4%. By what factor would the original sample size need to be multiplied if their sponsor ordered them to reduce the margin of error to 2%?
   a. 2   b. 4   c. 16   d. Impossible to determine without more information.   e. NOTA

27. The probability of a success on any one trial is .4 and you have a maximum of 10 trials in which to win. Which of the following expressions will calculate the probability that you will get a success on one of your first four attempts?

a. $_{10}C_4 (.4)^4 (.6)^4$

b. $.4 + .4(.6) + .4(.6)^2 + .4(.6)^3$

c. $.6 + .4(.6) + .4(.6)^2 + .4(.6)^3$

d. $4(.4)^4$

e. NOTA

28. A distribution that is approximately normal has a mean of 68 and a standard deviation of 9. Which of the following is true?
a. The IQR is approximately 18.
b. Approximately 95% of the data lies between 59 and 77.
c. It is impossible that the distribution contains a data value greater than 100.
d. There are more data values between 59 and 68 than between 77 and 86.
e. NOTA

29. The Chiles math team collected data on housing prices in the Tallahassee area. A histogram of the data revealed that the distribution was heavily skewed toward larger prices. The Chiles math team takes a SRS of 40 and computes the mean of each sample. Which of the following best describes the shape of the sampling distribution of the sample means?
a. Cannot tell because the mean is unknown
b. Skewed to the right
c. Normally distributed
d. Uniformly distribute
e. Skewed to the left

30. A company claims their light bulbs last 1028 hours in continuous use. A SRS of their light bulbs yields a sample mean of 1004 hours with a standard deviation of 84 hours. A t-test shows t = -2.119 and a P-value of .0387. If the test utilizes the degrees of freedom of 54, then the sample size must be?
a. 53
b. 54
c. 55
d. 56
e. The sample size cannot be determined without more information

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

Team

1. 19, Histograms
2. $\frac{8}{3} or 2\frac{2}{3}$
3. -.13
4. -.52
5. 131
6. −8.93
7. 25.331, Green
8. .99
9. 1437
10. $289\frac{1}{3}$
11. 3.59
12. $\frac{49}{45}$
13. 4368
14. Christina, Lei, Sue, and David-Must be in this order
15. 80, Frazer Fudge, Musketeers, and Mr. Good Bar

1. C-Because .008 is below any reasonable alpha level
2. B-Standard deviation can be computed for any distribution
3. D-If x doesn't take on values on or near zero the contextual information from the y-intercept will seem ridiculous
4. A-None are true
5. C- $\dfrac{5}{36} + \dfrac{30}{36} - \dfrac{4}{36} = \dfrac{31}{36}$
6. D- $\dfrac{1}{2} \otimes 2 \otimes \dfrac{1}{8} = \dfrac{1}{8}$
7. B-B is not center at the sample mean
8. D
9. E-Voluntary response
10. B-

	F	M	TOT
NW	.2	.1	.3
W	.45	.25	.7
TOT	.65	.35	1

11. B-Carry out a chi-square goodness of fit test and you will get a chi-square value of 19.26 and a P-value of .0007
12. B- $\dfrac{30-22}{\sigma} = \dfrac{2}{3} \rightarrow \sigma = 12$
13. C-Correlation does not change when you switch the variables
14. A- First four must fail so $.9^4 = .6561$
15. C-The purpose of a interval is to learn about the population. If you already know population it makes no sense to create an interval
16. C-Good experiments always have a control
17. D- $(8 \bullet 10 \bullet 10 - 1)(10^4)$
18. C-You can never draw causation from a observational study
19. E-The mean is increasing because np is growing, but the standard deviation increases when p gets closer to .5 not necessarily when it is increasing
20. D-Because zero is in the interval or T=.78 will yield a high P-value
21. B-With only five numbers it is difficult to ascertain the shape
22. C
23. E
24. B- $.4(60) + .3(80) = 48$
25. C-

X	P	XP
4	.4	1.6
6	.2	1.2
8	.4	3.2
	6	4x6=24

26. B-To cut margin of error in half you must increase the sample size by a factor of 4. This because n is located in the denominator and is under the square root sign
27. B-Geometric
28. D-59-68 is the same length but is closer to the mean so more likely to occur
29. C-Sample is large so population shape is irrelevant by the central limit theorem
30. C-   n-1=54   n=55

1. Since the data involves one variable you would use histograms to compare. The underclassmen have a median equal to 6 and a IQR equal to 7-4 = 3. The upperclassmen have a median of 7 and an IQR equal to 8-5 = 3. So 3 + 6 + 3 + 7= 19

2. The IQR is four-thirds standard deviations long. 1.5 times four-thirds equals two. . Two added to two-thirds equals two and two-thirds

3. A= .3 + .2 = .5
   B= .3 + .2 - .06 = .44
   C= .3
   D= .3 times .2 which equals .06
   E= 0 since disjoint means no intersection
   F= .2 times .25 which equals .05   .5 - .44 - .3 + .06 + .05 = -.13

4. $A= 1-(.9)^{10} =.65$

   $B=1-Binocdf(3,.1.10)=.01$

   C= compliment of B so .99

   $D=_{10}C_4(.1)^2(.9)^8 =.19$ .65 + .01 - .99 - .19 = -.52

5.   42   42   60   61   80   81   82   so A = 82

     49   50   51   74   75   85   85   so B = 49   82 + 49 = 131

6. A = 6.53
   B = -12.08
   C=-12.41
   D= 12-2=10
   E= -.97
   F= zero always   6.53 – 12.08 –12.41 +10 - .97 = -8.93

7.

	Observed	Expected	$\dfrac{(O-E)^2}{E}$
BR	320	300	1.3333
Y	210	200	.5
Red	170	200	4.5
Or	120	100	4
Gr	70	100	9
Bl	110	100	1
			20.33

   A= .001
   B= 20.33
   C= 6 – 1 = 5
   D= Green   .001 + 20.33 + 5 = 25.331, Green

8. Stochastics

Profit	Y	N	T
Y	55	35	90
N	19	33	52
T	74	68	142

$$A = \frac{19}{142} = .13$$

$$B = \frac{74 + 52 - 19}{142} = .75$$

$$C = \frac{19}{74} = .26$$

$$D = \frac{19}{52} = .37 \qquad .13 + .75 - .26 + .37 = .99$$

9. $n = \frac{Z^2 p(1-p)}{m^2} = 773 \qquad n = \left(\frac{Z\sigma}{m}\right)^2 = 664 \qquad 773 + 664 = 1437$

10. $n(.125) = 32 \rightarrow n = 256 \qquad \frac{1}{p} = 33\frac{1}{3} \qquad 256 + 33\frac{1}{3} = 289\frac{1}{3}$

11. $A = (.8)^4 = .41 \qquad B = np = 4(.8) = 3.2 \qquad C = \sqrt{npq} = \sqrt{4 \bullet .8 \bullet .2} = .8$

$D = .4096 + 4(.8)^3(.2) = .82 \qquad .41 + 3.2 + .8 - .82 = 3.59$

12. $X = \frac{8}{15} \qquad Y = \frac{5(.8)^4(.2)}{5(.8)^4(.2) + (.8)^5} = \frac{5}{9} \qquad \frac{8}{15} + \frac{5}{9} = \frac{49}{45}$

13. $X = 13_4C_3 12_4C_2 = 3744 \qquad Y = 13_4C_4 12_4C_1 = 624 \qquad 3744 + 624 = 4368$

14. $David = \frac{110 - 88}{9} = 2.44 \qquad Lei = \frac{120 - 96}{11} = 2.18$

$Christina = \frac{101 - 75}{12} = 2.17 \qquad Sue = \frac{87 - 68}{8} = 2.38$

Christina, Lei, Sue, and David

15. Load data into a list and then go to one-variable statistics to get Q_3 and Q_1.

$Q_3 - Q_1 = 130 - 110 = 20 \rightarrow 20 \otimes 1.5 = 30 \rightarrow 110 - 30 = 80 \rightarrow 130 + 30 = 160 \rightarrow 160 - 80 = 80$

Frazer Fudge. Musketeers, and Mr. Good Bar