Department of Commerce (CA)

COURSE : I M.Com (CA)

SEMESTER : II

SUBJECT : BUSINESS RESEARCH METHODS

SUBJECT CODE : 18MCC22C

PREPARED BY : DR.S.KALAVATHI

PHONE NO : 9842579433

SYLLABUS

UNIT-IV

Measures of central tendency- Standard Deviation – Correlation- Simple, Partial and Multiple correlation-regression models –OLS [Cordinary Least Square] Methods.

UNIT-IV

AVERAGE DEFINED-

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

- Clark

OBJECTIVES OF AVERAGING:-

There are two objectives of the study of averages:

(i) To Get Single Value that describes the characteristics of the entire group.

(ii) To Facilitate Comparison Measures of central value.

REQUISITES OF A GOOD AVERAGE:-

(i) Easy to Understand

(ii) Simple to Compute

(iii) Based on all the items

(iv) Not be Unduly Affected by Extreme Observations

(v) Rigidly Defined

(vi) Capable of Further Algebraic

(vii) Sampling Stability

MEASURES OF CENTRAL TENDENCY:-

TYPES OF AVERAGE:-

The following are the important types os averages:

(i) Arithmetic Mean

(ii) Median

(iii) Mode

(iv) Geometric Mean

(v) Harmonic mean

ARITHMETIC MEAN:

(a) Simple arithmetic Mean

(b) Weighted Arithmetic Mean

Simple Arithmetic Mean

Calculation of Simple Arithmetic Mean – Individual Observations:

The process of computing mean in case of individual observations.

Formula for computing mean is,

$$\overline{X} = \frac{\sum X}{N}$$

(i) $\sum X$= is add together all the values of the variable X and obtain the total.

(ii) N = Divide this total by the number of the observations.

Sum No 1

The following table gives the monthly income of 10 employees in the office:
Income (Rs.): 1,780 1,760 1,690 1,750 1,840 1,920 1,100 1,810 1,050 1,950

Calculation of Simple Arithmetic Mean – Discrete Series:

Direct Method

Formula for computing mean is,

$$\bar{X} = \frac{\sum fX}{N}$$

(i) Where f = Frequency; X = the variable in question; $\sum f$ (or) N* = Total number of observations.

(ii) $\sum fX$ - Multiply the frequency of each row with the variable and obtain the total.

Short-cut Method

Formula for computing mean is,

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

Where A = Assumed mean; d = (X-A); N = Total number of observations.

Sum No 2

From the following data of the marks obtained by 60 students of a class, calculate the arithmetic mean:

MARKS	20	30	40	50	60	70
NO. OF STIDENTS	8	12	20	10	6	4

Calculation of Simple Arithmetic Mean – Continuous Series:

Formula for computing mean is,

$$\bar{X} = \frac{\sum fm}{N}$$

Where m = mid-point of various classes f = The frequency of each class; N* = Total number of frequency.

(i) Obtain the mid-point of each class and denote it by **m.**

(ii) $\sum fX$ - Multiply these mid-points by the respective frequency of each class and obtain the total.

(iii) Divide the total obtained in step (i) by the sum of frequency. by N

Short-cut Method

Formula for computing mean is,

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

Where A = Assumed mean; d = deviation of mid points from assumed mean, (X-A); N = Total number of observations.

Sum No 3

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 STIDENTS	5	10	25	30	20	10

Calculation of Simple Arithmetic Mean – Combined Mean:

Formula for computing mean is,

$$\bar{X}_{12} = \frac{N_1 \bar{X}_1 + N_2 \bar{X}_2}{N_1 + N_2}$$

Sum No 4

The mean height of 25 male workers in a factory is 61 inches and the mean height of 35 female workers in the same factory is 58 inches. Find the combined mean height of 60 workers in the factory.

Weighted Arithmetic Mean:

$$\overline{X}_w = \frac{\sum WX}{\sum W}$$

Sum No 5

The train runs 25 miles at a speed of 30 mph., another 50 miles at a speed of 40 mph., then due to repairs of the track travels for 6 minutes at speed of 10 mph and finally covers the remaining distance of 24 miles at a speed of 24 mph. what is the average speed in miles per hour?

MEDIAN

The median by definition refers to the middle value in a distribution. The median is what is called a positional average.

Calculation of Median– Individual Observations:

Formula for computing median is,

$$\textbf{Med* = Size of } \frac{N+1}{2} \text{ th Item}$$

Sum No 6

From the following data of the wages of 7 workers compute the median wages:

Wages (in Rs.)	1100	1180	1080	1120	1200	1160	1400

Calculation of Median– Discrete Series:

Steps:

(i) Arrange the data in ascending (or) descending order of magnitude.

(ii) Find out the cumulative frequencies.

(iii) Apply the formula: Median = Size of $\frac{N+1}{2}$.

(iv) Now look at the cumulative frequency column and find that total which is either equal to $\frac{N+1}{2}$ (or) next higher to that and determine the value of the variable corresponding to it. That gives the value of median.

Sum No 7:

From the following data find the median:

Income (Rs.)	1000	1500	800	2000	2500	1800
No. of Persons	24	26	16	20	6	30

Calculation of Median– Continuous Series:

$$\text{Size of } \frac{N+1}{2} \text{ th Item}$$

$$\text{Median} = L + \frac{\frac{N}{2} - \text{c.f.}}{f} \times C$$

L = Lower limit of the median class.

c.f. = Cumulative frequency of the class preceding the median class (or) sum of the frequencies of all class lower than the median class.

F = Simple frequency of the median class.

C = the class interval of the median class.

Sum No 8:

Calculate the median for the following frequency distribution::

Marks	45-50	40-45	35-40	30-35	25-30	20-25
No. of Students	10	15	26	30	42	31

Marks	15-20	10-15	5-10
No. of Students	24	15	7

MODE:

The mode or the modal value is that value in a series of observations which occurs with the greatest frequency.

Calculation of Mode– Individual Observations:

For determining mode count the number of times the various values repeat themselves and the value occurring maximum number of times is the modal value.

Sum No 9:

Calculate the mode from the following data of the marks obtained by 10 students:

SI. No.	1	2	3	4	5	6	7	8	9	10
Marks Obtained	10	27	24	12	27	27	20	18	15	30

Calculation of Mode– Discrete Series:

In discrete series quite often mode can be determined just by inspection, i.e., by looking to that value of the variable around which the items are most heavily concentrated.

Sum No 10:

Calculate the value of mode for the following data

MARKS	10	15	020	25	30	35	40
FREQUENCY	8	12	36	35	28	18	9

Calculation of Mode– Continuous Series:

$$Mo = L + \frac{f_1 - f_0}{2f_x - f_0 - f_2} \text{ x } C$$

Where L = Lower limit of the model class ; f_1 = frequency of the model class; f_0 = frequency of the class preceding the modal class; f_2 = frequency of the class succeeding the modal class.

Mode = 2 Median – 2 Mean

Sum No 11:

Calculate mode from the following data

Marks	Above 0	Above 10	Above 20	Above 30	Above 40	Above 50
No. of Students	80	77	72	65	55	43

Marks	Above 60	Above 70	Above 80	Above 90	Above 100
No. of Students	28	16	10	8	0

STANDARED DEVIATION

Calculation of Standard Deviation – Individual Operations:

In case of individual operations standard deviation may be computed by applying any of the following two methods;

1. By taking deviations of the items from the actual mean
2. By taking deviations of the item from an assumed mean.

Deviations taken from actual mean:

$$\sigma = \sqrt{\frac{\sum x^2}{N}}$$

$$x = (x - \bar{X})$$

Deviations taken from Assumed mean:

$$\sigma = \sqrt{\frac{\sum d^2}{N} - \left(\frac{\sum d}{N}\right)^2}$$

Steps,

(i) $\sum d =$ Take the deviations of the items from an assumed mean.

(ii) $\sum d^2 =$ Square these deviations and obtain the total.

(iii) Substitute the values $\sum d^2, \sum d$ and N in the above formula.

Sum No 11:

Blood serum cholesterol levels of 10 persons are as under:
240, 260, 290, 245, 255, 288, 272, 263, 277, 251
Calculate standard with the help of assumed mean.

Calculation of Standard Deviation – Discrete Series:

Methods,

(i) Actual mean method

(ii) Assumed mean method

(iii) Step deviation method

Actual Mean Method:

$$\sigma = \sqrt{\frac{\sum fx^2}{N}}, \quad \text{were } x = (X - \overline{X})$$

Assumed Mean Method:

$$\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2}$$

were d = (X –A)

Steps,

(i) Take the deviations of the items from an assumed mean and denote these deviations by d,

(ii) Multiply these deviations by the respective frequencies and obtain the total, $\sum fd$.

(iii) Obtain the square of the deviations, d^2

(iv) Multiply the squared of the deviations by the respective frequencies, and obtain the total, $\sum fd^2$.

(v) Substitute the values in the above formula.

Sum No 12:

Calculate the standard deviation from the data given below:

Size of Item	3.5	4.5	5.5	6.5	7.5	8.5	9.5
Frequency	3	7	22	60	85	32	8

Step Deviation Method:

$$\sigma = \sqrt{\frac{\sum f d^2}{N} - \left(\frac{\sum f d}{N}\right)} \text{ x C}$$

Where, d = $\frac{(\text{X} - A)}{C}$

Sum No 13:

The annual salaries of a group of employees are given in the following table:

Salaries (in Rs. 000)	45	50	55	60	65	70	75	80
No. of Persons	3	5	8	7	9	7	4	7

Calculate the standard deviation of the salaries.

Calculation of Standard Deviation – Continuous Series:

$$\sigma = \sqrt{\frac{\sum f d^2}{N} - \left(\frac{\sum f d}{N}\right)} \text{ x C}$$

Where, d = $\frac{(\text{m} - A)}{C}$

Sum No 14:

Find the standard deviation of the following distribution:

Age	20-25	25-30	30-35	35-40	40-45	45-50
No. of Persons	170	110	80	45	40	35

Take assumed average = 32.5*.

Coefficient of Variation:

The corresponding relative measure is known as coefficient of variation.

$$C.V = \frac{\sigma}{\bar{X}} \times 100$$

CORRELATION

Definitions:

"Correlation analysis deals with the association between two (or) more variables."

TYPES:

1. Positive (or) negative.
2. Simple, Partial and Multiple.
3. Linear and non-linear.

Methods of Studying Correlation:

(i) Scatter Diagram Method'

(ii) Graphic Method

(iii) Karl Pearson`s Coefficient of Correlation.

(iv) Concurrent Deviation Method.

(v) Method of Least Squares.

KARL PEARSON`S COEFFICIENT OF CORRELATION:

$$r = \frac{\sum xy}{N\sigma_x \sigma_y}$$

$$x = (X - \bar{X}) \; ; \; y = (Y - \bar{Y})$$

σ_x = Standard deviation of series X

σ_y = Standard deviation of series Y

N = Number of pairs of observations

r = The (product moment) correlation coefficient.

This methods is to be applied only where deviations of items are taken from actual mean and not from assumed mean.

(or)

$$r = \frac{\sum xy}{\sqrt{\sum_x 2 \times \sum_y 2}}$$

$$x = (X - \bar{X}) \; ; \; y = (Y - \bar{Y})$$

Steps;

(i) Take the deviation of X series from the mean of X and denote these deviations by ***x***

(ii) Square these deviations and obtain the total. $\sum_x 2$

(iii) Take the deviation of Y series from the mean of Y and denote these deviations by ***y***

(iv) Square these deviations and obtain the total. $\sum_y 2$

(v) Multiply the deviations of X and Y series and obtain the total. $\sum xy$

(vi) Substitute the values of . $\sum x\,y$, $\sum_x 2$ and $\sum_y 2$ in the above formula

Direct Method of Finding out Correlation Coefficient:

$$\text{(vii)} \quad r = \frac{\sum X Y - (\sum X)(\sum Y)}{\sqrt{N \sum X^2 - (\sum X)^2}\ \sqrt{N \sum Y^2 - (\sum Y)^2}}$$

Sum No 15:

Calculate Karl Pearson`s coefficient of correlation from the following data and interpret its value:

Roll No. of Students	1	2	3	4	5
Marks in Accountancy	48	35	17	23	47
Marks in Statistics	45	20	40	25	45

REGRESSION ANALYSIS

DEFINITIONS;

"Regression is the measure of the average relationship between two or more variables in terms of the original units of the data"

REGRESSION EQUATIONS

There are two regression equations

(i) Regression Equation of Y on X

(ii) Regression Equations of X on Y

Regression Equation of Y on X

- Is used to describe the variations in the values of Y for given changes in X.

$$Y = a + bX$$

To determine the values of a and b, the following two normal equations are;

$$\sum Y = Na + b \sum X$$

$$\sum X Y = a \sum X + b \sum X^2$$

Regression Equations of X on Y

- Is used to describe the variations in the values of X for given changes in Y.

$$Y = a + bY$$

To determine the values of a and b, the following two normal equations are;

$$\sum X = Na + b \sum Y$$

$$\sum X Y = a \sum Y + b \sum Y^2$$

Sum No 16:

From the following data obtain the two regression equations:

X.	6	2	10	4	8
Y.	9	11	5	8	7

BOOKS REFERRED:

STATISTICAL METHODS, by S.P.GUPTA