STATISTICS

UNIT II –
MEASURES OF CENTRAL TENDENCY - MEAN

DR.S.ARULJOTHISELVI

ASSISTANT PROFESSOR

DEPARTMENT OF ZOOLOGY

03.09.2020

- **Mean** is an average of the given numbers: a calculated central value of a set of numbers.
- In simple words, it is the average. It's also the meanest because it takes the most math to figure it out.
- Measures of central tendency, or averages, are used in a variety of contexts and form the basis of statistics.
- The formula to calculate the mean is ratio of the total of observations to number of observations.

Mean is an essential concept in mathematics and statistics. The mean is the average or the most common value in a collection of numbers.

In statistics, it is a measure of central tendency of a probability distribution along median and mode. It is also referred to as an expected value.

Objectives/Functions of Averages

Averages occupy a prime place in the theory of statistical methods. That is why Bowley remarked, "Statistics is a science of averages." The following are the m ' objectives of an average:

1. Facilitates Comparison:

The foremost purpose of average is that it facilitates comparison. For instance, a comparison of the production of jute in Maharashtra a Punjab shows that production of jute in Maharashtra is much more as compared Punjab.

2. Formulation of Policies:

Averages are of great use in the formulation of various policy measures. For instance, when the Govt, finds that there is a fear of low product of sugar, it can formulate various policies to compensate the same.

3. Short Description:

Averages help to present the raw data in a brief a systematic manner.

4. Representation of Universe:

Average represents universe. According conclusions can be drawn in respect of the universe as a whole.

Arithmetic Mean in the most common and easily understood measure of central tendency. We can define mean as the value obtained by dividing the sum of measurements with the number of measurements contained in the data set and is denoted by the symbol $\bar{x}$.

to compute the **Arithmetic Mean** for three types of series:

- Individual Data Series
- Discrete Data Series
- Continuous Data Series

Arithmetic mean is the total of the sum of all values in a collection of numbers divided by the number of numbers in a collection. It is calculated in the following way:

$$\text{Arithmetic mean} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Individual Data Series

When data is given on individual basis. Following is an example of individual series:

Items	5	10	20	30	40	50	60	70

Discrete Data Series

When data is given alongwith their frequencies. Following is an example of discrete series:

Items	5	10	20	30	40	50	60	70
Frequency	2	5	1	3	12	0	5	7

Continuous Data Series

When data is given based on ranges alongwith their frequencies. Following is an example of continous series:

Items	0-5	5-10	10-20	20-30	30-40
Frequency	2	5	1	3	12

(i) Direct Method,

(ii) Short Cut Method,

(i) Direct Method:

Example 1. Find Mean for the following figures.

30	41	47	54	23	34	37	51	53	47

$\sum X = 30 + 41 + 47 + 54 + 23 + 34 + 37 + 51 + 53 + 47 = 417$; $N = 10$.

$$\therefore \quad \bar{X} = \frac{\Sigma X}{N} = \frac{417}{10} = 41.7.$$

(ii) Short Cut Method: *Assumed Mean Method*

Let, A = assumed mean
X = individual observations
N = total numbers of observations
d = deviation of assumed mean from individual observation, i.e. $d = X - A$

Then sum of all deviations is taken as $\Sigma d = \Sigma (X - A)$

Then find $\frac{\Sigma d}{N}$

Then add A and $\frac{\Sigma d}{N}$ to get $\overline{X}$

Therefore, $\overline{X} = A + \frac{\Sigma d}{N}$

You should remember that any value, whether existing in the data or not, can be taken as assumed mean. However, in order to simplify the calculation, centrally located value in the data can be selected as assumed mean.

மதிப்பெண்கள்	d=X-A (A=60)
53	-7
65	+5
70	+10
48	-12
55	-5
72	+12
65	+5
52	-8
63	+3
58	-2
61	+1
70	+10
	Σd=+12

$$\overline{X} = A + \frac{\Sigma d}{n}$$

$$= 60 + \left(\frac{12}{12}\right)$$

$$= 60 + 1$$

= 61 (கி.கிராம்)

DISCRETE DATA SERIES

Discrete series means where frequencies of a variable are given but the variable is without class intervals.

(i) Direct Method:

Here each frequency is multiplied by the variable, taking the total and dividing total by total number of frequencies, we get X.

Symbolically,

$X = \sum fx/N$

Where f = frequency,

X = the value of the variable

And N = the sum of frequency or $N = \sum f$

Example 1. Calculate A.M. from the following data

Marks obtained :	4	8	12	16	20
No. of students :	6	12	18	15	9

Solution:

Marks X	No. of students f	fX
4	6	24
8	12	96
12	18	216
16	15	240
20	9	180
	N = 60	$\Sigma fX = 756$

$$\text{As} \quad \bar{X} = \frac{\Sigma fX}{N}$$

$$\therefore \quad \bar{X} = \frac{756}{60} = 12.6$$

(ii) Short Cut Method:

Here Assumed Mean is taken and taking deviations of variable from it. We obtain X by using the following formula.

$$\overline{X} = A + \frac{\sum fd}{n} \quad (n = \sum f)$$

Where A = Assumed Mean

d. = (X-A);

f = frequency $\sum f$ or N = Total number of terms,

(Note :-This formula is often used when the variables are large in size or infractions and direct formula is not easy to use.)

(ii) Short Cut Method: *Assumed Mean Method*

Example : Calculate the Arithmentic Mean using short-cut method:

(அ) சுருக்கு முறை

கூற்றுச் சூத்திரம்

$$\overline{X} = A + \frac{\sum fd}{n}$$

மாதிரி இரண்டினை மேற்காணும் சூத்திரத்தைப் பயன்படுத்தி தீர்வு செய்வோம். ஊக சராசரி, A = 45 எனக் கொள்வோம்.

மதிப்பெண் X	மாணவர்கள் (f)	A = 45 (d = X - A)	fd
20	5	-25	-125
30	3	-15	-45
35	6	-10	-60
45	8	0	0
50	12	+5	+60
55	7	+10	+70
60	5	+15	+75
70	4	+25	+100
மொத்தம்	50		+75

$$\overline{X} = A + \frac{\sum fd}{n} \quad (n = \sum f)$$

$$= 45 + \frac{75}{50}$$

$$= 45 + 1.5$$

$$= 46.5$$

கூட்டுச்சராசரி = 46.5

Arithmetic Mean of Continuous Data Series

Formula

$$\bar{x} = \frac{f_1 m_1 + f_2 m_2 + f_3 m_3 ..... + f_n m_n}{N}$$

Where –

- N = Number of observations.
- $f_1, f_2, f_3, \ldots, f_n$ = Different values of frequency f.
- $m_1, m_2, m_3, \ldots, m_n$ = Different values of mid points for ranges.

1. நேரடி முறை:

ஆக்கிரமம்: கூட்டுச்சராசரி $X = \frac{\sum fm}{n}$

'm' என்பது பிரிவு இடைவெளியின் நடுமதிப்பைக் குறிக்கும்.

மதிப்பெண் X	மாணவர்களின் எண்ணிக்கை (f)	X-ன் நடுமதிப்பு mid X (M)	fm
20-30	5	25	125
30-40	8	35	280
40-50	12	45	540
50-60	15	55	825
60-70	6	65	390
70-80	4	75	300
மொத்தம்	$\sum f$=50		$\sum$fm=2460

$$\overline{X} = \frac{\sum fm}{n} = \frac{2460}{50} = 49.2$$

Example

Problem Statement:

Let's calculate Arithmetic Mean for the following continous data:

Items	0-10	10-20	20-30	30-40
Frequency	2	5	1	3

Solution:

Based on the given data, we have:

Items	**Mid-pt m**	**Frequency f**	fm
0-10	5	2	10
10-20	15	5	75
20-30	25	1	25
30-40	35	3	105
		$N = 11$	$\sum fm = 215$

Based on the above mentioned formula, Arithmetic Mean $\bar{x}$ will be:

$$\bar{x} = \frac{215}{11}$$
$$= 19.54$$

The Arithmetic Mean of the given numbers is 19.54.

(ii) Short Cut Method: *Assumed Mean Method*

Continuous Series

$$\overline{X} = A + \frac{\sum fd}{\sum f}$$

Where,

n = Total number of observations.

$\sum$ d = Total deviation value

d = Deviation of item from the assumed mean.

A = Assumed mean.

$\sum$ fd = Sum of products of deviations and their corresponding frequencies.

Calculate the average marks obtained by BCA students in Mathematics paper by short cut method

Class of Marks	0-10	10-20	20-30	30-40	40-50
No. of students	5	3	7	25	20

Solution:

Let A = 25

Marks of Students	No. of Students(f)	Mid-values(m)	d = m − 25	fd
0-10	5	5	-20	-100
10-20	3	15	-10	-30
20-30	7	25	0	00
30-40	25	35	10	250
40-50	20	45	20	400

$$X = A + \frac{\sum fd}{\sum f}$$

$$= 25 + \frac{520}{60} = 25 + 8.67 = 33.67$$

Advantages

1. Arithmetic mean is simple to understand and easy to calculate.
2. It is rigidly defined.
3. It is suitable for further algebraic treatment.
4. It is least affected fluctuation of sampling.
5. It takes into account all the values in the series.

Disadvantages

1. It is highly affected by the presence of a few abnormally high or abnormally low scores.
2. In absence of a single item, its value becomes inaccurate.
3. It can not be determined by inspection.