Anekant Education Society's

Tuljaram Chaturchand College of

Arts, Science and Commerce, Baramati

(Autonomous)

QUESTION BANK

FOR

F.Y.B.Sc. Sem - I

STATISTICS

STAT 1101: DESCRIPTIVE STATISTICS - I

(With effect from June 2019)

UNIT 1: Organization and Presentation of data

A) Questions for 1 mark

I] Choose the correct alternative

1. 'CSO' stands for ……..
   (a) Central Standard Office (b) Control Statistical Office
   (c) Central Statistical Office (d) None of the above
2. Central Government has established CSO for co-ordination of statistical organization with the states in ……
   (a) May, 1950 (b) May, 1951 (c) May, 1952 (d) none of the above
3. National Sample Survey (NSS) was set up in 1950 under the guidance of ……
   (a) C. R. Rao (b) P. V. Sukhatme (c) P. C. Mahalanobis (d) V. S. Huzurbazar
4. NSSO stands for ……..
   (a) National Service Scheme Office (b) National Scheme Service Office
   (c) National Statistical Service Office (d) National Sample Servey Office
5. The name of statistical institute which helps in preparing designs, programming, processing of information analysis, report writing and it is also centre of research and training etc. is…..
   (a) National Statistical Institute (b) National Sample Institute
   (c) Indian Statistical Institute (d) Indian Statistical Service
6. The name of Bureau exist in state government functioning for various activities such as (i) statistical co-ordination, (ii) state income, (iii) socio economic survey etc. is ……
   (a) Bureau of English and Statistics (b) Bureau of Economics and Statistics
   (c) Bureau of Economics & Sociology (d) None of the above
7. Statistical methods are treated as a branch of science which deals with …..
   (a) collection and presentation of data (b) analysis of data
   (c) interpretation of data (d) all the above
8. Statistics performs functions such as ……..
   (a) presentation of facts and figures (b) forecasting and planning
   (c) controlling and exploring (d) all the above
9. Statistics helps in almost all fields such as……..
   (a) decision-making, LIC, banks (b) government agencies and industries
   (c) business, trade, index numbers (d) all the above

B) Questions for 2 marks

1. Define statistics

C) Questions for 4 marks

1. Describe central Statistical organization. CSO.
2. Describe the functions of Bureau of economics and statistics.

D) Questions for 6 marks

1. Describe the importance and limitation of statistics

UNIT 2: Population and Sample

A) Questions for 1 mark

I] Choose the correct alternative

1. Primary data means…..

   (a) original data (b) it may be result of survey

   (c) it may be result of equity (d) all the above

2. Secondary data is…….

   (a) already collected by some other agency (b) a processed data

   (c) a finished data (d) all the above

3. Sampling is ……..

   (a) not always useful (b) not always possible

   (c) has number of advantages over census (d) the census

4. In which one of the following scales, data can be measured by using non-numeric notations or non-comparable symbols ?

   (a) nominal scale (b) ordinal scale (c) interval scale (d) ratio scale

5. The registration of two wheelers and four wheelers by allocating the series A, B, MH, GA or the sex-wise codification of 'males' or 'females' by M and F, are examples of ….

   (a) nominal scale (b) ordinal scale (c) interval scale (d) ratio scale

6. Which scale uses the concept of absolute zero ?

   (a) nominal scale (b) ordinal scale (c) interval scale (d) ratio scale

7. The measurements on height and weight are made on……

   (a) nominal scale (b) ordinal scale (c). interval scale (d) ratio scale

8. Which one of the following is not an example of nominal scale ?

   (a) Classification of individuals using blood groups.

   (b) Classification of students in various divisions of the same standard.

(c) Classification of individuals using sex, caste, nationality.

(d) Classification of students according to grades.

9. Which one of the following is an example of ordinal scale ?

(a) Group of students using weight such as light and heavy.

(b) Group of students using height such as short, medium, tall.

(c) Group of individuals according to income such as poor, middle class, rich.

(d) all of the above.

10. Which one of the following is not an example of ratio scale

(a) Fahrenheit scale of temperature measurement (b) Scale of type height (cm)

(c) Scale of type weight (kg) (d) Scale of type time (cm)

11. Attributes are measured using ……..

(a) normal scale only (b) ordinal scale only (c) both (a) and (b) (d) neither (a) nor (b)

12. Variables are measured using ……..

(a) nominal scale (b) ordinal scale

(c) interval scale and ratio scale (d) both (a) and (b)

13. Discrete variable is ……..

(a) a variable taking all possible values in a certain range

(b) a variable taking all values between 0 and

(c) a variable taking only particular values

(d) all of the above

14. Which one of the following is not an example of continuous variable ?

(a) Number of articles produced by a machine.

(b) Length of screw produced by a machine.

(c) Speed of a vehicle.

(d) Temperature at a certain place.

15. Sample is ……

(a) subset of population (b) part of population

(c) 5% of population (d) at least 50% of population

16. Samples selected by ……. method are non-overlapping.

(a) SRSWOR (b) SRSWR (c) stratified (d) systematic

17. If population is homogenous, then …… is better m, sampling.

(a) SRS (b) stratified (c) systematic (d) two stage

II] State whether the following statements are True or False :

1. The facts and figures obtained by an investigator himself is called as primary data.
2. Population census result is a classical example of secondary data.
3. Primary data are more reliable, but are expensive and time consuming.
4. The collection of data of production amount, in a particular industry is an example of primary data.
5. Primary data are more accurate than secondary data.
6. The data obtained from reports of municipal corporations and zilla parishad is an example of secondary data.
7. Examination result of a candidate can be recorded as pass or fail which is an example of quantitative variable.
8. We can express candidate's performance as percentage of marks which is an example of quantitative variable.
9. In an ordinal scale of measurement, 'distance' is an important factor in the measurements.
10. The zero point on an interval scale signifies absence of the property.
11. The zero point on a ratio scale signifies absence of the property.
12. Sampling method is useful to collect only primary data.
13. Samples drawn by SRSWOR are distinct.

III) Answer the following (1 mark each)

1. Define a variable.
2. What do you mean by an attribute?
3. Give an illustration of Nominal scale where it is used?
4. Define the term ordinal scale
5. Explain the term Nominal scale
6. Define the term Interval scale
7. Explain the term Ratio scale
8. Give Practical application of ordinal scale
9. Give an example of ratio scale where it is used?
10. Define discrete variable
11. Define continuous variable

B) Questions of 2 marks

1. What do you mean by population and statistical population?
2. What is a sample?
3. Draw a simple random sample with replacement of size n = 2, if the population units are 2, 3, 5.
4. Define sampling frame
5. In a population of size N = 6, the observations were 3, 4,7,9,10,11. Draw all possible SRSWOR of size 2.
6. If a population consists of 30 items then how many
   I) SRSWOR each of size 10
   II) SRSWR each of size 10 can be selected
7. Explain with illustration the term
   i) Finite population ii) Infinite population
8. Give illustrations of each of the following sampling methods
   i) SRSWR ii) SRSWOR

C) Questions for 4 marks

1. Explain the method of simple random sampling.
2. Describe the method of stratified random sampling
3. What do you mean by systematic sampling? Explain briefly.
4. State the advantages of sampling over census.
5. How does SRSWR differ from SRSWOR?
6. From the population containing 20 units obtain a systematic sample of size
7. Suppose 4^{th} unit is the first unit selected at random in the systematic sample. Also obtain all possible systematic samples. 6,18,10,12,16,14,15,13,9,13,12,17,16,15,8,12.
8. Explain the situations where sampling has larger scope as compared to census.
9. Distinguish between primary and secondary data with suitable examples

D) Questions for 6 marks

1. Discuss the advantages and disadvantages of sampling over census
2. What are the requirements of a good sample?
3. Explain the terms; population, sample and sampling unit.
4. Distinguish between Attributes and Variables
5. Explain the different methods of collecting primary data

6. Distinguish between primary data and secondary data and discuss the various methods of collecting primary data

UNIT 3: Univariate data analysis

A) Questions for 1 mark

I] Choose the correct alternative

1. Which limits are excluded in case of exclusive type of class intervals ?

   (a) lower limits (b) upper limits (c) either (a) and (b) (d) both (a) and (b)

2. A relative frequency distribution represents frequencies in terms of……

   (a) proportions (b) whole numbers

   (c) percentages (d) both (a) and (c)

3. The distribution

Marks below	10	20	30	40	50
No. of students	2	7	35	46	50

   (a) cumulative frequency distribution. (b) inclusive class type.

   (c) exclusive class type. (d) none of the above.

4. The following frequency distribution

Income more than T	No. of persons
500	100
1000	96
1500	92
2000	59
2500	28
3000	2

   is known as……

   (a) discrete frequency distribution,

   (b) continuous frequency distribution,

   (c) cumulative distribution in less than type,

   (d) cumulative distribution in greater than type

5. The class intervals of the grouped data :

5-9	10-14	15-19	20-24

   are of the type…….

   (a) inclusive class (b) discrete class (c) exclusive class (c) none of the above

6. In order to draw histogram for frequency distribution with classes of unequal width, we have to consider …….

(a) height of rectangles proportional to class intervals

(b) height of rectangles proportional to class frequency

(c) height of rectangles proportional to frequency density

(d) all the above

7. Which of the following statements is true for less than cumulative frequency curve ?

(a) The curve is non-decreasing in nature. (b) The curve is strictly decreasing in nature.

(c) The curve is non-increasing in nature. (d) The curve is strictly increasing.

8. In order to construct histogram for equal class interval, what is taken on the X-axis and Y-axis respectively ?

(a) lower limit and less than cumulative frequency.

(b) upper limit and less than cumulative frequency.

(c) class interval and frequency. (d) mid-points and frequency.

9. In order to construct a frequency polygon, equal class interval along X axis and Y-axis respectively is:

(a) lower limit and less than cumulative frequency

(b) upper limit and less than cumulative frequency.

(c) class interval and frequency.

(d) mid-points and frequency.

10. If the points (or vertices of frequency polygon) are joined by a smooth curves instead of straight lines we get a closed figure called as …….

(a) frequency distribution (b) trend

(c) cumulative distribution curve (d) frequency curve

11. If we plot values of upper boundaries on X-axis and points of less than type cumulative frequency on Y-axis and points are joined by a smooth curve we get figure called as …..

(a) frequency distribution. (b) less than type ogive

(c) greater than type ogive (d) none of the above.

12. Which one of the following is true for histogram ?

(a) It cannot be drawn for a frequency distribution with open end class.

(b) It is used to locate value of mode.

(c) It remains same if class width is changed.

(d) all of the above.

13. The base of histogram is ……

(a) class limits (b) class intervals

(c) class boundries (d) extended class intervals.

14. From Histogram we can prepare ……..

(a) frequency distribution (b) frequency polygon

(c) frequency curve (d) all the above

15. With the help of histogram, one can determine …….

(a) median (d) deciles (c) percentiles (d) mode

16. With the help of ogive curve, one cannot determine ……..

(a) median (b) deciles (c) percentiles (d) mode

17. Which one of the following is not a measure of central tendency ?

(a) Mean deviation (b) Mean (c) Median (d) Mode

18. When calculating the average rate of profit made by a company, the correct mean to use is the

(a) arithmetic mean (b) weighted mean

(c) geometric mean (d) either (a) or (c)

19. What is the major assumption we make while computing a mean from grouped data ?

(a) All values are discrete.

(b) Every value in a class is equal to the mid-point.

(c) No value occurs more than once

(d) Each class contains exactly the same number of values.

20. Which of the following is the first step in calculating the median of a data set ?

(a) Average the middle two values of the data set.

(b) Array the data in order.

(c) Determine the relative weights of the data values in terms of importance.

(d) Not necessary to arrange.

21. Which of the following is NOT correct ?

(a) Extreme values affect the median less strongly than they do affect the mean.

(b) A median can be calculated for qualitative descriptions.

(c) The median can be calculated for every set of data, even for all sets containing open ended classes.

(d) The median is adaptable for further mathematical manipulations.

22. Which one of the following is true for $\bar{x} = a + \frac{\Sigma(x-a)}{n}$ when 'a' is …..

(a) positive or negative (b) less than mean (c) greater than mean (d) any number

23. The formula for geometric mean G is ……

(a) $\frac{1}{n}\sum \log xi$ (b) $\log\left[\frac{1}{n}\sum xi\right]$ (c) antilog $\left[\frac{1}{n}\sum \log xi\right]$ (d) antilog $\left[\frac{1}{n}\log\sum xi\right]$

24. If x_1, x_2, ...x_n is a set of n observations, then harmonic mean of X is the reciprocal of……

(a) the-given observations and their arithmetic mean

(b) the arithmetic mean of the given observations

(c) the arithmetic mean of the reciprocals of the given observations

(d) reciprocal of mean

25. For n-observations the harmonic mean is given by …….

(a) $\frac{\sum 1/x}{n}$ (b) $\frac{n}{\sum 1/x}$ (c) $\frac{1/\sum x}{n}$ (d) $\frac{\sum\ 1/x}{1/n}$

26. With usual notations the formula of median for grouped data is…….

(a) $l_1 + \frac{N/c.f}{f} \times h$ (b) $l_2 + \frac{[N/c.f-1]}{f} \times h$

(c) $l_1 + \frac{[N/2-c.f]}{f} \times h$ (d) none of the above

27. Median for arranged data is……..

(a) mean of first and last value (b) most frequent value

(c) least frequent value (d) middle most value

28. Which one of the following is false ?

(a) Q_2 = Median (b) $Q_2 = 50^{th}$ percentile (c) $Q_2 = 25^{th}$percentile (d) $Q_2 = 5^{th}$decile

29. Which of the following measures of central tendency are not affected by the extreme values ?

(a) Geometric mean and harmonic mean (b) Arithmetic mean and median

(c) Median and mode (d) Arithmetic mean and mode

30. If x_1, x_2, ….x_n are n observations with mean $\bar{x}$, then $\sum(x_i - \bar{x})$ is ……..

(a) necessarily zero (b) necessarily non-negative

(c) may be non-negative (d) may be zero

31. The arithmetic mean of first n natural numbers is …..

(a) n (n+1)/2 (b) $(n^2+1)/2$ (c) $n(n^2 + 1)/2$ (d) $(n^2+n)/2n$

32. The value of 20th percentile is equal to ……

(a) second quartile (b) median

(c) second decile (d) fourth decile

33. The value of lower quartile is equal to ……

(a) P_{25} (b) D_3 (c) median (d) Q_3

34. The value of P_{50} …….

(a) cannot be equal to mode (b) cannot be equal to Q_1

(c) is equal to fifth decile (d) is equal to Q_3

35. Mode is ……

(a) middle most value (b) the minimum value

(c) most frequent value (d) the maximum value

36. Which one of the following is least affected by extreme values ?

(a) Arithmetic mean (b) Geometric mean

(c) Harmonic mean (d) Median

37. Median can be …..

(a) calculated when measurements are on the nominal scale

(b) affected by number and values of the observations

(c) affected by extreme values

(d) calculated for open end classes

38. Which one of the following cannot be found when terminal classes are open end?

(a) Mode (b) Median

(c) Arithmetic mean (d) Any one of the above

39. Which one of the following is true ? When a single extreme observation in the given data is missing, one can use …..

(a) mode and arithmetic mean (b) median and arithmetic mean

(c) median (d) any measure

40. Which measure of central tendency can be located by mere inspection for series of individual observations ?

(a) mode and arithmetic mean (b) median and arithmetic mean

(c) median and mode (d) mode only

41. Mode for data with frequencies vary haphazardly ……

(a) can be located by a formula

(b) can be found by grouping class interval

(c) can be found by graphical method

(d) cannot be found

42. Mean > Median > Mode valid for ….

(a) symmetrical distribution

(b) moderate negatively skewed distribution

(c) moderate positively skewed distribution

(d) both (a) and (b)

43. Given mean ($\bar{x}$) and median (Me), then the Mode (Mo) can befound by formula ($\bar{x}$ - Mo) = 3 ($\bar{x}$ - Me) …..

(a) in symmetrical distribution (b) in widely skewed distribution

(c) in moderately skewed distribution (d) in all cases

44. Which one of the relation is true?

(a) Mean = $\frac{1}{2}$ (Median – Mode)

(b) Mean – 3 (Mean – Median) = 2 Mode

(c) Median = Mode + 2/3 (Mode -$\bar{x}$)

(d) Mode = 2 Me – 3 $\bar{x}$

45. Usually one can find arithimetic mean, median and mode, when data are collected respectively on the scale …..

(a) nominal, ratio or interval (b) interval, ordinal or nominal

(c) ratio, nominal or interval (d) nominal, ratio or ordinal

46. Which one of the following is true for first n natural numbers ?

(a) $\frac{1}{n}\sum\frac{1}{x} < \frac{n}{\sum x}$ (b) $\frac{1}{n}\sum\frac{1}{x} > \frac{n}{\sum x}$ (c) $\frac{1}{n}\sum\frac{1}{x} = \frac{n}{\sum n}$ (d) $\frac{1}{n}\sum\frac{1}{x} < \frac{n}{\sum x}$

47. If $\overline{X}$ = simple arithmetic mean and $\overline{X}_w$= weighted mean of n observations, then which one of the following statement about the weighted arithmetic mean is false ?

(a) $\overline{X} = \overline{X}_w$ if weights are the same

(b) $\overline{X}_w$ is used to calculate standardised birth rate

(c) $\bar{X}_w$ is used to compute index numbers

(d) $\overline{X}_s < \overline{X}_w$ if smaller weights are assigned to the

48. Which one of the following statement is false ?

(a) Average changes if scale is changed.

(b) Average changes if origin is changed.

(c) In a negatively skewed distribution, the value of mode is greater than the arithmetic mean.

(d) Geometric mean is the square root of the product of all the observations.

49. Which one of the following statement is false for harmonic mean ?

(a) It is quite useful for the analysis of data pertaining to rates etc.

(b) It is not computed when even a single observation in a given data is zero.

(c) In a negatively skewed distribution, the value of mode is greater than the arithmetic mean.

(d) It gives small items more importance.

50. If 10 is subtracted from each observation, then mean of the set of n observations is …..

(a) not affected (b) zero (c) increased by 10 (d) decreased by 10

51. If we add 15 in each observation of a set, then arithmetic mean is ……..

(a) 15 times the original data (b) not affected

(c) increased by 15 (d) decreased by 15

52. The geometric mean of three number 7, 21 and 63 is equal to ……

(a) 30.3 (b) $\sqrt{91}$ (c) $\sqrt{9251}$ (d) 21

53. If x_1 and x_2 are two observations, then geometric mean to calculated only if ….

(a) x_1 and x_2 are equal to zero (b) x_1 and x_2 are negative

(c) x_1 and x_2 are positive (d) x_1 is negative and x_2 is positive

54. If each of the observation in a series is divided by 5, then the mean of new data set is…

(a) unchanged. (b) half of the old mean.

(c) decreased by 5 (d) the old mean is five times the new mean.

55. If each value in a data set of observations is multiplied by 20, the median of the new series will be …..

(a) increased by 20 (b) $\left(\frac{1}{20}\right)^{th}$ of the old median

(c) twenty times the old mean (d) unchanged

56. Which on the following measure is the crudest measure of dispersion ?

(a) range (b) quartile deviation

(c) mean deviation (d) standard deviation

57. In the following, which is not a measure of dispersion ?

(a) central deviation from mean (b) mean deviation

(c) quartile deviation (d) standard deviation

58. The wide application of 'range' is in the field of …….

(a) statistical quality control (b) economics

(c) psychometrics (d) none of the above

59. In the case when first and last class intervals are open end type, one can find…..

(a) range (b) mean deviation

(c) standard deviation (d) quartile deviation

60. The mean deviation can be obtained about ……..

(a) arithmetic mean (b) mode

(c) median (d) all the above

61. Let $x_1, x_2, \ldots\ldots, x_n$ a set of nobservations, then the formula of standard deviation is …..

(a) $\sqrt{\frac{\sum(xi-\bar{x})^2}{n}}$ (b) $\frac{\sum(xi-\bar{x}^2)}{N}$ (c) $\frac{\sqrt{\sum(xi-\bar{x})^2}}{n}$ (d) $\frac{\sum(xi-\bar{x}^2)}{n}$

62. Let $x_1, x_2 \ldots., x_n$ is a set of n observations, the formula of mean deviation from mean is …..

(a) $\sqrt{\frac{\sum|xi-\bar{x}|^2}{n}}$ (b) $\frac{\sum|xi-\bar{x}\ |}{n}$ (c) $\frac{\sum|xi-\bar{x}|^2}{n}$ (d) $\frac{\sum|xi-\bar{x}\ |}{\sqrt{n}}$

63. Let $x_1, x_2 \ldots., x_n$ is a set of n observations, L = largest and S = Smallest value in the data. Then the formula or range is ……

(a) R = S - L (b) $R = x_n - x_1$ (c) R = L - S (d) R = |S-L|

64. Mean deviation is the least when measured from ……..

(a) mean (b) median (c) mode (d) zero

65. The formula of quartile deviation or semi inter-quartile range is ...

(a) $Q_3 - Q_1$ (b) $(Q_3 - Q_1)/(Q_3 + Q_1)$

(c) $(Q_3 - Q_1)/2$ (d) $(Q_1 - Q_3)/2$

66. With usual notations, formula for coefficient of range of the set of observations $x_1, x_2, \ldots. x_n$ is ……..

(a) (L – S)/(L + S) (b) (L – S)/L

(c) (L + S) / (L – S) (d) L/S

67. Formula for coefficient of quartile deviation is equal to ……

(a) $(Q_3 - Q_1)/(Q_1 - Q_3)$ (b) $(Q_3 + Q_1)/(Q_1 - Q_3)$

(c) $(Q_3 + Q_1)/(Q_3 - Q_1)$ (d) $(Q_3 - Q_1)/(Q_3 + Q_1)$

68. In standard deviation, the deviations of the items xi are alwaystaken from ……

(a) mean (b) median (c) mode (d) zero

69. If the coefficient of range of series A is less than that of the series B, then standard deviation of A is …….

(a) series A is more consistent (b) series B is more consistent

(c) both series are equally consistent (d) none of the above

70. The concept of standard deviation was introduced by…….

(a) Karl Pearson (b) R. A. Fisher (c) Gauss (d) W. S. Gossett

71. Which one of the following relation between standard deviation (S.D.) and variance (Var (X)) of a variable X is true ?

(a) S.D. = Var (X) (b) $(S.D.)^2$ = Var (X)

(c) S.D. = $[\text{Var}(X)]^2$ (d) None of the above

72. Mean square deviation is minimum when deviations are taken from ……

(a) mean (b) mode (c) median (d) zero

73. Let $\{(x_i, f_i), i = 1, 2, \ldots\ldots n\}$ be a discrete frequency distributionhaving an average A and $\sum f_i = N$, then the mean deviation is …….

(a) $\frac{1}{N}\sum f_i(x_i - A)^2$ (b) $\frac{1}{N}\sum f_i(x_i - A)$

(c) $\frac{1}{N}\sum f_i|x_i - A|$ (d) $\frac{1}{N}\sum |f_i x_i - A|^2$

74. The value of standard deviation of a set of values will be…..

(a) zero if all observations are equal (b) always positive

(c) positive although the values are negative (d) all the above

75. If mean deviation = M.D. and standard deviation = S.D. of n valuesx_1, x_2 …., x_n then ….

(a) M. D. = S. D. (b) M.D. < S.D. (c) M.D. < S.D. (d) any one of the above

76. The quartile deviation (Q.D.), mean deviation (M.D.) and the standard deviation (S.D.) respectively are normally related by the expression …..

(a) M. D. ≤ Q.D. ≤ S. D. (b) S.D. ≤ Q.D. ≤ M.D.

(c) M.D. ≤ S.D. ≤ Q.D. (d) Q.D. ≤ M.D. ≤ S.D.

77. Which one of the following measures of dispersion ignores the signs of deviations from a central value ?

(a) quartile deviation (b) standard deviation

(c) range (d) mean deviation

78. The formula for quartile deviation or semi inter-quartile range is given by …..

(a) Q.D. = $(Q_3 - Q_1)$ /4 (b) Q.D. = $(Q_3 - Q_1)$ /2

(c) Q.D. = $(Q_3 - Q_1)$ x 2 (d) Q.D. =$(Q_3 + Q_1)$ /2

79. If a constant value 10 is subtracted from each observation of a set, the variance is ……

(a) increased by 100 (b) decreased by 10 (c) decreased by 100 (d) not changed

80. If each observation of a set is divided by 15, the S.D. of the coded set is ……..

(a) 15 times of S.D. of original set of observations

(b) unaltered

(c) $\frac{1}{225}^{\text{th}}$of S.D. of original set of observations

(d) $\frac{1}{15}^{\text{th}}$ of S.D. of original set of observations

[B] State Whether the following statements are True or False :

1. The classification of data invariably depends on the scale of measurement.
2. The frequency distribution has the advantage of representing data in compressed form.
3. Discrete data can be expressed only in whole numbers.
4. A univariate frequency distribution organizes data into groups of values describing one or more characteristics of the data.
5. The class widths of frequency distribution are of equal size.
6. A more than ogive is S-shaped.
7. The vertical scale of an ogive for a relative frequency distribution marks the fraction of the total number of observations that corresponds to respective class.
8. A histogram is a series of rectangles, with width proportional to class frequencies.
9. A 'less than' ogive is S-shaped in general.
10. Histogram cannot be drawn for a frequency distribution with open end class.
11. While drawing frequency curve it is not necessary to take care that the area under the curve is same as that of frequency polygon.
12. After plotting mid-points on X-axis and frequencies on Y-axis for frequency polygon, it is not necessary to join the mid values of preceding class and succeeding class with zero frequency.
13. In case the data items are three-digit, stem will be two-digit and leaf will be single digit at unit place.
14. If the observations are of the type 102.3, 102.8, 103.5 then integers such as 102, 103 can be taken as stem and fractional part as leaf.
15. Stem and leaf is another way of representing data, gives bar like structure infront of stem and no loss of data due to condensation.
16. Measures of central tendency in a data set refer to the value around which the observations are scattered.
17. If we arrange the observations in a data set from highest to lowest, the data point lying in the middle is the median of the data set.
18. With ungrouped data, the mode is most frequently used as the measure of central tendency.
19. The value most often repeated in a data set is called the arithmetic mean.
20. When working with grouped data, we may compute an approximate mean by assuming that each value in a given class is equal to its mid-point.
21. Extreme values in a data set have a strong effect on the median.

22. The arithmetic mean of a variable with values which are all equal, is same as their common value.
23. The sum of the deviations of given values of a variable from its measure of central tendency is necessarily zero.
24. Unlike geometric mean, harmonic mean cannot be computed when one or more observations in data are zero.
25. If a non-negative variable changes over time exponentially, then the appropriate average is harmonic mean.
26. Mode is the only measure of central tendency that can be obtained without calculation.
27. If we arrange 50 observations in ascending order, than the median will be the value of the 25th observation in the series
28. We can compute a mean for any data set.
29. The geometric mean of the ratio of x and y is the ratio of their geometric means.
30. The value of every observation in the data set is taken into accountwhen we calculate its median.
31. A mean calculated from grouped data alwaysgiven a good estimate of the true value, although it is seldom exact.
32. The geometric mean of a group of numbers is less than the arithmetic mean in all cases, except in the special case in which numbers are all the same.
33. Mean, median and mode never be same.
34. Simple arithmetic mean is less than the weighted arithmetic mean, if greater weights are attached to the smaller values and smaller weights are attached to the greater values.
35. In order to compute median if class intervals are unequal, the frequencies are ought to be adjusted to make the class intervals equal before proceeding with standard formulae.
36. The harmonic mean of n observations is the reciprocal of the arithmetic mean of the reciprocals of the numbers.
37. Harmonic mean becomes zero when one of the item is zero.
38. The arithmetic mean is an average which is such that if each value of the variable is replaced by it, the total remains the same.
39. Sachin travels by car for 4 days. He drove 10 hours each day. He drove first day at the rate of 45 km per hour, second day at the rate of 40 km per hour, third day at the rate of 38 km per hour, and fourth day at the rate of 37 km per hour. Then average speed of Sachin is 40.km per hour.
40. The mean will be the same whether the values of the variate are expressed in inches or in

cms.

41. Mean cannot be worked out if the class intervals are unequal.
42. Harmonic mean for 2, 7, 0, 12, 15 is positive.
43. Mode cannot be calculated if the distribution does not have frequency.
44. Sachin travels 8 km at 4 km per hour, 6 km at 3 km per hour and 4 km at 2 km per hour, the average rate per hour at which Sachin travelled is 6 km per hour.
45. The $\sum(x_i - \bar{x})^2 \leq (\bar{x} - a)^2$ always holds.
46. Arithmetic mean is dependent on change of origin and change of scale.
47. Sometimes arithmetic mean may be determined graphically.
48. Geometric mean is prefered for the situations like average population growth, average rate of returns on share, average bank interest rates etc.
49. The serious drawback of geometric mean is, it is not equal to zero, if any of the observation is zero.
50. Geometric mean is not applicable to qualitative type data.
51. Geometric mean cannot be computed if frequency distribution includes open end class and also cannot be determined graphically.
52. Since G.M. $\left[\frac{x}{y}\right] = \frac{\text{G.M.(x)}}{\text{G.M.(y)}}$then A.M. $= \frac{\text{A.M.(x)}}{\text{A.M.(y)}}$
53. Harmonic mean cannot be computed if frequency distributionincludes open end class and also cannot be determined graphically.
54. Arithmetic mean, geometric mean and also harmonic mean maynot be an actual observation in the data.
55. G.M. of x and y is G.M. of their A. M. and H.M.
56. Median can be computed for qualitative data.
57. Median is not capable of further mathematical, to find the treatment.
58. If the maximum frequency is repeated, to find the mode uniquely a method of grouping is adopted and a modal class is determined.
59. Mode can be determined if modal class is at the extreme.
60. Modal, premodal and postmodal classes should be of the same width.
61. Mode is applicable forqualitative and quantitative data.
62. Wherever we need to compute mean of ratios, percentage rates, etc., weighted averages are preferred.
63. In the construction of index numbers, death rates, birth rates weighted means are found to be suitable.
64. Average cannot give the idea about the nature of data.

65. If n = 10 and $\frac{\sum(x-5)}{5}$ = 18, then value of mean will be 12.

B) Questions for 2 marks

1. Define the term Class frequency with illustration
2. Define the term open end classes with illustration
3. Define the term class width with illustration
4. Define the term mid value with illustration
5. Define the term class boundaries with illustration
6. Define the term class limits with illustration
7. Define the term less than cumulative frequency with illustration
8. Define the term more than cumulative frequency with illustration
9. What do you mean by central tendency?
10. Define weighted A. M.
11. Define weighted G. M.
12. Define weighted H. M.
13. Show that the algebraic sum of deviations of observations from their A. M. is zero.
14. Find A. M. given that $\sum(x - 10) = 230$ and $n = 50$
15. State the empirical relation between A. M., Median, and Mode.
16. Explain the term trimmed mean.
17. Mention the situations in which the G. M. useful as average
18. Mention the situations in which the H. M. useful as average

C) Questions for 4 marks

1. Explain the different methods of classification
2. Distinguish between – Inclusive classes and Exclusive classes.
3. Explain the general guidelines of preparing a frequency distribution.
4. State requirements of a good measure of central tendency.
5. Discuss merits and demerits of A. M.
6. Discuss merits and demerits of Median.
7. Discuss merits and demerits of Mode.
8. Show that the sum of squares of deviations of all observations taken from A. M. is minimum.
9. Discuss the effect of change of origin and scale on A. M.
10. Find A.M. of first n natural numbers.

11. The mean monthly salary paid to 300 employees of a firm is Rs. 1470. There are 200 male employees & the remaining are females. If mean salary of males is Rs. 1505, obtain the mean salary of females.
12. A.M. of 50 items is 104. While checking, it was noticed that observation 98 was misread as 89. Find the correct mean.
13. If a & b are any two positive observations then prove that the corresponding A.M.≥G.M. ≥H.M.
14. If a & b are any two positive observations then prove that G.M. = (A.M. x H.M.)½.
15. A train travels the first half distance of its journey with a speed of 25 km./hr., the next one-fifth distance of a speed of 60 km./hr., the remaining one fourth distance at a speed of 40 km./hr. Find the average speed of the train during the journey.
16. Mean daily salary of 50 employees in a firm is Rs. 88.40 Frequency distribution of salaries of these employees in which some frequencies are missing, find it.

Salary	40–60	60–80	80 – 100	100 – 120	120 – 140
Frequency	6	---	17	---	5

17. A.M. of the following frequency distribution is 5, find the value of x

Variable	2	4	6	8
Frequency	x – 1	x+1	x+1	2x - 5

18. A variable takes values 7, 10, 13 … 50 terms, find its A.M. & median.
19. A variable takes values 2,3,4,…10. Find the mean deviation about median.
20. A variable takes values 1,2,…n. Find the mean and variance.
21. Two samples of sizes 40 & 50 have the same mean but different standard deviations 19 & 8 respectively. Find the standard deviation of the combined group.
22. If n=100, $\sum X = -20$, $\sum X^2 = 220$. Find S.D. & C.V.
23. Coefficients of variation of the two series are 60% & 80%. Their standard deviations are 20 & 16. What are their arithmetic means?
24. For a group of 30 male workers, the mean & S.D. of weekly overtime work in hours are 10 & 4 respectively, for 20 female workers the mean & S.D. are 5 & 3 respectively.

i) Calculate the mean for the two groups taken together.
ii) Is the overtime work more variable for the male group than for the female group? Explain.

C) Questions for 6 marks

1. Define A.M., Median, Mode and state the formula for each, in case of individual observations and frequency distribution.
2. A set of 10 values has A.M. 20. Find the A.M. if
   i) Each value is doubled & then increased by 2.
   ii) Each value is increased by 5 & then doubled.
3. Define Quartiles, Deciles, and Percentiles. Give the formula each for frequency distribution.
4. Explain relative & absolute measures of dispersion.
5. Distinguish between absolute & relative measures of dispersion.

UNIT 4: Moments, Skewness and Kurtosis

A) Questions for 1 mark

I] Choose the correct alternative

1. Which one of the following is a raw moment?
   a) Moment about a b) Moment about 4
   c) Moment about zero d) Moment about mean
2. The first order moment about origin is equal to ……..
   a) zero b) one c) three d) mean
3. In case of symmetric distribution the odd order central moments are….
   a) zero b) one c) positive d) negative
4. The central moments are invariant to the change of. ……..
   a) origin b) scale c) origin and scale d) scale but not on origin
5. The raw moments are invariant to ……
   a) change of scale only b) change of origin only
   c) change of origin and scale d) none of the above
6. The statement that ‘the variance is equal to the second central moment is. …..
   a) always true b) some times true
   c) never true d) ambiguous
7. The first order central moment about mean is equal to
   a) Zero b) One c) Three d) Mean

8. If the mean and S.D. are 2 & 4 resp. then the second moment about origin is

   a) 24 b) 20 c) 22 d) None of the above

9. With usual notion which one of the following is cored

   a) $\mu_4 = \mu_4' - 4\mu_3'\mu_1' - 6\mu_3'(\mu_1')^2 + (\mu_1^1)^4$ b) $\mu_4 = \mu_4^1 - 4\mu_3^1\mu_1^1 - 6\mu_2^1(\mu_1^1)^2 - 3(\mu_1^1)^2 - 3(\mu_1^1)^4$

   c) $\mu_2^1 = \mu_2 - \mu_1^1$ d) $\mu_3 = \mu_3 - 4\mu_2\mu_1 + \mu^3$

10. Karl Person's coefficient if skewness is

    a) $(\overline{X} - M_e)/\sigma$ b) $(\overline{X} - M_0)/\sigma$ c) $3(\overline{X} - M_0)/\sigma$ d) $3(\overline{X} - M_e)/\sigma$

11. Bowley's coefficient of skewness can have

    a) Values generally lying between +1 or –1 only.

    b) Any value

    c) No negative value

    d) No fractional value

B) Questions for 4 marks

1) In a certain distribution upper quartile exceeds the median by 10 units whereas the median exceeds the lower quartile by 7 units. Compute the coefficient of skewness.
2) For a frequency, Bowley's coefficient of skewness is 0.6. The sum of first and third quartiles is 100 and the median is 38, Find the two quartiles.
3) Find the c.v. of a frequency distribution given that its mean is 120, mode is 123 and Karl Pearson's coefficient of skewness is -0.3.
4) The first two moments of a distribution about the value 4 are 3 and 34. Find the mean and variance.
5) What is the effect of change of origin and scale on moments
6) For a frequency distribution, Bowley's coefficient of skewness is 0.6. The sum of the first and third quartiles is 100 and the median is 38. Find the two quartiles.
7) Define 'kurtosis'. What are the types of kurtosis?
8) Show that $\beta_2 \geq 1$, with usual notations.
9) The first four moments about 4 of a certain distribution are 1.5, 17, -30 and 308. Find kurtosis and interpret.

10) Given that $\beta_2 = 2.6$, $\beta_1 = 0.19$, $\mu_2 = 1.2$ Find μ_3 and μ_4.

11) Explain the term 'skewness' using suitable diagrams.

12) Explain the different measures of skewness.

13) For two distributions A and B following summary statistics are available.

	A	B
Median	20	24
Q_1	13	14
Q_3	30	31

Compare the skewness of two distributions using appropriate measures of skewness.

15) Given that $x = 1$, $\mu_2 = 3$, $\mu_3 = 0$ and $\mu_4 = 27$ find the first four raw moments.

16) Express first four central moments in terms of raw moments.

17) Find the third central moment of the following observations

18) The first four moments of a distribution about the value 4 of the variable are -1.5, 17, -30, and 108. Find the moments about the mean.

C) Questions for 6 marks

1. The first four moments of a distribution about the value '5' are 2, 20, 40, and 200 respectively. Find the first four central moments.
2. A distribution has mean 30, coefficient of variation 20% and coefficient of skewness is 0.3. Find its mode.
3. In a certain frequency distribution the sum of upper and lower quartiles is 45 and the difference between them is 15 . If the median is 20, find the coefficient of skewness
4. Define Karl-Pearson's coefficients γ_1 and γ_2 and discuss their utility in Statistics.
6. Given that $n = 100$, $\Sigma x^2 = 140$ $\Sigma x^3 = -40$, $\Sigma x^4 = 560$. Find β_1, β_2 and comment on the nature of skewness and kurtosis of the distribution.
7. Given that A.M. = 160 , mode = 157, 6 = 50 , find (i) Karl Pearson's coefficient of skewness (ii) median (iii) coefficient of variation.

UNIT 5: Theory of Attributes

A) Questions for 1 mark

I] Choose the correct alternative

1. Attribute is …….

   a) a quantitative characteristics b) a measureable characteristics

   c) a qualitative characteristics d) none of the above

2. Which of the following is not an example of attribute?

   a) literacy b) intelligence

   c) blindness d) number of pages in a book

3. If we divide the observations into more than two groups, the sub-division is called as…….

   a) Autonomous classification b) Dichotomous classification

   c) Manifold classification d) None of the above

4. Classification according to grades obtained in an examination is an example of ….

   a) Dichotomous classification b) Autonomous classification

   c) Manifold classification d) None of the above

5. Which of the following is an example of attribute characteristics?

   a) Height b) Weight c) Wages d) Blood group

6. The number of observations belonging to a particular class is known as…….

   a) Order of class b) Negative class c) Class frequency d) none of these

7. With three attributes the total number of class frequencies of all orders is equal to…..

   a) eight b) nine c) twenty seven d) none of the above

8. Ultimate class frequencies means the frequencies of the classes of……..

   a) zero order b) lowest order c) highest order d) none of the above

B) Questions for 2 marks

1. Define the term dichotomy & manifold classification
2. Comment on the association between the attributes A & B in each of the Following :-

   i) N= 100 (A) =75 (B)=60 (AB) =40 ii) N= 100 (A) =60 (B)=50 (AB) =20
3. Define variable and attribute
4. From the following ultimate class frequencies, compute remaining Frequencies (AB)=13, (Aβ) = 20 , (αB)= 15 , (αβ) =9

5. If N=200, (A) =150, (B) =100, (AB)=80 find (A) &(B).
6. Define positive and negative attribute.
7. Find the total no of class frequencies in case of two attribute.
8. Define the order of class.
9. If N=100 , (A) =550 , (B) =700 , &(AB) =300 find (αB) , (αβ).

C) Questions for 4 marks

1. Show that co-efficient of association lies between -1 to 1
2. Define method of operation N. with example

   ii) Explain i) positive and negative attribute ii) Ultimate class frequency
4. Define completely associated and dissociated of two attribute A& B
5. Define i) fundamental set of class frequency ii) Ultimate class frequency
6. Define association and independence of two attribute
7. If (A) = (B) = 4 , N = 8 obtain the co-efficient of association for each of the following cases i) (AB) =0 ii) (AB) = 1
8. If the attribute A&B are independent then show that i)A and β ii) α and β are independent.
9. If Q is the Yule's coefficient between two attribute interpret the values Q=0, Q=1, Q= -1
10. In a certain interview there were 126 candidates of which 70 were boys, 36 candidates were successful among them 20 were boys obtained the co-efficient of association between and attribute boy.
11. Find the co-efficient of association between education standard and Employment from data and comment on result.

	Employed	Unemployed
Graduate	572	96
Non-Graduate	900	432

12. Out of 600 persons in a locality 150 attacked by cholera. In all were inoculated against cholera of whom only 14 were attacked. Comment on the effectiveness of inoculation in preventing cholera As indicated by co-efficient of association

C) Questions for 6 marks

1. Define raw and central moments of series of individual observation.
2. Explain the following
   a) Positive and negative attribute
   b) Order of the class
   c) Ultimate class frequencies
   d)independence of two attribute
   e) Positive and negative association
   f) Fundamental set of class frequency
4. Define the coefficient of association Q and interpret the case Q = 0, Q = 1
5. Show that coefficient of association Q lies between -1 and 1
6. From the following ultimate class frequencies compute remaining frequency (AB) =13,(A β)=10 ,(αB)=9, (α β)=7.
7. The following are data on literacy and criminal tendency in a town Total population (in thousands)=224

   Total literates (in thousands) =40 Literate criminals (in hundreds) =03

   Illiterate criminals (in hundreds)=43

   Find the association between literacy and criminal tendency and comment.