M T W T F S Date:

CHP #5

⇒ Arithematic mean:- (Average)

1) A.M for ungrouped data:-

- If the data set represent a finite population of size N. Then A.M is denoted by "μ" and calculated as: $\mu = \frac{\sum X}{N}$

- If the data set represent a finite sample of size "n", then denote and calculated as:-

$$\bar{X} = \frac{\sum x}{n}$$

understanding by example

1(a) For population size:-

The number of employees at 6 different stores = 6, 8, 7, 8, 9, 10

$$\underset{x_1}{6},\ \underset{x_2}{8},\ \underset{x_3}{7},\ \underset{x_4}{8},\ \underset{x_5}{9},\ \underset{x_6}{10}$$

N = 6

$$\mu = \frac{\sum X}{N} = \frac{6+8+7+8+9+10}{6} = \frac{48}{6}$$

$$= 8$$

Date:

Example 2(a)

The daily Pocket money recieved by five first year students = 80, 120, 110, 90, 100

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5$

n = 5

$$\bar{X} = \frac{\Sigma X}{n} = \frac{80+120+110+90+100}{5} = \frac{500}{5} = 100$$

A.M For grouped data

If the values $x_1, x_2, x_3 - - - x_n$ occur with frequencies $f_1, f_2, f_3 - - - f_n$ Then A.M is

$$\bar{X} = \frac{\Sigma fx}{\Sigma f} \quad \text{or} \quad \bar{X} = \frac{\Sigma fx}{n}$$

where $n = \Sigma f$ the total frequencies

Example 3(a)

Calculate average age of first year student from the following frequencies

Age (year)	No. of students
13	2
14	5
15	13
16	7
17	3

Here age is a variable X and no. of students is frequencies f

X	f	fX
13	2	26
14	5	70

Scanned with CamScanner

M T W T F S Date:

15	13	195
16	7	112
17	3	51
Total	$\Sigma f = 30$	$\Sigma fx = 454$

$$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{454}{30} = 15.13 \text{ years}$$

Example 3(b)

Answers:-

$\therefore \bar{x} = \frac{\Sigma fx}{\Sigma f}$ where x is the class mark or mid point and f is no of student

So:-

Class interval	x	f	fx
100-104	102	5	510
105-109	107	15	1605
110-114	112	20	2240
115-119	117	25	2925
120-124	122	30	3660
125-129	127	20	2540
130-134	132	16	2112
135-139	137	14	1918
140-144	142	5	710
Total		$\Sigma f = 150$	$\Sigma fx = 18220$

Scanned with CamScanner

Date:

$\bar{X} = \frac{\Sigma fx}{\Sigma f} = \frac{18220}{150} = 121.466.$

The weighted arithematic mean:-

The mean obtained by assigning each observation a weight that reflect its importance. such

$$\bar{X}_w = \frac{\Sigma wx}{\Sigma w}$$

Example 4(a)

compute weight average

Sol:-

W(No. of emp)	X(Monthly ear)	(XW)
4	800	3200
22	45	990
20	100	2000
30	30	900
80	35	2800
300	15	4500
$\Sigma W = 456$		$\Sigma 14390$

$\bar{X}_W = \frac{\Sigma XW}{\Sigma W} = \frac{14390}{456} = 31.55$

Example 4(b)

subjects	Marks(X)	weights(w)	WX
English	73	4	292
Urdu	82	3	246

Scanned with CamScanner

Math	80	3	240
History	67	2	134
Science	62	2	124
Total:-		$\Sigma W = 14$	$\Sigma WX = 1036$

i) compute average marks according to the data.

$$\bar{X} = \frac{\Sigma WX}{\Sigma W} = \frac{1036}{14} = 74 \text{ marks}$$

ii) If equal weights are used than arithematic mean

$$\bar{X} = \frac{\Sigma X}{n} = \frac{73+82+80+67+62}{5} = 72.8$$

Example 6:

X	f(W)	WX
38.9	21.5	836.35
41.9	18.8	787.72
42.9	23.8	1021.02
Total	64.1	$\Sigma WX = 2645.09$

Mean Price paid per litre:-

$$X = \frac{\Sigma XW}{\Sigma W} = \frac{2645.09}{64.1} = 41.26 \text{ cents/l}$$

Combined A.M:

If n_1 values have mean $\bar{X}_1$, n_2 values have mean $\bar{X}_2$, — n_k

Scanned with CamScanner

Date:

...lues have mean $\bar{X}_k$ then combined
...M is

$$\bar{X}_c = \frac{n_1 x_1 + n_2 x_2 \cdots n_k x_k}{n_1 + n_2 + n_3 \cdots n_k} = \frac{\sum_{i=1}^{k} n_i x_i}{\sum_{i=1}^{k} n_i}$$

or $$\bar{X}_c = \frac{\sum n_i x_i}{\sum n_i}$$

Example 7(a)

Sections of mathematics = 3
class have no of students = 30, 35, 40
Averaged = 80, 85, 75
combined mean $\bar{X}_c = ?$

$$X_c = \frac{\sum n_i x_i}{\sum n_i} = \frac{n_1 x_1 + n_2 x_2 + n_3 x_3}{n_1 + n_2 + n_3}$$

$$= \frac{30(80) + 35(85) + 40(75)}{30 + 35 + 40}$$

$$= \frac{2400 + 2975 + 3000}{105} = \frac{8375}{105} = 79.76$$

Median:- (ungrouped) data)

For odd terms

7, 18, 19, 25, 78

$\boxed{\text{Median} = 19}$

by formula = $\frac{(n+1)}{2}$ th term
total terms

n = 5

Median = $\frac{(5+1)th}{2} = \frac{6th}{2} = 3th$

$\boxed{\text{Median} = 19} \Rightarrow \boxed{\tilde{x} = 19}$

symbol of median $\tilde{x}$

* For even terms:-

If data is given as

18, 19, 20, 25, 26, 28

$$\text{Median} = \frac{\frac{n}{2}th + \left[\frac{n}{2}+1\right]th\ term}{2}$$

total terms $n = 6$ so

$$\text{median} = \frac{3^{th} + 4^{th}\ term}{2} = \frac{20+25}{2} = 22$$

$\tilde{x} = 22.5$

Note: "If data is not arranged then first of all arrange it in ascending orders"

For grouped data

$$\tilde{x} = l + \frac{h}{f}\left(\frac{n}{2} - c\right)$$

median class

l = lower class boundary of $\tilde{x}$ class

h = size of class interval of $\tilde{x}$ class

f = frequency of median class

n = sum of frequencies = Σf

C = commulative frequency before the median class

Scanned with CamScanner

Date:

Example # 10:-

Sol:-

Time	f	C.B	C.F
118-126	3	117.5-126.5	3
127-135	5	126.5-135.5	8
136-144	9	135.5-144.5	17
145-153	12	144.5-153.5	29
154-162	5	153.5-162.5	34
163-171	4	162.5-171.5	38
172-180	2	171.5-180.5	40

$\Sigma f = 40$

$\therefore \Sigma f = n = 80$. $n/2 = 40/2 = 20$ in C.F Coloumn 20 lies under 29 so median lies in fourth Class, which is the median class -

$l = 144.5$ = l.c boundary

$h = 9$ = size of class interval

$f = 9$ = frequency of median class

$n = 40$ = sum of frequencies

$c = 17$ = C.F before median Class

So

$\tilde{x} = 144.5 + \frac{9}{}$

Scanned with CamScanner

b) Arranged in an array:-

119, 124, 126, 128, 132, 135, 135, 135, 136, 138, 138, 140, 140, 142, 142, 144, 144, 145, 145, 146, 146, 147, 147, 148, 149, 150, 150, 152, 153, 154, 156, 157, 158, 161, 163, 164, 165, 175, 178.

$$\tilde{X} = \frac{\left(\frac{n}{2}\right)^{th} + \left(\frac{n}{2}+1\right)^{th}}{2}$$

$$= \frac{\left(\frac{40}{2}\right)th + \left(\frac{40}{2}+1\right)th}{2} = \frac{20th + 21th}{2}$$

$$\tilde{X} = \frac{146+146}{2} = 146 LB$$

Mode

* For ungrouped data:

locating the maximum frequency

Example 11(a)

First of all arranged data

1, 1, 2, 3, 4, 5, 7, 8, 8, 8, 8, 10, 11, 12, 15, 17,

$\hat{\mu} = 8$ because 8 occurs 4 times and it is maximum frequency

ii) Mode for discrete frequency distribution:-

The mode exist alongside the highest frequency

Scanned with CamScanner

Date:

Example 12(a)

Size of Shoes X	No of Pair sold (f)
$5\frac{1}{2}$	15
5	25
$6\frac{1}{2}$	50
6	30
$7\frac{1}{2}$	20
8	~~20~~ 10

As the highest frequency is 50 so corresponding value $6\frac{1}{2}$ is the model value of shoes size.

iii) Mode for grouped data:-

The mode would lie in that class which has the maximum frequency. This class is called the modal class.

$$\hat{X} = L + \frac{f_2}{f_1 + f_2} \times h$$

l = lower class boundary

f_1 = frequency of class before model class.

f_2 = frequency of class ~~after~~ following model class

h = length of class interval.

Example #13:

Scanned with CamScanner

Calculate model and median heights

Sol:

X	f	C.B	C.F
61	5	59.5-62.5	5
64	18	62.5-65.5	23
67	42	65.5-68.5	65
70	27	68.5-71.5	92
73	08	71.5-74.5	100

$l = 65.5$

$f_1 = 18$

$f_2 = 27$

$h = 3$

$$\hat{X} = l + \frac{f_2}{f_1+f_2} \times h$$

$$= 65.5 + \frac{27}{18+27} \times 3$$

$$\hat{X} = 67.3$$

Empirical Relation b/w mean, median and mode:-

Difference b/w mean and median

$$\text{Mean} - \text{Median} = \frac{1}{2}(\text{median} - \text{mode})$$

$$2(\text{mean} - \text{median}) = (\text{median} - \text{mode})$$

$$2(\bar{X} - \tilde{X}) = \tilde{X} - \hat{X}$$

Scanned with CamScanner

$\hat{X} = \tilde{X} - 2\bar{X} + 2\tilde{X}$

Date:

$\hat{X} = 3\tilde{X} - 2\bar{X}$

Geometric Mean:-

$G = \sqrt[n]{(x_1 \cdot x_2 \cdot x_3 \ldots x_n)}$ where $x > 0$

if n is too long/large then easiest way:-

$G = \text{Antilog}\left(\frac{\sum \log X}{n}\right) \rightarrow$ for ungrouped data

and $G = \text{Antilog}\left(\frac{\sum (f \log x)}{\sum f}\right) \rightarrow$ for group data

Example 15(b)

calculate geometric mean

Sol:-

X	f	log X	f log X
13	2	1.1139	2.2278
14	5	1.1461	5.7305
15	13	1.1761	15.2893
16	7	1.2041	8.4287
17	3	1.2304	3.6912
	$\sum f = 30$		$\sum f \log X = 35.3675$

$G.M = \text{Antilog}\left[\frac{\sum (f \log X)}{\sum f}\right]$

$= \text{Antilog}\left[\frac{35.3675}{30}\right] = \text{Antilog}[1.1789]$

$= 15.09$

M T W T F S

Harmonic mean:-

"Reciprocal of the mean of the reciprocal of the numbers"

H = Reciprocal of mean ($\bar{x}$) of reciprocal of values

H = Reciprocal of $\bar{x}$ of $\left[\frac{1}{x_1}, \frac{1}{x_2} \cdots \frac{1}{x_n}\right]$

$$H = \left(\frac{1/x_1 + 1/x_2 + 1/x_3 \cdots 1/x_n}{n}\right)^{-1}$$

$$H = \frac{n}{1/x_1 + 1/x^2 + 1/x^3 \cdots 1/x_n}$$

$$= \frac{n}{\Sigma\, 1/x_i} \quad \text{or} \quad \frac{1}{H} = \frac{\Sigma\, 1/x_i}{n}$$

If $x_1, x_2, x_3 \cdots x_k$ represents class marks with frequencies $f_1, f_2 \cdots f_k$ then H.M.

$$H = \frac{n}{\Sigma f_i \left[1/x_i\right]} = \frac{\Sigma f}{\Sigma f\left[\frac{1}{x}\right]}$$

Example 16(a):

As the distance travel is same for each day only the speed per kilometer is changes

Days	a	b
1st	400	70 km/hr
2nd	400	90 km/hr
3rd	400	85 km/hr

Date:

$$H.M = \frac{3}{\frac{1}{70}+\frac{1}{90}+\frac{1}{85}} = 80.73\,kmh^{-1}$$

Relationship b/w A.M, G.M and H.M

A.M > G.M > H.M

If ($x_1 = x_2 = \cdots x_n$)

then H.M = G.M = A.M

Example # 18:-

consider a sample with data values of 10, 20, 12, 17, 16. prove A.M > G.M > H.M

X	log X	1/X
10	1	0.1
20	1.3010	0.05
12	1.079	0.0833
17	1.230	0.0588
16	1.204	0.0625
$\Sigma X = 75$	$\Sigma = 5.814$	$\Sigma 1/x = 0.3546$

$$A.M = \frac{\Sigma X}{n} = \frac{75}{5} = 15$$

$$G.M = \text{Antilog}\left(\frac{\Sigma \log x}{n}\right) = \text{Antilog}(1.1628) = 14.547$$

$$H.M = \frac{5}{0.3546} = 14.100$$

A.M > G.M > H.M

Quartile:

The three points dividing an arranged series of values in four equal parts are known as quartiles and denoted by Q_1, Q_2, Q_3

Quartiles for individual observations (ungrouped data)

Q_1 - value of $\left(\frac{n+1}{4}\right)$th term of array data

Q_2 = " " $2\left(\frac{n+1}{4}\right)$th " " "

Q_3 - " " $3\left(\frac{n+1}{4}\right)$th " " "

Quartiles for a frequency distribution of discrete data:

Commulative frequency

Q_1 = value of $\left[\frac{n+1}{4}\right]$th item under C.F

Q_2 = " " $2\left(\frac{n+1}{4}\right)$th = $\frac{n+1}{2}$th " "

Q_3 = " " $3\left(\frac{n+1}{4}\right)$th " "

Percenties:-

The ninety nine points which divide the data set in 100 equal parts and denoted by $P_1, P_2 — P_{99}$

Percentiles for ungrouped data:-

P_j = value of $j\left(\frac{n+1}{100}\right)$ item of array data

Percentiles for descrete frequency distribution:

P_j = value of $j\left(\frac{n+1}{100}\right)$th under C.F

Example 19 (a):-

The weight measurement of 15 onions in grams are given
65, 68, 60, 78, 65, 74, 58, 54, 61, 57, 63, 64, 72, 66, 67

Find three quartiles D_6, P_{40} and P_{85}.

Sol:-

Array X = 54, 57, 58, 60, 61, 63, 64, 65, 65, 66, 67, 68, 72, 74, 78

$Q_1 = \left(\frac{n+1}{4}\right)$th term $= \left[\frac{15+1}{4}\right] =$ 4th term

$Q_1 = 60$

$Q_2 = 2\left(\frac{n+1}{4}\right)$th term $= 2\left(\frac{16}{4}\right) =$ 8th term

$Q_2 = 65$

$Q_3 = 3\left(\frac{16}{4}\right) =$ 12th term $= 68$

$D_6 = 6\left[\frac{n+1}{10}\right] = 6\left[\frac{16}{10}\right] = 9.6$th

= 9th term + 0.6 (10 - 9th term)

= 65 + 0.6

$\boxed{D_6 = 65.6}$

$P_{40} = 40\left[\frac{n+1}{100}\right] = 2\left[\frac{16}{5}\right] = 6.4$th term

= ~~6th + 0.4(5th - 4th)~~

= 63 + 0.4(61 - 60)

= 63.4

Measure of Dispersion:-

The degree to which numerical data tend to spread on average value is called the dispersion or variation

Types of measure of Dispersion:

i) Range (R) ii) Quartile Deviation (Q.R)
iii) Mean Deviation (M.D) iv) Standard deviation(S)

Relative measure of Dispersion:

$$\text{Relative dispersion} = \frac{\text{Absolute Dispersion}}{\text{Average}}$$

Range (R)

$$R = \underbrace{X_m}_{\text{largest observation}} - \underbrace{X_0}_{\text{smallest observation}}$$

coefficient of range:-

$$\text{Coefficient of range} = \frac{X_m - X_0}{X_m + X_0}$$

Example# 20:-

108, 110, 112, 124, 127, 118, 113

Range $R = X_m - X_0 = 127 - 108 = 19$

$$\text{Coefficient of Dispersion} = \frac{X_m - X_0}{X_m + X_0}$$

$$= \frac{19}{235} =$$

Date:

Quartile deviation and coefficient of Quartile deviation:

Quartile range = Q.R = $Q_3 - Q_1$

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

coefficient of " " = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

Example 21(a):

55, 42, 47, 56, 49, 46, 51, 60, 46 are marks of 9 students. Find median, Q.D and coefficient of Q.D

Solution:-

Data array = 42, 46, 46, 47, 49, 51, 55, 56, 60

median $\tilde{x} = \left[\frac{n+1}{2}\right]$ th term

$= \left[\frac{10}{2}\right]$ th term = 5th term

$\tilde{x} = 49$.

$Q_1 = \left[\frac{n+1}{4}\right]$ th item $= \left(\frac{10}{4}\right)$ th term

$= 2.5$th

$= 2$th term $+ 0.5(3^{rd} - 2^{nd})$

$= 46 + 0.5(46 - 46)$

$= 46$

$Q_2 = 2\left[\frac{n+1}{4}\right] = 2\left[\frac{10}{4}\right] = 5$th term

$= 49$

Scanned with CamScanner

M T W T F S

$Q_3 = 3\left[\frac{n+1}{4}\right]$th item $= 3\left[\frac{10}{4}\right]$th item

$= 7.5$th item

$= 7$th $+ 0.5(8^{th} - 7^{th})$

$= 55 + 0.5(56 - 55)$

$Q_3 = 55.5 \approx 56$

$Q.D = \frac{Q_3 - Q_1}{2} = \frac{55 - 46}{2} = 4.5$

coefficient of $Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{4.5}{101}$

$= 0.0445$

Mean deviation and coefficient of M.D.

$\rightarrow$ $M.D\ \bar{X} = \sum |X - \bar{X}|$ / n

mean deviation about mean

other names: average deviation, Absolute mean deviation

$\rightarrow$ $M.D\ \hat{X} = \sum |X - \hat{X}|$ / n

Mean deviation about mode

$\rightarrow$ $M.D\ \tilde{X} = \sum |X - \tilde{X}|$ / n

mean deviation about median

about situation $(\bar{X}, \hat{X}, \tilde{X})$

for grouped data $= \sum f\ |X - \text{Average}$

Coefficient of M.D (about mean) $= \frac{M.D\bar{X}}{\bar{X}}$

" " " (about mode) $= \frac{M.D\hat{X}}{\hat{X}}$

Date:

" " (about median) = $M.D\tilde{x}$

Example 22(a)

Find the mean deviation

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

$$\bar{X} = \frac{12+6+7+3+15+10+18+5}{8} = \frac{76}{8} = 9.5$$

now M.D about mean

$$M.D\bar{x} = \frac{\sum|x - \bar{x}|}{n}$$

X	$X-\bar{X}$	$\lvert X-\bar{x}\rvert$
12	2.5	2.5
6	-3.5	3.5
7	-2.5	2.5
3	-6.5	6.5
15	5.5	5.5
10	0.5	0.5
18	8.5	8.5
5	-4.5	4.5
	0	$\Sigma = 34$

$$M.D\bar{x} = \frac{34}{8} = 4.25$$

Example 22(b)

Marks	f	x	fx	$\lvert x-\bar{x}\rvert$	$f\lvert x-\bar{x}\rvert$
0-9	2	4.5	9	41.67	83.
10-19	3	14.5	43.5	31.67	95.
20-29	8	24.5	196.0	21.67	173.
30-39	24	34.5	828.0	11.67	280.
40-49	27	44.5	1201.5	1.67	45.
50-59	40	54.5	2180	8.33	33
60-69	11	64.5	709.5	18.33	201.
70-79	5	74.5	372.5	28.33	141.

$\Sigma f = 120$ $\quad \Sigma fx = 5540$ $\quad \Sigma f\lvert x-\bar{x}\rvert = 1353.3$

$$\bar{X} = \frac{\Sigma fx}{\Sigma f} = \frac{5540}{120} = 46.17$$

$$M.D = \frac{\Sigma f\lvert x-\bar{x}\rvert}{n} = \frac{\cancel{120}(1353.36)}{120}$$

$$= 11.278$$

Standard Deviation:-

The standard deviation is a measure of the amount of variation of data set

$$S = \sqrt{\frac{\Sigma(X-\bar{X})^2}{n}} \rightarrow \text{for sample standard deviation}$$

$$\sigma = \sqrt{\frac{\Sigma(X-\mu)^2}{N}} \rightarrow \text{for population " "}$$

Date:

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

$$\sigma = \sqrt{\frac{\sum f(x-\mu)^2}{\sum f}}$$

Example 23(a)

1) 12, 6, 7, 3, 15, 10, 18, 5

$$\bar{x} = \frac{12+6+7+3+15+10+18+5}{8} = 9.5$$

x	$x-\bar{x}$	$(x-\bar{x})^2$
12	2.5	6.25
6	-3.5	12.25
7	-2.5	6.25
3	-6.5	42.25
15	5.5	30.25
10	0.5	0.25
18	8.5	72.25
5	-4.5	20.25
	0	$\sum(x-\bar{x})^2 = 190$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} = \sqrt{\frac{190}{8}} = \sqrt{23.75}$$

$$= 4.87$$

Example 23(b)

M T W T F S Date:

Height	f	Class mark X	fx	$X-\bar{x}$	$(X-\bar{x})^2$	$f(x-\bar{x})^2$
60-62	5	61	305	-6.45	41.60	208.0[illegible]
63-65	18	64	1152	-3.45	11.90	214.2[illegible]
66-68	42	67	2814	-0.45	0.20	8.5[illegible]
69-71	27	70	1890	2.55	6.50	175.[illegible]
72-74	8	73	584	5.55	30.80	246.[illegible]
	$\Sigma f = 100$		$\Sigma f(x) = 6745$			$\Sigma f(x-\bar{x})^2 = 852.$[illegible]

$$X = \frac{\Sigma fx}{\Sigma f} = \frac{6745}{100} = 67.45 \text{ inches}$$

$$S = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{\Sigma f}} = \sqrt{\frac{852.75}{100}}$$

$$= 2.92 \text{ inches}$$