MEASURES OF DISPERSION

Birinder Singh, Assistant Professor, PCTE

Dispersion

- Dispersion refers to the variations of the items among themselves / around an average.
- Greater the variation amongst different items of a series, the more will be the dispersion.
- As per Bowley, *"Dispersion is a measure of the variation of the items"*.

Birinder Singh, Assistant Professor, PCTE

Objectives of Measuring Dispersion

- To determine the reliability of an average
- To compare the variability of two or more series
- For facilitating the use of other statistical measures
- Basis of Statistical Quality Control

Birinder Singh, Assistant Professor, PCTE

Properties of a Good Measure of Dispersion

- Easy to understand
- Simple to calculate
- Uniquely defined
- Based on all observations
- Not affected by extreme observations
- Capable of further algebraic treatment

Measures of Dispersion

Absolute

- Expressed in the same units in which data is expressed
- Ex: Rupees, Kgs, Ltr, Km etc.

Relative

- In the form of ratio or percentage, so is independent of units
- It is also called **Coefficient of Dispersion**

Birinder Singh, Assistant Professor, PCTE

Methods of Measuring Dispersion

- **Range**
- **Interquartile Range & Quartile Deviation**
- **Mean Deviation**
- **Standard Deviation**
- **Coefficient of Variation**
- **Lorenz Curve**

Birinder Singh, Assistant Professor, PCTE

RANGE (R)

- It is the simplest measures of dispersion
- It is defined as the difference between the largest and smallest values in the series

$$R = L - S$$

R = Range, L = Largest Value, S = Smallest Value

- Coefficient of Range = $\frac{L - S}{L + S}$

Practice Problems – Range

Q1: Find the range & Coefficient of Range for the following data: 20, 35, 25, 30, 15

Ans: 20, 0.4

Q2: Find the range & Coefficient of Range:

X	10	20	30	40	50	60	70
F	15	18	25	30	16	10	9

Ans: 60, 0.75

Q3: Find the range & Coefficient of Range:

Size	5-10	10-15	15-20	20-25	25-30
F	4	9	15	30	40

Ans: 25, 5/7

Birinder Singh, Assistant Professor, PCTE

Range

MERITS	DEMERITS
- Simple to understand - Easy to calculate - Widely used in statistical quality control	- Can't be calculated in open ended distributions - Not based on all the observations - Affected by sampling fluctuations - Affected by extreme values

Birinder Singh, Assistant Professor, PCTE

Interquartile Range & Quartile Deviation

- ***Interquartile Range*** is the difference between the upper quartile (Q_3) and the lower quartile (Q_1)
- It covers dispersion of middle 50% of the items of the series
- Symbolically, Interquartile Range = $Q_3 - Q_1$

- ***Quartile Deviation*** is half of the interquartile range. It is also called Semi Interquartile Range
- Symbolically, Quartile Deviation = $\frac{Q_3 - Q_1}{2}$

- ***Coefficient of Quartile Deviation***: It is the relative measure of quartile deviation.
- Coefficient of Q.D. = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

Birinder Singh, Assistant Professor, PCTE

Practice Problems – IQR & QD

Q1: Find interquartile range, quartile deviation and coefficient of quartile deviation:

28, 18, 20, 24, 27, 30, 15 *Ans: 10, 5, 0.217*

Q2:

X	10	20	30	40	50	60
F	2	8	20	35	42	20

Ans: 10, 5, 0.11

Q3:

Age	0-20	20-40	40-60	60-80	80-100
Persons	4	10	15	20	11

Ans: 14.33, 0.19

Birinder Singh, Assistant Professor, PCTE

MEAN DEVIATION (M.D.)

- It is also called Average Deviation
- It is defined as the arithmetic average of the deviation of the various items of a series computed from measures of central tendency like mean or median.
- M.D. from Median = $\frac{\Sigma\,|X - M|}{N}$ or $\frac{\Sigma\,|d_M|}{N}$
- M.D. from Mean = $\frac{\Sigma\,|X - \bar{X}|}{N}$ or $\frac{\Sigma\,|d_{\bar{x}}|}{N}$
- Coefficient of M.D.$_M$ = $\frac{M.D._M}{Median}$
- Coefficient of M.D.$_{\bar{X}}$ = $\frac{M.D._{\bar{x}}}{Mean}$

Birinder Singh, Assistant Professor, PCTE

Practice Problems – Mean Deviation

Q1: Calculate M.D. from Mean & Median & coefficient of Mean Deviation from the following data: 20, 22, 25, 38, 40, 50, 65, 70, 75

Ans: 17.78, 17.22, 0.39,0.43

Q2:

X:	20	30	40	50	60	70
f:	8	12	20	10	6	4

Ans: 10.67, 10.33, 0.26, 0.26

Q3: Calculate M.D. from Mean & its coefficient:

X:	0-10	10-20	20-30	30-40	40-50
f:	5	8	15	16	6

Ans: 9.44, 0.349

Birinder Singh, Assistant Professor, PCTE

Mean Deviation – Short Cut Method

- If value of the average comes out to be in fractions, the calculation of M.D. by $\frac{\Sigma |X - \bar{X}|}{N}$ would become quite tedious. In such cases, the following formula is used:
- M.D. = $\frac{(\Sigma fX)_A - (\Sigma fX)_B - [(\Sigma f)_A - (\Sigma f)_B] \bar{X} \text{ or } M}{N}$

Practice Problems – Shortcut Method

Q4: Calculate M.D. from Mean & Median using shortcut method: 7, 9, 13, 13, 15, 17, 19, 21, 23

Ans: 4.25, 4.22

Q5: Calculate M.D. from Mean & Median & coefficient of Mean Deviation from the following data:

X:	0-10	10-20	20-30	30-40	40-50
f:	6	28	51	11	4

Ans: 6.57, 0.29, 6.5, 0.28

Birinder Singh, Assistant Professor, PCTE

Mean Deviation

Merits

- Simple to understand
- Easy to compute
- Less effected by extreme items
- Useful in fields like Economics, Commerce etc.
- Comparisons about formation of different series can be easily made as deviations are taken from a central value

Demerits

- Ignoring '±' signs are not appropriate
- Not accurate for Mode
- Difficult to calculate if value of Mean or Median comes in fractions
- Not capable of further algebraic treatment
- Not used in statistical conclusions.

Birinder Singh, Assistant Professor, PCTE

Standard Deviation

- Most important & widely used measure of dispersion
- First used by Karl Pearson in 1893
- Also called root mean square deviations
- It is defined as the square root of the arithmetic mean of the squares of the deviation of the values taken from the mean
- Denoted by σ (sigma)
- $\sigma = \sqrt{\frac{\Sigma(X-\bar{X})^2}{N}}$ or $\sqrt{\frac{\Sigma x^2}{N}}$ where $x = X - \bar{X}$
- Coefficient of S.D. $= \frac{\sigma}{\bar{X}}$

Birinder Singh, Assistant Professor, PCTE

Calculation of Standard Deviation

Individual Series

- Actual Mean Method
- Assumed Mean Method
- Method based on Actual Data

Discrete Series

- Actual Mean Method
- Assumed Mean Method
- Step Deviation Method

Continuous Series

- Actual Mean Method
- Assumed Mean Method
- Step Deviation Method

Birinder Singh, Assistant Professor, PCTE

Standard Deviation – Individual Series Actual Mean Method

- $\sigma = \sqrt{\frac{\Sigma(X-\bar{X})^2}{N}}$ or $\sqrt{\frac{\Sigma x^2}{N}}$ where $x = X - \bar{X}$

Q1: Calculate the SD of the following data:

16, 20, 18, 19, 20, 20, 28, 17, 22, 20

Ans: 3.13

Birinder Singh, Assistant Professor, PCTE

Standard Deviation – Individual Series Assumed Mean / Shortcut Method

- $\sigma = \sqrt{\frac{\Sigma d^2}{N} - \left(\frac{\Sigma d}{N}\right)^2}$ where $d = X - A$

Q2: Calculate the SD of the following data:

7, 10, 12, 13, 15, 20, 21, 28, 29, 35

Ans: 8.76

Standard Deviation – Individual Series Method Based on Use of Actual Data

- $\sigma = \sqrt{\frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2}$

Q3: Calculate the SD of the following data:

16, 20, 18, 19, 20, 20, 28, 17, 22, 20

Ans: 3.13

Birinder Singh, Assistant Professor, PCTE

Standard Deviation – Discrete Series Actual Mean Method

- $\sigma = \sqrt{\frac{\Sigma f(X-\bar{X})^2}{N}}$ or $\sqrt{\frac{\Sigma f x^2}{N}}$ where $x = X - \bar{X}$

Q4: Calculate the SD of the following data:

X:	3	4	5	6	7	8	9
F:	7	8	10	12	4	3	2

Ans: 1.602

Standard Deviation – Discrete Series Assumed Mean / Shortcut Method

- $\sigma = \sqrt{\frac{\Sigma f d^2}{N} - \left(\frac{\Sigma f d}{N}\right)^2}$ where $d = X - A$

Q5: Calculate the SD of the following data:

X:	3	4	5	6	7	8	9
F:	7	8	10	12	4	3	2

Ans: 1.602

Birinder Singh, Assistant Professor, PCTE

Standard Deviation – Discrete Series Step Deviation Method

- $\sigma = \sqrt{\frac{\Sigma f d'^2}{N} - \left(\frac{\Sigma f d'}{N}\right)^2} \times i$ where $d' = \frac{X - A}{i}$

Q6: Calculate the SD of the following data:

X	10	20	30	40	50	60	70
F:	3	5	7	9	8	5	3

Ans: 16.5

Birinder Singh, Assistant Professor, PCTE

Standard Deviation – Continuous Series Step Deviation Method

- $\sigma = \sqrt{\frac{\Sigma f d'^2}{N} - \left(\frac{\Sigma f d'}{N}\right)^2} \times i$ where $d' = \frac{X - A}{i}$

Q7: Calculate the Mean & SD of the following data:

X	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
F:	5	10	20	40	30	20	10	4

Ans: 39.38, 15.69

Birinder Singh, Assistant Professor, PCTE

Variance

- It is another measure of dispersion
- It is the square of the Standard Deviation
- Variance = $(SD)^2 = \sigma^2$

Q8: Calculate the Mean & Variance:

X:	0-10	10-20	20-30	30-40	40-50
F:	2	7	10	5	3

Ans: 25, 118.51

Birinder Singh, Assistant Professor, PCTE

Combined Standard Deviation

- It is the combined standard deviation of two or more groups as in case of combined arithmetic mean
- It is denoted by $\sigma_{12} = \sqrt{\frac{N_1\sigma_1{}^2+N_2\sigma_2{}^2+N_1d_1{}^2+N_2d_2{}^2}{N_1+N_2}}$

where σ_{12} = Combined SD

σ_1 = SD of first group

σ_2 = SD of second group

$d_1 = \overline{X_1} - \overline{X_{12}}$

$d_2 = \overline{X_2} - \overline{X_{12}}$

Practice Problems

Q9: Two samples of sizes 100 & 150 respectively have means 50 & 60 and SD 5 & 6. Find the Combined Mean & Combined Standard Deviation.

Ans: 56, 7.46

Birinder Singh, Assistant Professor, PCTE

IMPORTANT PRACTICE PROBLEMS

Q10: The mean weight of 150 students is 60 kg. The mean weight of boys is 70 kg with SD of 10 kg. The mean weight of girls is 55 kg with SD of 15 kg. Find the number of boys & girls and their combined standard deviation.

Ans: 50, 100, 15.28

Q11: Find the missing information from the following:

	Group I	Group II	Group III	Combined
Number	50	?	90	200
SD	6	7	?	7.746
Mean	113	?	115	116

Ans: 60, 120, 8

Important Practice Problems

Q12: For a group of 100 observations, the mean & SD were found to be 60 & 5 respectively. Later on, it was discovered that a correct item 50 was wrongly copied as 30. Find the correct mean & correct SD.

Ans: 60.20, 4.12

Q13: The mean, SD and range of a symmetrical distribution of weights of a group of 20 boys are 40 kgs, 5 kgs and 6 kgs respectively. Find the mean & SD of the group if the lightest and the heaviest boys are excluded.

Ans: 40, 5.17

Q14: The mean of 5 observations is 4.4 and the variance is 8.24. If three observations are 4,6 and 9, find the other two.

Ans: 1, 2

Coefficient of Variation (C.V.)

- It was developed by Karl Pearson.
- It is an important relative measure of dispersion.
- It is used in comparing the variability, homogeneity, stability, uniformity & consistency of two or more series.
- Higher the CV, lesser the consistency.
- C.V. $= \frac{\sigma}{\bar{X}} \, x \, 100$

Practice Problems

Q1: The scores of two batsmen A & B in ten innings during a certain match are:

A	32	28	47	63	71	39	10	60	96	14
B	19	31	48	53	67	90	10	62	40	80

Ans: B, B

Q2: Goals scored by two teams A & B in a football session were as follows:

No. of goals scored	0	1	2	3	4
No. of matches by A	27	9	8	5	4
No. of matches by B	17	9	6	5	3

Ans: B

Q3: Sum of squares of items is 2430 with mean 7 & N = 12. Find coefficient of variation. *Ans: 176.85%*

Birinder Singh, Assistant Professor, PCTE