Course ID : 31610   **Course Code : SP/ECO/304/SEC-1**

Course Title : Data Analysis

Time: 2 Hours   **Full Marks: 40**

Answer *any 20* of the following.   20×2=40

যে কোনো কুড়িটি প্রশ্নের উত্তর দাও।

1. A tabular summary of a set of data showing the fraction of the total number of items in several classes is a

   a. Frequency distribution
   b. Relative frequency distribution
   c. Frequency density
   d. Cumulative frequency distribution

2. What are the types of statistical data?

a. Primary and private data
b. Primary and secondary data
c. Sample and primary data
d. Sample and secondary data

3. Since the mode is the most frequently occurring data value, it

   a. can never be larger than the mean
   b. is always larger than the median
   c. is always larger than the mean
   d. None of the above answers is correct

4. Which one of these statistics is unaffected by outliers?

   a. Mean
   b. Interquartile range
   c. Standard deviation
   d. Range

5. The sum of the percent frequencies for all classes will always equal
   a. One
   b. 100
   c. the number of items in the study
   d. None of the above answers is correct.

6. Which of the following is **not** a measure of dispersion?

   a. the variance
   b. the 50th percentile
   c. the standard deviation
   d. the interquartile range

7. The sum of deviations of the individual data elements from their mean is

a. always greater than zero
b. always less than zero
c. sometimes greater than and sometimes less than zero, depending on the data elements
d. always equal to zero

8. If a data set has an even number of observations, the median

a. cannot be determined
b. is the average value of the two middle items
c. must be equal to the mean
d. is the average value of the two middle items when all items are arranged in ascending order

9. The measure of dispersion that is influenced most by extreme values is

a. the variance
b. the standard deviation
c. the range
d. the interquartile range

10. The difference between the largest and the smallest data values is the

a. variance
b. interquartile range
c. range
d. coefficient of variation

11. Which of the following is not a measure of central location?

a. mean
b. median
c. variance
d. mode

12. The value that has half of the observations above it and half the observations below it is called the
a. range
b. median
c. mean
d. mode

13. The most frequently occurring value of a data set is called the

a. range
b. mode
c. mean
d. median

14. If for a distribution difference of first quartile and median is greater than difference of median and third quartile then distribution is classified as
a. absolute open ended

b. positively skewed
c. negatively skewed
d. not skewed at all

15. If beta one is 9, beta two is 11 then coefficient of skewness is
    a. 0.589
    b. 0.689
    c. 0.489
    d. None of the above

16. Moment about mean which is indication whether distribution is symmetrical or asymmetrical is considered as
    a. first moment
    b. third moment
    c. second moment
    d. fourth moment

17. Frequency distribution is considered as negatively skewed if all values of distribution moves to
    a. lower tail
    b. median tail
    c. variance tail
    d. upper tail

18. If for a distribution difference of first quartile and median is less than difference of median and third quartile then distribution is classified as
    a. negatively skewed
    b. not skewed at all
    c. absolute open ended
    d. positively skewed

Exhibit-1

The following data show the number of hours worked by 200 students.

Number of Hours	Frequency
0 - 9	40
10 - 19	50
20 – 29	70
30 - 39	40

19. Refer to Exhibit 1. The class width for this distribution

    a. is 9
    b. is 10
    c. is 11

d. None of the above answers is correct.

20. Refer to Exhibit 1. The number of students working 19 hours or less

   a. is 40
   b. is 50
   c. is 90
   d. can not be determined without the original data

21. Refer to Exhibit 1. The relative frequency of students working 9 hours or less

   a. is 0.2
   b. is 0.45
   c. is 40
   d. can not be determined from the information given

22. The correlation coefficient is used to determine:
   a. A specific value of the y-variable given a specific value of the x-variable
   b. A specific value of the x-variable given a specific value of the y-variable
   c. The strength of the relationship between the x and y variables
   d. None of these

23. Regression modeling is a statistical framework for developing a mathematical equation that describes how
   a. one explanatory and one or more response variables are related
   b. several explanatory and several response variables response are related
   c. one response and one or more explanatory variables are related
   d. All of these are correct

24. In regression analysis, the variable that is being predicted is the
   a. response, or dependent, variable
   b. independent variable
   c. intervening variable
   d. is usually x

25. If the correlation coefficient is a positive value, then the slope of the regression line
   a. must also be positive
   b. can be either negative or positive
   c. can be zero
   d. can not be zero

26. A least squares regression line of the form $y = a + b x$ is fitted to the data set below.

x	25	15	10	5
y	10	10	15	25

The equation of the line is:
   a. $y = -0.69 + 24.4x$
   b. $y = 24.4 - 0.69x$
   c. $y = 24.4 + 0.69x$
   d. None of the above