Unit-5

Measures of Central Tendency and dispersion

Introduction:

An average is a single value which is used to represent all of the values in the series. The average lies in between two Extremes. That is largest and smallest items. It is some time called as measure of "Central Tendency".

Objective of Central Tendency:

-> To determine are single value that may be used to discribe the characteristics of Entire Series.

-> To facilates comparission

-> To facilitates statistical inference

-> To help in daission making process

Characteristics of an ideal average :-

-> It should be Easy to understand.

-> It should be simple and comparable.

-> It should be based on all the items of the series

-> It should not be affected by Extreme items.

-> It should be rigidly defined by a mateamatical formula.

-> It should be capable of further all algebric treatment

-> It should have Sampling stability.

Definition of an average:

Average is an attempt to find one single figure to describe whole of the figures.

Unit-3

Measures of Central Tendency and dispersion

Introduction:

An average is a single value which is used to represent all of the values in the series. The average lies in between two extremes. That is largest and smallest items. It is some times called as measure of "Central Tendency".

Objective of Central Tendency:

-> To determine one single value that may be used to describe the characteristics of entire series

-> To facilates comparission

-> To facilitates statistical inference

-> To help in daission making process

Charcteristics of an ideal average:-

-> It should be Easy to understand.

-> It should be simple and comparable.

-> It should be based on all the items of the serie's

-> It should not be affected by extreme items.

-> It should be rigidly defined by a mateamatical foumula.

-> It should be capable of furtuer all algebric treatment

-> It should have sampling stability.

Definition of an average:

Average is an attempt to find one single figure to describe whole of the figures.

Types of Averages:

1. Arithmatic mean
   - i) Simple arithmatic mean
     - -> individual observation
     - -> discrete series
   - ii) weighted arithmatic mean
2. Median
3. Mode
4. Geometric mean
5. Harmonic mean.

1. Arithmatic mean:

The most popular and widely used measure of representing the entire data by one value is called as an average and statisticians call it as "arithmatic mean".

It is obtained by adding together all the items and by dividing this total by the number of items

Simple arithmatic mean:-

individual Observation:

$$\bar{x} = \frac{x_1 + x_2 + x_3 - - - - - + x_n}{N}$$

$$\bar{x} = \frac{\Sigma x}{N}$$

Where $\bar{x}$ = Arithmatic mean

Σx = Sum of all values of x variable $(x_1 + x_2 + x_3 - - - - x_n)$

N = Number of observations

steps: 1

Add together all the values of variable x and explain the total that is Σx

step-2 Divide the total by the number of observation by the N

The following table gives the monthly income of 10 employees in an office.

income: 14780 ; 15760 ; 26690 ; 27750 ; 24840 ; 24720 ; 19100 ; 17810 ; 27050 ; 26950

calculat A.M

Employee	income
1	14780
2	15760
3	26690
4	27750
5	24840
6	24720
7	19100
8	17810
9	27050
10	26950
N = 10	Σx = 2,25,650

$$\bar{x} = \frac{\Sigma x}{N}$$

$$\bar{x} = \frac{2,25,650}{10}$$

$$\bar{x} = 2,25,650$$

Hence, The average of 10 employee is ∴ 2,25,650-----

Calculate mean from the following data:

Income pA (in lacks): 10, 20, 30, 40, 50, 60, 70, 80

Employee	income data
1	10
2	20
3	30
4	40
5	50
6	60
7	70
8	80
N = 8	$\Sigma x = 360$

$$\bar{x} = \frac{\Sigma x}{N}$$

$$\bar{x} = \frac{360}{8}$$

$$\bar{x} = 45$$

Hence the average income is

$\therefore$ 45,00,000/-

Shortcut method:

$$\bar{x} = A + \frac{\Sigma d}{N}$$

where A = Assumed mean

d = Deviations of assumed means

Deviations of items & assumed mean

$$d = (x - A)$$

Steps:

Take an assumed mean

Step 2 Take the deviations of items from the assumed mean

Step 3 Obtain the sum of these deviations

Step 4 apply the formula.

Calculate the Arithmatic mean by shortcut method

Employee	Income	$d = x - A$
1	14780	14780 - 24920 = -10.140
2	15760	15760 - 24920 = -9160
3	26,690	26690 - 24920 = 1770
4	27,750	27,750 - 24920 = 2830
5	24840	24840 - 24920 = -80
6	24920 M	24920 - 24920 = 0
7	19,100	19100 - 24920 = -5820
8	17,810	17810 - 24920 = -7110
9	27,050	27050 - 24920 = 2130
10	26,950	26950 - 24920 = 2030
N = 10	= 2,25,650	Σd = 23550

8760

32310

-23550

$\bar{x} = A + \frac{\Sigma d}{N}$

$= 24920 + \frac{23550}{10}$

$= 24920 + 2355$

$= 22.565$ where Arthmaticmean

From the following data calculate A.M by direct and and shortcut method.

Roll	1	2	3	4	5	6
mark	5	15	25	35	45	55

Roll	Marks
1	5
2	15
3	25
4	35
5	45
6	55
N = 6	$\Sigma x = 180$

$$\bar{x} = \frac{\Sigma x}{N}$$

$$\bar{x} = \frac{180}{6}$$

$$\bar{x} = 30/-$$

∴ Where the Arthamatic Mean

Roll	Marks	$d = x - A$
1	5	-20
2	15	-10
3	25 = A	0
4	35	10
5	45	20
6	55	30
N = 6		$\Sigma d = 30$

$$\bar{x} = A + \frac{\Sigma d}{N}$$

$$\bar{x} = 25 + \frac{30}{6}$$

$$\bar{x} = 25 + 5$$

$$\bar{x} = 30$$

∴ AM = 30

Calculation of A.M in discrete series

Direct method:

$$\bar{x} = \frac{\Sigma fx}{N}$$

f = frequency

x = Variable

N = Total number of observations

Step 1 :- Multiply the frequency of each row with a variable and obtain the total Σfx

Step 2 divide the total obtained by step 1 by the number of observations.

Calculate A.M for the following data

Marks :	20	30	40	50	60	70
No. of students	8	12	20	10	6	4

Marks	No. of stu	Σfx
20	8	20×8 = 160
30	12	30×12 = 360
40	20	800
50	10	500
60	6	360
70	4	280
	N = 60	Σfx = 2460

$$\bar{x} = \frac{\Sigma fx}{N}$$

$$\bar{x} = \frac{2460}{60}$$

$$\boxed{\bar{x} = 41}$$

Short cut method :

$$\bar{x} = A + \frac{\Sigma fd}{N}$$

Where A = Assumed mean

f = frequency

d = diviation of item from the assumed mean

N = Total number of observations.

Step:1 Take an assumed mean

Step 2 Take the deviations of the variable x from the assumed

Step 3 Multiply these deviations with the respective frequency and take the total Σfd

Calculate AM by Short cut method.

Mark (x)	No. of student (f)	d = x - A	Σfd
20	8	-20	8 × 20 = -160
30	12	-10	-120
40 (A)	20	0	0
50	10	10	100
60	6	20	120
70	4	30	120
	N = 60	$\Sigma d = 30$	$\Sigma fd = 60$

$$\bar{x} = A + \frac{\Sigma fd}{N}$$

$$\bar{x} = 40 + \frac{60}{60}$$

$$\bar{x} = 40 + 1$$

$$\boxed{\bar{x} = 41}$$

Calculate AM for discrete series in direct and shortcut method.

Marks	5	15	25	35	45	55
No. of students	10	20	30	50	40	30

direct method:

$$\bar{x} = \frac{\Sigma fx}{N}$$

$$\bar{x} = \frac{6300}{180}$$

$$\bar{x} = 35$$

shortcut method:

$$\bar{x} = A + \frac{\Sigma fd}{N}$$

$$\bar{x} = 25 + \frac{1800}{180}$$

$$\bar{x} = 25 + 10$$

$$\bar{x} = 35$$

(x) Marks	(f) No. of Students	fx	d = x - A	fd
5	10	50	-20	-200
15	20	300	-10	-200
25 (A)	30	750	0	0
35	50	1750	10	500
45	40	1800	20	800
55	30	1650	30	900
	N = 180	$\Sigma fx = 6300$	d = 30	$\Sigma fd = 1800$

Calculation of Arthmatic mean in continuous method

$$\bar{x} = \frac{\Sigma fm}{N}$$

where f = frequency

m = mid point various classes

N = Total number of observations

Step 1:- Obtain midpoint of each class (m)

Step 2: Multiply these mid points by the respective frequency of each class and obtain the total Σfm

Step 3: Divide the total obtained in step 2 by the sum of frequency

$$\text{midpoint (m)} = \frac{\text{upper limit} + \text{lower limit}}{2}$$

From the following data compute Arithmetic mean by direct method.

Marks	0-10	10-20	20-30	30-40	40-50	50-60
No. of students	5	10	25	30	20	10

Marks	No. of students	Midpoint	fm
0-10	5	5	25
10-20	10	15	150
20-30	25	25	625
30-40	30	35	1050
40-50	20	45	900
50-60	10	55	550
	N = 100		Σfm = 3300

$\bar{x} = \Sigma fm / N$

$\bar{x} = 3300 / 100$

$\bar{x} = 33$

Shortcut method:

$$\bar{x} = A + \frac{\Sigma fd}{N}$$

A: Assumed mean

f: frequency

d: deviation of mid points from assumed mean

$d = M - A$

N: Total Number of observations

Step 1: Take an assumed mean

Step 2: Take the deviation of mid points from assumed mean

Step 3: multiply the respective frequency of each class by the deviations and obtain the total Σfd

Apply the formula

Marks	No. of stu (f)	Midpoint $\frac{u+l}{2}$	$d = m - A$	fd
0-10	5	5	-30	-150
10-20	10	15	-20	-200
20-30	25	25	-10	-250
30-40	30	35 (A)	0	0
40-50	20	45	10	200
50-60	10	55	20	200
	N = 100			$\sum fd = -200$

600
400
-200

$$\bar{x} = A + \frac{\sum fd}{N}$$

$$35 + \frac{-200}{100} = 35 + (-2) = 35 - 2 = 33$$

$$\therefore \bar{x} = 33$$

In order to specify calculations divide the deations by class interval's that is calculate $\frac{m-A}{i}$ and then multiply by i. the formula is as follows

$$\bar{x} = A + \frac{\sum fd}{N} \times i$$

A = Assumed mean

f = frequency

d = deviation of midpoints from assumed mean

$\frac{m-A}{i}$ -0.2 -2

N = Total no. of observations

i = class interval

Marks	No. of stu	Midpoint $\frac{u+l}{2}$	$d = \frac{m.A}{i}$	Σfd
0-10	5	5	$\frac{5-35}{10} = \frac{-30}{10} = -3$	-15
10-20	10	15	$\frac{15-35}{10} = -2$	-10
20-30	25	25	-1	-25
30-40	30	35 (A)	0	0
40-50	20	45	1	20
50-60	10	55	2	20
	N = 100			$\Sigma fd = -20$

$\bar{x} = A + \frac{\Sigma fd}{N} \times i$

$= 35 + \frac{-20}{100} \times 10$

$\bar{x} = 35 + (-0.2) \times 10$

$\bar{x} = 35 - 2$

$\boxed{\bar{x} = 33}$

Calculate AM by direct and shortcut method

marks less than	10	20	30	40	50	60
No of stu	10	30	60	110	150	180

Marks less than	No. of stu	Midpoint $\frac{u+l}{2}$ d	fd = mA
0-10	10	5	50
10-20	30	15	450
20-30	60	25	1500
30-40	110	35	3850
40-50	150	45	6750
50-60	180	55	9900
	N = 540		$\Sigma fd = 22500$

$\bar{x} = \frac{\Sigma fM}{N}$

$\bar{x} = \frac{22500}{540}$

$\boxed{\bar{x} = 41.66}$

graph.

Q:-

i = class intervals

Marks	No. of Stu	Mid point	$d = \frac{x_i - A}{i}$	fd.
0-10	10	5	$\frac{5-25}{10} = -2$	-20
10-20	30	15	$\frac{15-25}{10} = -1$	-30
20-30	60	25 (A)	0	0
30-40	110	35	1	110
40-50	150	45	2	300
50-60	180	55	3	540
	N = 540			fd = 900

$$\bar{x} = A + \frac{\Sigma fd}{N} \times i$$

$$\bar{x} = 25 + \frac{900}{540} \times 10$$

$$\bar{x} = 25 + 1.66 \times 10$$

$$\bar{x} = 25 + 16.6$$

$$\bar{x} = 41.66$$

Calculate AM for the following Continuous Series

Marks	0-10	10-30	30-60	60-100
No of sto	5	12	25	08

$\bar{X} = A + \frac{\Sigma fd}{N} \times C$ C: common factor

Marks	(f) No of sto	Midpoint $\frac{U+l}{2}$	$d = \frac{m-A}{c}$	fd.
0-10	5	5 (C)	$\frac{5-45}{5} = -8$	-40
10-30	12	20	$\frac{20-45}{5} = -5$	-60
30-60	25	45 (A)	$\frac{45-45}{5} = 0$	0
60-100	08	80	$\frac{80-45}{5} = 7$	56
	N = 50			$\Sigma fd = -44$

$\bar{X} = A + \frac{\Sigma fd}{N} \times C$

$\bar{X} = 45 + \frac{-44}{50} \times 5$

$\bar{X} = 45 + (-4.4)$

$\bar{X} = 40.6$

Correcting - Incorrect Items:

The mean marks of 100 student's was found to be 40 later on it was found that the score of 53 was misread as 83. find the correct mean corresponding to the correct score.

$(\bar{x})$ Mean = 40

N = 100

$\bar{x} = \frac{\Sigma x}{N}$

$\bar{x} = \frac{\Sigma x}{N} = 40 = \frac{\Sigma x}{100}$

$= 40 \times 100$

$\Sigma x = 4000/-$

It is incorrect.

Coorrect Σx = Incoorrect Σx - woraig Num + Coorrect Num.

= 4000 - 8.3 + 53

Σx = 4000 - 30

Σx = 3970//

$\bar{x} = \frac{\Sigma x}{N}$

$= \frac{3970}{100} = 39.7$//.

mean of 100 observations is found to be 40, if at the time of computation Two items were woraigly taken as 30 and 27 instead of 3 and 72 find the coorrect mean.

$\bar{x}$ = 40

N = 100

$\bar{x} = \frac{\Sigma x}{N}$

$40 \times \frac{\Sigma x}{100}$

= 4000//

① Coorrect Σx = Incoorrect - woraig num + coorrectn

4000 - 30 - 27 + 3 + 72

= 4000 - 30 + 27 + 75

= 4000 - 57 + 75

= 4000 + 18

= 4018//

$\bar{x} = \frac{\Sigma x}{N}$

$= \frac{4018}{100}$

= 40.18//

where it is incoorrect.

Merits: Advantage:

-> It is Easy to understand and easy to calculate.

2. It is based on all the observations.

3. It is rigiddy displain define

4. It is the sender of gravity that balances tere values on a Either side of the meter.

5. It is relatively reliable

Demerits:

1. It can't be determined by inspiration

2. It can't be located practically

3. It can't use qualitative charectaristic like honesty, unity

and intelligence.

5. It is affected very much by Extreme values

6. It can't be optained by a single observation.

Weighted arthamatic mean

$$\bar{x} = \frac{\Sigma xw}{\Sigma w}$$

if a candidate optains the following marks English-75, stats-60, maths-59, physics-55, Chemistry-63, find the weighted mean if weights 2, 1, 3, 3, and 1 respectively are alloted to the subjects.

Subject	Marks (x)	weights (w)	xw
English	75	2	150
stats	60	1	60
maths	59	3	177
Physics	55	3	165
Chemistry	63	1	63
		$\Sigma w = 10$	$\Sigma xw = 615$

$\bar{x} = \frac{\Sigma xw}{\Sigma w}$

$\bar{x} = \frac{615}{10}$

$\bar{x} = 61.5$

Calculate simple arthemetic mean and weighted arthamatic mean foorthe following data.

Course	(x) pass percentage	w No. of stu	xw
MCA	71	3	213
MBA	83	4	332
Mcom	73	5	365
BA	74	2	148
Bcom	65	3	195
BSc	66	3	198
N = 6	432	20	$\Sigma xw = 1451$

$$\bar{x} = \frac{\Sigma x}{N}$$

$$= \frac{432}{6}$$

$$= 72$$

$$\bar{x} = \frac{\Sigma xw}{\Sigma w}$$

$$= \frac{1451}{20}$$

$$= 72.55$$

from the following data calculate simple Arithmatic and weighted arithmatic mean for the colleges A, B, C and find out the performence of best college

College name	A		B		C	
	No. %	No of Students	Pass %	No of Stud	Pass %	No of Stud
MA	71	3	82	2	81	2
Mcom	83	4	76	3	76	3.5
BA	73	5	73	6	74	4.5
B.com	74	2	76	7	58	2
Bsc	65	3	65	3	70	7
Msc	66	3	60	7	73	2

	A			B			C		
Cource	No.of.Pl. (x)	No.of.stu (w)	ExW	No.of.Pl. (x)	No.of.stu (w)	ExW	No.of.Pl. (x)	No.of.stu (w)	ExW
MA	71	3	213	82	2	164	81	2	162
M.com	83	4	332	76	3	228	76	3.5	266
BA	73	5	365	73	6	438	74	4.5	333
B.com	74	2	148	76	7	532	58	2	116
BSc	65	3	195	65	3	195	70	7	490
MSE	66	3	198	60	7	420	73	2	146
N = 6	x = 432	w = 20	ExW = 1451	x = 432	w = 28	xw = 1977	x = 432	w = 21	xw = 1513

$\bar{x} = \frac{\Sigma x}{N}$

$\bar{x} = \frac{432}{6}$

$\bar{x} = 72$

$\bar{x} = \frac{\Sigma xw}{\Sigma w}$

$\bar{x} = \frac{1451}{20}$

$\bar{x} = 72.5$

$\bar{x} = \frac{\Sigma x}{N}$

$\bar{x} = \frac{432}{6}$

$\bar{x} = 72$

$\bar{x} = \frac{\Sigma xw}{\Sigma w}$

$\bar{x} = \frac{1977}{28}$

$\bar{x} = 70.6$

$\bar{x} = \frac{\Sigma x}{N}$

$\bar{x} = \frac{432}{6}$

$\bar{x} = 72$

$\bar{x} = \frac{\Sigma xw}{\Sigma w}$

$\bar{x} = \frac{1513}{21}$

$\bar{x} = 72.04$

Performence of Best College is A

The weighted A.M for college A is higher than college B & C ∴ college A is considered as best college.

Median

Median is a value of the variable which divides the data into two equal parts. it refers to the middle value of a distribution

calculation of Median Individual series;

step 1: Arrange the size of in ~~can~~ accending order and decending order. Apply $\left(\frac{N+1}{2}\right)^{th}$ item calculate median.

from the following data of wages of 7 workers calculate the median wage.

wage in ₹: 1100, 1170, 1080, 1120, 1200, 1160, 1400

S.NO	Accending order
1	1080
2	1100
3	1120
4	1150
5	1160
6	1200
7	1400

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

$M = \frac{7+1}{2} = 8/2$

$M = 4$

$\therefore$ Median = 4

size of 4th item is ₹ 1150

$\therefore$ Median wage = 1150

Obtain the value of median from the following data.

391, 384, 591, 407, 672, 522, 777, 753, 2488, 1490

S.NO	Accending order	S.NO	Accending order
1	384	6	672
2	391	7	753
3	407	8	777
4	522	9	1490
5	591	10	2488

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

$M = \frac{10+1}{2} = 5.5$ item

size of 5.5 5^{th} & $5+6$ item.

$= \frac{591+672}{2}, \frac{1263}{2}$

$\therefore$ Median = 631.5

Calculation of Median discrete series

Step 1: Arrange the size of items in Assending order (or) descending order

2: Calculate Cumulative frequency

3: Apply $\frac{n+1}{2}$ formula

4: find out the cumulative frequency which includes $\left(\frac{n+1}{2}\right)^{th}$ item calculate median.

Median :- Size of item corresponding to the cumulative frequency which includes $\left(\frac{n+1}{2}\right)^{th}$ item.

Marks	45	55	25	35	5	15
No. of stu	40	30	30	50	10	20

Marks	No. of students (f)	C.f
5	10	10
15	20	30
25	30	60
35	30	90
45	40	(130) M
55	50	180.
	N = 180	

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

$\frac{180+1}{2} = \frac{181}{2} = 90.5$

$\therefore$ 90.5 item = 130

Median 130 of c.f lies in the marks of 45

Form the following data find out the value of median.

Income	1000	1500	800	2000	2500	1800
No. of persons	24	26	16	20	06	30

Income	No of persons	c.f
800	6	6
1000	16	22
1500	20	42
1800	24	66
2000	26	92
2500	30	122
	Ef = 122	
	N =	

c.f

Median: $\left(\frac{N+1}{2}\right)^{th}$ item

$= \frac{122+1}{2} = \frac{133}{2}$

$= 61.5$

M.: 61.5 item : x = 66

Median 660f c.f in the income = 1800

Calculation of Median Continuous Series

Step 1: Calculate cumulative frequency

2: Ascertain cumulative frequency which includes $\left(\frac{N}{2}\right)^{th}$ item. the corresponding and lower limit(L) of the class and (I) of the class interval of the corresponding class and cumulative frequency of the preceding ~~class~~ class

3: Calculate median

$$\text{Median} = L + \frac{\frac{N}{2} - cf}{f} \times i$$

$$L + \frac{\frac{N}{2} \cdot cf}{f} \times i$$

L = lower class

cf = cumulative frequency

f = frequency of the class

i = class interval.

N = Ef = Total number of frequency.

-> class intervals only arranging in assending order.

Calculate Median

Marks	0-10	10-20	20-30	30-40	40-50	50-60
No. of Students	10	20	30	50	40	30

Marks	No. of. Students	C.f
0-10	10	10
10-20	20	30
20-30	30	60 cf
30-40	50 f	110
40-50	40	150
50-60	30	180
	N = 180	

$\left(\frac{n}{2}\right)^{th} = \frac{180}{2} = 90$

$M = L + \frac{\frac{N}{2} - cf}{f} \times i$

$M = 30 + \frac{90 - 60}{50} \times 10$

$M = 30 + \frac{30}{50} \times 10$

$M = 30 + 0.6 \times 10$

$M = 30 + 6$

$\therefore$ Median = 36

Calculate Median for the following frequency distribution.

Freq. Marks	45-50	40-45	35-40	30-35	25-30	20-25	15-20	10-15	5-10
No. of stud	10	15	26	30	42	31	24	15	07

$\left(\frac{n}{2}\right)^{th} = \frac{206}{2} = 100$

$M = L + \frac{\frac{n}{2} - cf}{f} \times i$

Marks	No. of. stu	C.f
5-10	7	7
10-15	15	22
15-20	24	46
20-25	31	77 Cf
25-30	42 f	119
30-35	30	149
35-40	26	175
40-45	15	190
45-50	10	200

$M = 25 + \frac{100 - 77}{42} \times 5$

$M = 25 + \frac{23}{42} \times 5$

$M = 25 + 0.5 \times 5$

$M = 25 + 2.5$

$M = 27.5$

Merits of Median

1. It is rigidly defined.
2. It is easy to understand and easy to calculate.
3. It is not at all affected by extreme values.
4. It can be calculated for distributions with open ended class
5. Median is the only average to be used while dealing with qualitative data.
6. It can be determined graphically.

Demerits of median

1. In case of even number of observations median cannot be determined exactly.
2. It is not based on all the observations.
3. It is not capable of further mathematic treatment.

mode

It is the value in a series of observations which occures with greatest frequency

Individual series

find mode for the following data.

3, 5, 8, 5, 4, 5, 9, 3 (unimodel)

5 is repeated more time

∴ mode = 5 (unimode)

find mode for the following data

2, 2, 4, 6, 8, 10, 12, 12, 14, 16.

∴ mode = 2 and 12

(Bimode)

Descreat series

find the mode from the following frequency distribution

size	1	2	3	4	5	6	7	8	9	10	11	12
frequency	3	8	15	23	35	40	32	28	20	45	14	06

steps for grouping table:.

step 1: The frequency of coloumn 1 are the orginal frequency

step 2: Coloumn is uptained by combining the frequency two by two

step 3: in coloumn 3 leave the first frequency and combine the remaining frequency two by two.

step 4: in coloumn 4 combined the frequency three by three.

step 5: in coloumn leave the first frequency and combine remaining frequency three by three