DESCRIPTIVE STATISTICS I (101)

QUESTIONS

1. C.U. 1997:

(a) What do you mean by Skewness of a frequency distribution? Give two measures of Skewness and find the limits within which they lie.

2. C.U. 1998:

(a) Show that $b_1 = \dfrac{{m_3}^2}{{m_2}^3}$ and $b_2 = \dfrac{m_4}{{m_2}^2}$ can be used as measures of Skewness and Kurtosis of a frequency distribution, where m_2, m_3 and m_4 are the second, third and fourth Central Moments. Also prove that, $b_2 \geq b_1 + 1$ and explain the situation under which equality holds.

(b) Write short notes on: Sheppard's Correction for Moments.

3. C.U. 2002:

(a) Distinguish between:
- **(i)** Primary Data and Secondary Data
- **(ii)** Bar Diagram and Pie Diagram
- **(iii)** Frequency Data and Non-frequency Data
- **(iv)** Discrete Variable and Continuous Variable

(b) Show that the Standard Deviation S of a set of observations $x_1, x_2, \ldots\ldots, x_n$ is given by,

$$S^2 = \frac{1}{2n^2} \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{j=1}^{n} (x_i - x_j)^2$$

Hence or otherwise examine the consequence of adding a constant value to all the observations.

(c) What do you mean by Kurtosis of a frequency distribution? Distinguish between the Leptokurtic, Mesokurtic and Platykurtic distributions. What can you say about the tails of the above three distributions when they are symmetric and have same variance?

(d) Write short notes on: Sheppard's Correction for Moments.

4. C.U. 2003:

(a) What is a Ratio Chart? Discuss its uses and point out its advantages over a simple Line Diagram.

(b) What is a Frequency Curve? Give the broad categories under which frequency distribution may be put, indicating in each case the nature of the Frequency Curve.

(c) In a frequency table, the upper class boundary of each class interval has a constant ration to the lower class boundary. Show that, the Geometric Mean 'G' can be expressed by the formula:

$$\log G = x_0 + \frac{C}{N} \sum_i (i-1)\, f_i$$

where, 'x_0' is the logarithm of the mid-value of the first class interval, 'C' is the logarithm of the ratio between upper and lower class boundaries of a class and 'N' is the total frequency.

(d) Define Skewness and give its two measures, the first one based on the arithmetic mean and median and second one based on the three quartiles. Show that, the first measure lies between -3 and 3 and second measure lies between -1 and 1.

5. C.U. 2004:

(a) Give an account of the different modes of diagrammatic representation of statistical data. **6**

(b) Describe situations where the three measures Mean, Median and Mode of Central Tendency are most appropriate. **6**

(c) Express the r^{th} order Central Moment in terms of the r^{th} and lower order Raw Moments. **4**

6. C.U. 2005:

(a) Distinguish between the following:

(i) Attribute and Variable

(ii) Discrete Variable and Continuous Variable

(iii) Frequency Polygon and Frequency Curve **5**

(b) How do Ordinal Data differ from Nominal Data? Explain through illustrations. **5**

(c) Distinguish between Absolute and Relative measures of Dispersion. What are the Percentiles of a set of data? **5**

(d) Is the Median of the logarithms of a set of positive real numbers equal to the logarithm of the Median? What will be your answer for Arithmetic Mean? Justify your answers. **5**

(e) Define Pearson's co-efficient of Skewness. Show that it lies between -1 and 1. **5**

(f) Describe how one can use 'Box Plots' to compare two frequency distributions. **5**

(g) **(i)** Discuss the concept of Trimmed Mean as a measure of Central Tendency. Why and how is it used? **4**

(ii) Show that, the Standard Deviation of a set of observations, apart from a numerical constant, may be regarded as the root mean square of all possible pairs of differences of the observations. **3**

(iii) Show that, for any set $(x_1, x_2, ..., x_n)$ of observations,

$$\begin{vmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 \\ \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i^4 \end{vmatrix} \geq 0$$

Hence establish the inequality related to Pearson's b_1 and b_2 coefficients. **8**

7. C.U. 2006:

(a) Distinguish between:

(i) Interval Scale and Ratio Scale of measurement

(ii) Cross-sectional Data and Time-series Data **5**

(b) Discuss two merits and demerits of a Box Plot. **5**

(c) Comment on the following:

(i) In a study of the ages and professions of deceased men, it was found that the profession with the lowest average age of death was “Student”. So it appears that student profession is very dangerous.

(ii) A study on fighting in bars in which someone was killed found that in 90% of the cases, the person who started fighting was the one who died. **5**

(d) How can you find the first quartile for an ungrouped data of size ‘n’? Explain with examples for n = 10 and n = 11. **5**

(e) Let $\bar{x}$, M and s be the Mean, Median and Standard Deviation respectively, for a set of data. Show that,

$$s \geq \left| \bar{x} - M \right|$$

What is the implication of the above result? **5**

(f) **(i)** How would you design and administer a Questionnaire? **7**

(ii) Illustrate, with an example, the use of Stem and Leaf Display to summarize data. Discuss its advantages over Histogram. **8**

(g) **(i)** Discuss the method of obtaining Mode for grouped data when the class widths are equal. Indicate two merits and two demerits of Mode. **7**

(ii) Define Interquartile Range. What does it represent? Why is this useful? Where do we use it? **4**

(iii) Why Sheppard’s Corrections for Moments are required? State two conditions that are necessary for these corrections to be valid. **4**

8. **C.U. 2007:**

(a) Briefly discuss the differences between statistics as numerical facts and statistics as a field or discipline of study. **5**

(b) Explain the following statement:

“Blindly using any data that happen to be available can lead to misleading information and bad decisions.” **5**

(c) Distinguish (with examples) between:

(i) Population and Sample

(ii) Nominal Scale and Ordinal Scale **5**

(d) Suppose a train moves ‘n’ equal distances each of ‘b’ kms, say, with speeds V_1, V_2, ... , V_n kms per hour. Also suppose that another train moves for ‘n’ equal time intervals, each of length ‘t’ hours, say, with the above speeds. Obtain the average speed in each case. **5**

(e) Define Skewness and suggest two measures for it. Find the range of any one of the measures. **5**

(f) **(i)** How would you design a Questionnaire? Discuss with an example. **8**

(ii) Describe the situations where you would recommend the use of **(i)** a Multiple Bar Diagram and **(ii)** a Pie Diagram. Give an example for each. **7**

(g) **(i)** Distinguish with examples between a ‘Histogram’ and a ‘Stem and Leaf Display’. Which one do you prefer and why? **8**

(ii) Explain the concept of a Box Plot. Describe how one can use Box Plots to compare two frequency distributions. **7**

9. **C.U. 2008:**

(a) Write a short note on abuses of Statistics with examples. **5**

(b) Discuss the relevance of Statistics to agriculture. **5**

(c) Distinguish (with examples) between:

(i) Primary Data and Secondary Data

(ii) U-shaped and J-shaped frequency distributions **5**

(d) Point out the advantages of using a Ratio Chart over a simple Line Diagram. **5**

(e) Explain the concept of a Controlled Experiment. Discuss its uses. **5**

(f) Suppose that you are a purchasing agent for a large manufacturing firm and that you regularly place orders with two different suppliers. After several months of operation, you find the mean number of days required to fill orders is 10 days for both the suppliers but the dispersions are different. Which supplier would you prefer and why? Give one reason in favour of the supplier you are not prefering. **5**

(g) Define Kurtosis. Give a measure of Kurtosis using Moments. When this measure will be appropriate? **5**

(h) **(i)** Distinguish between Textual and Tabular presentation of data with an example. **7**

(ii) Define the term "frequency distribution". How can you select the number of classes, width of each class and class limits in developing a frequency distribution for data on a continuous variable? **8**

(i) Define p-th Percentile. Describe a method of obtaining p-th Percentile when there are 'n' number of observations. Discuss the objectives of computing Percentiles. **7**

10. **C.U. 2009:**

(a) Distinguish giving suitable examples, between:

(i) Qualitative Data and Quantitative Data

(ii) Discrete Variable and Continuous Variable **5**

(b) What is Asymmetric distribution? Discuss the use of Box Plot in this connection. **2+3**

(c) What is Median? How is it computed from discrete and continuous frequency distributions? **1+4**

(d) What is Coefficient of Variation? Cite a situation where such a measure is appropriate. What is the drawback of such measure? **2+2+1**

(e) Write a note on Sheppard's Corrections on Moments. **5**

(f) What are Range (R) and Standard Deviation (s) as measures of dispersion? What are the defects of the measure R? What are the advantages of the measure 's' over 'R'? Show that, for 'n' observations,

$$\frac{R^2}{2n} \le s^2 \le \frac{R^2}{4}$$

Discuss the boundary cases. **2+2+1+2+5+3**

(g) **(i)** Starting from an appropriate data set, define Frequency, Cumulative Frequency (both types), Relative Frequency and Frequency Density.

(ii) Discuss the use of Histogram for the study of a continuous frequency distribution.

(1+3+1+2)+3

11. **C.U. 2010:**

(a) Explain with suitable examples the distinction between Primary Data and Secondary Data. **5**

(b) What are Ogives (less-than and greater-than types)? Show that Median of variable is the abscissa of the point of intersection of the two Ogives. **2+3**

(c) Define Geometric Mean (g) and Harmonic Mean (h) and show that $g \geq h$. When does equality hold? **2+2+1**

(d) What is Trimmed Mean? Discuss a situation where it is used. Suggest any other measure for the same situation. **2+2+1**

(e) What is Mode? Derive its working formula for a continuous frequency distribution. **2+3**

(f) Define Quartiles of a frequency distribution and discuss their uses in measuring the Skewness of the frequency distribution. **2+3**

(g) **(i)** Define Moments of a frequency distribution and discuss their uses in defining Kurtosis measure of the distribution. What is the defect of this Kurtosis measure?

(iii) If $g_r = \left\{ \frac{1}{n} \sum_{i=1}^{n} |x_i|^r \right\}^{1/r}$, study the behaviour of g_r as a function of 'r' (≥ 0). **(2+2+1)+5**

12. C.U. 2011:

(a) Distinguish between different types of measurement scales giving an example of each. **5**

(b) Distinguish between a Questionnaire and a Schedule giving examples where each of them can be used. **5**

(c) What is a Box Plot? Discuss its uses. **5**

(d) Show that the Mean Deviation about the Median is less than the Standard Deviation. **5**

(e) Let 'x' be a variable taking the values 1,2,….,m with frequencies f_i , $i = 1,2,...,m$ and $\sum_{i=1}^{m} f_i = n$. Let F_i and $F_i^{\;*}$ be respectively the corresponding less than type and greater than type cumulative frequencies, $i = 1,2,...,m$.

(i) Suggest how the data can be displayed.

(ii) Describe how you will plot the cumulative frequencies and determine the Median from them.

(iii) Obtain the mean and variance as a function of $F^{\,*}s$. **3+4+8**

(f) **(i)** Give two measures of Skewness – one based on Moments and one based on Quantiles and discuss their relative merits.

(ii) Find the limit of the Quantile-based measure in (i).

(iii) The Mean of a distribution is greater than its Median. What can be said about the Skewness of the distribution? Justify.

(iv) Prove that $b_2 \geq b_1 + 1$, where the symbols have their usual meaning. **5+2+4+4**

13. C.U. 2012:

(a) Distinguish (with an example in each case) between:

(i) Frequency Data and Non-frequency Data

(ii) a Discrete Variable and a Continuous Variable **5**

(b) Agricultural production in a country comprises paddy, wheat, pulses and other items. Suggest a diagrammatic method for studying how the total agricultural production of the country, as also the above four components, change over a given period, indicating the method of drawing. **5**

(c) Illustrate, with an example, the use of Stem and Leaf Display to summarize data. Point out its advantages over a Histogram. **5**

(d) In a set of 50 observations, the maximum and the minimum values are 85 and 25 respectively. Find the maximum possible value of the Standard Deviation of these observations. Establish the relevant result. **5**

(e) Suppose that the frequency distribution of radii of 'n' spheres is symmetric. Show that the frequency distribution of the volumes of those spheres is positively skew. **5**

(f) Deduce the relation between the Quartile Deviation Q(x) and Q(y) of the variables 'x' and 'y' respectively, where ax+ by = c ; a, b, and c being constants. **5**

(g) A variable assumes 3 equidistant values $x_0 - h$, x_0 and $x_0 + h$ with relative frequencies p, 1 – 2p and p, respectively. Calculate Pearson's b_2 coefficient for the distribution. Find the limiting values of this measure as $p \to 0$ and $p \to 1/2$. Will you recommend b_2 as a measure of unimodality? **5**

(h) **(i)** Comment why the Arithmetic Mean is considered as an inappropriate measure in the following cases:

(a). To find the average of student-teacher ratio of 'n' different schools.

(b). To find the average speed of a sprinter who runs the first 100 m at a speed of v_1 m/s and the second 100 m at a speed of v_2 m/s.

(ii) What are the problems of comparing two or more data sets in respect of Dispersion? Mention two measures that take care of these problems, indicating how.

(iii) Indicate how two or more data sets of similar nature can be compared in respect of Central Tendency, Dispersion and Skewness using Box-Plots. How can a Box-Plot be modified to detect possible outliers in a given data set? **4+4+7**

14. **C.U. 2013:**

(a) Distinguish (with one example in each case) between:

(i) Cross-sectional Data and Time-series Data

(ii) Interval Scale and Ratio Scale **5**

(b) Indicate how Central Tendency, Dispersion and Skewness of a given data set can be examined with the help of a single diagram. Can you use the same diagram to detect possible Outliers in the data set? **5**

(c) An assessee depreciated the machinery of his factory by 10% each in the first two years and by 40% in the third year and thereby claimed 21% average depreciation relief from income tax department, whereas the income tax officer objected and allowed only 20%. Which assessment would you like to support and why? **5**

(d) The mean and the standard deviation of life-time of bulbs produced by a factory are 'm' hours and 's' hours, respectively. Provide a lower bound for the percentage of bulbs having life-time between m – 2s and m + 2s hours. Clearly state the relevant result. **5**

(e) Prove, by a geometrical argument, that for a J-shaped frequency distribution with its longer tails towards the higher values of the variable, the median is nearer to the first quartile than to the third. What is the significance of this result? **5**

(f) Show that the Mean Deviation about Mean cannot exceed the Standard Deviation. When are the two equal? **5**

(g) **(i)** Given the estimates of percentage of national income going to five different income groups (bottom 20%, next 20%, middle 20%, next 20% and top 20%) in a country in the years 2000, 2005 and 2010, suggest a suitable diagrammatic method to represent the data and provide a rough sketch.

(ii) Which measure of the average would you consider to be the most suitable in each of the following cases:

(a). size of ready-made shoes

(b). percentage growth rates of an economy over consecutive years

Justify your answers.

(iii) Show that the Mean Deviation about A (MD_A) based on 'n' values may be obtained by the formula:

$$n.MD_A = (S_2 - S_1) + A(n_1 - n_2)$$

where 'S_1' is the sum of the values that are less than 'A' and 'n_1' is the number of such values, while 'S_2' is the sum of the values that are greater than 'A' and 'n_2' is the number of such values. Hence, show that MD_A is minimum when 'A' is the Median. **4+4+7**

(h) **(i)** Explain two different real-life cases where you would like to recommend Coefficient of Variation rather than Standard Deviation as the appropriate measure of Dispersion.

(ii) Show that the Standard Deviation 's' of a set of 'n' values $x_1, x_2, \ldots\ldots, x_n$ is given by,

$$n.s^2 = \sum_{i=2}^{n} \frac{i}{i-1}(x_i - \overline{x_i})^2$$

where, $\overline{x_i} = \dfrac{\sum_{j=1}^{n} x_j}{i}$ $for\ i = 2(1)n$

(iii) Prove that $b_2 \geq 1$ and $b_2 - b_1 - 1 \geq 0$, where b_1 and b_2 are the usual Pearson's measures of Skewness and Kurtosis, respectively. Also, examine the cases where $b_2 = 1$ and $b_2 - b_1 - 1 = 0$. **3+4+8**

15. **<u>C.U. 2014</u>:**

(a) How do 'Ordinal' Data differ from 'Nominal' Data? Explain with one example of each kind. **5**

(b) Describe the situations where you would recommend the use of **(i)** a Multiple Bar Diagram and **(ii)** a Pie Diagram. Illustrate by suitable examples. **5**

(c) What do you mean by a Symmetric frequency distribution? Show that for a frequency distribution symmetric about a value 'm', the Arithmetic Mean equals 'm'. **5**

(d) The Arithmetic Mean and the Median of the radii of 50 circular discs are 3.1 cm. and 3.2 cm. respectively. How much can be said about the Arithmetic Mean and the Median of the surface area of these discs? Justify your answers. **5**

(e) If, for a collection of spheres, the frequency distribution of radius is Symmetric, show that the frequency distribution of volume is Positively Skew. **5**

(f) If a 3-point frequency distribution has the values $x_0 - h$, x_0 and $x_0 + h$ with respective relative frequencies $(1 - p) / 2$, p and $(1 - p) / 2$, calculate Pearson's b_2 coefficient and find the limiting values of b_2 as $p \to 0$ and $p \to 1$. Comment on your findings. **5**

(g) **(i)** Indicate how Central Tendency, Dispersion and Skewness of a set of data on a continuous variable can be examined with the help of its cumulative frequency diagram.

(ii) Is the Median of the logarithm of a set of positive real numbers equal to the logarithm of the Median? What will be your answer for the Arithmetic Mean? Justify your answers.

6+4

(h) **(i)** In a school, a class is divided into two sections of sizes 'n_1' and 'n_2' students, respectively, with different Mathematics teachers for the two sections. Show that the Variance (S^2) of the Mathematics marks (x) for the whole class can be expressed as:

$$(n_1 + n_2)S^2 = (n_1 s_1^{\ 2} + n_2 s_2^{\ 2}) + [n_1 (\overline{x_1} - \overline{x})^2 + n_2 (\overline{x_2} - \overline{x})^2] \ = \ S_A^{\ 2} + S_B^{\ 2} \text{, say,}$$

where, '$\overline{x}$' denotes the mean marks for the whole class and the other symbols have their usual meaning.

How will you interpret this result in the above context? Also explain the significance of the two cases: **(a).** $S_A^{\ 2} = 0$ **(b).** $S_B^{\ 2} = 0$

(ii) You are to compare the dispersions of two sets of 50 recordings on minimum daily temperature recorded in each of Kolkata and Darjeeling during 50 consecutive winter days. Suggest a suitable measure you would like to use and justify. **6+3**

16. C.U. 2015:

(a) Distinguish (with an example in each case) between:

(i) Qualitative Data and Quantitative Data

(ii) Frequency Data and Non-frequency Data **5**

(b) What do you mean by 'Cross-validation' of data? Illustrate your answer with two suitable examples. **5**

(c) Mineral production in a region comprises coal, iron and manganese. Suggest a diagrammatic method for studying how the total mineral production of the region, as also the above three components, change over a given period, indicating the method of drawing. **5**

(d) In a set of 50 observations, the maximum and the minimum values are 70 and 20, respectively. Find the minimum possible value of the Standard Deviation of these observations. Establish the relevant result. **5**

(e) If two variables 'x' and 'y' follow a linear relation y = 3x + 1, examine the relation between their Quartile Deviations. **5**

(f) **(i)** Indicate how two or more data sets of the same kind can be compared in respect of their Central Tendency, Dispersion and Skewness using Box-plots.

(ii) How can Box-plots be modified to detect 'Outliers' in a data set?

(iii) If 'b_1' and 'b_2' are Pearson's Co-efficients of Skewness and Kurtosis, respectively, show that, $b_2 \geq 1$ and $b_2 - b_1 - 1 \geq 0$. Examine the cases where $b_2 = 1$ and $b_2 - b_1 - 1 = 0$, stating the significance of your findings. **6+3+6**

(g) Let $\overline{x}$ and 's' be the Mean and Standard Deviation of the data set $x_1, x_2, \ldots\ldots, x_n$. If N(k) be the number of x_i's such that $\left| x_i - \overline{x} \right| \geq k.s$ for some k > 0, show that, $\dfrac{N(k)}{n} \leq \dfrac{1}{k^2}$.

Explain the significance of this result for k = 3. **6**

17. C.U. 2016:

(a) Distinguish between **(i)** Interval and Ratio and **(ii)** Ordinal and Nominal scales of measurement with appropriate examples. **5**

(b) One intends to draw suitable diagrams corresponding to the data of yield of paddy per hectare in 5 countries for the last 20 years. Although joining of consecutive points by straight line is valid in one case, may be invalid in the dual case – Justify the statement by stating the underlying assumption of joining the points. If drawing of Line Diagrams is invalid in one case, suggest the most suitable diagram to represent the data and justify its use. **5**

(c) Positive integers from 10 to 54, both inclusive, are placed in 5 groups of 9 each. Find the highest possible Arithmetic Mean of the Medians of these 5 groups. **5**

(d) Suppose the observations 'x_i' are bounded above and below by two numbers 'b' and 'a' respectively, i = 1, 2, ….., n. Find the attainable upper bound of the Variance of the observations in terms of 'a' and 'b'. **5**

(e) Obtain the relationship between Variance of the first 'n' odd positive integers and the first 'n' even positive integers. Write the general result corresponding to Variance as obtained through the above example and prove it. **5**

(f) For two linearly related variables 'x' and 'y', given by the equation y = a + bx, a, b > 0; let 'Q_{xj}' and 'Q_{yj}', j = 1,2,3 denote the respective Quartiles. Find the relationship between the measures of Skewness of the variables based on the Quartiles and justify the relationship. **5**

(g) Define Mean Deviation around a number 'A' for a set of 'n' observations. When it will attain the minimum value? Prove the corresponding result. **5**

(h) **(i)** Box Plot is commonly known as 5-number summary of the data – Critically examine the statement and explain how these numbers are computed.

(ii) Show that Mean Deviation about Mean cannot exceed the Standard Deviation. Explain the case of equality.

(iii) State and prove Liaponouv's Inequality. **5+4+6**

18. **<u>C.U. 2017</u>:**

(a) Distinguish between Cross Sectional and Time Series data with appropriate examples. **5**

(b) Explain, with an example, how a Histogram is constructed and how a Frequency Polygon may be obtained from it, justifying the choices of abscissa and ordinate. Indicate the shape of the Polygon, if total number of classes is extremely large. **5**

(c) Indicate, with an example, how a Stem and Leaf display is constructed with special reference to choice of the number of stems and reading of individual data from the display. What is the major criticism of such display? **5**

(d) Consider a set S = {2, 4, 6, 8, x, y} with distinct elements. If 'x' and 'y' are both prime numbers and 0 < x < 40 and 0 < y < 40, then examine, with proper justification, the validity of the following sentences:

(i) The maximum possible Range of the set is greater than 33.

(ii) The Median can never be an even number.

(iii) If y = 37, the Average of the set will be greater than the Median. **5**

(e) Find the Mean and Variance of a positive integer valued variable in terms of Cumulative Frequencies of appropriate types. **5**

(f) Three positive integers 'a', 'b' and 'c' are such that their Average is 20 and a ≤ b ≤ c. If the Median is (a + 11), what is the least possible value of 'c'? – Justify your answer. **5**

(g) Find the Standard Deviation of a set of observations $x_1, x_2, \ldots\ldots, x_n$ in terms of the deviation of 'x_i' from its cumulative mean $\overline{x_i} = \frac{1}{i}\sum_{j=1}^{i} x_j \,;\, i = 2,3,\ldots\ldots,n$. **5**

(h) **(i)** The daily expenditures of 120 families are given below:

Expenditure in (Rs.)	100 – 120	120 – 140	140 – 160	160 – 180	180 – 200
No. of families	12	x	38	y	14

If the Median is 150, find 'x' and 'y'.

(ii) State and prove the AM – GM inequality.

(iii) Two groups with 'n_1' and 'n_2' observations have AMs '$\overline{x_1}$' and '$\overline{x_2}$', SDs 's_1' and 's_2'. Derive the formula for combined Standard Deviation based on these quantities and discuss each of the following special cases **(a)** $\overline{x_1} = \overline{x_2}$, **(b)** $n_1 = n_2$, **(c)** $\overline{x_1} = \overline{x_2}$, $n_1 = n_2$, **(d)** $\overline{x_1} = \overline{x_2}$, $n_1 = n_2$, $s_1 = s_2$. **5+4+6**

(i) **(i)** If a Positively Skewed distribution has a Median of 50, justify the validity of the following statements by drawing relevant figure: **(a)** Mean is greater than 50, **(b)** Mean is less than 50, **(c)** Mode is less than 50, **(d)** Mode is greater than 50.

(ii) Define 'b_1' and 'b_2' coefficients and indicate their role for measuring Skewness and Kurtosis. Show that **(a)** $b_2 \geq 1$ and **(b)** $b_2 - b_1 - 1 \geq 0$. Critically examine the cases when the equalities hold.

(iii) Define Bowley's measure of Skewness for a variable 'x' given by Sk_x. Let $y = \alpha + \beta x$; $\alpha, \beta > 0$. Find the value of Sk_y in terms of Sk_x. **4+7+4**