STA 122

Instruction: Answer 10 questions from each of the five sections

Time Allowed: 40 Minutes

Section A: Descriptive Statistics

1. Which of the following is a raw score that has been transformed into standard deviation units?
    A. μ-score
    B. *z*-score +++
    C. σ-score
    D. t-score
2. All these are examples of continuous variable except:
    A. Height
    B. Examination score
    C. Length
    D. None of these +++
3. Statistics is the study of the following:
    A. collection or design of data
    B. analysis and interpretation of data
    C. presentation and organization of data
    D. All of the above +++
4. Commonly used statistical methodologies in data analysis include
    A. Descriptive statistics
    B. Inferential statistics
    C. All of the above +++
    D. None of the above
5. Measure of dispersion (or variability) includes the following except
    A. standard deviation
    B. variance
    C. coefficient of variation
    D. mode +++
6. Which of the following is a correct use or application of Statistics?
    A. Statistics helps in providing a better understanding and exact description of a phenomenon of nature.

B.

B. It helps in proper and efficient planning of a statistical inquiry in any field of study.

C. It helps in collecting an appropriate quantitative data.

D. All of the above +++

7. Which of the following is a correct use or application of Statistics?

A. It helps in presenting complex data in a suitable tabular, diagrammatic and graphic form for an easy and clear comprehension of the data.
It helps in understanding the nature and pattern of variability of a phenomenon through quantitative observations.

C. It helps in drawing valid inference, along with a measure of their reliability about the population parameters from the sample data D. All of the above +++

8. With the help of histogram we can draw

A. frequency polygon B. frequency curve

C. frequency distribution

D. all the above +++

9. The first quartile divides a frequency distribution in the ratio

A. 4 : 1 B. 1 :4 C. 3 : 1

D. 1 : 3 +++

10. If the minimum value in a set is 9 and its range is 57, the maximum value of the set is

A. 33

B. 66 +++

C. 48

D. 24

11. For a negatively skewed distribution, the correct inequality is

A. Mode < median

B. mean < median

C. mean < mode +++

D. None of the above

12. The numbers of faults occurring on a production line in a 9-week period are given as: 32 29 27 26 29 39 33 29 37. The third quartile value is:

A. 29 +++

B. 35

C. 31

D. 28

13. _____ is the fractional part of the population which is also a representative part of that population.

A. sampling

C.

B. survey

C. sample +++

D. census

14. Given the set of data 1, 5, 6, 10, 3 the range and the mean deviation respectively are

A. 9, 6 B. 6,
10 C. 9,4

D. 9, 4.5 +++

15. Which of the following statistics is robust (i) mean (ii) median (iii) interquatile range (iv) standard deviation (v) Variance

A. I &ii

B. ii & v
ii & iii +++

D. iii & v

16. What is the class size of a class whose lower limit is 5 and upper limit is 10?

A. 10

B. 5

C. 5.5

D. 7.5 +++

17. What is the class mark of a class whose upper limit is 11.5 and lower limit is 7.2?

A. 18.7

B. 9.35 +++

C. 4.3

D. -3.4

18. What is the class size of a class whose upper boundary is 10.5 and lower boundary is 4.5?

A. 14

B. 7 +++

C. 15 D. 6

19. The harmonic mean of five figures is 3.505843, if four of them are 7, 5, 3 and 2, what is the fifth figure (to the nearest whole number)?

A. 4
+++

B. 3
C.
2

D. 5

20. The table below presents the marks obtained by students who took an examination. Compute the mean score.

D.

Score	50	60	40	80	70
Number of Student	8	15	2	5	10

A. 32

B. 62 +++

C. 72

D. 50

21. If the arithmetic mean and geometric mean of two values are 10 and 8 respectively, find the values.

A. 20, 4

B. 5, 2

C. 16, 4 +++ D. 10, 8

22. If the first and third quartiles of a set of observations are 33.5 and 43 respectively. What is semi interquartile range?

A. 4.22

B. 4.75 +++

C. 5.5 D. 2.2

23. The following are the measure of dispersion except

A. Range

B. Mean Deviation

C. Harmonic mean +++
Standard Deviation

24. If N=10, $\bar{X} = 12$ and $\sum x^2 = 1530$. What is the coefficient of Variation?

A. 25
+++

B. 15 C.
40

D. 22

25. What is the value of $\sum(x - \bar{x})^2$ if the sample variance is s_x^2 = 148.8 and the sample size is 5?

A. 595.2 +++

B. 29.76

C. 143.8

D. 144

26. Find the co-efficient of variation of a variable whose mean is 10 and variance is 25

A. 15% B.
36%

C. 25%

D. 50% +++

27. One of the following is NOT a statistic

E.

A. μ +++
B. $\bar{x}$
C. s
D. s^2

28. One of the following is NOT a measure under categorical variable
    a. Nominal
    b. Ordinal
    c. Interval +++
    d. None of the above
29. A statistician collected data on Income per person 200 individuals. He has a mean of 12482.3 and median 6975.14. Which of these is likely to be true? A. The distribution is left skewed.
    B. The distribution is right skewed. +++
    C. The distribution is symmetric.
    D. None of the above.
30. For the set of numbers 2,3,5,6,7,7,8 A. The median is greater than the mode
    B. The mean is greater than the mode
    C. The mean is greater than the median
    D. The mean is less than the median. +++

31. The differences between the upper and lower boundary of class is called _____
    A. Class size +++
    B. Class limit
    C. Class interval
    D. Class boundary

32. Given the following set of data: 9,8,5,*x*,6,7. If the arithmetic mean is 6.5, find *x*.
    A. 6
    B. 5
    C. 4 +++
    D. 3

33. The most frequently occurring number in a set of values is called the _____
    A. Frequency
    B. Mode +++
    C. Mean
    D. Media

34. The goal of _____ is to focus on summarizing and explaining a specific set of data.
    A. inferential Statistics
    B. descriptive Statistics +++
    C. Bayesian statistics
    D. All of the above

35. The _____ is often the preferred measure of central tendency if the data are severely skewed
    A. Mean
    B. Mode
    C. Median +++
    D. Range

36. Which measure of central tendency takes into account the magnitude of score?
    A. Mean +++
    B. Median
    C. Mode
    D. Range

37. If sample mean is 20, population standard deviation is 3 and sample size is 64, find the interval estimate of the mean.
    a. (19.02, 21.50)
    b. (19.00, 20.50)
    c. (19.27, 20.74)+++
    d. (19.22, 21.50)

38. If the largest value of a data set is doubled, which of the following is false? A. The mean increases.
    B. The standard deviation increases
    C. The inter-quartile range increases +++
    D. The median remains unchanged

39. The following list is a set of data ordered from smallest to largest: 2, 12, *y, y, y*, 18, 18, 19. All values are integers and "*y*" is neither 12 nor 15.

    I. The median and the first quartile cannot be equal.

    II. The mode is 18

    III. The median is 12

    Which of the following statement(s) is/are true about the data set?

    A. I only

    B. II only +++

    C. III only

    D. None of the statements

40. A professor scaled the score on an exam by multiplying the student's raw score by 1.2 then adding 15 points. If the mean and standard deviation of the scores before scaling were 51 and 5, respectively. Then the mean and standard deviation of the scaled scores are respectively?

    A. 76.2 and 21

    B. 76.2 and 6 +++

    C. 76.2 and 5

    D. Cannot be without the raw score

41. Suppose the mean of scores obtained by students in a test is 14. If a mean of 217 is obtained when the scores are squared, what is the standard deviation of the scores.

    A. 14.73

    B. 14.25

    C. 3.94

    D. 4.58 +++

42. Constructing a frequency distribution

    A. a. is one of the most common means of summarizing data.

    B. b. begins by recording the number of times a particular value occurs.

    C. c. is the basis for construction of a percentage distribution.

    D. d. all of the above +++

43. The General Social Survey asked respondents what region of the country they lived in at age 16. The level of measurement of this variable is:

    A. nominal. +++

    B. ordinal.

    C. interval/ratio.

    D. none of the above.

44. Generally, ordinal variables are best displayed visually with:

    A. cumulative percentages.

    B. collapsed values.

    C. pie charts.

□□ bar graphs. +++□

45. A mode:

    A. can be found only for discrete variables.
    B. is usually less than the mean.
    C. is the value of the most frequently occurring score. +++
    D. is usually less for ordinal than for interval/ratio variables.

46. The variance measures deviation around the:

    A. mode.
    B. median.
    C. mean. +++ D. sum of squares.

47. The variance:

    A. is the square of the standard deviation.
    B. can be 0.
    C. is never negative.
    D. all of the above. +++

48. Gender, State of Origin, Religion and Type of diseases are examples of:

    A. Nominal scale +++
    B. Ordinal scale
    C. Interval scale
    D. Ratio Scale

49. The measure that deals with the spread of the data is known as measure of

    A. Location
    B. Dispersion +++
    C. Skewness
    D. Data

50. μ and β are examples of

    A. statistic
    B. parameters +++
    C. statistics values
    D. None of the above

51. Which of the following is an example of discrete variable?

    A. Height

B. Weight
C. Length
D. Number of Oranges +++

52. Which of the following is **not** an example of a ratio scale variable?

A. Age (in days)
B. Speed (in seconds)
C. Height (in inches)
D. Colour of urine +++

53. As _____ is to population so is _____ sample.
A. Statistics; parameters
B. Parameters; statistics
C. Descriptive; inferential
E. Parameter; statistic +++

SECTION B: Elements of Probability

1. Sample space has probability
   A. 0.0
   B. 0.5
   C. 1.0+++
   D. None of the above
2. Sample space is defined as
   A. Set of all experimental outcomes+++

B. Set of some experimental outcomes

C. Set of one experimental outcomes

D. None of the above

3. A probability near 1 indicates an event is

A. Almost certain to occur+++

B. Not certain to occur

C. Impossible

D. None of the above

4. Consider the experiment of drawing 1 card from a deck of 52 cards and observing whether or not the card is a spade. What is the probability of drawing a spade?

A. 1/3

B. 1/4+++

C. 0

D. 1/8

5. Two events *A* and *B* are mutually exclusive if

A. $P(A \cap B) = 1$

B. $P(A \cup B) = P(A)$

C. $P(A \cap B) = P(A)$

D. $P(A \cap B) = 0$ +++

6. Suppose that we have a sample space $S = \{E_1, E_2, E_3\}$, where E_1, E_2 and E_3denote the experimental outcomes and the following probability assignments are given for the sample events

$$P(E_1) = 0.3, P(E_2) = 0.25, P(E_3) = 0.45.$$

Let $A = \{E_1, E_2\}$. Find$P(A)$.

A. 0.55+++

B. 0.30 C. 0.75 D. 0.90

7. A couple decides to have children until either (i) they have both a boy and a girl or (ii) they have three children. What is the set of outcomes? Write B for boy, G for girl, and write them in birth order.

A. (BG, GGG, BB, GB, BBB, GGB)

B. (BG, GB, BBG, BBB, GGB, GGG) +++

C. (BG, GB, GGG, BBB)

D. (BG, GB, GGB, BBG)

8. A fair die is thrown twice and the two outcomes are added. What is the probability of getting an odd number? Hint: There are 36 outcomes in all. A. 1/36

B. 2/36

C. 18/36 +++

D. 1

9. There are two car factories, A and B. Each year, factory A produces 1000 cars, of which 10 are lemons. Factory B produces 2 cars, which are lemons. All cars go to a single lot, where they are thoroughly mixed up. I buy a car. What is the probability that it is a lemon? A. 0.009

   B. 0.002
   C. 0.012 +++
   D. 0.05

10. There are two car factories, A and B. Each year, factory A produces 1000 cars, of which 10 are lemons. Factory B produces 2 cars, which are lemons. All cars go to a single lot, where they are thoroughly mixed up. I buy a car. What is the probability it came from factory B?

   A. 0.002 +++
   B. 0.009
   C. 0.05 D. 0.012

11. How many subcommittees of 3 people can be chosen from a committee of 8 people?
   A. 28 +++
   B. 10
   C. 16
   D. 82

12. A class has 24 students. 4 can represent the class at an exam board. How many combinations are possible when choosing this group?

   A. 12650
   B. 10626+++
   C. 42504
   D. 255024

13. Suppose $P(A) = 0.70$, $P(B) = 0.20$ and $P(A \cap B) = 0.15$. Find $P(A \cup B)$.

   A. 0.75++
   +
   B. 0.90 C. 0.35
   D. 0.85

14. Suppose $P(A) = 0.70$, $P(B) = 0.20$ and $P(A \cap B) = 0.15$. Find $P(A \cap B^c)$.

   A. 0.55+++
   B. 0.75
   C. 0.05
   D. 0.35

15. In a study of 100 students that had been awarded Ondo State scholarship, it was found that 40 had part-time jobs, 25 had made the Dean's list the previous semester, and 15 had both a part

time job and had made the Dean's list. What is the probability that a student had a part-time or was on the Dean's list?

A. 0.40 B. 0.25

C. 0.10

D. 0.50+++

16. A production department has 35 similar milling machines. The number of breakdowns on each machine averages 0.06 per week. Determine the probabilities of less than three machines breaking down in any week.

A. 0.2700 B. 0.2572

C. 0.1225

D. 0.6497 +++

17. The number of samples of five light-bulbs that can be selected from a lot of 20 bulbs is?

A. 1860480 B. 15504 +++

C. 4845

D. 116280

18. If the probability of producing a defective screw is $p = 0.01$, what is the probability that a lot of 100 screws will contain more than 2 defectives?

A. 7.09% B. 91.97%

C. 92.91%

D. 8.03% +++

19. One of the following is the mathematical definition of a conditional probability of event A given event B has occurred

E. $P(A|B) = \frac{P(A \cap B)}{P(A)}$

F. $P(A|B) = \frac{P(A \cap B)}{P(B)}$ +++

G. $P(B|A) = \frac{P(A \cap B)}{P(A)}$

H. $P(B|A) = \frac{P(A \cap B)}{P(B)}$

20. A random variable y has a pdf given below $f(y) = k(y^2 + 4) \ \ for\ 0 < y < 1$. Find the value of k that makes $f(y)$ a valid pdf?

A. 3/13 +++

B. 3/12

C. 7/13

D. 7/12

21. If two events are independent, then

A. they must be mutually exclusive

B. the sum of their probabilities must be equal to one

C. their intersection must be zero +++

D. None of the alternative

22. Two events, A and B, are mutually exclusive and each have a nonzero probability. If event A is known to occur, the probability of the occurrence of event B is

A. One

B. any positive value

C. Zero +++

D. any value between 0 and 1

23. Sample space is itself an

A. Event +++

B. Subject

C. Probability

D. Statistics

24. Let $P(A) = 0.60$, $P(B) = 0.45$ and $P(A \cap B) = 0.30$. Find $P(A|B)$.

A. 0.50

B. 0.667 +++

C. 0.30

D. 0.75

25. Suppose $P(A_1) = 0.6$ and $P(A_2) = 0.4$. If $P(A_1 \cap A_2) = 0$, are A_1 and A_2 independent events?

A. No +++

B. Yes

C. Neither Yes nor No

D. Undecided

26. Suppose $P(A_1 \cap A_2) = 0$. Suppose further that $P(A_1) = 0.60$, $P(A_2) = 0.40$, $P(B|A_1) = 0.70$ and $P(B|A_2) = 0.20$. Find $P(A_1|B)$.

A. 0.42 B. 0.12

C. 0.28

D. 0.84 +++

27. Probability can be defined as

A. measure of the likelihood that an event will not occur

B. numerical measure of the likelihood that an event will occur +++

C. numerical measure of the likelihood that all events will occur

D. None of the above

28. Suppose two events E and F with nonzero probabilities are mutually exclusive. Then

A. E and F are independent

B. E and F are dependent +++

C. All of the above

D. None of the above

29. An experiment has four equally likely outcomes: E_1, E_2, E_3, and E_4. What is the probability that E_2 occurs?

A. 1/2

B. 2/4

C. 1/4 +++

A. 1/10

30. Let ${C_n}^N$ be the number of combinations of N objects taken n at a time and ${P_n}^N$ be the number of permutations of N objects taken n at a time. What is the relationship between number of permutations and combinations of N objects taken n at a time?

A. ${C_n}^N = n! \times {P_n}^N$

B. ${P_n}^N = n! \times {C_n}^N$ +++

C. $C_n^N = n \times P_n^N$

D. ${P_n}^N = n \times {C_n}^N$

31. An experiment has four equally likely outcomes: E_1, E_2, E_3, and E_4. What is the probability that any three of the outcomes occur (e.g., E_1 or E_2or E_4)? A. 1/24 +++

B. 1/12

C. 1/4

D. 3/4

32. When are events said to be collectively exhaustive? This is when

A. The compliment of the union of events is the entire sample space

B. The intersection of some of the events is a null set

C. The union of events is the entire sample space+++

D. None of the above

33. A fair coin is flip twice. The outcomes are {HH, HT, TH, TT}. Each has the same probability. Which of the following options satisfies that the event H1(where the first flip comes up heads) is independent from the event H2 (where the second flip comes up heads).

E. (H1∪H2)

F. P(H1 ∩ H2)

G. P(H2|H1) = P(H2) +++

H. P(H1) = P(H2)

34. A company receives regular deliveries of raw materials from a supplier. The supplies do not always arrive on time. Over the last 100 delivery days, supplies have been late on 13 occasions. What is the probability that the supplies will be on time on the next delivery day?
A. 1/100

B. 87/100
+++
C. 13/100 D.
0.012

35. A die is rolled and a coin is tossed, find the probability that the die shows an odd number and the coin shows a head.
A. 1
B. 1/6
C. 1/3
D. ¼ +++

36. Suppose that 24% of companies in a certain sector of the economy have announced plans to expand in the next year (and the other 76% will not). In a sample of twenty companies chosen at random drawn from this population, where n=20. Find the probability that the number of companies which have announced plans to expand in the next year will be three or more.
A. 0.8915 +++
B. 0.1484
C. 0.5331
D. 0.7321

37. Suppose that 24% of companies in a certain sector of the economy have announced plans to expand in the next year (and the other 76% will not). In a sample of twenty companies chosen at random drawn from this population, find the probability that the number of companies which have announced plans to expand in the next year will be precisely five.
A. 0.2012 +++
B. 0.2514
C. 0.2381
D. 0.2325

38. The volume of water in commercially supplied fresh drinking water containers is approximately Normally distributed with mean 70 litres and standard deviation 0.75 litres. Estimate the proportion of containers likely to contain. P(X > 70.9)

A. 0.\251%
B. 0.511
C. 0.111
D. 0.1151 +++

39. Two events A and B are said to be INDEPENDENT if . . . . . .
A. $P(A|B) = P(A) \cdot P(B)$
B. $P(A|B) = P(A \cap B)$
C. $P(A|B) = P(A)$ +++ D. $P(A|B) = 0$

40. The probability density function for a binomial distribution is given as . . . . . .
A. $P(X = x) = \binom{x}{n} P^x (1-P)^{n-x}$
B. $P(X = x) = \binom{n}{x} P^n (1-P)^{n-x}$
C. $P(X = x) = \binom{n}{x} P^x (1-P)^{n-x}$ +++

D. $P(X = x) = \binom{n}{x} P^x (1 - P)^{x-n}$

41. The probability mass function for a Poisson distribution is given as . . . . . .

A. $P(x, \lambda) = \frac{e^{-\lambda} x^{\lambda}}{x!}$

B. $P(x, \lambda) = \frac{e^{-\lambda} \lambda^{x}}{\lambda!}$

C. $P(x, \lambda) = \frac{e^{-x} x^{\lambda}}{x!}$ +++

D. $P(x, \lambda) = \frac{e^{-\lambda} \lambda^{x}}{x!}$

42. How many ways can four children be lined up?

A. 16 B. 4

C. 12

D. 24 +++

43. Suppose out of 10 objective questions in STA 122, five are to be answered. How many different sets of questions are possible?

A. 50

B. 30240 +++

C. 1200

D. 24000

44. The set of all possible values of a random variable X is called

A. Values of X

B. Support of X +++

C. Random support of X

D. Density of X

45. A random variable X is a discrete random variable if

A. there are a finite number of possible outcomes of X.

B. there are a countably infinite number of possible outcomes of X. C. all of the above. +++

D. none of the above.

46. The expected value of *X* having binomial distribution with parameters *n* and *p* is

A. np +++

B. n/p

C. p/n

D. 1/(np)

47. Tay-Sachs disease is a rare but fatal disease of genetic origin occurring chiefly in infants and children, especially those of eastern European extraction. If a couple are both carriers of Taysachs disease, a child of theirs has probability 0.25 of being born with the disease. If such a couple has four children, what is the probability that no child will have the disease?

A. 0.316 +++

B. 0.114

C. 0.322

D. 0.631

48. Tay-Sachs disease is a rare but fatal disease of genetic origin occurring chiefly in infants and children, especially those of eastern European extraction. If a couple are both carriers of Taysachs disease, a child of theirs has probability 0.25 of being born with the disease. If such a couple has four children, what is the probability that all the children will have the disease? A. 0.4

    B. 0.04

    C. 0.004 +++

    D. 0.0004

49. Let X be the number of borehole a company drill with a mean of 3 borehole per town in Nigeria. What is the probability that the company drill at most one borehole in a randomly selected town? A. 0.0498

    B. 0.4980

    C. 0.1992 +++

    D. 0.4000

50. Let X be the number of borehole a company drill with a mean of 3 borehole per town in Nigeria. What is the probability that the company drill at least one borehole in a randomly selected town A. 0.9000 B. 0.9800

    C. 0.5000

    D. 0.9502 +++

51. Suppose scores on an IQ test are normally distributed. If the test has a mean of 100 and a standard deviation of 10, what is the probability that a person who takes the test will score greater than 120?

    A. 0.0228 +++

    B. 0.228

    C. 0.822

    D. 0.0822

52. Suppose scores on an IQ test are normally distributed. If the test has a mean of 100 and a standard deviation of 10, what is the probability that a person who takes the test will score between 90 and 110?

    A. 0.0050 B. 0.0500

    C. 0.6800

    D. 0.6826 +++

53. Which of the following is correct?

    A. The expected value of X having exponential distribution with rate parameter λ is $1/\lambda$..

    B. The exponential distribution is memoryless.

    C. The exponential distribution is often used to model life time of electrical components or electronic device.

    D. All of the above. +++

SECTION C: Estimation and Test of Hypothesis

1. If sample mean is 20, population standard deviation is 3 and sample size is 64, find the interval estimate of the mean
   a. (19.02, 21.50)
   b. (19.00, 20.50)
   c. (19.27, 20.74) +++
   d. (19.22, 21.50)
2. Statistical inference consists of two methods, these are:
   A. Classical and interval
   B. Classical and Bayesian +++
   C. Point and interval
   D. Point and Bayesian
3. Use 90% confidence and sample data to construct a confidence interval for the population proportion p with n = 140, x = 81:
   A. $0.514 < p < 0.644$
   B. $0.509 < p < 0.649$ C. $0.510 < p < 0.648$ +++
   D. $0.513 < p < 0.645$
4. 50 people are selected randomly from a certain population and it is found that 13 people in the sample are over 6 feet tall. What is the point estimate of the true proportion of people in the population who are over 6 feet tall?
   A. 0.50
   B. 0.74
   C. 0.26 +++ D. 0.19
5. Of 112 randomly selected adults, 34 were found to have high blood pressure. Construct a 95% confidence interval for the true percentage of all adults that have high blood pressure.
   A. $19.1\% < p < 41.6\%$
   B. $23.2\% < p < 37.5\%$
   C. $21.8\% < p < 38.9\%$ +++
   D. $20.2\% < p < 40.5\%$
6. Use 99% confidence level and sample data to find a confidence interval for estimating the population μ with test scores: $n = 95$, $\bar{x} = 73.8$, $\sigma = 7.4$
   A. $72.5 < \mu < 75.1$
   B. $71.8 < \mu < 75.8$ +++
   C. $72.3 < \mu < 75.3$
   D. $72.0 < \mu < 75.6$
7. A laboratory tested twelve chicken eggs and found that the mean amount of cholesterol was 208 milligrams with s = 18.2 milligrams. Construct a 95% confidence interval for the true mean cholesterol content of all such eggs.
   A. $196.4 < \mu < 219.6$ +++
   B. $198.6 < \mu < 217.4$
   C. $196.3 < \mu < 219.7$
   D. $196.5 < \mu < 219.5$
8. The mean and standard deviation of the diameters of a sample of 250 rivet heads manufactured by a company are 0.72642 in and 0.00058 in, respectively. Find the 90% confidence limits for the mean diameter of all the rivet heads manufactured by the company.

A. $\mu = 0.64672 \pm 0.000060$
B. $\mu = 0.92642 \pm 0.000072$
C. $\mu = 0.72642 \pm 0.000060$ +++
D. $\mu = 0.26427 \pm 0.000060$ 9. A two-tailed test is one where:
A. results in only one direction can lead to rejection of the null hypothesis
B. negative sample means lead to rejection of the null hypothesis
C. results in either of two directions can lead to rejection of the null hypothesis +++
D. no results lead to the rejection of the null hypothesis

10. A statement or claim that an experimental wants to test is called:
A. hypothesis +++
B. alternative hypothesis
C. optional hypothesis
D. null hypothesis

11. All the following are procedures for testing hypothesis except:
A. Choose level of significance
B. Define basic statistical terms +++
C. Formulate hypothesis
D. Make statistical decision

12. A manufacturer of multi-vitamin tablet claims that riboflavin content of his tablets is greater than 2.49mg. A check by NAFDAC using 82 tablets shows a mean riboflavin content of 2.52mg with standard deviation of 0.18mg, calculate the value of the test statistic.
A. 1.04
B. 2.002
C. 1.04
D. 1.507 +++

13. The Probability of committing an error by the most unbiased neutral judge taking the position that frees the guilty is:
A. Type II error
B. Level of significance +++
C. Type I error
D. Power of the test

14. An error committed by most unbiased neutral judge taking the position that frees the guilty is:
A. Type I error +++
B. Margin of error
C. Type II error
D. Power of the test

15. The null and alternative hypotheses divide all possibilities into:
A. two sets that overlap
B. two non-overlapping sets +++
C. two sets that may or may not overlap
D. as many sets as necessary to cover all possibilities

16. If we reject the null hypothesis when it is, in fact, true then we have:
A. made a Type I error +++

B. made a Type II error
C. made scientific progress
D. both (b) and (c) above

17. Which of the following is true of the null and alternative hypotheses?
    A. Exactly one hypothesis must be true +++
    B. both hypotheses must be true
    C. It is possible for both hypotheses to be true
    D. It is possible for neither hypothesis to be true
18. Out of 300 households in a town 123 have T.V sets. Find 95% confidence limits to the true value of the proportion of the households with T.V. sets in the whole town.
    A. 0.410 – 0.055
    B. 0.410 – 0.355
    C. 0.355 – 0.465 +++ D. 0.410 – 0.465
19. Of type I and type II errors, which is traditionally regarded as more serious?
    A. Type I +++
    B. Type II
    C. They are equally serious
    D. Neither is serious
20. A two-tailed test is one where:
    A. results in only one direction can lead to rejection of the null hypothesis
    B. negative sample means lead to rejection of the null hypothesis
    C. results in either of two directions can lead to rejection of the null hypothesis +++
    D. no results lead to the rejection of the null hypothesis
21. The value set for □ is known as:
    A. the rejection level
    B. the acceptance level
    C. the significance level +++
    D. the error in the hypothesis test
22. The form of the alternative hypothesis can be:
    A. one-tailed
    B. two-tailed
    C. neither one nor two-tailed
    D. one or two-tailed +++
23. If a teacher is trying to prove that new method of teaching Mathematics is more effective than traditional one, he/she will conduct a:
    A. one-tailed test +++
    B. two-tailed test
    C. point estimate of the population parameter
    D. confidence interval
24. If our sample size is < 30, the suitable test statistic is:
    A. Z-statistic
    B. Chi-square
    C. Null hypothesis D. t-statistic +++

25. A statistical statement which may or may not be true concerning one or more populations is called: A. Type I error

    B. Critical region

    C. Statistical hypothesis +++

    D. Chi-square distribution

26. A quantity from the given information which when compared with the tabulated value to make decision about hypothesis is known as:

    A. Alternative hypothesis
    B. Null hypothesis
    C. One tail test
    D. Test statistic +++

27. Find the critical value $Z\propto_{/2}$ that corresponds to a degree of confidence of 98%.

    A. 2.575

    B. 2.05

    C. 2.33 +++ D. 1.75

28. An urn contains an unknown proportion of red and white marbles. A random sample of 60 marbles selected with replacement from the urn showed that 70% were red. Find the 95% confidence limits for the actual proportion of red marbles in the urn.

    A. $0.68 < p < 0.82$ B. $0.65 < p < 0.78$

    C. $0.25 < p < 0.65$

    D. $0.58 < p < 0.82$ +++

29. Find the critical value $Z\propto$ $_{/2}$ that corresponds to a degree of confidence of 91%.

    A. 1.75

    B. 1.70 +++

    C. 1.645

    D. 1.34

30. One-tailed alternatives are phrased in terms of: A. □

    B. < or > +++

    C. □ or =

    D. □ or □

31. In constructing confidence interval when the sample size is small, we use:

    A. F-distribution

    B. t-distribution +++

    C. z-distribution

    D. p-distribution

32. In constructing confidence interval when the sample size is large, we use:

    A. F-distribution

B. t-distribution

C. z-distribution +++

D. p-distribution

33. Find the point estimate of the true proportion of people who wear hearing aids if, in a random sample of 496 people, 73 people had hearing aids.

A. 0.128

B. 0.853

C. 0.147 +++ D. 0.145

34. A survey of 865 voters in one state reveals that 408 favour approval of an issue before the legislature. Construct the 95% confidence interval for the true proportion of all voters in the state who favour the approval.

A. $0.438 < p < 0.505$ +++

B. $0.444 < p < 0.500$

C. $0.471 < p < 0.472$

D. $0.435 < p < 0.508$

36. A random sample of 194 full-grown lobsters had a mean weight of 21 ounces and a standard deviation of 3.8 ounces. Construct a 98% confidence interval for the population mean μ.

A. $20 < \mu < 23$

B. $21 < \mu < 23$

C. $20 < \mu < 22$ +++ D. $19 < \mu < 21$

37. Given the following summary from test scores: n = 95, $\bar{x}$ = 73.8, σ = 7.4, what is the 99% confidence interval for population mean μ?

A. $72.5 < \mu < 75.1$

B. $71.8 < \mu < 75.8$ +++

C. $72.3 < \mu < 75.3$ D. $72.0 < \mu < 75.6$

38. A sociologist develops a test to measure attitudes about public transportation, and 27 randomly selected subjects are given the test. Their mean score is 76.2 and their standard deviation is 21.4. Construct the 95% confidence interval for the mean score of all such subjects.

A. $74.6 < \mu < 77.8$

B. $67.7 < \mu < 84.7$ +++

C. $64.2 < \mu < 88.2$ D. $69.2 < \mu < 83.2$

39. Two independent sampling stations were chosen for this study, one located downstream from the acid mine discharge point and the other located upstream. For 12 monthly samples collected at the downstream station, the species diversity index had a mean value $\bar{x}_1$ = 3.11 and a standard deviation s_1 = 0.771, while 10 monthly samples collected at the upstream station had a mean index value $\bar{x}_2$ = 2.04 and a standard deviation s_2 = 0.448. Find a 90% confidence interval for the difference between the population means for the two locations.

A. $0.45 \leq \mu_1 - \mu_2 \leq 1.05$ B.
$1.05 \leq \mu_1 - \mu_2 \leq 1.55$

C. $0.59 \leq \mu_1 - \mu_2 \leq 1.05$

D. $0.59 \leq \mu_1 - \mu_2 \leq 1.55$ +++

40. A professor of statistics claims that the proportion of independent voters in Minnesota is at greater than 40%. To test the claim which of the following alternative hypothesis should we use? a. H_1: $\mu = 0.40$

b. H_1: $P = 0.40$

c. H_1: $P > 0.40$ +++

d. H_1: $P \neq 0.40$

41. The ability of a test to reject wrong hypothesis is called

A. Power of a Test +++

B. Power function

C. Decision function

D. Test Statistic

42. The Hypothesis H_0:$\mu_A = \mu_B$ vs H_1:$\mu_A \neq \mu_B$ is

A. Directional

B. Two- tailed test

C. Non- directional

D. B and C +++

43. …….. is the study of relationship existing between a population and samples drawn from the same population.

A. Statistics
B. Distribution
C. Sampling theory +++
D. Sample

44. The probability distribution of a sample ……. is called sampling distribution

A. Parameter
B. Statistic +++
C. Population
D. Member

45. …… approximation is considered suitable when the sample size is large.

A. t-distribution
B. Normal distribution +++
C. Binomial distribution
D. p-distribution

46. When is a random sample of size n said to be small?
    A. $n < 20$
    B. $n \leq 20$
    C. $n < 30$ +++
    D. $n \leq 30$
47. A random sample of 100 oranges was taken from a big basket and 13 of them were found to be bad. What is the standard error of the proportion of the bad ones?
    A. 0.1300
    B. 0.0295 +++
    C. 0.000871
    D. 0.6700

48. …….. is the statistic that is used to infer the value of an unknown population parameter in a statistical model
    A. Estimate
    B. Parameter
    C. Estimator +++
    D. Statistics

49. An estimator is a …… while estimate is a value. A. Function +++
    B. Statistics
    C. Distribution
    D. Proportion

SECTION D: Simple Regression and Correlation

1. Which phrases i-iv below applies to the simple correlation coefficient?

    (i) Measures linear association between two variables,

    (ii) It is never negative,

    (iii) It has positive slope,

    (iv) It depends on the units of measurement of X and Y.

    A. iii only

    B. i and ii

C. i only +++

D. iii and iv

2. Which of the following expressions denotes a covariance?

A. $\frac{1}{\mathrm{n}}\sum_{i=1}^{\mathrm{n}}(x-\bar{x})(y-\bar{\mathrm{y}})$ +++

B. $\frac{1}{\mathrm{n}}\sum_{i=1}^{n}(x-\bar{x})(\bar{\mathrm{y}}-\mathrm{y})$

C. $\frac{1}{\mathrm{n}}\sum_{i=0}^{n}(x-\bar{x})(y-\bar{\mathrm{y}})$

D. $\frac{1}{\mathrm{n}}\sum_{i=0}^{n}(x-\bar{x})(\bar{\mathrm{y}}-\mathrm{y})$

3. Which of the following expressions denotes a covariance?

A. $\frac{1}{n}(\sum_{i=1}^{1}xy-nx\bar{y})$

B. $\frac{1}{n}\sum_{i=1}^{1}xy-n\overline{xy}$

C. $\frac{1}{n}(\sum_{i=1}^{1}xy-\overline{xy})$

D. $\frac{1}{n}\sum_{i=1}^{1}xy-\overline{xy}$ +++

4. Find the covariance for the following data

Height (m) x	1.71	1.60	1.64
Weight (kg) y	60	57	53

A. 0.186.87

B. -186.87

C. 0.127 +++

D. -0.127

5. Which of the following is not a type of correlation?

A. Pearson Product Moment

B. Spearman Ranks

C. Pearson Product Multiple +++

D. Kendall Ranks

6. Suppose the value of the correlation coefficient obtained for two continuous variables *x* and *y* using is -0.682, which of the following statement is true?

A. There is a perfect weak negative correlation

B. Both values of *x* and *y* decreases

C. As the values of *x* increases, the values of *y* decreases +++

D. Both values of *x* and *y* increases.

7. A student produces a correlation of +1.3. This is

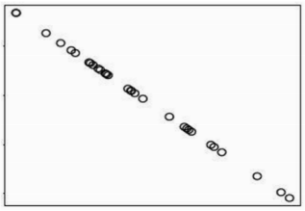
A. a high positive correlation

B. a significant correlation

C. an impossible correlation +++

D. only possible if N is large

8. The Spearman correlation is used with
   A. measured data
   B. quantitative data
   C. categorical data +++
   D. scaled data
9. The regression model can be used to
   A. summarize a relationship between variables
   B. predict a value of one variable from its relationship to another related variable
   C. analyze a relationship between two variables
   D. all of the above +++
10. The scatter diagram below displays:



   A. Strong positive correlation
   B. Weak Positive Correlation
   C. Weak negative Correlation
   D. Strong negative correlation +++
11. Given below are data points for variables x and y. Calculate the Pearson correlation coefficient for these data?

X	20	18	24	20	22	14	18
Y	16	12	18	17	21	10	10

   A. 0.56
   B. 0.79
   C. 0.84 +++
   D. 0.77
12. It is known that length of snakes do affect their weight. As part of study, the following data were recorded from the free snakes ranging in an area.

Length	3	9	10	11	13	15
Weight	4	7	8	7	9	13

Predict the weight of a snake for a length value of 20?

A. 14.49 +++
B. 16.7
C. 20.4 D. 6.7

13. Which correlation is the strongest: +0.65 or -0.70?
    A. -0.70 +++
    B. +0.65
    C. Depends on N
    D. The information given is incomplete.

14. The nature and strength of relationship between two or more variables is known as _____
    A. Regression analysis
    B. correlation analysis +++
    C. probability analysis
    D. chi-square analysis

15. _____determines the association between two variables but it does not prove that one particular variable alone cause the change in the other.
    I. Variance ratio
    J. Correlation coefficient
    K. Correlation +++
    L. Coefficient of determination

16. The formula to calculate spearman rank correlation is _____
    a. $r_s = 1 - \frac{6\square^2}{n(n^2-1)}$ +++
    b. $r_s = 1 - \frac{6\square^2}{n(n^2+1)}$
    c. $r_s = 1 + \frac{6\square^2}{n(n^2-1)}$
    d. $r_s = 1 - \frac{6\square^2}{n(n^2+1)2}$

17. The distinction between simple and multiple correlation is based upon _____
    A. number of variables +++
    B. number of variance
    C. number of average
    D. correlation coefficient

18. In the linear equation of a least square method Y = a + bx, Y represents _____
    A. Dependent variable +++
    B. Independent variable
    C. Constant variable
    D. Slope variable

19. In regression analysis, the best known method for having a straight line equation is known as _____

A. Least square method +++

B. Linear regression method

C. Non-linear method

D. partial regression method

20. For a normal good, if price increases then the quantity demanded decreases. What type of correlation co-efficient would you expect in this situation? A. 0

B. positive

C. negative +++

D. unknowable

21. Suppose the regression line for a set of data, $\hat{y} = 3x + b$, passes through points (2,5). If $\bar{x}$ and $\bar{y}$ are the sample means of the *x* and *y* values respectively, then $\bar{y}$ =

A. $\bar{x}$

B. $\bar{x} - 2$

C. $\bar{x} + 5$

D. $3\bar{x} - 1$ +++

22. The equation of the least squares regression line for a set of points in a scatterplot is given by $\hat{y} = 1.3 + 0.27x$. The point (3, y_1) is one point on this scatterplot. What is the value of y_1

A. 0.27

B. 1.3

C. 2.11 +++

D. 0.81

23. In the equation of a straight line, $Y = \beta_0 + \beta_1 x$ the term, β_1 is the:

A. Slope +++

B. Independent variable

C. Dependent variable

D. Intercept

24. Fit a linear model on the following data

X	22	18	31	25
Y	100	121	110	130

A. Y = 122.18 - 0.30X +++

B. $Y = 122.18 + 0.30X$

C. $Y = 0.30 + 122.18X$

D. $Y = 0.30 - 122.18X$

25. Which of the following equation is not correct?

A. $r = \frac{1}{n\sigma_x\sigma_y}\sum_{i=1}^{n} xy - \overline{xy}$

B. $r = \frac{\sum_{i=1}^{n} xy - \overline{xy}}{\sqrt{(\sum_{i=1}^{n} x^2 - \overline{x}^2)(\sum_{i=1}^{n} y^2 - \overline{y}^2)}}$

C. $r = \frac{\overline{xy} - \sum_{i=1}^{n} xy}{\sqrt{(\overline{x}^2 - \sum_{i=1}^{n} x^2)(\overline{y}^2 - \sum_{i=1}^{n} y^2)}}$

D. $r = \frac{\frac{1}{n}\sum_{i=1}^{n}(x-\bar{x})(y-\bar{y})}{\sum_{i=1}^{n}(x-\bar{x})\sum_{i=1}^{n}(y-\bar{y})}$ +++

26. For the regression equation *Y* = 3*X* - 2, if the mean of *Y* is 10, what is the mean of *X*?

A. 8

B. 28 +++

C. 4

D. Cannot be determined

27. The following statements are true except one

A. Regression shows the relationship between two variables

C. Covariance of Standardized data is Correlation

B. Correlation measures the level of association of two variables

D. The intercept coefficient in simple linear regression is a measure of correlation +++

28. Given that *X* = 1.50 + .50*Y*, what is the predicted value for a *Y* value of 6?

A. 3.00 B.
3.50

C. 4.00

D. 4.50 +++

29. The coefficient of determination is defined as _____ A. Ratio of explained variance to the total variance +++

B. Ratio of total variance to the explained variance

C. Ratio of unexplained variance to the total variance

D. Ratio of explained variance multiply by total variance

30. In regression analysis, a straight line relationship between variable x and y can be represented by an equation _____

A. y = a + bx +++

B. y= a – bx

C. y = a + bx^2

D. y = a - bx^2

31. In spearman rank correlation, 'd' represent _____

A. The difference between the ranks +++
B. The difference of the rank
C. The difference with the rank
D. None of the above

32. The coefficient of variation can be calculated by

A. $\frac{\text{d}}{\bar{x}} \times \frac{100}{1}$
B. $\frac{\sigma}{100} \times \bar{x}$
C. $\frac{\sigma}{\bar{x}} \times \frac{n}{100}$
D. $\frac{\sigma}{\bar{x}} \times 100$+++

33. In regression analysis, a straight line relationship between variable x and y can be represented by an equation _____

A. y = a + bx +++
B. y= a – bx
C. y = a + bx^2
D. y = a - bx^2

34. If the slope of the regression line is calculated to be 2.5 and the intercept 16 then the value of *Y* when *X* is 4 is:

A. 2.5
B. 66.5
C. 16
D. 26 +++

35. R^2 is the statistical notation for

A. Pearson's Co-efficient of Correlation
B. The Co-efficient of Variation
C. Spearman's Co-efficient of Rank Correlation
D. The Co-efficient of Determination +++

36. All correlation coefficients share in common the property that they range between

A. +1 and 0
B. +1.00 and -1.00. +++
C. + 0.1 and –0.1
D. +1.96 and -1.96

A. Height ..... A
D. 24 ..... B
E. $P(A|B) = \frac{P(A \cap B)}{P(A)}$ 12

B. 0 +++
C. +1

D. +0.1

38. In a statistics course, a linear regression equation was computed to predict the final exam score based on the score on the first test of the semester. The equation was: $\hat{y}=25+0.7x$ where *y* is the final exam score and *x* is the score on the first test. Funke scored 80 on the first test. On the final exam Funke scored 85. What is the value of her residual?
    B. -4
    C. 4 +++
39. A straight line of the for , $\hat{y}=a+bx$ is fitted to the 5 data points (1,0), (0,3), (3,1), (2,-3) and (4,2) by the method of least squares. What is the value of *b?*
    A. -3
    B. -2/3
    C. -1/10 +++
    D. 4
40. Suppose the regression equation, $\hat{y}=6.4-0.12x$ represents the relationship between years of schooling completed and the number of job offers. Where *x* is the number of years completed in school and *y* is the number of job offers. What is the estimated change in the number of job offers that corresponds to the completion of additional 10 years of school?
    A. a decrease of 5.2
    B. an increase of 7.6
    C. a decrease of 7.6
    D. a decrease of 1.2 +++
41. The following data is regressed with least squares regression $Y = \beta_0 + \beta_1 x$.

X	1	20	30	40
Y	1	400	800	1300

The value of β_1 most nearly is

    A. 27.48
    B. 28.97
    C. 32.63 +++ D. 40.00
42. At an agricultural show 10 shetland sheep were ranked by a qualified judge and by a trainee judge. Their rankings are shown in the table. Calculate the correlation coefficient for these data.

Qualified Judge	1	2	3	4	5	7	8	9	5	10
Trainee Judge	1	3	2	6	7	3	9	10	3	8

    A. 0.685
    B. 0.827 +++
    C. 0.797
    D. -0.685

43. In a linear regression $Y = a + bX$, the fitted or predicted value of Y is? A. The value of X associated with a particular value of Y.

B. The value of X predicted by the regression equation.

C. The values of the parameters predicted by the estimators

D. The values of Y obtained by substituting specific values of X into the estimated sample regression equation +++

44. Which phrases i-iv below apply to the sample correlation coefficient?

(i) Measures linear association between two variables,

(ii) Is never negative,

(iii) Has positive slope,

(iv) Depends on the units of measurement of X and Y .

A. iii only
B. i and ii
C. i only +++ D. iii and iv

45. A study was conducted to find out if men and women differ in their attitudes toward college-level mathematics. Twenty men and thirty-eight women selected at random from those identified as being at high risk of failure participated in the study. Each student was asked to respond to a series of questions, and the answers were combined to obtain a math anxiety score. What type of variables is presented in the study?

A. Two categorical independent variables and a numerical dependent variable B. A categorical dependent variable and a numerical independent variable C. Two categorical dependent variables and a numerical independent variable

D. A categorical independent variable and a numerical dependent variable +++

46. Which of the following denotes the nature of relationship between two variables?

A. The value of the correlation coefficient (r)

B. The sign of the correlation coefficient (r) +++

C. The covariance between the two variables

D. All of the above

47. The value of a correlation coefficient signifies

A. the nature of relationship that exits between two variables

B. that there is a perfect relationship between two variables

C. that there is a positive relationship between two variables

D. the strength of relationship that exits between two variables +++

SECTION F: Design and Analysis of Experiment

1. The following are the basic principles of Design of Experiment except ----------
    A. Randomization
    B. Replication
    C. Blocking
    D. Robustness +++
2. ---------- is a method of increasing the precision of experiment.
    A. Randomization
    B. Replication +++
    C. Confounding
    D. Blocking
3. A way of controlling the uncertainty in our results during experiments is called ----------
    A. Replication +++
    B. Randomization
    C. Blocking
    D. Factorial
4. A method of controlling uncertainty in the results from experimental design is called ----------
    A. Replication +++
    B. Randomization
    C. Blocking
    D. Factorial
5. A technique useful in controlling known sources of variation in the design of experiments is ----
    -

A. Blocking +++

B. Replication

C. Robustness

D. Confounding

6. The first step in planning, conducting and analyzing an experiment is ---------

A. Recognition and statement of the problems +++

B. Choice of factors

C. Choice of designs

D. Conducting the experiment

7. $Y=\beta_0+\beta_1X_1+\beta_2X_2+\beta_{12}X_1X_2+\in$. In this equation, β_{12} represents------

A. Output variables

B. Response variables

C. Interaction between effects +++

D. Input variables

8. Examples of Designs of Experiment are except ----------

A. Completely randomized block design

B. Randomized block design

C. Organized block design +++

D. Response surface design

9. Well-chosen experimental designs maximize the amount of --------- that can be obtained in experimental effect.

A. Data

B. Information +++

C. Mean

D. Median

10. The first era of experimental design was introduced by ----------

A. Box and Wilson

B. R.A. Fisher +++

C. Montgomery D. C.

D. Bodunwa O.K

11. An experiment involving two levels of temperature and three surgical procedures conducted on a set of animals was intended to measure the number of days taken for the experiment to be completed. The experiment was repeated ten times. This experiment is an example of

    A. 3 X 2 factorial design +++
    B. 2 X 3 X N factorial design
    C. Many factors experiment
    D. Experiment involving temperature, surgical procedure and animals

12. A combination of levels of factors in an experiment is known as

    A. Response variable
    B. Replicate
    C. Treatment +++
    D. Experimental design

13. For an ANOVA model given as $y_{ij} = \mu + \tau_i + e_{ij}$, the y_{ij} term represents

    A. Response +++
    B. Observation effect for the i^{th} treatment term
    C. Random error term
    D. Treatment effect

14. In an experiment involving four levels of a factor was such that level I was replicated 5 times, level II replicated 7 times, level III replicated 7 times and level IV was replicated

    10 time. Such an experiment is said to be

    A. an ANOVA design
    B. an experimental design
    C. an unbalance design +++
    D. a balance design

15. An experiment involves k > 2 treatments with N total observations. The F-ratio arising from the analysis has an F-distribution with what degrees of freedom?

    A. N – k
    B. k – 1
    C. (k – 1)(N – k) +++
    D. (N – k)(k – 1)

16. An experiment involving k > 2 treatments with *N* total observations has ---- degrees of freedom for the error term.

    A. N – k +++
    B. k – 2
    C. (k – 1)(N – k)
    D. (N – k)(k – 1)

17. Results of an analysis of variance are declared significant if:
    A. the value of F-ratio is less than the F value obtained from statistical table
    B. the value of F-ratio is greater than the F value obtained from statistical table +++ C. the value of the mean square error (MSE) is greater than the value obtained from statistical table
    D. the means of the treatments are considered significant

18. In an experiment involving 3 treatments with each replicated 5 times, what would be the value of the F distribution to be used in deciding the significance of the means assuming

    □□ 0.05?

    A. 5.41
    B. 6.94
    C. 3.89 +++
    D. 6.93

19. Which of the following statistics is used to compute sum of square due to treatment in analysis of variance?

    A. $\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y})^2$
    B. $\sum_{i=1}^{k} n_i(\bar{y}_i - \bar{y})^2$ +++
    C. $\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2$
    D. $\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2$

20. The smallest unit to which treatment can be assigned in an experiment is called?

    A. Experimental unit +++
    B. Level unit
    C. Design unit
    D. Factor unit

21. An experiment involving three different flow rates was performed such that each flow rate was replicated 6 times and the response variable measured. It was desired to test the mean difference of the responses based on the flow rates using ANOVA. What would be the degree of freedom for the error term?

    A. 6
    B. 18
    C. 12 +++
    D. 5

22. Which of the following best characterizes the null hypothesis for chi-square tests? A. The frequency of people in each category will be a specific proportion +++

    B. The expected frequencies will be significantly lower than obtained frequencies
    C. The group means expected at the population level will not reliably differ
    D. All of the above

23. Given the following responses from and experiment and it was desired to perform an ANOVA test, compute the sum of square for total (SST).

    Treatment I: 12, 16, 9
    Treatment II: 15, 16, 21, 20, 18
    Treatment III: 8, 6, 9, 6

    A. 316.0 +++
    B. 258.58
    C. 320.0
    D. 129.29

24. How many eras do we have in history of design of experiments?

    A. Two
    B. Three
    C. Four +++
    D. Five

25. The essence of performing an experiment is to ----------

    A. Draw conclusion and make recommendations +++
    B. Selection of the response variable
    C. Statistical analysis
    D. Choice of design

26. Design of Experiment is a ---------- and ---------- ways of conducting and analyzing controlled tests to evaluate the factors that are affecting the response variables

A. Structured, organized +++
B. Conclusion, recommendation
C. Organized, conclusion
D. Unstructured, recommendation

27. Given the following extract from an ANOVA table, what would be the value of F ratio?

Source of variation	Degree of freedom	Sum of square
Treatment	4	258.70
Error	19	158.99
Total	23	417.69

A. 7.729 +++
B. 0.129
C. 64.675
D. 8.368

28. What are the values of the following: $F_{3,11;0.05}$ and $F_{2,9;0.01}$?

A. 5.12 and 9.55
B. 3.52 and 6.70
C. 3.59 and 8.02 +++
D. 4.45 and 6.11

29. Which of the following statistics is used to compute sum of square total in analysis of variance?

A. $\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y})^2$ +++

B. $n\sum_{i=1}^{k}(\bar{y}-\bar{y})^2$

C. $\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i})^2$

D. $\sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij2}$

30. Given the following extract from an ANOVA table, what would be the value of the mean square error (MSE)?

Source of variation	Degree of freedom	Sum of square	Mean square

Treatment	7	75.54	
Error	----	---	
Total	34	127.02	

A. 0.907
B. 1.907 +++
C. 10.791
D. 5.660

31. Name one method of randomization

A. Fish method
B. Hat method +++
C. Drop & pick method
D. None of the above

32. An Experiment is pre-arranged from the beginning. True or False?
A. True +++
B. False
C. I don't know
D. All of the above

33. Let the yield (kg) per plot from an experiment be as follows

Variety	Observations		
A	14.3	11.6	11.8 14.2
B	20.7	21.0	
C	32.6	32.1	33.0
D	24.3	25.2	24.8

What is the grand mean of the data?
A. 23.12
B. 22.13 +++
C. 19.33
D. 20.13

34. ANOVA test enables us to compare several population means separately. True or False?
A. True
B. False +++
C. I don't know
D. All of the above

35. One of the following is an assumption of ANOVA
A. Observations are dependent
B. Each treatment is non-normal
C. There is homogeneity of variance +++
D. There is heteroscedasity

36. Which of the following is not a method of removing extraneous sources of variation in an experimental design?

A. Balancing
B. Homogeneous grouping
C. Blocking
D. Randomization +++

37. ---------- is a method of removing all extraneous sources of variation
A. Randomization
B. Replication
C. Local control +++
D. Completely randomized

38. By homogeneity of variance, we mean A. The treatments have common variance +++
B. The treatments are taken from the same experiment
C. The variance several treatments are not too big
D. None of the above

39. Examples of experimental design include the following except
A. Complete randomize block design
B. Split-plot design
C. Latin square design
D. Youden square design +++

40. Given that the treatment and error degrees of freedom are respectively 2 and 9, and that the F calculated from the ANOVA is 6.00 tabulated what would be the conclusion from the analysis at □□ 0.05?
A. The null hypothesis will be rejected +++
B. The null hypothesis will be accepted
C. The alternate hypothesis will be determined
D. More information is required to reach a conclusion

41. Calculate the sum of squares treatment for the following set of data with treatments A, B, C as follows: Treatment A (3, 4, 5, 4), Treatment B (2, 4, 3, 3) and Treatment C (4, 6, 5, 5)

A. 20
B. 8 +++
C. 5 D. 11

42. The following are the assumption of ANOVA except?

A. Independent observation
B. Normality
C. Homogeneity of variance
D. Heteroskedacity +++

43. When similar experimental units are collected together, it is called ----

A. ANOVA

B. Blocking +++

C. Objectiveness

D. Correction

44. What basic principle of experimental design is used to remove extraneous sources of variation?

A. Replication

B. Randomization

C. Local Control +++ D. ANOVA

45. To increase precision in an experiment, what basic principle is needed?

A. Randomization

B. Replication +++

C. Local Control

D. None of the above

46. What principle of experimentation reduces the standard error of mean to increase precision?

A. Replication +++

B. Local Control

C. Randomization

D. ANOVA