Measures of disperson: The numerical values by which we measures the disperson or variability of a data set are called measures of disperson.

There are two kinds of measures of disperson. They are

i) Absulate measures of disperson

ii) Relative " " "

Absulate measures of disperson are

i) Range
ii) Quartile deviation
iii) Mean deviation
iv) Variance and Standard deviation.

Relative measures of disperson are

i) co-efficient of range
ii) Co-efficient of quartile deviation.
iii) co-efficient of mean deviation.
iv) co-efficient of variation.

Measures of Dispersion

Dispersion: Dispersion measures the variability of a set of observations among themselves or about some central values.

Purposes of dispersion:

Measures of dispersion is needed for four basic purposes

i) To determine the reliability of an average.

ii) To serve as a basis for control of the variability.

iii) To compare two or more series with regard to their variability.

iv) To facilitate the computation of other statistical measures.

Properties of a good measures of dispersion:

i) It should be simple to understand.

ii) It should be easy to compute.

iii) It should be rigidly defined.

iv) It should be based on all the observations.

v) It should be sampling stability.

vi) It should be suitable for further algebric treatment.

vii) It should not be affected by extreme observations

Range: Range is the difference between the largest and smallest observations in a set of data.

Range = $X_L - X_S$; X_L = largest observation
X_S = smalles "

Co-efficient of range: The relative measures corresponding to range, called the co-efficient of range is computed by the following formula

$$\text{Coefficient of range} = \frac{X_L - X_S}{X_L + X_S} \times 100$$

Quartile deviation: if Q_3 and Q_1 are third and first quartiles of a data set, quartile dieviation denoted by, Q.D is given by

$$Q.D = \frac{Q_3 - Q_1}{2}$$

small quartile deviation means high informity and large quartile deviation means large variation among the central observations.

Co-efficient of quartile deviation: ~~quo~~

The relative measures corresponding to quartile deviation is called co-efficient of quartile deviation.

$$\text{Co-efficient of } Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

Mean deviation: Mean deviation is obtained by calculating the absulate deviations of each observation from mean or median or mode and averaging these deviations by taking their arithmatic mean.

So we can define mean deviation in three ways, namely

i) mean deviation about mean

ii) " " " median

iii) " " " mode.

Suppose $x_1 x_2 \dots x_n$ are n observations of a data set and $\bar{x}$ is the mean, then mean deviation about mean is

$$M.D(\bar{x}) = \frac{\Sigma|x-\bar{x}|}{n}$$

Mean deviation about median is

$$M.D(Me) = \frac{\Sigma|x-Me|}{n}$$

Mean deviation about mode is

$$M.M(Mo) = \frac{\Sigma|x-Mo|}{n}$$

Here Me, Mo are median and mode of ~~of~~ the observations.

Co-efficient of mean deviation:

The relative measures corresponding to the mean deviation is called co-efficient of mean deviation. The co-efficient of mean deviation can be computed by the following three formulae.

Co-efficient of mean deviation about

$$\text{mean} = \frac{M.D(\bar{x})}{\bar{x}} \times 100$$

Co-efficient of mean deviation about median

$$= \frac{M.D(Me)}{Me} \times 100$$

Co-efficient of mean deviation about

$$\text{mode} = \frac{M.D(Mo)}{Mo} \times 100$$

vvi

Example: The following data refer to number of years worked by 9 employees of a factory:

7, 4, 10, 9, 15, 12, 7, 9, 7

Compute the mean deviation (i) from mean (ii) Median, (iii) Mode,
(iv) show that the mean deviation from median is minimum. (v) Also calculate co-efficient of mean deviation from mean, median, mode and comment.

Solution:

$$\text{Mean} = \frac{7+4+10+9+15+12+7+9+7}{9} = 8.9 \text{ years.}$$

$$\text{Median} = \left(\frac{n+1}{2}\right)^{th} \text{ observation, Here } n = 9 \text{ (odd)}$$

$$= \left(\frac{9+1}{2}\right)^{th} \text{ "}$$

$$= (5)^{th} \text{ observation.}$$

Now arrange ascending order

4, 7, 7, 7, $\underline{9}$, 9, 10, 12, 15

5th observation is 9.

Hence median = 9 years.

Mode = 7 years, since 7 has highest frequency.

computation of mean deviation

No of years worked	Deviation from mean $\lvert D\rvert = \lvert x-\bar{x}\rvert$	Deviation from median $\lvert D\rvert = \lvert x-Me\rvert$	Deviation from mode $D = \lvert x-Mo\rvert$
7	$\lvert 7-8.9\rvert = 1.9$	$\lvert 7-9\rvert = 2$	$\lvert 7-7\rvert = 0$
4	4.9	5	3
10	1.1	1	3
9	0.1	0	2
15	6.1	6	8
12	3.1	3	5
7	1.9	2	0
9	.1	0	2
7	1.9	2	0
$\Sigma x = 80$	$\Sigma\lvert D\rvert = 21.1$	$\Sigma\lvert D\rvert = 21$	$\Sigma\lvert D\rvert = 23$

Mean deviation about mean $(M.D(\bar{x})) = \frac{\Sigma\lvert x_i-\bar{x}\rvert}{n} = \frac{21.1}{9} = 2.34$ years

" " " median $\{M.D(Me)\} = \frac{\Sigma\lvert x_i - Me\rvert}{n} = \frac{21}{9} = 2.33$ years

" " " Mode $\{M.D(Mo)\} = \frac{\Sigma\lvert x_i - Mo\rvert}{n} = \frac{23}{9} = 2.56$ years co-efficient of mean deviation about mean is $= \frac{MD(\bar{x})}{\bar{x}} \times 100$

$= \frac{2.34}{8.9} \times 100$

$= 26.29\%$

" " " " median $= \frac{M.D(Me)}{Me} \times 100$

$= \frac{2.33}{9} \times 100$

$= 25.89\%$

" " " " Mode $= \frac{M.D(Mo)}{Mo} \times 100$

$= \frac{2.56}{7} \times 100$

$= 36.57\%$

From these calculations it is clean that both mean deviation about median and co-efficient of mean deviation about median is minimum.

Variance and Standard deviation:

Variance: The arithmatic mean of the squares of the deviation of the observations from their arithmatic mean is known as variance.

Standard deviation: The positive square root of variance is called standard deviation.

Population variance: Suppose $x_1\, x_2 \cdots x_N$ are N observations of a population and μ is the mean of population, then population variance denoted by σ^2 is given by

$$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N}$$

and population standard deviation is the positive square root of population variance. It is defined by the formula

$$\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}}$$

Sample variance: Suppose $x_1, x_2 \cdots x_n$ are n values of a sample data and $\bar{x}$ its sample mean, then sample variance denoted by s^2 is defined by

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1}$$

The sample standard deviation denoted by s is the positive square root of the sample variance.

Hence $s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}}$

Example: 12, 15, 17, 20

Soln: sample mean, $\bar{x} = \frac{12+15+17+20}{4} = 16$

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1}$$

$$= \frac{(12-16)^2 + (15-16)^2 + \cdots + (20-16)^2}{4-1}$$

$$= \frac{16+1+1+16}{3} = 11.33$$

Sample standard deviation

$$s = \sqrt{11.33} = 3.37.$$

Sample variance for grouped data:

let $x_1 x_2 \cdots x_k$ be k values of a variable or k mid points of k classes with corresponding frequencies $f_1 f_2 \cdots f_k$ then Sample variance is defined by

$$s^2 = \frac{\sum f_i (x_i - \bar{x})^2}{n-1} \qquad n = \sum f_i$$

for convenience and simplicity m, the divisor n in place of n-1

$$\therefore s^2 = \frac{\sum f_i (x_i - \bar{x})^2}{n}$$

computating for-mula for sample variance

$$s^2 = \frac{1}{n-1}\left[\sum f_i x_i^2 - \frac{(\sum f_i x_i)^2}{n}\right]$$

$$s = \sqrt{\frac{1}{n-1}\left[\sum f_i x_i - \frac{(\sum f_i x_i)^2}{n}\right]}$$

vvi Example:

No. of worked per month	30-55	55-80	80-105	105-130	130-155	155-180	180-205
Number of workers	3	4	6	9	12	11	5

calculate variance and standard deviation of the frequency distribution.

Soln:

class interval	Midpoint x	Frequency f	$f_i x_i$	$f_i x_i^2$
30-55	42.5	3	127.5	5418.75
55-80	67.5	4	270	18225.00
80-105	92.5	6	555	51337.50
105-130	117.5	9	1057.5	124256.25
130-155	142.5	12	1710	243675
155-180	167.5	11	1842.5	308618.75
180-205	192.5	5	962.5	185281.25

$\Sigma f_i = 50$ $\Sigma f_i x_i = 6525$ $\Sigma f_i x_i^2 = 936812.5$

Sample variance:

$$s^2 = \frac{1}{n-1}\left[\Sigma f_i x_i^2 - \frac{(\Sigma f_i x_i)^2}{n}\right]$$

$$= \frac{1}{49}\left[936812.50 - \frac{(6525)^2}{50}\right]$$

$$= 1740.82$$

sample standard deviation

$$s = \sqrt{1740.82}$$

$$= 41.72 \text{ hours.}$$

CamScanner

co-efficient of variation:

~~Start~~ The corresponding relative measure of standard deviation is known as co-efficient of variation.

co-efficient of variation for population:

if μ is the mean and σ is the standard deviation of a population data set, then co-efficient of variation denoted by c.v is defined by $c.v = \frac{\sigma}{\mu} \times 100$

co-efficient of variation for a sample:

if $\bar{x}$ is the sample mean and s is the standard deviation of sample data set, the co-efficient of variation is defined by

$$c.v = \frac{s}{\bar{x}} \times 100$$

Example: 6.9.2 Page- 6-19

Examples: lives of two models of refrigerators in a recent survey were found as follows

life (No of years)	Model A	Model B
0-2	5	2
2-4	16	7
4-6	13	12
6-8	7	19
8-10	5	9
10-12	4	1

i) what is the average life of each model of these refrigerators?

ii) which of the two models shows more uniformity?

iii) A person wants to buy a new refrigerator; which one will prefer?

Soln:

class interval	mid-points x_i	Model A f	Model A $f_i x_i$	Model A $f_i x_i^2$	Model B f	Model B $f_i x_i$	Model B $f_i x_i^2$
0-2	1	5	5	5	2	2	2
2-4	3	16	48	144	7	21	63
4-6	5	13	65	325	12	60	300
6-8	7	7	49	343	19	133	931
8-10-~~12~~	9	5	35	405	9	81	729
10-12	11	4	44	484	1	11	121
		50	256	1706	50	308	2146

~~Comp~~

Computations of mean, Variance, and co-efficient of variation of lifetimes for two models are shown below.

	Model-A	Model-B
Arithmatic mean	$\bar{x}_A = \frac{256}{50} = 5.12$ years	$\bar{x}_B = \frac{304}{50} = 6.16$ years.
	$s_A^2 = \frac{1}{n_1 - 1}\left[\Sigma f_i x_i^2 - \frac{(\Sigma f_i x_i)^2}{n}\right]$	$s_B^2 = \frac{1}{49}\left[2146 - \frac{308^2}{50}\right]$
	$= \frac{1}{49}\left[1706 - \frac{256^2}{50}\right]$	$= 5.08$
	$= 8.07$	
	$s_A = \sqrt{8.07}$	$s_B = \sqrt{5.08}$
	$= 2.84$ years	$= 2.25$ years
	$C.V(A) = \frac{2.84}{5.12} \times 100$	$C.V(B) = \frac{2.25}{6.16} \times 100$
	$= 55.47\%$	$= 36.53\%$

i) Average lifetimes of refrigerator of model A is 5.12 years while of model B is 6.16 years

ii) Since $C.V(B) < C.V(A)$, hence refrigerators of model B shows greater uniformity.

iii) Due to greater uniformity in life time, the person will prefer model B.

vvi Example: Two cricketers scored the following runs in randomly selected 10 one-day matches.

Player A	42	32	40	45	17	83	59	64	76	72
Player B	95	3	28	70	31	14	82	0	59	108

i) who is the better run-getter?

ii) who " " consistent player?

iii) A prize is given to the best player. who will get the prize.

Soln: In order to find out who is better run-getter, we will compare the average runs scored and to find out who is more consistent, we will compare the co-efficient of variation.

calculation of mean and co-efficient of variation

Cricketer A: x_i	x_i^2	cricketer B: x_i	x_i^2
42	1764	95	9025
32	1024	3	9
40	1600	28	784
45	2025	70	4900
17	289	31	961
83	6889	14	196
59	3481	82	6724
64	4096	0	0
76	5776	59	3481
72	5184	108	11664
530	32127	490	37744

Cricketer A

$$\bar{x}_A = \frac{\Sigma x_i}{n} = \frac{530}{10} = 53$$

$$s_A^2 = \frac{\Sigma (x_i - \bar{x})^2}{n-1}$$

$$= \frac{1}{n-1}\left[\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}\right]$$

$$= \frac{1}{10-1}\left[32127 - \frac{530^2}{10}\right]$$

$$= 448.56$$

$$s_A = 21.18$$

$$c.v_A = \frac{s_A}{\bar{x}_A} \times 100$$

$$= \frac{21.18}{53} \times 100$$

$$= 39.96\%$$

Cricketer B

$$\bar{x}_B = \frac{\Sigma x_i}{n} = \frac{490}{10}$$

$$= 49$$

$$s_B^2 = \frac{1}{n-1}\left[\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}\right]$$

$$= \frac{1}{9}\left[37744 - \frac{490^2}{10}\right]$$

$$= 1526$$

$$s_B = 39.06$$

$$c.v(B) = \frac{s_B}{\bar{x}_B} \times 100$$

$$= \frac{39.06}{49} \times 100$$

$$= 79.71\%$$

(i) Since the average score for the player A is more than Player B. so A is a better run getter.

(ii) The co-efficient of variation of run for player A is less than player B; hence player A is more consistent.

(iii) The player A will ~~be~~ get the prize.

CamScanner