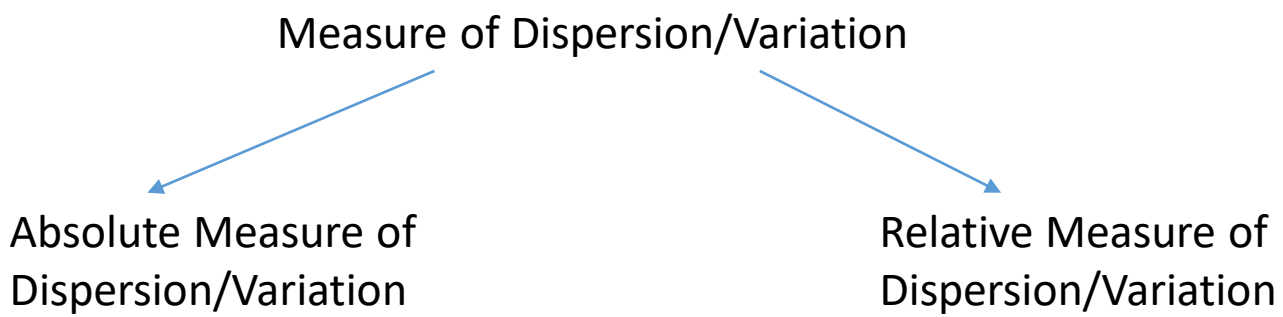
Probability and Statistics

Slide Set # 4


Measure of Dispersion/Variation
Absolute Measure of Dispersion/Variation
Relative Measure of Dispersion/Variation

Absolute Measure of Dispersion/Variation

Uses for comparison of variability of two or more data set of same units

Types of absolute measure of dispersion/variation

- Range
- Variance
- Standard Deviation

Types of absolute measure of dispersion/variation

Range: (worst measure)

"Difference b/w maximum and minimum value of data"

Variance:

"Average squared deviation from mean"

Standard Deviation: (ideal measure)

"Square root of variance"

Types of absolute measure of dispersion/variation

Variance: (Computation)

Def: Average squared deviation from mean

Formula:

$\sigma^2 = \frac{\sum(x-\bar{x})^2}{n}$ or $\sigma^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2$ (population variance)

$s^2 = \frac{\sum(x-\bar{x})^2}{n-1}$ or $s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$ (sample variance)

$\sigma^2 = \frac{\sum f(x-\bar{x})^2}{\sum f}$ or $\sigma^2 = \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2$ (population variance, for grouped data)

$s^2 = \frac{\sum f(x-\bar{x})^2}{\sum f-1}$ or $s^2 = \frac{\sum fx^2 - \frac{(\sum fx)^2}{\sum f}}{\sum f-1}$ (sample variance, for grouped data)

Types of absolute measure of dispersion/variation

Standard Deviation (S.D): (Computation)

Def: Square root of variance

Formula:

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

$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$ or $s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}}$ (sample S.D)

$\sigma = \sqrt{\frac{\sum f(x-\bar{x})^2}{\sum f}}$ or $\sigma = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$ (population S.D, for grouped data)

$s = \sqrt{\frac{\sum f(x-\bar{x})^2}{\sum f-1}}$ or $s = \sqrt{\frac{\sum fx^2 - \frac{(\sum fx)^2}{\sum f}}{\sum f-1}}$ (sample S.D, for grouped data)

Example of variance and S.D of raw data:

Data Set I	41	44	45	47	47	48	51	53	58	66
Data Set II	20	37	48	48	49	50	53	61	64	70

Which data has more variation / less consistency from above two sample data?

Working of data set I:

x	mean of x	(x-mean)^2
41	50	81
44		36
45		25
47		9
47		9
48		4
51		1
53		9
58		64
66		256
500	50	494

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{494}{9} = 54.88$$
$$s = 7.408$$

or

x	x^2
41	1681
44	1936
45	2025
47	2209
47	2209
48	2304
51	2601
53	2809
58	3364
66	4356
500	25494

$$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{\mathrm{n}}}{\mathrm{n}-1} = \frac{25494 - \frac{500^2}{10}}{9} = 54.88$$
$$s = 7.408$$

Example of variance and S.D of raw data:

Data Set I	41	44	45	47	47	48	51	53	58	66
Data Set II	20	37	48	48	49	50	53	61	64	70

Working of data set II:

x	mean of x	(x-mean)^2
20	50	900
37		169
48		4
48		4
49		1
50		0
53		9
61		121
64		196
70		400
500	50	1804

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{1804}{9} = 200.44$$

$$s = 14.157$$

or

x	x^2
20	400
37	1369
48	2304
48	2304
49	2401
50	2500
53	2809
61	3721
64	4096
70	4900
500	26804

$$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{\mathrm{n}}}{\mathrm{n}-1} = \frac{26804 - \frac{500^2}{10}}{9} = 200.44$$

$$s = 14.157$$

Example of variance and S.D of grouped data:

Class Interval	Class Boundaries	Frequency (f)
30-39	29.5-39.5	3
40-49	39.5-49.5	1
50-59	49.5-59.5	8
60-69	59.5-69.5	10
70-79	69.5-79.5	7
80-89	79.5-89.5	7
90-99	89.5-99.5	4
		40

Compute variance and S.D of above data

Class Interval	Class Boundaries	Frequency (f)	x	fx	fx^2
30-39	29.5-39.5	3	34.5	103.5	3570.75
40-49	39.5-49.5	1	44.5	44.5	1980.25
50-59	49.5-59.5	8	54.5	436	23762
60-69	59.5-69.5	10	64.5	645	41602.5
70-79	69.5-79.5	7	74.5	521.5	38851.75
80-89	79.5-89.5	7	84.5	591.5	49981.75
90-99	89.5-99.5	4	94.5	378	35721
		40	451.5	2720	195470

$$\sigma^2 = \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2 = 262.75$$

$$\sigma = 16.209$$

Relative Measure of Dispersion/Variation

Uses for comparison of variability of two or more data set of different units

Types of relative measure of dispersion/variation

- Co-efficient of variation (ideal)

Types of relative measure of dispersion/variation

Coefficient of variation:

Def: ratio of s.d(x) to mean(x)

Formula:

$$C.V = \frac{S.D(x)}{mean(x)} * 100\%$$

Example of coefficient of variation

Price (in Rs.)	100	230	240	234	250
life (in months)	30	32	33	35	38

Which data has less variation?

Solution : (Do it yourself by using calculator)

Book : Neil.Weiss
Exercise : 3.62-3.78
Page: 110-111