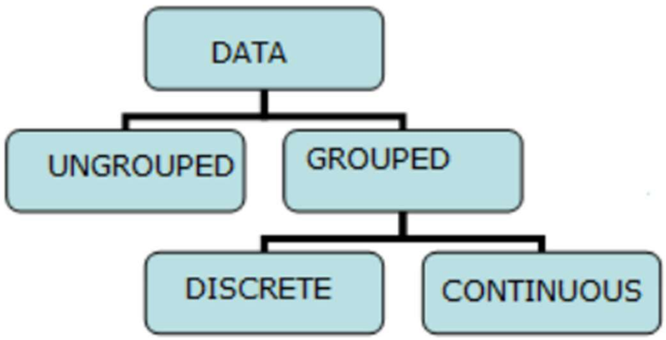
MATHEMATICS

Chapter 13: STATISTICS

STATISTICS

Key Concepts

1. Statistics deals with the collection presentation, analysis and interpretation of data.
2. Data can be either ungrouped or grouped. Further, grouped data can be categorized into
   a. Discrete frequency distribution.
   b. Continuous frequency distribution.

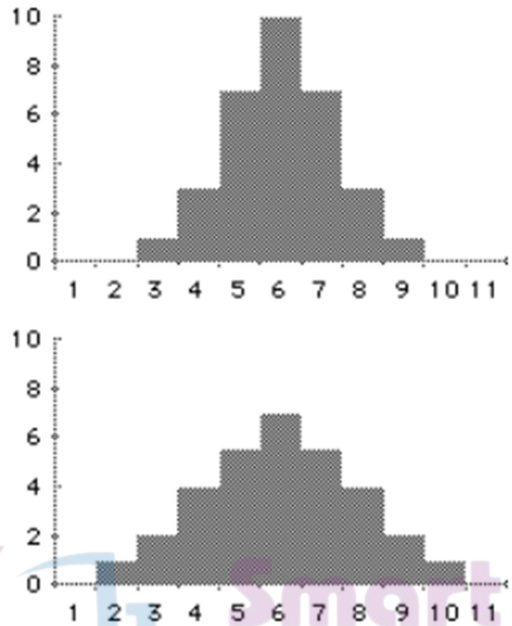


3. Data can be represented in the form of tables or in the form of graphs. Common graphical forms areBar charts, pie diagrams, histograms, frequency polygons, ogives etc.
4. First order of comparison for the given data is the measure of central tendencies. Commonly usedmeasures are (i) Arithmetic mean (ii) Median (iii) Mode.
5. Arithmetic mean or simply mean is the sum of all observations divided by the number of observations.It cannot be determined graphically. Arithmetic mean is not a suitable measure in case of extreme values in the data.
6. Median is the measure which divides the data into two equal parts. The median is the middle termwhen the data is sorted.

   In case of odd observations, the middle observation is median. In case of even observations, themedian is the average of the two middle observations.
7. Median can be determined graphically. It does not consider all the observations.
8. The mode is the most frequently occurring observation. For a frequency distribution, mode may ormay not be uniquely defined.
9. Measures of central tendencies namely mean, median and mode provide us with a single value, which is the representative of the entire data. These three measures try to condense the entire datainto a single central value.
10. Central tendencies indicate the general magnitude of the data.

11.Two frequency distributions may have the same central value but still they have a different spread orthey vary in their variation from the central position. So it is important to study how the other observations are scattered around this central position.

12.Two distributions with the same mean can have different spread as shown below.



13.Variability or dispersion captures the spread of data. Dispersion helps us in differentiating the datawhen the measures of central tendency are the same.

14.Like ‘measures of central tendency’ gives a single value to describe the magnitude of data. The **measures of dispersion** give us a single value to describe variability.

15.The dispersion or scatter of a dataset can be measured from two perspectives:

i.Taking the order of the observations into consideration, the two measures are

a. Range
b. Quartile deviation

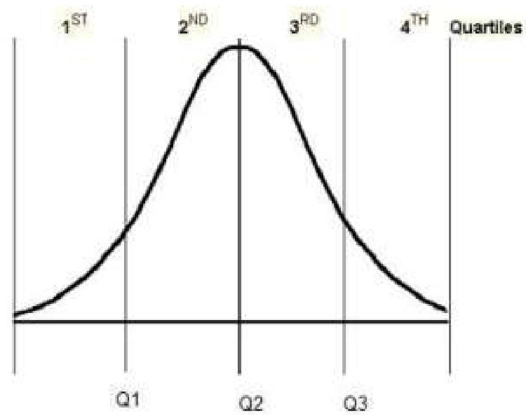
ii.Taking the distance of each observation from the central position yields two measures:

a. Mean deviation
b. Variance and Standard deviation

16.**Range** is the difference between the highest and lowest observation in the given data. Greater the range for data, far more scattered are its observations compared to data having asmaller range.

17.The range at best gives a rough idea of the variability or scatter

18. There are three quartiles namely Q_1, Q_2 and Q_3, which divide the data into four equal parts. Here, $Q_2$is the median of the data.



19. The quartile deviation is half of the difference between the upper quartile and lower quartile.

20. If x_1, x_2, ... x_n are the set of points and point a is the mean of the data. Then the quantity x_i – a is called the deviation of xi from mean a. Then the sum of the deviations from mean is always zero.

21. In order to capture the average variation, we must get rid of the negative signs of deviations.There are two remedies

    Remedy I: Take the absolute values of the deviations.

    Remedy II: Take the squares of the deviation.

22. Mean of the absolute deviations about a gives the 'mean deviation about a', where a is the mean. Itis denoted as MD (a). Therefore,

    MD(a) = Sum of absolute values of deviations from the mean 'a' divided by the number of observations. Mean deviation can be calculated about median or mode or any other observations.

23. Merits of mean deviation

    (1) It utilizes all the observations of the set.

    (2) It is least affected by the extreme values.

    (3) It is simple to calculate and understand.

24. Mean deviation is the least when calculated about the median.

    If the variations between the values is very high, then the median will not be an appropriate centraltendency representative.

25.Limitations of Mean Deviation

i. The foremost weakness of mean deviation is that in its calculations, negative differences areconsidered positive without any sound reasoning.

ii. It is not applicable to algebraic operations.

iii. It cannot be calculated in the case of open end(s) classes in the frequency distribution.

26.Measure of variation based on taking the squares of the deviation is called the variance.

27.Let the observations be $x_1, x_2, x_3,..,x_n$

Let mean = $\bar{x}$

Squares of deviations: $d\,(x_i - \bar{x})^2$

Case 1: The sum d_i is zero. This will imply that all observations are equal to the mean x bar.

Case 2: The sum d_i is relatively small. This will imply that there is a lower degree of dispersion.

Case 3: The sum d_i is large. There seems to be a high degree of dispersion.

28.Variance is given by the mean of squared deviations. If the variance is small, then the data pointsare clustering around mean otherwise they are spread across.

29.Standard deviation is simply expressed as the positive square root of variance of the given data set.Standard deviation of a set of observations does not change if a non-zero constant is added or subtracted from each observation.

30.Variance considers the square of the deviations.
Hence, the unit of variance is in square units of observations.
For standard deviation, its units are the same as that of the observations. That is the reason whystandard deviation is preferred over variance.

31.Standard deviation can help us compare two sets of observations by describing the variation from the 'average', which is the mean. It is widely used in comparing the performance of the two data setssuch as two cricket matches or two stocks.

In finance, it is used to access the risk associated with a particular mutual fund.

32.Merits of standard deviation

i. It is based on all the observations.

ii. It is suitable for further mathematical treatments.

iii. It is less affected by the fluctuations of sampling.

33. A measure of variability which is independent of the units is called the coefficient of variation and isdenoted as C.V.

It is given by the ratio of standard deviation (σ) and the mean ($\bar{x}$) of the data.

34. It is useful for comparing data sets with different units and wildly varying mean. However, mean should be non-zero. If mean is zero or even if it is close to zero, then the coefficient of variation failsto help.

35. Coefficient of variation—a dimensionless constant that helps in comparing the variability of twoobservations with the same or different units.

36. The distribution having a greater coefficient of variation has more variability around the central valuethan the distribution having smaller value of the coefficient of variation.

Key formulae

1. **Arithmetic mean**

(a) Raw data

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

(b) Discrete data

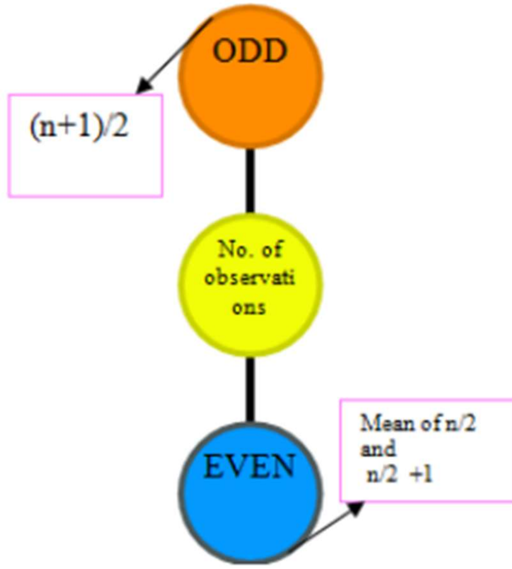
$$\bar{x} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i} = \frac{1}{N}\sum_{i=1}^{n} x_i f_i$$

(c) Step Deviation Method: Let 'a' be the assumed mean and 'h' be the class size.

$$\bar{x} = a + \frac{\sum_{i=1}^{n} f_i d_i}{N} \times h$$

2. **Median**

(a) Median of Ungrouped Data



(b) Median = $l + \left(\frac{\frac{n}{2} - cf}{f}\right) \times h$

Where l = the lower limit of median class.
cf = the cumulative frequency of the class preceding the median class.
f = the frequency of the median class.
h = the class size.

3. Mode for a grouped data is given by

$$\text{Mode} = l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times h$$

l = lower limit of the modal class.
h = size of the class interval.
f_1 = frequency of the modal class.
f_0 = frequency of the class preceding the modal class.
f_2 = frequency of the class succeeding the modal class.

4. Mean Deviation about mean $MD(\bar{x}) = \frac{\sum_{i=1}^{n} f_i |x_i - \bar{x}|}{\sum_{i=1}^{n} f_i} = \frac{1}{N}\sum_{i=1}^{n} f_i |x_i - \bar{x}|$

5. Mean Deviation about median $MD(M) = \frac{1}{N}\sum_{i=1}^{n} f_i |x_i - M|$

6. Variance

(a) For ungrouped data

$$\sigma^2 = \frac{\sum_i (x_i - \bar{x})^2}{n}$$

$$\text{Also, } \sigma^2 = \text{Var}(X) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2$$

(b) For grouped data.

$$\sigma^2 = \frac{\sum_i f_i(x_i - \bar{x})^2}{n}$$

$$\text{Also, } \sigma^2 = \text{Var}(X) = \frac{1}{n}\sum_{i=1}^{n} f_i x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} f_i x_i\right)^2$$

$$\text{Or, } \sigma^2 = \text{Var}(X) = h^2\left[\frac{1}{n}\sum_{i=1}^{n} f_i u_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} f_i u_i\right)^2\right], \text{ where } u_i = \frac{x_i - a}{h}$$

7. Standard Deviations

(a) For ungrouped data $\sigma = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n}}$

(b) For grouped data $\sigma = \sqrt{\frac{1}{N}\sum_{i=l}^{2} f_i(x_i - \bar{x})^2}$, where $\bar{x}$ is the mean of the distribution and $N = \sum_{i=l}^{n} f_i$.

(c) Short-Cut Method $\sigma = \frac{h}{N}\sqrt{N\sum_{i=l}^{n} f_i y_i^2 - \left(\sum_{i\ l}^{n} f_i y_i\right)^2}$

8. Coefficient of Variation: $CV = \frac{\sigma}{\bar{x}} \times 100, \bar{x} \neq 0,$

MIND MAP : LEARNING MADE SIMPLE CHAPTER - 15

Statistics

Variance & Standard Deviation of a discrete frequency distribution

riance of a discrete frequency distribution

Var $\left(\sigma^2\right)=\frac{1}{N}\sum f_i\left(x_i-\bar{x}\right)^2$

ndard Deviation (S.D.) of a discrete frequency stribution

S.D $(\sigma)=\sqrt{\frac{1}{N}\sum f_i\left(x_i-\bar{x}\right)^2}$

where $N=\sum f_i$ and $\bar{x}$ =mean

Variance and Standard Deviation for ungrouped data

Variance of ungrouped data

Var $\left(\sigma^2\right)=\frac{1}{n}\sum\left(x_i-\bar{x}\right)^2$

Standard Deviation (S.D.) for un grouped data:

S.D. $(\sigma)=\sqrt{\frac{1}{n}\sum\left(x_i-\bar{x}\right)^2}$

Coefficient of Variation

Coefficient of variation (C.V.)

$=\frac{\sigma}{\bar{x}}\times 100\,;\ \bar{x}\neq 0$

where $\bar{x}$ is mean

Measures of Dispersion

The degree to which numerical data tend to spread about an average value is called the dispersion of the data. The four dispersions are: Range, Mean deviation, Standard deviation and Variance.

Range

The difference between the highest and the lowest element of a data called its range.

$range\ = x_{\max}-x_{\min}$

Eg: Find the range of the given data:

4, 7, 8, 9, 10, 12, 13, 17

Here, $x_{\max}=17$ and $x_{\min}=4$

Range = 17–4 =13

Shortcut method to find variance and standard deviation

iance

$)=\frac{h^2}{N^2}\left[N\sum f_iy_i^2-\left(\sum f_iy_i\right)^2\right]$

ndard Deviation

$=\frac{h}{N}\sqrt{N\sum f_iy_i^2-\left(\sum f_iy_i\right)^2}$

here $y_i=\frac{x_i-A}{h}$

$N=\sum f_i$

A=assumed mean

h=width of class interval

Mean Deviation

Mean Deviation (M.D.) for ungrouped data

- M.D. about mean i.e., $M.D.(\bar{x})=\frac{\sum\left|x_i-\bar{x}\right|}{n}$, where $\bar{x}$ is mean and n= no. of terms
- M.D. about median i.e., $M.D.(M)=\frac{\sum\left|x_i-M\right|}{n}$, where M is median and n= no. of terms

Mean Deviation (M.D.) for grouped data

M.D. about mean i.e., $M.D.(\bar{x})=\frac{\sum f_i\left|x_i-\bar{x}\right|}{n}$

M.D. about median i.e., $M.D.(M)=\frac{1}{N}\sum f_i\left|x_i-M\right|$

where $N=\sum f_i$

Step Deviation or Shortcut Method

$M.D.(\bar{x})=\frac{1}{N}\sum f_i\left|x_i-\bar{x}\right|$

here $\bar{x}=a+\frac{\sum_{i=1}^{n} f_id_i}{N}\times h$

nd $d_i=\frac{x_i-a}{h}$

Here, a = assumed mean

h = common factor

N = sum of frequencies

$M.D.(M)=\frac{1}{N}\sum f_i\left|x_i-M\right|$

here, M(median) $=l+\frac{\frac{N}{2}-C}{f}\times h$

Median class

Median class is the class interval whose C.F. is just greater than or equal to $\frac{N}{2}$, f, h and C are respectively the lower limit, the frequency, width of the median class and C the cumulative frequency (C.F.) of class just preceding to median class.

Variance & Standard Deviation of a Continuous frequency distribution

Variance of a continuous frequency distribution

Var $(\sigma^2)=\frac{1}{N}\Sigma f_i\left(x_i-\bar{x}\right)^2$

Standard deviation of a continuous frequency distribution

S.D $(\sigma)=\frac{1}{N}\sqrt{N\Sigma f_ix_i^2-\left(\Sigma f_ix_i\right)^2}$

Important Questions

Multiple Choice questions-

Question 1. If the varience of the data is 121 then the standard deviation of the data is

(a) 121

(b) 11

(c) 12

(d) 21

Question 2. The mean deviation from the mean for the following data: 4, 7, 8, 9, 10, 12, 13 and 17 is

(a) 2

(b) 3

(c) 4

(d) 5

Question 3. The mean of 1, 3, 4, 5, 7, 4 is m the numbers 3, 2, 2, 4, 3, 3, p have mean m – 1 and median q. Then, p + q =

(a) 4

(b) 5

(c) 6

(d) 7

Question 4. If the difference of mode and median of a data is 24, then the difference of median and mean is

(a) 12

(b) 24

(c) 8

(d) 36

Question 5. The coefficient of variation is computed by

(a) S.D/.Mean × 100

(b) S.D./Mean

(c) Mean./S.D × 100

(d) Mean/S.D.

Question 6. The geometric mean of series having mean = 25 and harmonic mean = 16 is

(a) 16

(b) 20

(c) 25

(d) 30

Question 7. When tested the lives (in hours) of 5 bulbs were noted as follows: 1357, 1090, 1666, 1494, 1623. The mean of the lives of 5 bulbs is

(a) 1445

(b) 1446

(c) 1447

(d) 1448

Question 8. Mean of the first n terms of the A.P. a + (a + d) + (a + 2d) + ……… is

(a) a + nd/2

(b) a + (n – 1)d

(c) a + (n − 1)d/2

(d) a + nd

Question 9. The mean of a group of 100 observations was found to be 20. Later on, it was found that three observations were incorrect, which was recorded as 21, 21 and 18. Then the mean if the incorrect observations are omitted is

(a) 18

(b) 20

(c) 22

(d) 24

Question 10. If covariance between two variables is 0, then the correlation coefficient between them is

(a) nothing can be said

(b) 0

(c) positive

(d) negative

Very Short Questions:

1. In a test with a maximum marks 25, eleven students scored 3, 9, 5, 3, 12, 10, 17, 4, 7, 19, 21 marks respectively. Calculate the range.
2. Coefficient of variation of two distributions is 70 and 75, and their standard deviations are

28 and 27 respectively what are their arithmetic mean?

3. Write the formula for mean deviation.
4. Write the formula for variance.
5. Find the median for the following data.

   x_i 579101215

   f_i 862226
6. Write the formula of mean deviation about the median
7. Write the formula of mean deviation about the median.
8. Find the mean of the following data 3,6,11,12,18.
9. Express in the form of a + ib (3i-7) + (7-4i) – (6+3i) + i^{23}
10. Find the conjugate of $\sqrt{-3} + 4i^2$

Short Questions:

1. The mean of 2,7,4,6,8 and p is 7. Find the mean deviation about the median of these observations.
2. Find the mean deviation about the mean for the following data!

   x_i 1030507090

   f_i 42428168
3. Find the mean, standard deviation and variance of the first natural n numbers.
4. Find the mean variance and standard deviation for following data.
5. The mean and standard deviation of 6 observations are 8 and 4 respectively. If each observation is multiplied by 3, find the new mean and new standard deviation of the resulting observations.

Long Questions:

1. Calculate the mean, variance and standard deviation of the following data:

Classes	30-40	40-50	50-60	60-70	70-80	80-90	90-100
Frequency	3	7	12	15	8	3	2

2. The mean and the standard deviation of 100 observations were calculated as 40 and 5.1 respectively by a student who mistook one observation as 50 instead of 40. What are the correct mean and standard deviation?
3. 200 candidates the mean and standard deviation was found to be 10 and 15 respectively. After that if was found that the scale 43 was misread as 34. Find the correct mean and correct S.D

4. Find the mean deviation from the mean 6,7,10,12,13,4,8,20

5. Find two numbers such that their sum is 6 and the product is 14.

Assertion Reason Questions:

1. In each of the following questions, a statement of Assertion is given followed by a corresponding statement of Reason just below it. Of the statements, mark the correct answer as.

   Assertion (A) : In order to find the dispersion of values of x from mean $\bar{x}$, we take absolute measure of dispersion.

   Reason (R) : Sum of the deviations from mean ($\bar{x}$) is zero.

   (i) Both assertion and reason are true and reason is the correct explanation of assertion.

   (ii) Both assertion and reason are true but reason is not the correct explanation of assertion.

   (iii) Assertion is true but reason is false.

   (iv) Assertion is false but reason is true.

2. In each of the following questions, a statement of Assertion is given followed by a corresponding statement of Reason just below it. Of the statements, mark the correct answer as.

   Assertion (A) : The mean deviation about the mean for the data 4, 7, 8, 9, 10, 12, 13, 17 is 3.

   Reason (R) : The mean deviation about the mean for the data 38, 70, 48, 40, 42, 55, 63, 46, 54, 44 is 8.5.

   (i) Both assertion and reason are true and reason is the correct explanation of assertion.

   (ii) Both assertion and reason are true but reason is not the correct explanation of assertion.

   (iii) Assertion is true but reason is false.

   (iv) Assertion is false but reason is true.

Answer Key:

MCQ

1. (b) 11
2. (b) 3
3. (d) 7
4. (a) 12

5. (b) S.D./Mean
6. (b) 20
7. (b) 1446
8. (c) a + (n − 1)d/2
9. (b) 20
10. (b) 0

Very Short Answer:

1. The marks can be arranged in ascending order as 3,3,4,5,7,9,10,12,17,19,21.

   Range = maximum value – minimum value

   =21-3

   = 18

2. Given C.V (first distribution) = 70

   Standard deviation $= \sigma_1 = 28$

   C.V $\frac{\sigma 1}{\bar{x}1} \times 100$

   $= 70 = \frac{28}{\bar{x}1} \times 100$

   $\bar{x} = \frac{28}{70} \times 100$

   $\bar{x} = 40$

   Similarly for second distribution

   C.V $= \frac{\sigma_2}{x_2} \times 100$

   $75 = \frac{27}{\bar{x}_2} \times 100$

   $\bar{x}_2 = \frac{27}{75} \times 100$

   $\bar{x}_2 = 36$

3.

$$\text{MD}\left(\bar{x}\right) = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i} = \frac{1}{x} \sum f_i |x_i - \bar{x}|$$

4.

Variance $\sigma^2 = \frac{1}{n}\sum f_i\left(x_i - \bar{x}\right)^2$

5.

x_i	**5**	**7**	**9**	**10**	**12**	**15**
f_i	8	6	2	2	2	6
$c.f$	8	14	16	18	20	26

n = 26Median is the average of 13th and 14th item, both of which lie in the c.f 14

$\therefore x_i = 7$

$$\therefore \text{ median} = \frac{\text{13observation} + \text{14th observation}}{2}$$

$$= \frac{7+7}{2} = 7$$

6.

$$MD.(M) = \frac{\sum f_i \,| x_i M |}{\sum f_i} = \frac{1}{n}\sum f_i \,| x_i - M |$$

7. Range = maximum value – minimum value

= 113-4

= 9

8.

$$\text{Mean} = \frac{\text{sun of observation}}{\text{Total no of observation}}$$

$$= \frac{50}{5} = 10$$

9. Let

$$Z = \cancel{3i} - \cancel{7} + \cancel{7} - 4i - 6 - \cancel{3i} + \left(i^4\right)^5 . i^3$$

$$= -4i - 6 - i \left[\begin{array}{l} \because i^4 = 1 \\ i^3 = -i \end{array}\right.$$

$$= -5i - 6$$

$$= -6 + (-5i)$$

10.

Let $z = \sqrt{-3} + 4i^2$

$=\sqrt{3}\ i - 4$

$\bar{z} = -\sqrt{3}\ i - 4$

Short Answer:

1. Observations are 2, 7, 4, 6, 8 and p which are 6 in numbers n = 6

The near of these observations is 7

$$\frac{2+7+4+6+8+p}{6} = 7$$

$= 27 + p = 42$

$= p = 15$

Arrange the observations in ascending order 2,4,6,7,8,15

$$\therefore \text{Medias (M)} = \frac{\frac{n}{2}\text{ th observation} + \left(\frac{n}{2}+1\right)\text{ th observation}}{2}$$

$$= \frac{\text{3rd observation} + \text{4th observation}}{2}$$

$$= \frac{6+7}{2} = \frac{13}{2}$$

$= 6.5$

Calculation of mean deviation about Median.

xi	xi-M	|xi-M|
2	-4.5	4.5
4	-2.5	2.5
6	-0.5	0.5
7	0.5	0.5
8	1.5	1.5
15	8.5	8.5
Total		18

$\therefore$ Media's deviation about median $= \frac{{}^{3}\cancel{18}}{\cancel{6}} = 3$

2. To calculate mean, we require $f_i x_i$ values then for mean deviation, we require $|x_i - \bar{x}|$ values and $f_i|x_i - \bar{x}|$ values.

x_i	f_i	$f_i x_i$	$\|x_i - \bar{x}\|$	$f_i\|x_i - \bar{x}\|$

10	4	4	40	160
30	24	720	20	480
50	28	1400	0	0
70	16	1120	20	320
90	8	720	40	320
	80	4000		1280

$n = \sum f_i = 80 \quad \sigma d \sum f_i xi = 4000$

$\bar{x} = \frac{\sum f_i xi}{n} = \frac{4000}{80} = 50$

Mean deviation about the mean

$\text{MD}\ (\bar{x}) = \frac{\sum f_i \,|xi - \bar{x}|}{n} = \frac{1280}{80} = 16$

3. The given numbers are 1, 2, 3, ……, n

Mean

$\bar{x} = \frac{\sum n}{n} = \frac{\frac{n(n+1)}{2}}{n} = \frac{n+1}{2}$

Variance

$\sigma 2 = \frac{\sum xi^2}{n} - \bar{x}$

$= \frac{\sum n^2}{n} - \left(\frac{n+1}{2}\right)^2$

$= \frac{n(n+1)(2n+1)}{6n} - \frac{(n+1)^2}{4}$

$= (n+1)\left[\frac{2n+1}{6} - \frac{n+1}{4}\right]$

$= (n+1)\left(\frac{n-1}{12}\right) = \frac{n^2-1}{12}$

$\therefore$ Standard deviation $\sigma = \frac{\sqrt{n^2-1}}{12}$

4.

x_i	**4**	**8**	**11**	**17**	**20**	**24**	**32**

f_i	3	5	9	5	4	3	1

Note: - 4th, 5th and 6th columns are filled in after calculating the mean.

xi	f_i	$f_i x_i$	$xi-\bar{x}$	$\left(xi-\bar{x}\right)^2$	$f_i x_i\left(xi-\bar{x}\right)^2$
4	3	12	-10	100	300
8	5	40	-6	36	180
11	9	99	-3	9	81
17	5	85	3	9	45
20	4	80	6	36	144
24	3	72	10	100	300
32	1	32	18	324	324
Total	30	402			1374

Here $n=\sum f_i=30,\quad \sum f_i x_i=42$

$\therefore$ Mean $\bar{x}=\dfrac{\sum f_i x_i}{n}=\dfrac{420}{30}=14$

$\therefore$ Variance $\sigma^2=\dfrac{1}{n}\sum f_i\left(x_i-\bar{x}\right)^2$

$=\dfrac{1}{30}\times 1374$

$=45.8$

$\therefore$ Standard deviation $\sigma=\sqrt{45.8}$

$=6.77$

5. Let $x_i, x_2 \ldots\ldots x_6$ be the six given observations

Then $\bar{x}=8$ and $\sigma=4$

$\bar{x}=\dfrac{\sum x_i}{n}=8=\dfrac{x_1+x_2+\ldots\ldots+x_6}{6}$

$x_1+x_2+\ldots\ldots x_6=48$

Also $\sigma^2 \frac{\sum x_1^2}{n} - \left(\bar{x}\right)^2$

$= 4^2 = \frac{x_1^2 + x_2^2 ...... + x_6^2}{6} - (8)^2$

$= x_1^2 + x_2^2 + ... x_6^2$

$= 6 \times (16 + 64) = 480$

As each observation is multiplied by 3, new observations are

$3x_1, 3x_2, ... 3x_6$

New near $\overline{X}$ = $\frac{3x_1 + 3x_2 + ... 3x_6}{6}$

$= \frac{3(x_1 + x_2 + ... x_6)}{6}$

$= \frac{3 \times 48}{6}$

$= 24$

Let σ_1 be the new standard deviation, then

$\sigma_1^2 = \frac{(3x_1)^2 + (3x_2)^2 + ... + (3x_6)^2}{6} - \left(\overline{X}\right)^2$

$= \frac{9(x_1^2 + x_2^2 + ... x_6^2)}{6} - (24)^2$

$= \frac{9 \times 480}{6} - 576$

$= 720 - 576$

$= 144$

$\sigma_1 = 12$

Long Answer:

1.

Classes	**Frequency**	**Mid Point**	$f_i xi$	$\left(x_i - \bar{x}\right)^2$	$f_i\left(x_i - \bar{x}\right)^2$
30-40	3	35	105	729	2187
40-50	7	45	315	289	2023
50-60	12	55	660	49	588
60-70	15	65	975	9	135

70-80	8	75	600	169	1352
80-90	3	85	255	529	1587
90-100	2	95	190	1089	2178
Total	50		3100		10050

Here $n = \sum f_i = 50, \sum f_i x_i = 3100$

$\therefore$ Mean $\bar{x} = \frac{\sum f_i x_i}{n} = \frac{3100}{50} = 62$

Variance $\sigma^2 = \frac{1}{n}\sum f_i\left(xi - \bar{x}\right)^2$

$= \frac{1}{50} \times 10050$

$= 201$

Standard deviation $\sigma = \sqrt{201} = 14.18$

2. Given that n = 100

Incorrect mean $\bar{x} = 40$

Incorrect S.D $(\sigma) = 5.1$

As $\bar{x} = \frac{\sum x_i}{n}$

$40 = \frac{\sum x_i}{100} = \sum x_i = 4000$

= incorrect sum of observation =4000

= correct sum of observations = 4000 – 50 + 40

= 3990

So correct mean = $\frac{3990}{100} = 39.9$

Also $\sigma = \sqrt{\frac{1}{n}\sum x_i^2 - \left(\bar{x}\right)^2}$

Using incorrect values,

$5.1 = \sqrt{\frac{1}{100}\sum x_i^2 - (40)^2}$

$= 26.01 = \left[\frac{1}{100}\sum x_i^2 - 1600\right]$

$= \sum x_i^2 = 2601 + 160000$

= 162601

$\sum x_i^2 = 162601$

= incorrect

= correct $\sum x_i^2 = 162601 - (50)^2 + (40)^2$

= 162601 - 2500 + 1600 = 161701

$\therefore$ Correct $\sigma = \sqrt{\frac{1}{100}\text{correct}\sum x_i^2 - \left(\text{correct } \bar{x}\right)^2}$

$= \sqrt{\frac{1}{100}(161701) - (39.9)^2} = \sqrt{1617.01 - 1592.01}$

$= \sqrt{25} = 5$

Hence, correct mean is 39.9 and correct standard deviation is 5.

3.

$n = 200, \overline{X} = 40, \sigma = \overline{15}$

$\overline{X} = \frac{1}{n}\sum x_i = \sum x_i = n\overline{X} = 200 \times 40 = 8000$

Corrected $\sum x_i$ = Incorrect $\sum x_i$ – (sum of incorrect +sum of correct value)

= 8000-34+43 = 8009

$\therefore$ Corrected mean = $\frac{\text{corrected}\sum x_i}{n} = \frac{8009}{200} = 40.045$

$\sigma = 15$

$15^2 = \frac{1}{200}\left(\sum x_i^2\right) - \left(\frac{1}{200}\sum x_i\right)^2$

$225 = \frac{1}{200}\left(\sum x_i^2\right) - \left(\frac{8000}{200}\right)^2$

$225 = \frac{1}{200} \times 1825 = 365000$

Incorrect $\sum x_i^2 = 365000$

Corrected $\sum x_i^2$ = (incorrect $\sum x_i^2$) – (sum of squares of incorrect values) + (sum of square of correct values)

$= 365000 - (34)^2 + (43)^2 = 365693$

Corrected $\sigma = \sqrt{\frac{1}{n}\sum x_i^2 - \left(\frac{1}{n}\sum x_i\right)^2} = \sqrt{\frac{365693}{200} - \left(\frac{8009}{200}\right)^2}$

$\sqrt{1828.465 - 1603.602} = 14.995$

4. Let $\overline{X}$ be the mean

$$\overline{X} = \frac{6+7+10+12+13+4+8+20}{8} = 10$$

x_i	$\lvert d_i \rvert = \lvert x_i - \overline{X} \rvert = \lvert x_i - 10 \rvert$
6	4
7	3
10	0
12	2
13	3
4	6
8	2
20	10
Total	$\sum d_i = 30$

$\sum d_i = 30$ and n = 8

$$\therefore MD = \frac{1}{n}\sum \lvert d_i \rvert = \frac{30}{8} = 3.75$$

$$\therefore MD = 3.75$$

5. Let x and y be the no.

x + y = 6

xy = 14

$$x^2 - 6x + 14 = 0$$

$$D = -20$$

$$x = \frac{-(-6) \pm \sqrt{-20}}{2\times 1}$$

$$= \frac{6 \pm 2\sqrt{5}\,i}{2}$$

$$= 3 \pm \sqrt{5}\,i$$

$$x = 3 + \sqrt{5}\,i$$

$$y = 6 - \left(3 + \sqrt{5}\,i\right)$$

$= 3 - \sqrt{5}\,i$

when $x = 3 - \sqrt{5}\,i$

$y = 6 - \left(3 - \sqrt{5}\,i\right)$

$= 3 + \sqrt{5}\,i$

Assertion Reason Answer:

1. (i) Both assertion and reason are true and reason is the correct explanation of assertion.
2. (iii) Assertion is true but reason is false.