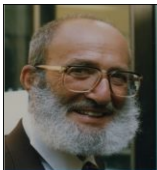

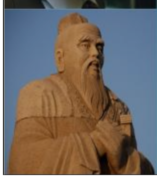

The only way to learn mathematics is to do mathematics.
— Paul Halmos —

MEASURES OF DISPERSION

I hear and I forget.
I see and I remember.
I do and I understand.
— Confucious —

INTRODUCTION

The second important characteristic of a distribution is given by dispersion. Two distributions may be identical in respect of its first important characteristic i.e. central tendency and yet they may differ on account of scatterness.

Dispersion for a given set of observations may be defined as the amount of deviation of the observations, usually, from an appropriate measure of central tendency.

Measures of dispersion may be broadly classified into

1. Absolute measures of dispersion
2. Relative measures of dispersion (unit free measure)

INTRODUCTION

Absolute measures of dispersion are classified into

i. Range
ii. Mean Deviation
iii. Quartile Deviation
iv. Standard Deviation

Likewise, we have the following relative measures of dispersion :

i. Coefficient of Range
ii. Coefficient of Mean Deviation
iii. Coefficient of Quartile Deviation
iv. Coefficient of Variation

RANGE

For a given set of observations, range may be defined as the difference between the largest and smallest of observations.

$$Range = Max - Min$$

The corresponding relative measure of dispersion, known as coefficient of range, is given by

$$Coefficient\ of\ Range = \frac{Max - Min}{Max + Min} \times 100$$

RANGE

Example 1: Find range and coefficient of range for the following:

$$12, 6, 7, 3, 15, 10, 18, 5$$

Solution:

$$Range = 18 - 3 = 15$$

$$Coefficient\ of\ Range = \frac{18 - 3}{18 + 3} \times 100 = 71.43$$

RANGE

Example 2: Find range and coefficient of range for the following:

Size	Frequency
8	14
9	21
10	17
11	12
12	6

Ans: 4, 20

Weight (lb)	Frequency
118–126	3
127–135	5
136–144	9
145–153	12
154–162	5
163–171	4
172–180	2

Ans: 63, 21.14

RANGE

Result: Range remains unaffected due to a change of origin but affected in the same ratio due to a change in scale i.e., if for any two constants a and b, two variables x and y are related by y = a + bx,

$$Range(y) = |b| \times Range(x)$$

Example: If the relationship between x and y is given by 2x+3y=10 and the range of x is 15, what would be the range of y?

Solution: Given that $2x + 3y = 10 \Rightarrow y = \frac{10}{3} - \frac{2}{3}x$

$$\therefore Range(y) = \frac{2}{3} Range(x) = 10$$

MEAN DEVIATION

1. MD from Mean:
   - For raw data: $MD\ from\ Mean = \frac{1}{N}\sum|x - \bar{x}|$
   - For frequency data: $MD\ from\ Mean = \frac{1}{N}\sum f|x - \bar{x}|$
2. MD from Median:
   - For raw data: $MD\ from\ Median = \frac{1}{N}\sum|x - M|$
   - For frequency data: $MD\ from\ Median = \frac{1}{N}\sum f|x - M|$

- $Coefficient\ of\ MD = \frac{MD\ about\ A}{A} \times 100$

MEAN DEVIATION

Result 1: Mean deviation takes its minimum value when the deviations are taken from the median.

Result 2: Mean deviation remains unchanged due to a change of origin but changes in the same ratio due to a change in scale i.e. if y = a + bx, a and b being constants, then $MD(y) = |b| \times MD(x)$.

Example: If x and y are related as 4x+3y+11 = 0 and mean deviation of x is 5.40, what is the mean deviation of y?

Answer: $MD(y) = 7.20$.

MEAN DEVIATION

Example 1: Find MD from mean, MD from median and corresponding coefficients of MD for the following:

12, 6, 7, 3, 15, 10, 18, 5

Solution:

For MD from mean:

$$\bar{x} = \frac{76}{8} = 9.5$$

$$MD\ from\ \bar{x} = \frac{2.5 + 3.5 + 2.5 + 6.5 + 5.5 + 0.5 + 8.5 + 4.5}{8} = 4.25$$

$$Coefficient\ of\ MD = \frac{4.25}{9.5} \times 100 = 44.74$$

MEAN DEVIATION

Example 1: Find MD from mean, MD from median and corresponding coefficients of MD for the following:

12, 6, 7, 3, 15, 10, 18, 5

Solution:

Cont….

For MD from median:

$$M = 8.5$$

$$MD\ from\ M = \frac{3.5 + 2.5 + 1.5 + 5.5 + 6.5 + 1.5 + 9.5 + 3.5}{8} = 4.25$$

$$Coefficient\ of\ MD = \frac{4.25}{8.5} \times 100 = 51.83$$

MEAN DEVIATION

Example 2: Compute MD from mean and coefficient of MD for the following:

x	f	fx	$\|x - \bar{x}\|$	$f\|x - \bar{x}\|$
6	3	18	2.12	6.36
7	5	35	1.12	5.6
8	7	56	0.12	0.84
9	6	54	0.88	5.28
10	4	40	1.88	7.52
	25	203		25.6

$\bar{x} = \frac{203}{25} = 8.12$

$MD\ from\ \bar{x} = \frac{25.6}{25}$

$= 1.024$

$CMD = \frac{1.024}{8.12} \times 100$

$= 12.61$

MEAN DEVIATION

Example 3: Compute MD from median and coefficient of MD for the following:

CI	Frequency	cf	x	$\lvert x-M \rvert$	$f\lvert x-M \rvert$
0 - 10	5	5	5	26.67	133.35
10 - 20	10	15	15	16.67	166.7
20 - 30	12	27	25	6.67	80.04
30 - 40	18	45	35	3.33	59.94
40 - 50	9	54	45	13.33	119.97
50 - 60	6	60	55	23.33	139.98
	60				699.98

$M = 31.67$

$$MD\ from\ M = \frac{699.98}{60} = 11.67$$

$$\text{CMD} = \frac{11.67}{31.67} \times 100 = 36.85$$

QUARTILE DEVIATION

- Quartile deviation is given by

$$QD = Q_3 - Q_1$$

- Coefficient of quartile deviation is given by

$$CQD = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

- Semi-Interquartile Range:

$$SQR = \frac{Q_3 - Q_1}{2}$$

QUARTILE DEVIATION

Example 1: Find quartile deviation and coefficient of quartile deviation for the following:

CI	Frequency	cf
0 - 10	5	5
10 - 20	10	15
20 - 30	12	27
30 - 40	18	45
40 - 50	9	54
50 - 60	6	60
	60	

QUARTILE DEVIATION

Example 2: Find quartile deviation and coefficient of quartile deviation for the following:

Rainfall in cms	No. of Cities	CF
0 - 10	10	10
10 - 20	15	25
20 - 30	20	45
30 - 40	10	55
40 - 50	5	60
	60	

QUARTILE DEVIATION

- Quartile deviation provides the best measure of dispersion for open-end classification.
- It is also less affected due to sampling fluctuations. Like other measures of dispersion.
- Quartile deviation remains unaffected due to a change of origin but is affected in the same ratio due to change in scale.

Example: If the quartile deviation of x is 6 and 3x + 6y = 20, what is the quartile deviation of y?

Answer: $QD(y) = 3$.

STANDARD DEVIATION

Although mean deviation is an improvement over range so far as a measure of dispersion is concerned, mean deviation is difficult to compute and further more, it cannot be treated mathematically. The best measure of dispersion is, usually, standard deviation which does not possess the demerits of range and mean deviation.

Standard deviation for a given set of observations is defined as the root mean square deviation when the deviations are taken from the AM of the observations.

STANDARD DEVIATION

➢For raw data:

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{N}} = \sqrt{\frac{\sum x^2}{N} - \bar{x}^2}$$

➢For frequency data:

$$s = \sqrt{\frac{\sum f(x-\bar{x})^2}{N}} = \sqrt{\frac{\sum fx^2}{N} - \bar{x}^2}$$

The square of standard deviation, known as variance, is regarded as a measure of dispersion.

STANDARD DEVIATION

A relative measure of dispersion using standard deviation is given by coefficient of variation (cv) which is defined as the ratio of standard deviation to the corresponding arithmetic mean, expressed as a percentage.

$$Coefficient\ of\ variation\ (CV) = \frac{SD}{AM} \times 100$$

STANDARD DEVIATION

Example 1: Find the standard deviation and the coefficient of variation for the following numbers:

12, 6, 7, 3, 15, 10, 18, 5

Solution:

$$\bar{x} = \frac{76}{8} = 9.5$$

$$s = \sqrt{\frac{912}{8} - 9.5^2} = 4.87$$

$$cv = \frac{4.87}{9.5} \times 100 = 51.26$$

x	x²
3	9
5	25
6	36
7	49
10	100
12	144
15	225
18	324
76	912

STANDARD DEVIATION

Example 2: Find the standard deviation and the coefficient of variation for the following:

Solution:

$$\bar{x} = \frac{1840}{60} = 30.67$$

$$s = \sqrt{\frac{68300}{60} - 30.67^2}$$

$$= 14.06$$

$$cv = \frac{14.06}{30.67} \times 100 = 45.84$$

CI	Frequency	x	fx	fx²
0 - 10	5	5	25	125
10 - 20	10	15	150	2250
20 - 30	12	25	300	7500
30 - 40	18	35	630	22050
40 - 50	9	45	405	18225
50 - 60	6	55	330	18150
	60		1840	68300

PROPERTIES OF STANDARD DEVIATION

1. If all the observations assumed by a variable are constant i.e. equal, then the SD is zero. This means that if all the values taken by a variable x is k, say, then s = 0. This result applies to range as well as mean deviation.
2. SD remains unaffected due to a change of origin but is affected in the same ratio due to a change of scale i.e., if there are two variables x and y related as y = a+bx for any two constants a and b, then SD of y is given by

$$s_y = |b|s_x$$

PROPERTIES OF STANDARD DEVIATION

3. If there are two groups containing n_1 and n_2 observations, $\overline{x_1}$ and $\overline{x_2}$ as respective AMs, s_1 and s_2 are respective SDs, then combined SD is given by

$$s = \sqrt{\frac{n_1({s_1}^2 + {d_1}^2) + n_2({s_2}^2 + {d_2}^2)}{n_1 + n_2}}$$

$$where, d_1 = \bar{x} - \overline{x_1}$$

$$d_2 = \bar{x} - \overline{x_2}$$

$$and\ \bar{x} = \frac{n_1\overline{x_1} + n_2\overline{x_2}}{n_1 + n_2}$$

STANDARD DEVIATION (CODING METHOD EXAMPLE)

Example 3: Find the standard deviation for the following:

Height (in)	Frequency	x	$u=\frac{x-67}{3}$	fu	fu²
60 - 62	5	61	-2	-10	20
63 - 65	18	64	-1	-18	18
66 - 68	42	67	0	0	0
69 - 71	27	70	1	27	27
72 - 74	8	73	2	16	32
	100			15	97

STANDARD DEVIATION (CODING METHOD EXAMPLE)

Example 3: Continue…

$$s_u = \sqrt{\frac{\sum fu^2}{N} - \bar{u}^2}$$

$$s_x = \sqrt{\frac{\sum fu^2}{N} - \bar{u}^2} \times C$$

$$s_x = \sqrt{\frac{97}{100} - \left(\frac{15}{100}\right)^2} \times 3 = 2.92$$

STANDARD DEVIATION

Example 4: If AM and coefficient of variation of x are 10 and 40 respectively, what is the variance of (15–2x)?

Solution: Let $y = 15 - 2x \quad \Longrightarrow s_y = 2s_x$

$$CV(x) = \frac{s_x}{\bar{x}} \times 100$$

$$40 = \frac{s_x}{10} \times 100$$

$$\therefore\ s_x = 4 \qquad \Longrightarrow s_y = 2(4) = 8$$

$$\therefore var(y) = 8^2 = 64$$

STANDARD DEVIATION

Example 5: For a group of 50 male workers the mean and standard deviation of their daily wages are 63 dollars and 9 dollars respectively. For a group of 40 female workers these values are 54 dollars and 6 dollars respectively. Find the mean and variance of the combined group of 90 workers.

Solution:

Here $n_1 = 50, \overline{x_1} = 63, s_1 = 9$ and $n_2 = 40, \overline{x_2} = 54, s_2 = 6$

$$\bar{x} = \frac{50 \times 63 + 40 \times 54}{50 + 40} = 59$$

$$d_1 = -4, d_2 = 5$$

$$s = \sqrt{\frac{50(81+16) + 40(36+25)}{50+40}} = 9 \qquad \Longrightarrow s^2 = variance = \$81$$

STANDARD DEVIATION

Example 6: For a group of 50 male workers the mean and standard deviation of their daily wages are 63 dollars and 9 dollars respectively. For a group of 40 female workers these values are 54 dollars and 6 dollars respectively. Examine consistency of both the group.

Solution:

Here $n_1 = 50, \overline{x_1} = 63, s_1 = 9$ and $n_2 = 40, \overline{x_2} = 54, s_2 = 6$

$$CV(M) = \frac{9}{63} \times 100 = 14.29$$

$$CV(F) = \frac{6}{54} \times 100 = 11.11$$

Therefore, female group has more consistent wage.

BASIC QUESTIONS

Q.1 Which measure of dispersion is most useful?

Ans: Standard deviation

Q.2 Which measure of dispersion is the easiest to calculate?

Ans: Range

Q.3 If $3x + 2y = 40$ and range(x) = 12 then range(y) = ?

Ans: $y = \frac{40}{2} - \frac{3x}{2} \qquad \Longrightarrow range(y) = \frac{3}{2} range(x) = \frac{3}{2} \times 12 = 18.$

BASIC QUESTIONS

Q.4 If $3x - 4y = 20$ and MD(x) = 8 then MD(y) = ?

Ans: $y = \frac{3x}{4} - \frac{20}{4} \quad \Rightarrow MD(y) = \frac{3}{4} MD(x) = \frac{3}{4} \times 8 = 6.$

Q.5 If $3x + y = 12$ and QD(x) = 15 then QD(y) = ?

Ans: $y = 12 - 3x \quad \Rightarrow QD(y) = 3\, QD(x) = 3 \times 15 = 45.$

Q.6 If $2x + y = 10$ and SD(x) = 4 then SD(y) = ?

Ans: $y = 10 - 2x \quad \Rightarrow SD(y) = 2\, SD(x) = 2 \times 4 = 8.$

BASIC QUESTIONS

Q.7 If $4x - 3y = 15$ and Var(x) = 3 then Var(y) = ?

Ans: $y = \frac{4x}{3} - \frac{15}{3} \quad \Rightarrow SD(y) = \frac{4}{3} SD(x)$

$$\Rightarrow Var(y) = \left(\frac{4}{3}\right)^2 \times Var(x) = \frac{16}{9} \times 3 = 5.33.$$

Q.8 If mean = 20 and variance = 4 then coefficient of variation = ?

Ans: Coefficient of Variation $= \frac{SD}{Mean} \times 100 = \frac{2}{20} \times 100 = 10\%.$