Biyani's Think Tank

Concept based notes

Quantitative Methods for Economics
Economics II

(Bachelor of Arts)

Dr. Archana Dusad
Shalini
Deptt. of Arts
Biyani Girls College, Jaipur

Published by :
Think Tanks
Biyani Group of Colleges

Sector-3, Vidhyadhar Nagar,
Jaipur-302 023 (Rajasthan)
Ph : 0141-2338371, 2338591-95 Fax : 0141-2338007
E-mail : acad@biyanicolleges.org
Website :www.gurukpo.com; www.biyanicolleges.org

Edition : 2012

Leaser Type Setted by :
Biyani College Printing Department

Preface

I am glad to present this book, especially designed to serve the needs of the students. The book has been written keeping in mind the general weakness in understanding the fundamental concepts of the topics. The book is self-explanatory and adopts the "Teach Yourself" style. It is based on question-answer pattern. The language of book is quite easy and understandable based on scientific approach.

Any further improvement in the contents of the book by making corrections, omission and inclusion is keen to be achieved based on suggestions from the readers for which the author shall be obliged.

I acknowledge special thanks to Mr. Rajeev Biyani, *Chairman* & Dr. Sanjay Biyani, *Director* (*Acad.*) Biyani Group of Colleges, who are the backbones and main concept provider and also have been constant source of motivation throughout this endeavour. They played an active role in coordinating the various stages of this endeavour and spearheaded the publishing work.

I look forward to receiving valuable suggestions from professors of various educational institutions, other faculty members and students for improvement of the quality of the book. The reader may feel free to send in their comments and suggestions to the under mentioned address.

Author

Syllabus

Section-A

Nature and uses of statistics Methods of collection and presentation of data. Graphical and diagrammatic representation of data. Measures of Central Tendency and dispersion.

Section - B

Correlation and rank correlation, Regression analysis (ungrouped data, Method of least squares with one independent variable).

Analysis of time series—components of time series and determination of trend. Index Numbers, Interpolation (Binomial expansion and Newton's Methods).

Section - C

Logarithms, Determinants and Matrices, Quadratic Equation, Arithmetic, Geometric and Harmonic Progressions, Binomial Theorem. Simple differentiation involving one independent variable only.

(**Note** : Use of non-programmable calculator is permitted).

Chapter -1
Statistics: An Introduction

Q-1 What do you mean by 'Statistics'?

Ans . Statistics has been defined in two ways, first it is defined as statistical data (Plural sense) while second is 'Statistical Methods' (Singular sense)

Statistical Numerical Data (plural sense) ↓

Numerical data which have following characteristic that called statistics

Statistics are aggregate of a facts
Statistics are numerical statement of facts
Affected by multiplicity of causes
Reasonable standard of accuracy
Statistics are collected for predetermined Purpose
Statistics must be collected in a
Statistics must be collected in a systematic manner
They should be placed in relation to each other

Statistical Methods (Singular Sense) ↓

Statistical in this sense has been defined as a science which provide tools for- ↓

Collection ↓

Classification ↓

Tabulation ↓

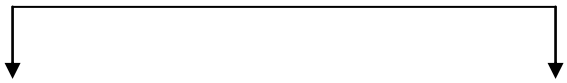
Presentation ↓

Analysis ↓

Interpretation

Q-2 What the subject matter of statistics?

Ans. Subject matter of statistics divided in two parts:-



Principle of Statistics	Applied statistics
1- Collection of Data	1. Descriptive statistics [Average, Dispersion, skewness, Index Number, trends etc.]
2- Classification	2 . Scientific Applied statistic [association of attributes correlation, Regression]
3- Tabulation	3. Business Statistic: [Study of various problem relating to business through Analysis and their solutions)
4- Presentation	
5- Analysis	
6- Interpretation	
7- Forecasting	

Q.3 What is the object of statistics?

Ans : The main objective of statistics is to make sense out facts and figures to prove the unknowns and to throw light on various problems.

Multiple Choice Questions

Q-1 Which of the following statement is true ?

(a) Statistics is derived from the French word 'Statistik'
(b) Statistics is derived from the Latin word 'Status'
(c) Statistics is derived from the Italian word 'Statista'

(d) None of the above.

Q-2 Statistics is concerned with:
(a) Qualitative information
(b) Quantitative information
(c) Either (A) or (B)
(d) Both (A) or (B)

Q-3 Drinking habit of a person is :
(a) An attribute
(b) A discrete variable
(c) A variable
(d) A Continuous variable

Q-4 Statistics is defined in terms of numerical data in the
(a) Singular sense
(b) Plural sense
(c) Either (A) or (B)
(d) Both (A) or (B)

Q-5 Marks of a student is an example of :
(a) A continuous variable
(b) A discrete variable
(c)An attribute
(d) None of these.

Q-6 Annual income of a person is:
(a) A continuous variable
(b) A discrete variable
(c) An attribute
(d) None of these.

Q-7 An attribute is :
(a) A qualitative characteristic
(b) A measureable characteristic
(c) A quantitative characteristic
(d) All these

Q-8 Nationality of a student is
(a) An attribute
(b) A discrete variable
(c)A continuous variable
(d) Either (A) or (C)

Q-9 Data collected on religion from the census reports are
(a) Sample data

(b) Secondary Data
(c) Primary data
(d) Either (A) or (B)

Q-10 Which method of collection of data covers the widest area?
(a) Direct Interview method
(b) Mailed questionnaire method
(c) Telephone interview method
(d) All of these

ANSWERS KEY:

1-A	2.D	3.A	4.A	5.B	6.B	7.A	8.A	9.B	10.B

Chapter -2

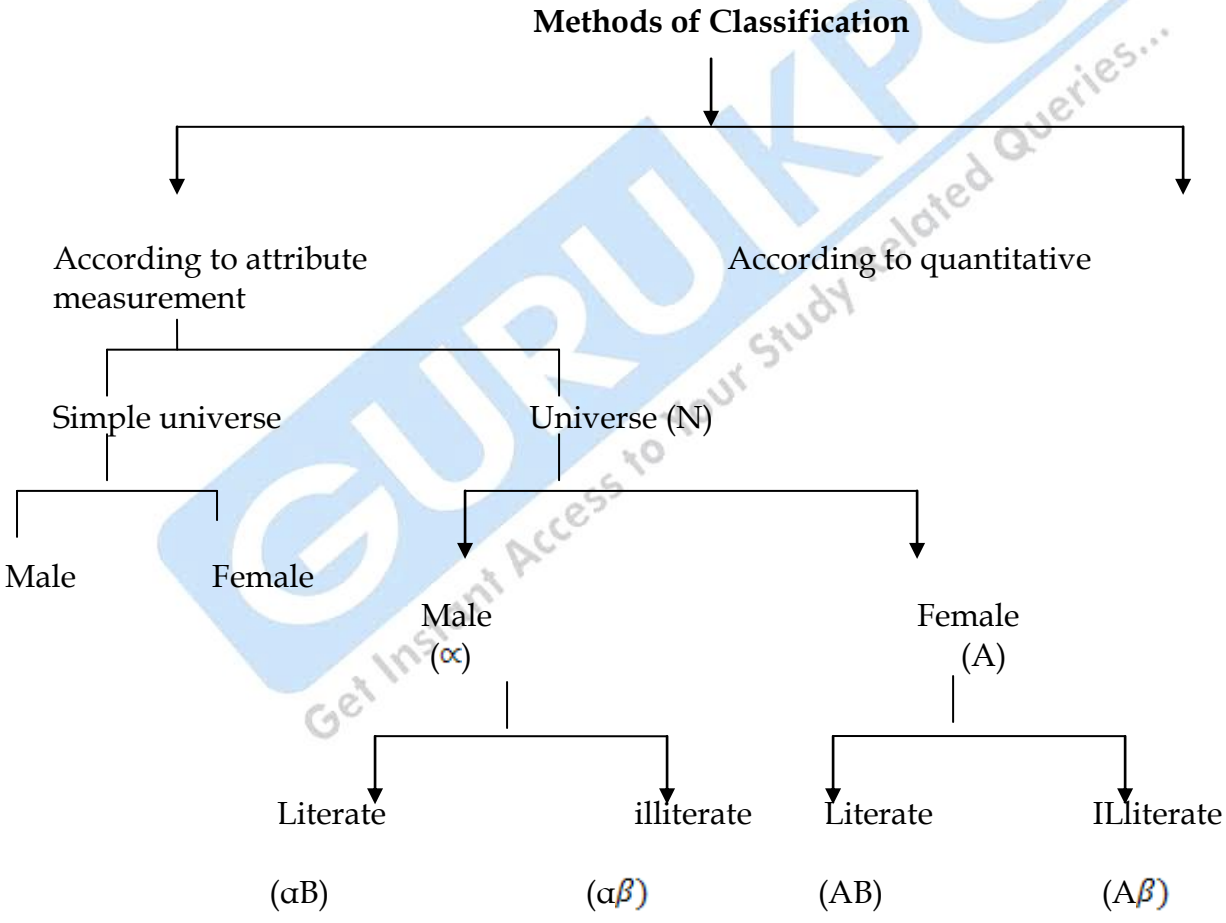
Classification

Q.1 Define classification.

Ans The process of arranging data in groups or classes according to resemblances and similarities in technically called classification. It is necessary to provide the collected data a form or structure to be appropriate for analysis and draw inference. To serve the purpose, the collected data have to be classified into classes and sub classes according to their characteristics. This process is called classification.

Q.2 Narrate the method of classification.

Ans.



Q-3 What is the meaning of frequency Distribution?

Ans. A frequency distribution is a series where a number of items with similar values are put in separate groups. It possess two part, on its left, then are

values or size or measurement and an its right the number of times the value is repeated which is called number of items or frequency.

Q-4 Describe the main technical terms of continuous frequency distribution?

Ans Main technical terms of continuous frequency distribution are as following:

1- **Class limit:-** The two values which determine the limit of a group are called class limit. The lower value is called lower limit(l_1) whereas the higher value is called upper limit (l_2)

Example: - l_1 l_2 , l_1 l_2

10 - 20 , 20 - 30

2- **Magnitude of class Interval:-** The difference between the upper and lower limit of a class is called magnitude of class interval it is expressed by symbol is (i).

3- **Mid value:** The central place of class limit is called mid value.

$$\text{Mid value} = \frac{\text{upper limit} + \text{lower limit}}{2}$$

4- **Class Frequency: -** The number of items or observation corresponding to a particular class is known as frequency of that class.

Example: - Frequency of (0-10) class is 5

Q 5 How many methods of classifying the data according to class interval/discuss it?

There are two methods of classifying the data according to class interval

(i) Exclusive Method

(ii) Inclusive Method

(i) **Exclusive Methods :** When the class interval are so fixed that the upper limit of are class is the lower limit of the next class, it is called exclusive method:

Example:

l_1	l_2
10-	20
20-	30
30-	40

(ii) Inclusive Methods:

When lower and upper limit both are included in that its :

Example:

0-09

10-19

20-29

Q-6 **State the formula of Prof. H.A. Struges to find out the number of classes.**

Ans. According to struggles rule,

Class interval (i) = $\frac{L-S}{1+3.322\ (\log N)}$

Where,
L= largest + value
S= Smallest +Value

Q-7 **Explain the types of statistical series**

Ans. **Types of statistical series**

On the Basis of Object

1. **Time series:** When data are observed over a period of time, it is called time series. In such series the information is presented for each unit of time for example: - For each year, each month, per minute, per hour.
2. **Spatial Series:** Collected information is presented are the geographical or Locational differences like states, cities, regious, zones, areas, countries etc.
3. **Condition Series:**When the collected data are capable of numerical measurement and are presented and arranged according to some condition such as marks obtained. **Height in cms. , ages in** years etc.

ON THE BASIS OF CONSTURCTION

Individual Series: In this case every item is independent. No item is placed under any group and is kept fully independent.

Example:

S No	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Marks	**8**	**17**	**28**	**17**	**16**	**23**	**30**	**5**	**9**	**21**

Discrete Series: Discrete series is one in which exact measurement are possible in whole numbers. In this series size or value as well as items both are given.

Example:

Marks	0	1	2	3	4	5	6	7	8
No. of Students	2	4	5	3	7	4	2	8	1

Continuous series: In this series the size of items is not definite but it lies between the numbers.

This become necessary in case of such variable which can take any fractional value and in whose case an exact measurement is not possible.

Example:-

Weight in kgs	40-45	45-50	50-55	55-60	60-65	65-70	70-75
No. of Students	7	10	12	15	18	09	2

Q-8 What is Bi-variate/two way frequency Distribution?

Ans : When a frequency distribution involve only one variable it is called un vicariate Frequency distribution but when data are classified on the basis of two variables it is called B_1-Variate/Two way frequency distribution.

Example:

1- Age is wives and Age of husbands.

2- Marks in Statistics and marks in economics

3- Heights of Students and Weights of Students.

Multiple Choice Question

Q-1 Which part of table is known as heart of table?

(a) Foot notes (b) Source

(c) Main body of the table (d) Table number & heading

Q-2 To show the figures shown is a table more clearly and explains than more clearly which part of table is used

(a) Source (b) Sub-heading
(c) Railing (d) Foot notes

Q-3 The data gathered from primary table like total, averages of the data are presented in tabular form along the original data this is known as :
(a) Primary Data (b) General Purpose Data
(c) Simple Data (d) Secondary Data

Q-4 A table showing data of more than one characteristics or attributes is called:
(a) Simple table (b) Complex table
(c) Double table (d) joiple table

Q-5 Systematic arrangement of data in columns and rows is known as :
(a) Statistical series (b) Interpretation
(c) Tabulation (d) Classification

Q-6 When the tabulation does is accomplished with the help of machines
(a) When there is small number of data (b) When there is large number of data
(c) Both of these (d) None of these

Q-07 Mechanical tabulation is more economical then tabulation by hand
(a) False (b) True
(c) Partly false (d) partly true

ANSWER KEY

1-C	**2-D**	**3-D**	**4-B**	**5-C**	**6-B**	**7-B**		

Chapter-3
Tabulations

Q-1 Define Tabulation?

Ans. Tabulation is one of the most important devices for presenting data in a condensed and systematic form to furnish the maximum information in minimum possible space without dispensing with the usefulness of the data.

Q-2 Describe the objects of tabulation?

Ans. Major objects of tabulation are:-

1. To Simplify statistical data:
2. To clarify similarity and dissimilarity
3. To facilitate comparison
4. To make data concise
5. To clarify the characteristics of data
6. To present facts in Minimum Space
7. To Detect errors in the facts

Q-3 Enumerate different rules of Tabulation?

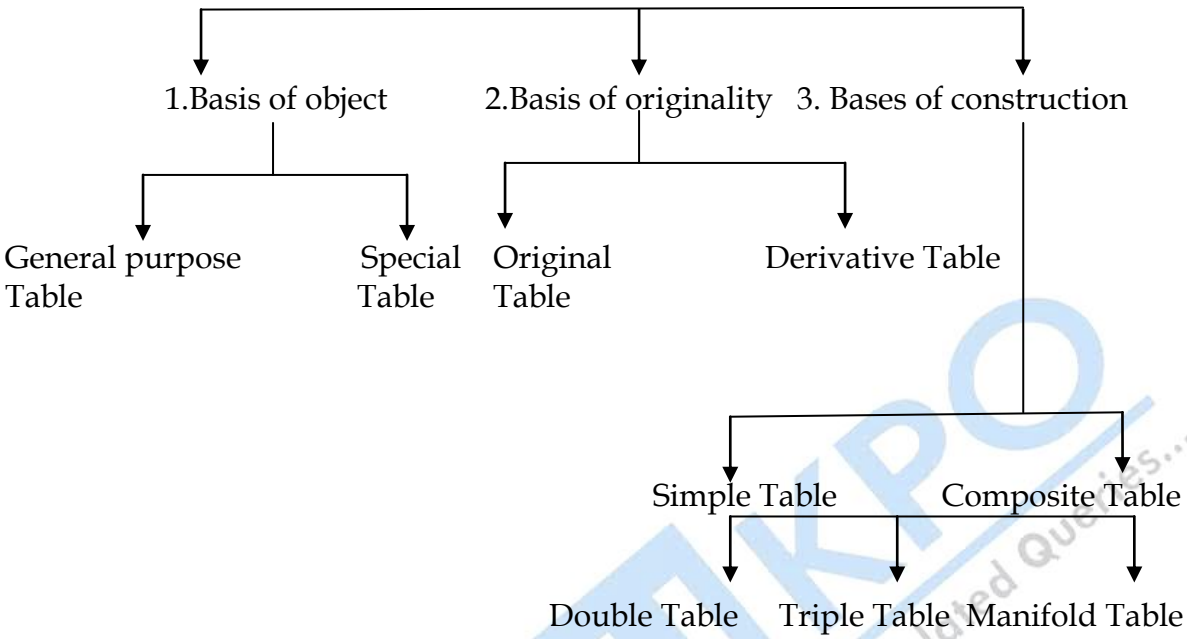
Ans. There cannot be any hard and fast rules for tabulation similar in all cases, certain rules for guidance to prepare a table are as follows:

1- Each statistical table must have a number
2- A proper title of a table must be placed above it.
3- Wording of caption should be as brief as possible. Caption means the heading of vertical columns.
4- More vital data with long description are written in rows.
5- For a neat, tidy and attractive looking, there should be proper rulings and spacing in a table.
6- The table should contain a separate column for sub-totals of each class and a separate column for total of combined classes.
7- The presentation of data is a table should be as comprehensive as possible relevant to the object of investigations.
8- Proper footnotes are used only if there is need to call attentions to some figures or headings which may not be understandable without these notes.
9- The sources from which the data have been taken be mentioned either on the left hand side or right hand side of the base live.

Q-4 State the types of tables.

Ans.

STATISTICAL TABLE



Multiple Choice Question

Q-1 Which part of table is known as heart of table?

(b) Foot notes (b) Source

(c) Main body of the table (d) Table number & heading

Q-2 To show the figures shown is a table more clearly and explains than more clearly which part of table is used

(b) Source (b) Sub-heading

(c) Railing (d) Foot notes

Q-3 The data gathered from primary table like total, averages of the data are presented in tabular form along the original data this is known as :

(b) Primary Data (b) General Purpose Data

(c) Simple Data (d) Secondary Data

Q-4 A table showing data of more than one characteristics or attributes is called:
(b) Simple table (b) Complex table
(c) Double table (d) joiple table

Q-5 Systematic arrangement of data in columns and rows is known as :
(b) Statistical series (b) Interpretation
(c) Tabulation (d) Classification

Q-6 When the tabulation does is accomplished with the help of machines
(b) When there is small number of data (b) When there is large number of data
(c) Both of these (d) None of these

Q-07 Mechanical tabulation is more economical then tabulation by hand
(b) False (b) True
(c) Partly false (d) partly true

ANSWER KEY

1-C	**2-D**	**3-D**	**4-B**	**5-C**	**6-B**	**7-B**		

Chapter-4

Measure of Central Tendency

Q-1 What do you mean by measures of central tendency?

The central tendency of a variable means typical value around which other values tend to concentrate which can be measured. Such concentration of the values in the central part of distribution is referred as measure of central tendency also known as averages.

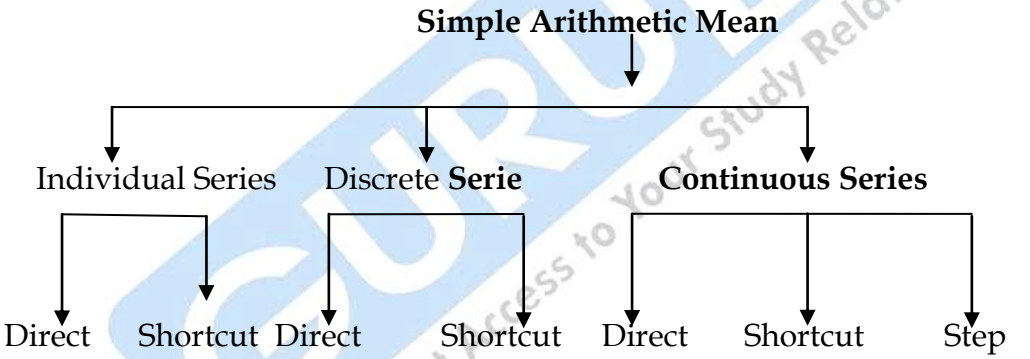
Q-2 Define Arithmetic mean? How many kinds of arithmetic mean? Give formulas for calculate arithmetic mean in all series.

The arithmetic mean of a series of items is obtained by adding the value of the items and dividing by the number of items. It is the center of gravity of a distribution.

The arithmetic mean is of two kinds:-

(a) Simple Arithmetic Mean

(b) Weighted Arithmetic Mean



$\bar{x} = \frac{\sum X}{N}$; $\bar{x} = A + \frac{\sum dx}{N}$; $\bar{x} = \frac{\sum fx}{N}$ $\bar{x} = A + \frac{\sum fdx}{N}$; $\bar{x} = \frac{\sum fx}{N}$ $\bar{x} = A + \frac{\sum fdx}{N}$ $\bar{x} = A + \frac{\sum fdx'}{N} \times i$

Where:

$\bar{x}$ = Arithmetic Mean

$\sum X$ = Total of Values

N = Number of Items

$\sum fX$ = Sum of the product of values with respective frequency

$\sum dx$ = Sum of deviation from assumed mean

A = Assumed Mean

$\sum fdx$ = Sum of the product of deviation from assumed mean with respective frequencies

i = Common factor

dx = Step deviation

fdx' = Total of product of step deviation with its respective class frequencies

Weighted Arithmetic mean $\bar{X}_w$

Direct Method

$$\overline{X_w} = \frac{\sum xw}{\sum w}$$

Shortcut Method

$$\overline{X_w} = A_w + \frac{\sum wdx}{\sum w}$$

Where: $\overline{x_w}$ = Weighted arithmetic mean

$\sum W$ = Total of weights

$\sum xw$ = Sum of product of values with their respective weights

A_w = Assumed weighted mean

dx = deviation of values from assumed weighted mean

$\sum Wdx$ = Sum of the product of deviation with respective weights.

Q.3 Give the definition of median? Have to compute median.

Ans. Median is a position Average. It is the value of middle item of a variable when the items are arranged in ascending or descending order. It divides the group into two equal parts, one part comprising all values greater and the other all values less than the median.

Computation of Median (M)

Individual Series
Continuous Series

→ Arrange in ascending or descending order → $m = \left(\frac{N}{2}\right)^{th} item$

Discrete Series

→ Find cum. Frequency → M = Value of $\left(\frac{N+1}{2}\right)^{th} item$ → Find cum

→ M = Value of $\left(\frac{N+1}{2}\right)^{th}$ *item*

$M=l_1+\frac{i}{f}(m-c)$

in descending order

M= $l_2+\frac{i}{f}(m-c)$

Where

l_1 = lower limit; l_2 = Upper limit

I = class interval of median group

F = frequency of median group

$m\left(\frac{N}{2}\right)^2$ item

C = cum frequency of preceeding median group

Q. 4 Define Mode? How to calculate mode?

Ans. Mode is the value of that observation which occurs with the greatest frequency and thus it is the most fashionable value.

Calculation of Mode (Z)

Individual series Series	**Discrete Series**	**Continous**
1-By Inspection value Inspection → Which occurs most frequency	1-By Inspection	1- By
2 - By Change in Discrete series	2- By grouping process → Find the most usual group	2- By grouping process → The value of most usual item

$$z = l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

$$\Delta_1 = f_1 - f_0$$

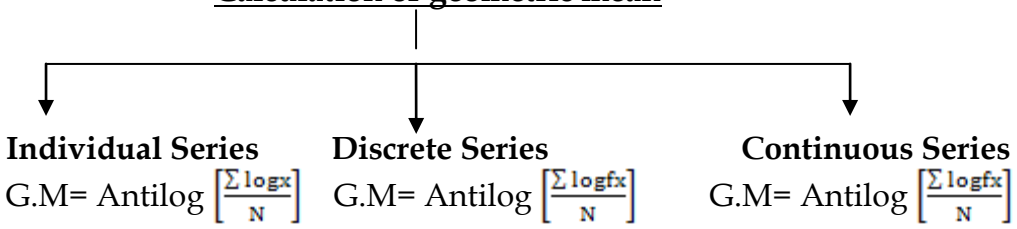
$$\Delta_2 = f_1 - F_2$$

l_1 = lower limit

I = class interval of mode group

Q.5 How to calculate Geometric Mean?

Ans. **Calculation of geometric mean**



Weighted Geometric Mean

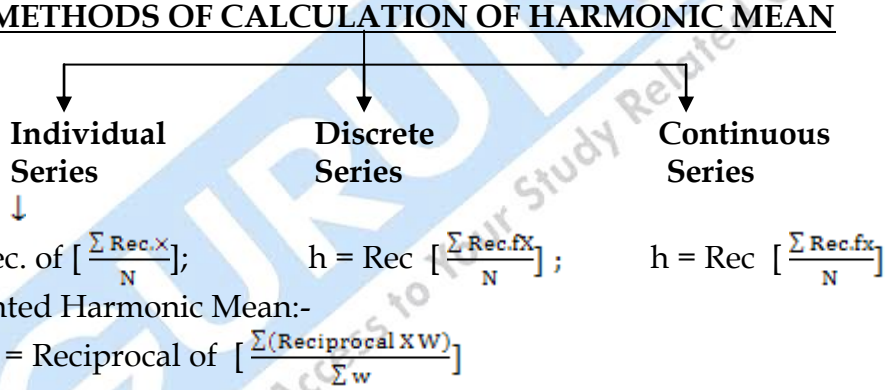
GM $_{(w)}$:- Antilog $\frac{\Sigma(\log \times W)}{\Sigma W}$

Q-6. When Harmonic Mean is used? Give the Methods of Calculation of Harmonic Mean?

Ans. Uses of harmonic Mean:-

1. To compute Average Speed
2. To find out average of prices
3. To study about Velocity

→ **METHODS OF CALCULATION OF HARMONIC MEAN**



h = Rec. of $\left[\frac{\Sigma \text{Rec.} \times}{N}\right]$; h = Rec $\left[\frac{\Sigma \text{Rec.fx}}{N}\right]$; h = Rec $\left[\frac{\Sigma \text{Rec.fx}}{N}\right]$

Weighted Harmonic Mean:-

$H_{(w)}$ = Reciprocal of $\left[\frac{\Sigma(\text{Reciprocal} \times W)}{\Sigma w}\right]$

Q-7. What is relationship between Arithmetic mean, Geometric mean and Harmonic mean?

Ans. Relationship between Arithmetic Mean, Geometric Mean, and Harmonic Mean:-

i. The geometric mean of two values is equal to geometric mean of their arithmetic mean and harmonic mean.

G.M. = $\sqrt{\bar{x}. H. M.}$ or $(G.M.)^2 = \bar{x}$. H.M.

ii. If all the items of a variable are the same, the arithmetic mean, geometric mean and harmonic mean are equal.

$\bar{x}$ = GM = HM

iii. If the values or measurement vary the arithmetic mean is greater than GM and the G.M. is greater than H.M.

X > G,M. > H.M.

Multiple Choice Question

Q-1 Measure of central tendency are known as :

(a) difference (b) average

(c) both (d) None of these

Q-2 Measure of central tendency for a given set of observation measures:

(a) The scatterness of the observation

(b) The central location of the observation

(c) both (A) or (B)

(d) None of these

Q-3 The average has relevance for

(a) Homogenous population (b) Heterogeneous population

(c) both(A) or (B) (d) None of these

Q-4 A measure of central tendency tries to estimate the

(a) Central value (b) lower value

(c) upper value (d) None of these

Q-5 The number of measures of central tendency is

(a) Two (b) three

(c) four (d) five

Q-6 A.M is never less than G.M

(a) True (b)False

(c) both (a) or (b) (d) None of these

Q-7 While computing the A.M from a grouped frequency distribution, we assume that

(a) The classes are of equal length

(b)The classes have equal frequency

(c) None of the above

(d) All the values of a class are equal to the mid-value of that class

Q-8 Median is unaffected by extreme values

(a) True (b) False

(c) Both (d) None of these

Q-9 When the distribution is symmetrical mean, median and mode

(a) Coincide (b) do not coincide
(c) both (d) None of these

Q-10 _____is the value of the variable corresponding to the highest frequency
(a) Mode (b) Mean
(c) Median (d) None of these.

ANSWERS KEY :

1-B	**2-B**	**3-B**	**4-A**	**5-B**	**6-A**	**7-C**	**8-A**	**9-A**	**10-A**

Chapter-5
Measure of Dispersion

Q-1 What do you mean by dispersion?

Ans. The term dispersion refers to the variability in the size of items. It indicates that the size of items in a series is not uniform. The values of various items differ from each other.

Q-2 What is the difference between Absolute and relative measure of dispersion?

Ans. Dispersion may be measured in two ways:-

1) Absolute Measure
2) Relative Measure

Difference between Absolute and Relative Measure are as following:-

1. Absolute Measure:-

The measure of dispersion which can be expressed in terms of the original units of a series.Such as Rupees, Meters, Kilograms] etc. are termed as absolute measure. Such measure may be erroneous for comparison if the series one originally in different units.

2. Relative Measure:-

The measure of dispersion which can be expressed as a percentage or ratio of the average are called relative measure. For comparing the variability of two or more series it is the relative measure that has to be taken into account.

Q-3 Explain various measures of dispersion?

Ans.= **METHODS OF MEASURING DISPERSION**

METHODS OF LIMITS	METHODS OF AVERAGING DEVIATION
1. Range	1. Mean Deviation
2. Inter Quartile Range	2. Standard Deviation
3. Percentile Range	3. Other Measures
4. Quartile Deviation	

1. Range:-

Range = H - L

Where

H = Highest Value

L = Lowest Value

Coefficient of Range:- $\frac{H-L}{H+L}$

2. Inter–Quartile Range:-

I.O.R. = $Q_3 - Q_1$

Where

Q_3 = Upper Quartile

Q_1 = Lower Quartile

3. Percentile Range:-

P.R. = $P_{90} - P_{10}$

Where

P_{90} = Percentile 90

P_{10} = Percentile 10

4. Quartile Deviation :-

Q.D. → $\frac{Q_3-Q_1}{2}$

Where

Q3 = Upper Quartile

Q1 = Lower Quartile

Coefficient of Quartile Deviation → $\frac{Q_3-Q1}{Q^3+Q1}$

Mean Deviation (δ) —

In Individual Series

	Direct Method	Short Cut Method
1. From Mean	$\delta x = \frac{\sum \lvert d\bar{x} \rvert}{N}$	$\delta x = \frac{\sum Xa - \sum Xb - (Na-Nb)\bar{x}}{N}$
2. From Median	$\delta M = \frac{\sum \lvert d\tilde{M} \rvert}{N}$	$\delta M = \frac{\sum Ma - \sum Mb}{N}$
3. From Mode	$\delta Z = \frac{\sum \lvert d\hat{z} \rvert}{N}$	

Where :-

δ = Mean Deviation

dx = Deviation from Mean

d_M = Deviation from Median

d_Z = Deviation from Mode

N = Number of Items

| | = Ignoring ± signs

Where :-

$\sum Xa$ = Total of value above Mean

$\sum X_b$ = Total of value below Mean

Na = Number of item above mean

Nb = Number of item below mean

$\sum Ma$ = Total of value above median

M_b = Total of value below median

N = Total Number of Item

Mean Deviation
In Discrete Series

Direct Method

1. From mean $\delta x = \frac{\sum fld\bar{x}l}{N}$
2. From Median $\delta M = \frac{\sum fld\bar{M}l}{N}$
3. From Mode $\delta Z = \frac{\sum fldzl}{N}$

Short Cut Method

$\delta x = \frac{\sum fxa - \sum fxb - (Na - Nb)X}{N}$

$\delta_m = \frac{\sum fMa - \sum fMb - (Na - Nb)M}{N}$

$\delta z = \frac{\sum fZa - \sum fZb - (Na - Nb)Z}{N}$

Where:-

$fld\bar{x}l$ = sum of product of deviation with their respective frequency

$\sum fXa$ = Total product of size and frequencies above the average
$\sum fXb$ = Total of product of size and frequency below average
$\sum Na$ = Total frequencies above mean
$\sum Nb$ = Total frequencies below mean
$\sum fMa$ = Total of product of size and frequencies above Media
$\sum fMb$ = Total of product of size and frequencies below Median

Mean Deviation
In Continuous Series

Direct Method

1. From mean $\delta x = \frac{\sum fld\bar{x}l}{N}$
2. From Median $\delta M = \frac{\sum fld\bar{M}l}{N}$
3. From Mode $\delta Z = \frac{\sum fldzl}{N}$

Short Cut Method

$\delta x = \frac{\sum f\ dx' + (X-A)(N_b - N_a)}{N}$

$\delta m = \frac{\sum fldm'l + (m-m')(Nb-Na)}{N}$

Where:-

$\sum fld\ M'$ = sum product of deviation from assumed median and frequencies
M & M′ = Actual Median and Assumed Median
Nb = Total Number of items below Actual Median
oNa = Total Number of items above Actual Median
$\sum fldx'l$ = Sum of the product of deviation from assumed Mean with respective frequencies.

5. STANDARD DEVIATION

In Individual Series

Direct Method

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

Where:-

σ = Standard Deviation

$\sum d^2$ =sum of the square of deviation from Mean or $\sum d^2 = \sum (X - \overline{X})^2$

N = Number of item

N= Total Number of items

Short Cut Method

$$\sigma = \sqrt{\frac{\sum d^2x - (\bar{x}-A)^2}{N}}$$

σ= Standard Deviation

$\sum dx$ = sum of deviation from assumed Mean

$\sum d^2 x$ = sum of square of deviation from assumed Mean

Standard Deviation

In Discrete Series

Direct Method

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

σ = Standard Deviatio

N = Number of item

$\sum fd^2$ = Total of product of

N = Total No. of items

dx = (X-A)

d = $(X - \overline{X})$

d^2 = Square of each deviation

Short Cut Method

$$\sigma = \sqrt{\frac{\sum fd^2x - (\bar{x}-A)^2}{N}}$$

σ = Standard Deviation

X = Actual Mean

A = Assumed Mean Square of deviation

And frequencies

fdx = [f × dx]

fd^2x = [fdx ×dx]

Standard Deviation

In Continuous Series

Direct Method

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

Where:-

σ = Standard Deviation

N = Number of items

$\sum fd^2$ = Total of product of Square of deviation & frequency

Short Cut Method

$$\sigma = \sqrt{\frac{\sum fd^2x - (X-A)^2}{N}}$$

σ = Standard Deviation

dx = (X-A)

fdx = [f × dx]

fd^2x = fdx × dx

N = Total No. of items

Coefficient of $\sigma = \frac{\sigma}{\bar{x}}$

Coefficient of variation $= \frac{\sigma}{\bar{x}} \times 100$

Variance = σ^2

or

$\sigma = \sqrt{\text{Variance}}$

Q.4 **How to calculate combined standard deviation?**

Ans. Calculation process of combined standard deviation is as follows:-

$$\sigma_{12\text{------}n} = \sqrt{\frac{N1\,(\sigma1^2 + d1^2) + N2\,(\sigma2^2 + d2^2) + \ldots..Nn\,(\sigma^2n + d^2n]}{N1 + N2\ldots Nn}}$$

Where:-

N1, N2, …Nn= Number of items of different groups

$\sigma, \sigma2, \ldots \sigma n$ = Standard Deviation of different groups

d1, d2 ….dn = Difference between the combined mean and the means of different groups.

Multiple Choice Question

Multiple Choice Question:

Q-1 Measure of dispersion determines the data of:

(a) Variability (b) Center

(c) Division (d) all of the above

Q-2 The degree to which the data is spread about an

(a) Center (b) deviation

(c) Dispersion (d) all of the these

Q-3 Measures of dispersion measures the dispersion

(a) Correlation coefficient (b) Average

(c) Standard deviation (d) None of these

Q-4 Which of the following are uses of measures of dispersion?
(a) Determine reliability of average (b) Compare Viability
(c) Devising system of quality control (d) All of these

Q-5 How many types of dispersion are there?
(a) 0 (b) 1
(c) 2 (d) 3

Q-6 Measures of dispersion represented in same statistical unit are :
(a) Absolute (b) Relative
(c) Qualitative (d) All of these

Q-7 The measure of dispersion calculated by dividing absolute measure y measure of central tendency is _____measure of dispersion
(a) Absolute (b) Relative
(c) Qualitative (d) All of these

Q-8 _____Measures of dispersion are called as coefficient of dispersion
(a) Absolute (b) Relative
(c) Qualitative (d) All of these

Q-9 _____Measures of dispersion are suitable for comparison of sample of same population.
(a) Relative (b) Qualitative
(c) Absolute (d) Quantitative

Q-10 If the items are in different units, absolute measure of dispersion is a good measure
(a) True (b) False
(c) Partly true (d) don't say

ANSWERS KEY:

2-C	3-B	4-D	5-C	6-A	7-B	8-B	9-C	10-B

Chapter-6
Probability

Q.1 Explain the meaning of Probability?

Ans. The word 'Probability' is related with chance of happening or not happening of an event. For instance, many students pursue MBA course after graduation because they feel that MBA degree improve their chances for a higher lifetime income. A Sales Manager may decide to spend more on advertising he feels that it will increase sale of his products. These are all personal decision that are made in the face of uncertainty. Probability theory is frequently used as an aid in making decision in the face of uncertainty.

Q.2 Explain various approaches to the concepts of probability?

Ans. There are three approaches of probability. This are-

1) The classical or mathematical approach.
2) The empirical or the statistical approach.
3) The axiomatic approach.

1) The classical or mathematical approach:

The basic approach for the classical theory is that outcomes of a random experiment are equally likely.

Let a random experiment result in n equally likely, mutually exclusive and exhaustive simple event and let m (out of n) be the number of outcomes favorable to the happening of an event A, then the probability of happening of A, denoted by P (A) is defined as-

$$P(A) = \frac{m}{n} = \frac{\text{Number of outcomes favorable to A}}{\text{Total Number of outcomes}}$$

Ex:- If a coin is tossed, there are two equally likely results a head or a tail, hence the probability of head is ½

2. Empirical or Statistical Approach:-

In Empirical approach the probability of an event represents the proportion of times under identical circumstances the result can be expected to occur. The main assumption are-

a) The experiment are random
b) There is large number of experiment.

Let an event A occurs m times in N repetitions of a random experiment. Then the ratio m/N, gives the relation frequency of the event A and it will not vary appreciably from one trial to another. In the limiting case when N becomes

sufficiently large it more or less settles to a number which is called the probability of A or

$$P(A) = \lim \frac{m}{N}$$

$$N = \infty$$

Ex:- If an unbiased coin tossed, then the classical approach gives the probability of a head as ½. This means if we toss an unbiased coin, say 20 times, this classical approach suggests that we should get head 10 times. However, in practice this will not generally be true. In fact, in 20 throws of a coin, we may get no head at all or 1 or 2 heads. However, the empirical definition of probability suggests that if a coin is tossed a large number of times say a 500 times, we should except on an average 250 heads and 250 tails.

1. Axiomatic Approach:-

The axiomatic probability includes the concept of both the classical as well as the empirical definition of probability. In this approach the probability of an event ranges from 0 to 1. The probability of the entire sample space is 1 and for mutually exclusive events the probability of the happening of either A or B denoted by P (A U B).

P (A U B) = P (A) + (B)

For events of simultaneous occurrence or the probability of both A and B happening together is denoted by P (A ∩ B)

P (A∩B) = P (A) . P (B)

Q.3 What do you mean by permutation?

Ans. Permutation refers to the different arrangement of objects in a set where all elements are different and distinguishable. These arrangements are to be done without repetition of any individual object.

Permutation of 'n' Different objects taken 'r' at a time:-

$${}^{n}P_{r} = \frac{n!}{(n-r)!}$$

⇨ Permutation of 'n' different objects and 'r' will be equal to 'n'

$${}^{n}P_{n} = \frac{n!}{(n-n)!} = \frac{n!}{0!}\frac{n!}{1} = n!$$

⇨ Permutation of 'r' objects out of 'n' different objects when particular object is always taken

${}^{r}P_{1} \times {}^{n-1}P_{r-1}$

⇨ Permutation excluding 'P' particular items:-

${}^{n-p}Pr$

⇨ Number of Permutation of such items in which some items are common:-

$$\frac{n!}{p!q!r!}$$

⇨ Permutation when out of given n items, r items are taken

together when each item can be taken in any arrangement

n^r ways

(ii) Multiplication Theorem:-

a) When events are Independent –

$P(A \cap B) = P(A) \times P(B)$

b) When events are dependent

A) Conditional probability

$$P(A/B) = \frac{N(A \cap B)}{N(B)} = \frac{N(A \text{ and } B)}{N(B)}$$

$$P(B/A) = \frac{N(A \cap B)}{N(A)} = \frac{N(A \text{ and } B)}{N(B)}$$

Or

$P(A \cap B) = P(A).\ P(B/A)$

$P(A \cap B) = P(B).\ P(A/B)$

B) Marginal probability :-

$$P(A) = \frac{N(A)}{N(S)}$$

$$P(B) = \frac{N(B)}{N(S)}$$

C) Joint probability :-

$$P(A \cap B) = \frac{N(A \cap B)}{N(C)} = \frac{N(A \text{ and } B)}{N(S)}$$

Q.4 What do you mean by Bernoulli's Theorem?

Ans. If a trial of an experiment can result in success with probability P and failure with probability q (1-p) the probability of exactly 'r' success in n trials can be determined by Bernoulli's Theorem.

$P(r) = {}^nC_r \cdot p^r \cdot q^{n-r}$

Here n = Number of experiments

r = Number of success

p = Probability of success

q = 1 – p = probability of unsuccessful

Q.5 Explain Mathematical Expectation.

Ans. Probability of happening of an event and getting money by a person on happening of that event, the product of both is known a mathematical expectation.

Ex:- Probability of happening an event is ¼ and Shyam gets Rs.320 on that event then

M.E. $= 320 \times \frac{1}{4} = 80$ Rs.

M.E.(E) = P. X

Here:-

X = Money to be received on success

P = Probability of event

Multiple Choice Question

1. Initially , probability was a branch of
   (a) Mathematics (b) Statistics
   (c) Physics (d) Economics
2. All possible outcomes of a random experiment form the
   (a) Sample space (b) events
   (c) both (d) none of the above
3. The terms 'chance' and probability are synonymous.
   (a) True (b) False
   (c) Both (d) None of the above
4. The limitations of the classical definition of probability is that:
   (a) It is applicable if the if the elementary events are equally likely
   (b) It is applicable when the total number of elementary events is finite
   (c) It is applicable if the elementary events are mutually independent
   (d) Both(a) or (b)
5. The classical definition of probability is based on the feasibility at subdividing the possible outcome of the experiments into :
   (a) Exhaustive and equally likely cases
   (b) Mutually exclusive and exhaustive cases
   (c) Mutually exclusive and exhaustive cases
   (d) mutually exclusive, exhaustive and equally likely cases.

6. Two broad divisions of probability are :
   (a) Subjective probability and objective probability
   (b) Statistical probability and Mathematical probability
   (c) Deductive probability and non-deductive probability
   (d) None of these

7. When a dice is tossed, the sample space is :
   (a) S=(1,2,3,4,5)
   (b) S=(1,2,3,4)
   (c) S=(1,2,3,4,5,6)
   (d) S=None of these

8. According to the statistical definition of probability, the probability of an event A is the
   (a) Limiting value of the ratio of the number of times the event A occurs to the number of times the experiment is repaired
   (b) The ratio of the frequency of the occurrences of A to the non-occurrence of A
   (c) The ratio of the frequency of the occurrences of A to the total frequency
   (d) The ratio of the favorable elementary events to A to the total number of elementary events.

9. Probability mass function is always:
   (a) 0 (b) greater than 0
   (c) Greater than equal to 0 (d) equally likely

10. If one of the outcomes cannot be expected to occur in preference to the other in an experiment the events are :
   (a) Compound events (b) simple events
   (c) favorable events (d) equally likely events

Q-11 There are two coins. On one face of each coin 1 is written and on the other face 24 written. Two coins are tossed simultaneously. The expected value of the total on the coins is:
   (a) 4 (b) 2
   (c) 3 (d) 7

Q-12 There are 2 white and 4 black balls in a box. A person takes 3 balls at random from the box. If he receives Rs.10 for each white ball and receive Rs.5 for each black ball, then the expected value of the amount received by him (in rupees) is :
   (a) 20 (b)30
   (c) 15 (d) 25

Q-13 There are 10 electric bulbs in a box in which 3 are defective bulbs. If 3 bulbs are selected at random from the box, then the expected number of defective bulbs is :
   (a) 0.7 (b)0.9
   (c) 0.6 (d) 0.5

Q-14 Two coins are tossed simultaneously. A person receives Rs.8 for each head and looses Rs.10 for each tail. The expected value of the amount gained by him (in rupees) in
   (a) 2 (b)-2
   (c) 3 (d) -4

Q-15 There are 5 white and 3 black bulbs in a box. 3 bulbs are taken at random from the box. The expected number of black balls is :

(a) 7/8 (b)9/8
(c) 11/8 (d) 5/8

Q-16 There are 3 black balls and 2 white balls in a box. 2 balls are taken from it. Rs.24 is given for each black ball. What amount (in rupees) should be charged for each white ball so that the game is fair.
(a) 26 (b) 36
(c) 16 (d) 46

Q-17 There are 5 tickets in a box numbered 1, 1,2,2,2 respectively. Two tickets are taken at random from it, then the expectation of the total of the numbers on the tickets is :
(a) 11/5 (b)13/5
(c) 16/5 (d) 7/5

Q-18 There are 100 tickets in a lottery o Re. 1 each. There is only one ticket in the lottery bearing a prize of Rs.80. A person purchased on e ticket. Hits expectation is :
(a) -0.20 (b)-0.80
(c) 0.20 (d) 0.80

Q-19 A player tosses 3 fair coins , He wins Rs. 5 if 3 heads appear, Rs. 3 if 2 heads appear, Re. 1 if 1 head appears. On the other hand he losses Rs. 15 if 3 tails appear. His expected gain (in rupees) is :
(a) 0.15 (b) 0.25
(c) -0.25 (d) -0.15

Q-20 Two tickets are taken at random from 5 tickets number 1 to 5. The expected value of the sum obtained on the two tickets is :
(a) 5 (b)3
(c) 6 (d) 7

ANSWERS KEY :

1-A	**2-A**	**3-A**	**4-D**	**5-D**	**6-A**	**7-C**	**8-A**	**9-C**	**10-D**
11-C	**12-A**	**13-B**	**14-B**	**15-B**	**16-B**	**17-C**	**18-D**	**19-A**	**20-D**

Chapter 7

Probability Distribution

Q.1 What is meant by Theoretical Frequency distributions?

Ans. For systematic presentation of data, we used the frequency distribution based on actual observation, which is called as observed frequency distribution. It is constructed by collecting, classifying, tabulating the data obtained from statistical investigation.

Distribution which are not obtained by actual observation or experiments, but are deduced mathematically under certain predetermined hypothesis or assumptions or by following certain probability rules are called theoretical frequency distribution. These may be called as expected or model frequency distributions.

Q-2 Narrate the main features of Binomial Distribution?

Ans: Binomial distribution was discovered by swiss mathematician James Bernoulli so this distribution is called as 'Bernoulli Distribution' also. This is a discrete frequency distribution assumption:-

(i) The random experiment is performed repeatedly with a fixed and finite number of trials is denoted by 'n'.

(ii) There are two mutually exclusive possible outcomes on each trail. Which are known as 'Success' and 'failure' Success is denoted by p whereas failure is denoted by q.

(iii) The outcomes of any given trial does not affect the outcomes on subsequent trails, means the trials are independent.

(iv) The probability of success and failure remains constant from trial to trial.

Q-3 Discuss the silent features of the poisson distribution?

Ans: Poisson distribution was discovered by French mathematician Simon Denis Poisson. It is a discrete probability distribution.

The poison distribution is applicable where the successful events in the total events are few situations where poisson distribution is applicable.

(i) Number of defective blades out of total blades produced in a factory

(ii) Number of goals scored at a football match, where in number of attempts (n) may be a lot of but the success (p) are few.

(iii) Number of mistakes found in the pages of a book.

(iv) Number of telephone calls done during every 5 minute by a businessman.

(v) Number of typing errors per page in a typed material.

Q-4 What are the chief properties of Normal Distribution?

Ans: Normal distribution is a continuous probability distribution. When 'n' the number of trials is very large or infinite ($n \to \infty$) and neither 'p' nor 'q' is very small or nearer to equal them binomial distribution tends to be a normal distribution. This distribution is a bell shaped curve.

Properties of normal curve:

(i) Shape: It is perfectly symmetrical and bell shaped.

(ii) Position: of mean, mode and median: Mean, mode and median remains equal in normal distribution, they are found in the mid of the distribution and distribute the area of curve in equally two parts.

(iii) Asymptotic: - As the distance of the curve from the mean increases , the curve comes closer and closer to the axis but never touches it.

(iv) Unimodel :- It has only one mode so it is uni-modal.

(v) Continuous Distribution: It is a continuous variable distribution.

(vi) Equidistance of quartiles:- The difference between 'third quartile median' and median and first quartile are equal.

(vii) Quartile Deviation and Probable Error: Quartile deviation is equal to probable error, which is about 2/3 of standard deviation.

(viii) Mean Deviation: The mean deviation about mean is $4/5\sigma$

(ix) Points of Inflection: The points where the curveity of normal curve changes its direction are termed as ' Points of inflections'

(x) Standard Deviation determines the width of curve: If the value of standard deviation is less the width of curve will also be less and vice-versa.

(xi) Parameters: These are two parameters of normal distribution. Arithmetic mean and standard deviation.

(xii) Constants:- The following are the constants of normal distribution:

Mean – $\bar{X}$ or u
Standard deviation σ
Variance = $\sigma^2 = u^2$
Third central moment = $u_3 = 0$
Fourth central moment = u_4

Moment coefficient of kurtosis $\beta_2 = \frac{u_4}{u_2^2} = 3$

Moment coefficient of skewness $\beta_1 = \frac{u_3^2}{u_2^3} = 0$

Multiple Choice Question

Q-1 Probability distribution may be:
(a) discrete (b) Continuous
(c) Either (a) or (b) (d) Both (a) or (b)

Q-2 Theoretical distribution is a :
(a) Probability distribution (b) standard distribution
(c) Random distribution (d) None of the these

Q-3 A theoretical probability distribution is :
(a) Discrete (b) continuous
(c) Either (a) or (b) (d) both (b) or (c)

Q-4 In discrete case the probability of the entire space is :
(a) -1 (b) 1
(b) 0 (d) none of these

Q-5 The discrete probability distribution is :
(a) Poisson distribution (b) Normal distribution
(c) Binomial (d) both (a) or (b)

Q-6 Binomial distribution is a :
(a) Continuous (b) discrete
(c) both (a) or (b) (d) none of these

Q-7 An important continuous probability distribution is :
(a) Normal (b) Poisson distribution
(c) Geometric distribution (d) Binomial

Q-8 Binomial distribution is symmetrical if :
(a) P>q (b) p<q
(c) p=q (d) none of the these

Q-9 Parameter is a characteristic of :
(a) Population (b) sample
(c) Probability distribution (d) Both (a) and (b)

Q-10 An example of a parameter is :
(a) Sample mean (b) population mean
(c) Population S.D (d) Both(b) and (c)

ANSWRS KEY

1-D	**2-C**	**3-A**	**4-B**	**5-A**	**6-B**	**7-D**	**8-C**	**9-A**	**10-B**

Chapter 8
Theory of Estimation

Q-1 Explain the meaning of estimation and estimates.

Ans. Estimator:- Any sample statistic which used to estimate a population parameter is called an estimator. The sample means (X) can be an estimator of the population mean (u) and the sample proportion can be used as an estimator of the population proportion.

Estimates: An estimate is defined as a specific numerical value of the estimator. An estimate is a specific observed value of statistic.

Q-2 How many kinds of Estimates? Explain it.

Ans. Estimates are classified into the following two categories:

(i) Point Estimates (ii) Interval Estimates

(i) Point Estimates:- A point Estimate is a single number which is used to estimate an unknown population parameter. The sample mean (X)is the best estimator of the population mean (u).

(ii) Interval Estimates: - An interval estimate describes a range of values within which a population parameter is likely to be.

Q-3 What are the quantities of a Good Estimator?

Ans. An estimator should have good quantities. Some statistics are better estimator then others. A good estimator should have following qualities.

Unbiasedness: The term unbiasedness refers to the act that a sample mean is an unbiased estimator of a population mean because the mean of the sampling distribution of sample means taken from the sample population is equal to the population means itself.

(ii)Efficiency: Efficiency refers to the size of the standard error of the statistic. If we compare any two statistics from a sample o the same size and try to decide which the more efficient estimator is, we would select the statistic which has the smallest standard error or standard deviation of the sampling distribution.

(i) Consistency: The consistency is also an important quality of an estimator. A Statistic is a consistent estimator of a population parameter.

(ii) Sufficiency: The sufficiency plays an important role in estimation. An estimator is sufficient if it makes so much use of the given information in

the sample that no other estimator could extract from the sample additional information about the population parameter being estimated.

Chapter 9

Sampling Theory

Q-1. Discuss the methods of collection of data?

Ans. The required data may be obtained by following methods:-

(a) Census Method

(b) Sampling Method

(a) Census Method: Under this method, data are collected for each and every unit of the population or universe which are of interest in any particular situation. This method is used where the size of the universe is not big and there is a need for accurate results.

(b) Sampling Method: Under this method some representative units are selected from the universe and studied thereafter conclusions are drawn for the entire universe or that basis. The representative units selected from the universe for the purpose of investigation are termed as sample. This method is used widely at present time.

Q-2. Define sampling and what are the essentials of good sampling?

Ans. Any scientific study, instead of every unit of the universe only a part of it is studied and the inferences are drawn on that basis for the entire universe. The part of universe which we select for the purpose of investigation is termed as 'sample' and the method for selecting that part is termed as "sampling method"

Essentials of Good Sampling:

A good sample should have following characteristics:-

(i) A sample should be truly representative of the universe

(ii) The size of sample should be adequate

(iii) The sample should be free from all bias

(iv) General knowledge and logic should be used in the selection of sample.

(v) Each and every unit of the universe must have independent chance of selecting sample.

(vi) There should be homogeneity and stability in the process of selection of each unit of the sample.

(vii) The size of sample should be adequate enough in proportion to the universe.

(viii) The success of sampling depends upon the lack of bias tendency.

Q-3. What is the Basis of Sampling?

Ans. The theory of sampling is based on the two important law which are as follows:-

(i) Law of Statistical Regularity: - have of statistical regularity that if a sample is taken at random from a universe, it is likely to possess the same characteristic as that of the population. This lane is reliable where sample is chosen at random and the size of sample is adequately large.

Random selections means a selection where each and every item of the universe has an equal chance of being selected in the sample

(ii) Have of Inertia of large Numbers:- The low of inertia of large numbers is a corollary of the lane of statistical regularity. This law states large aggregates are relatively more stable than small ones. This law is based on the fact that when large number is considered the variations in the component p arts tend to balance each other so that the total change would be insignificant.

Q-4. Discuss the various methods of sampling?

Ans: There are various methods of sampling suitable for various situations. Some important methods of sampling are given below:-

(i) Deliberate or purposive method
(ii) Random sampling
(iii) Mixed sampling
(iv) Other methods

(i) Deliberate/Purposive sampling: - An investigation may select typical units which he thinks are representative of the universe in accordance with some purposive principle.

Under this method, some criterion for selection of units is laid down first and the units are selected is accordance with it.

(ii) Random sampling:- In random sampling each and every unit of the population has an equal chance of being selected in the sample. The selection of units in sample is not affected by the bias of the investigator and is just a matter of chance.

A random sample can generally be selected in the following ways:-

(a) Lottery Method : Under this method, all the members of the population or universe are serially numbered on small slips of a paper. They are put in drum and thoroughly mixed by vibrating the drum. After mixing

the numbered slips are drawn out of the drum one by one according to the size of the sample. The number of slips so drawn constitute a random sample.

(b) Random number method : In this method, sampling is conducted on the basis of the random numbers which are available from the random number table.

(iii) Mixed Sampling: Under this category those sampling methods are included which have combined characteristic of both purpose and random sampling methods of mixed sampling as follow:

(a) Stratified sampling: in stratified random sampling, the population is divided into strata(groups) before the sample is drawn.

(b) Multi stage sampling: in this sampling method, sample of elementary units is selected in stages. Firstly a sample of cluster is selected and among them a sample of elementary units is selected.

(c)

(iv) Other methods: Other methods can be describe as follows:

(a) Extensive sampling :- In this method, number of units selected in the sample is very large

(b) Convenience sampling: Under this method, a sample is obtained by selecting such units from the universe which are collected neither by probability more by judgment but by convenience of the investigator.

(c) Quota Sampling:- In this method, first of all the universe is divided into certain parts. The number of units which are to be selected from each part is decided and termed as 'quota'. Each investigator is then told to contact or examine a certain number of people which constitutes his quota. The selection of sample items within the quota depends on personal judgment of the investigator.

(d) Area Sampling: In this method area it is sampling unit and the units regarding which information is to e collected are termed as elementary units.

(e) Sequential sampling: Under this method sample size is determined on the basis of estimated sampling error. Sample is selected from lots inspite of units. Which lot is to be selected or which is to be rejected is decided on the basis of accepted sampling of the lot.

(f) Self selected sample: Under this method, units of the sample are not selected by the investigator but the units of the universal are given freedom to be a part of sample or not.

Q-5 State the various steps in tests of significance.

Ans. To test o f significance following procedure is adopted:

(i) Formation of Hypothesis: To test of significance me can formulate two hypothesis.

(ii) Null hypothesis: It is denoted by H_o and used as an important tool for testing the significance of difference. Null hypothesis assumes that there is no significant difference in the sample and population in a specific matter under consideration.

$E_x = H_O = \mu_1 = \mu_2$

(iii) Alternative Hypothesis: It is denoted by H_1 and when we reject the null hypothesis that means alternative hypothesis is selected. The alternative hypothesis specifies that the difference between sample statistic and population parameter is significant Example:-

H_O = μ=200

H_1 = $\mu \neq 200$

H_1 = $\mu > 200$

H_1 = $\mu > 200$

(iii) Determination of a suitable level of significance: Next step is to check the validity of hypothesis at a certain level of significance. The rejection or acceptance of null hypothesis is depends on the significance level. Generally 1% or 5% level of significance are used to test the hypothesis 5% level of significance is adopted that means 5%

chances of rejecting the null hypothesis and 95% chances of accepting the null hypothesis. The level of significance is denoted by e(alpha)

(iv) Determination of a suitable test statistic:- Next step is to determine a test statistic. A null hypothesis is accepted or rejected on the basis of test statistic. For large sample 2 test , small samples t test and in some specific cases f or x^2 test can be used.

(v) Determination of critical region:- The null hypothesis is accepted if the value of sample statistic is neither acceptance region and rejected if the value of sample statistic is under the rejection region. Division of region is based on level of significance and alternative hypothesis.

On the basis of normal distribution the critical values at various level of significance are presented in the following table:

Level of Significance	Two tail	Right Tail	Left Tail
.10 (10%)	± 1.645	1.28	-1.28
.05 (5%)	± 1.96	1.645	-1.645
.01(1%)	± 2.58	2.33	-2.33

(vi) Interpretation or Conclusion:- It is final step which refer to the acceptance or rejection of null hypothesis. The decision is made on the basis of the comparison of computed value of test statistic with the values in critical region. If computed value falls in the acceptance region, we accept the null hypothesis otherwise it is rejected.

Q-6 Define type I and Type II Error:-

Ans: **Type I Error (α):** It refers to the rejection of a null hypothesis when it is true.

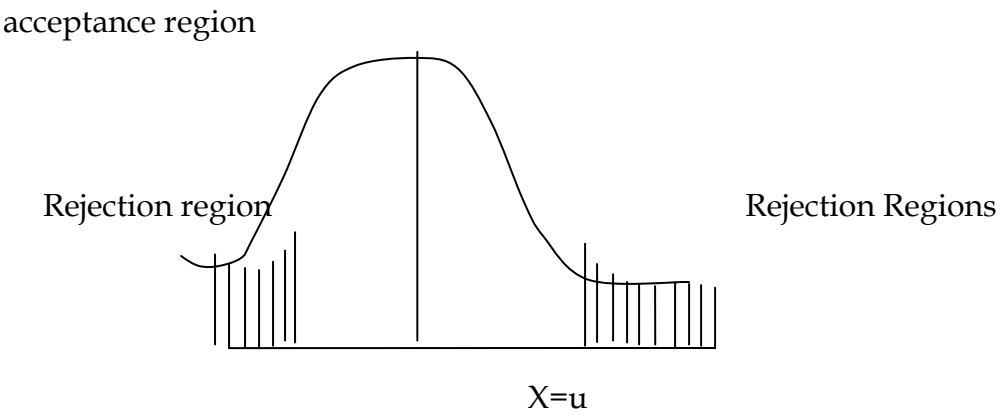
Type II Error (β) : Accepting a null hypothesis when it is false is called type II erro which is symbolized by (β) beta.

Decision

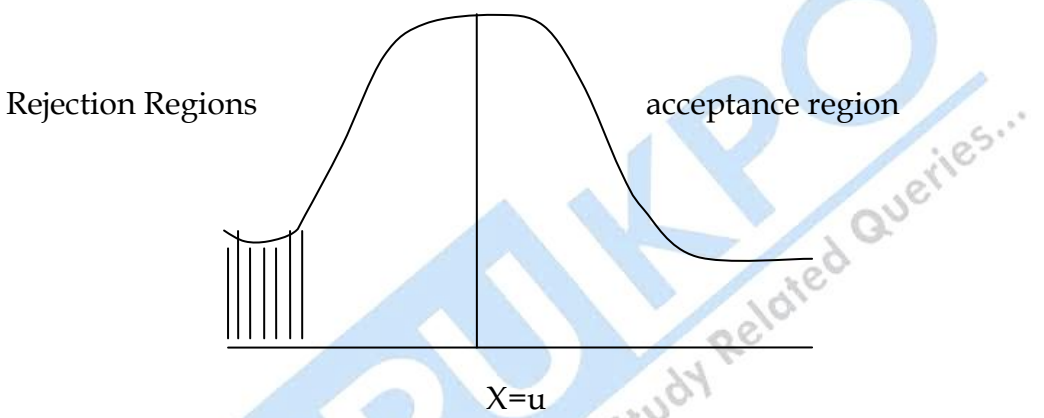
Condition	accept H_O	Reject H_o
H_0 is false	Incorrect decision(Type II error)	Correct decision.
H_O is true	Correct decision	incorrect decision(type I error)

Q-7 Discuss two tailed and one tailed tests of hypothesis.

Ans: **Two Tail Test:** In two tail test null hypothesis is rejected if the sample value is significantly higher or lower than the hypothesized value of the population parameter. The two fail test mainly includes two rejections across at both the ends of sampling distribution



2- One tail test :- In one tail test the rejection region is located only at one end of the distribution either at the left hand side or at right hand side.



Q-8 Discuss the test of significance in special reference to large sample.

Ans: one the basis of size of sample and nature of sample the test of significance ins divided into two parts:-

(i) Test of significance in large variables and

(ii) Test of significance in small variable

A sample is regarded as large only when its size is equal to 30 or it exceeds 30($n \geq 30$) and a sample is regarded small when the size of sample is less than 30 that is $n<30$ large samples are further divided into two parts:-

a- Sampling of attributes

b- Sampling of variables

(a) Sampling of Attributes: Attributes refers to the qualitative characteristic of a phenomenon which cannot be measured numerically. The sampling of attributes is regarded as the selection of sample from a population from the view point of presence and absence of a specific characteristic. The presence of attributes is symbolized by a while the absence by α.

The selection of a unit from population is known as event denoted by 'n'. The presence of attribute is regarded as success (P) whiles the absence of attribute as failure (q). The total no of items are denoted by N

The various tests of significance for attributes are discussed under the following hands:-

(i) Tests for number of success
(ii) Tests for proportion of successes
(iii) Tests for difference between two proportions(same universe)
(iv) Test for difference between two proportions (different universe)
(v) Comparing a sample proportion with a pooled proportions
(vi) Tests for Number of success

$$Z = \frac{Difference}{S.E.}$$

Where
Difference = Difference between observed and expected frequency
S.E = $\sqrt{npq}$
n= Size of sample
p= Probability of success
q= Probability of failure
(iii) Tests for proportion of success

$$Z = \frac{Difference}{S.E}$$

Difference = Difference between actual & expected proportion of success

$$S.E = \sqrt{\frac{pq}{n}}$$

(iv) Test for Difference between two proportions (same universe)

$$Z = \frac{P_1 - P_2}{\sqrt{p_o q_o \left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

(iv) Test for difference between two proportions (different universe)

$$Z = \frac{P_1 - P_2}{\sqrt{\left[\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}\right]}}$$

(v) Comparing a sample proportion with a pooled proportion

$$Z = \frac{P_0 - P}{\sqrt{P_O Q_O \frac{n_2}{n_1(n_1+n_2)}}}$$

P_O = Combined proportion
N_1+N_2 = Size of population

N_2 = (Size of population – n_1)

(b) Sampling of variables: - Variable refer to quantitative data which can be measured numerically such as height, weight, age , salary etc.

Test of significance in variables is carried out for the following purpose.

(a) To find out the estimated rate for parameter

(b) Comparison of observed and expected values.

(c) Applying the test of significance for tested the reliability of the estimates.

Various tests for significance for variables are discussed under the following heads:

1- Test of significance of mean

$Z = \frac{x-u}{S.E}$

$S.E = \frac{S}{\sqrt{n}}$

S= Standard deviation

N= Size of sample

2- Test of significance of difference between Two means:-

To test the significance of difference between two sample mean following formulas can be used

(i) Both sample have been drawn from same population :-

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma P\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

(ii) Both Samples are drawn from different populations:-

$$= \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(\sigma_1)^2}{n_1} + \frac{(\sigma_2)^2}{n_2}}}$$

(iii) If standard deviation of population is not known:-

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}}$$

(iv) Mean of first sample is compared with combined mean:-

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{n_2}{n_1(n_1+n_2)} + \frac{(s_2)^2}{n_2}}}$$

Comparing mean of a sample with the pooled mean of two samples:-

$$Z = \bar{X}_1 - \overline{X_{1.2}}$$

$$\sqrt{\sigma^2 \frac{n_2}{n_1 (n_1+n_2)}}$$

Multiple Choice Question

Select the correct alternative out of the given ones:

1- Sampling errors are :

(a) Caused by inaccurate measurement

(b) the result of the chance selection of the sampling units

(c) of no great concern

(d) larger for a census than for a sample

2. Non-sampling errors are

(a) Caused by inaccurate measurement

(b) the result of the chance selection of sampling units

(c) of no great concern

(d) always larger for a census than for a sample

3. If μ_2 is the population mean, and and σ_x^2 is the population variance, then the mean and variance of a sample are equal to :

(a) μ_2 and σ_x^2 (b) μ_2 /n and σ_x^2 /n

(c) μ_2 /n and σ_x^2 /n^2 (d) μ_2 and σ_x^2 /n

4. A sample consists of

(a) all units of the population (b) 50per cent units of the population

(c) 5 per cent units of the population (d) any fraction of the population

5. If we sample without replacement,

(a) it is important to consider the size of the

(b) A larger sample relative to the size of the population
Sample relative to the size of the population is preferred because it will reduce the sampling error.

(c) the sample size is not important

(d) use a smaller sample.

6. If we sample without replacement and the sample is large relative to the population,

(a) the sample variance will be small. (b) the sample mean will be large.

(c) the sample variance will large (d) the sample variance must be adjusted

7. Probability of selection varies at each subsequent draw in:

(a) Sampling without replacement (b) sampling with replacement

(c) both (a) and (b) (d) neither (a) or (b)

8. The number of possible samples of size n from a population of N units with replacement is :
(a) n^2 (b)N^2
(c) ∞ (d) none of the above

9. A function of variates for estimating a parameter is called:
(a) an estimate (b) an estimator
(c) a frame (d) a statistic

10. Let the standard error of an estimator T under srswor is more than the standard error of T under stratified randomly sampling. Then T under stratified sampling as compared to under srswor is :
(a) More reliable. (b) Equally reliable
(c) less reliable (d) not comparable

11. Which of the following basis distinguishes cluster sampling from stratified sampling?
(a) A sample is always drawn from each stratum
(b) clusters are preferably heterogeneous whereas strata
Where as no sample of elementary units is are taken as homogeneous as possible
drawn from cluster
(c) all of the above
(d) Small size clusters are better whereas there is no such restriction for stratum size

12. Non-response in surveys mean:
(a) Non-return of questionnaire by the respondents
(b) non-availability of respondents.
(c) Refusal to give information by the respondents
(d) all of the above

13. Choose the pair of symbols that best completes this sentence: _____is a parameter, wheras_____is a static.
(a) N, μ (b) N, n
(c) σ ,s (d) all of these

14. Regarding the number of strata, which statement is true?
(a) More the number of strata, poorer it is.
(b) Lesser the number of strata, better it is
(c) More the number of strata, better it is
(d) Not more than ten items should be there in a stratum.

15. In random sampling, we can describe mathematically how objective our estimates are. Why is this?
(a) We always know the chance that a
(b) Every sample always has an equal chance of being Population element will be included in the
(c) (a) and (b) but not (c)
Sample
(d) All the samples are of exactly the same size and can be counted.

16. The magnitude of the standard error of an estimate is an index of its:
(a) Accuracy (b) precision
(c) efficiency (d) all of these

17. An estimate based on a fixed set of values of a sample always posses:
(a) a single value (b) any value
(c) a value equal to one (d) all of these

18. If the sample values are 1,3,5,7,9, the standard error of sample mean is :
(a) S.E.=$\sqrt{2}$ (b) S.E.=1$\sqrt{2}$
(c) S.E=2.0 (d) S.E.=1/2.

19. Which of the following statements does not hold good in case of stratified sampling?
(a) Stratified sampling is always good. (b) Stratified sampling is convenient.
(c) Reduces error for fixed cost (d) enables to gather information about different stratum separately.

20. What precaution(s) make(s) cluster sampling more efficient?
(a) Choosing clusters having largest within
(b) By taking clusters of small size.Variation.
(c) Choosing cluster having least variation cluster
(d) All of the above

ANSWER KEY:

1-B	**2-A**	**3-D**	**4-D**	**5-A**	**6-D**	**7-A**	**8-D**	**9-B**	**10-A**
11-D	**12-B**	**13-C**	**14.D**	**15-D**	**16-B**	**17-A**	**18-A**	**19-B**	**20-D**

Chapter-10

Tests of significance - 't' test & 'z' test

Q.1 What do you understand by 't' test?

Ans. 't' test is used to carry out the test of significance for small samples. Use of 't' distribution for the purpose of estimation is required whenever the sample size is 30 or less than 30 and standard deviation of population is not known.

Q.2 What is Degree of freedom?

Ans. The degree of freedom refers to the number of values in a sample we can specify freely. It can be obtained by deducting number of constraints from the total number of observations in the sample. To calculate the degree of freedom following formula can be used:-
Degree of freedom (df) = (C-1) (r-1)
Where –
C = No. of columns
r = No. of rows

Q.3 What procedure adopted for testing the significance of the mean in small samples?

Ans. Procedure is as follows:-

- ⇨ Null Hypothesis – The t-distribution is a continuous distribution where the value of mean, mode and median can be zero.
  HO = $\bar{x} = \mu = 0$
  H1 = $\bar{x} = \mu \neq 0$
- ⇨ Level of significance – Usually the hypothesis is tested at 5% or 1% level of significance.
- ⇨ Unbiased Estimate of population – The standard deviation of population is unknown. Therefore in place of it unbiased estimate $\hat{s}$ is used.

$$\hat{s} = \sqrt{\frac{\sum(X-\bar{x})^2}{n-1}}$$

- ⇨ Calculation of t – statistic:-

$$t = \frac{\bar{x}-\mu}{\hat{s}/\sqrt{n}} \text{ or } \frac{\bar{x}-\mu}{\hat{s}/\sqrt{n-1}}$$

Where:-

$$S = \sqrt{\frac{\sum(X-\bar{x})^2}{n}}$$

⇨ Critical value of t – The critical or tabulated value of t is obtained at a specified level (∝-0.05 or 0.01) of significance for (n-1) degree of freedom.

⇨ Decision – If the calculated value of t is more than the tabulated value, it falls in rejection region and the null hypothesis is rejected and the calculated value of t is less than the tabulated value, we accept the null hypothesis and the difference is insignificant.

Q.5 Discuss Fisher's Z test.

Ans. To test the significance of correlation of coefficient in small samp0le Prof. Ronald Fisher originated a specific device according to which the coefficient of correlation is transformed into Z statistics.

$$Z = \frac{1}{2}\log\left(\frac{1+r}{1-r}\right)$$

Q.6 What do you understand by 'F' test?

Ans. The F test is used to test the significance of difference between two variances. By using the F test it is ascertained whether the two samples can be regarded as drawn from the normal population having the same variance.

$$F = \frac{\text{Larger Estimate of the population variance}}{\text{Smaller Estimate of the population variance}}$$

Where:-

$$S_1^2 = \frac{\sum(X - \bar{x}1)}{n_1 - 1}$$

$$S_2^2 = \frac{\sum(X - \bar{x}2)^2}{n_2 - 1}$$

Multiple Choice Question

Q-1 Student 't' Distribution was founded in L:

(a) 1908 (b) 1900

(c) 1896 (d) 1902

Q-2 Student "t" distribution was propounded by:

(a) Bowley (b) Connor

(c) W.S.Gosset (d) Tippett

Q-3 The shape of 't' distributin is :

(a) Bell (b) Round

(c) Square (d)flat

Q-4 What are the formula of 't' distribution

(a) $t\ \frac{\bar{x}-u}{\sigma\bar{x}}$ (b) $\frac{\bar{x}-\sigma}{\bar{x}}$

(c) $\frac{\bar{x}}{u}\ x\ \sigma\ \bar{x}$ (d) $\frac{\bar{x}-6}{\sigma\ \bar{x}}$

Q-5 'z' test is propounded by :

(a) Ronald Fisher (b) Connor

(c) borssen (d)Bowely

Q-6 What are the formula of Degrees freedom :

(a) (df)=(C-2)(r-2) (b) (df)=(c-1)(r-1)

(c) (df)=(c-c)(r-3) (d)(df)=(r-3)(c-2)

Q-7 What is d-

(a) $\bar{x}$-x (b) $\bar{x}$-u

(c) $\bar{X}$-x (d) X-$\bar{X}$

Q-8 What is $\hat{s}$ -

(a) *Combined estimate of standard deviation* (b) *Mean*

(c) *Assume Mean* (d) *Actual Mean*

Q-9 What is formula $\bar{D}$ –

(a) $\frac{\sum n}{d}$ (b) $\frac{Dn}{\sum n}$

(c) $\frac{\sum D}{N}$ (d) $\frac{\sum D^2}{N}$

Q-10 What is the full form of df-

(a) Degree of freedom (b) Deviation of freedom

(c) Deviation of mean (d) Distribution of freedom

ANSWER KEY

1-A	**2-C**	**3-A**	**4-A**	**5-A**	**6-B**	**7-D**	**8-A**	**9-C**	**10-A**

Chapter 11

Chi-Square Test (X^2 Test)

Q-1 What do you mean by chi-square! Discuss its Characteristics.

Ans. The chi-square test is used to determine if the two attributes are independent of each other. It is a non-parametric test where no assumption is mode about the parameters of population. It is a measure to evaluate the difference between observed frequencies and expected frequencies to examine whether the difference so obtained is due to a chance factor or due to sampling error.

Characteristic of X^2 test

(i) Chi-square test is based on frequencies not on parameters.
(ii) It is a non-parametric test where no parameters are required.
(iii) Additive property is also found in chi-square test.
(iv) It is useful to test the hypothesis about the independence of attributes.
(v) It can be used in complex contingency tables.
(vi) It is very widely used for research purposes in behavioral science.

Q-2. Define Assumption of X^2 test!

Ans. (i) The sample should be selected randomly.
3- All items of sample should be independent of each other.
4- The frequency of any cell should not be less than 5 otherwise it will lead us to make incorrect inferences. If the frequency is less than 5, we apply yate's correction.
5- The total no of items should be more than 50
6- The constraints on cell frequencies should be linear.

Q-3. How to calculate the value of X^2?

Ans. To test the independence of attributes in chi-square test following procedure is adopted :-

(i) Null hypothesis:- Null hypothesis is adopted and it is assumed that both the attributes are independent and the difference between observed frequencies and expected frequencies is NIL

$H_o = f_o = f_e$

(i) Calculation of expected frequencies: - With the help of observed frequency, calculate the expected frequencies.

Expected frequency : $\frac{Total\ of\ related\ column\ X\ total\ of\ related\ Row}{Total\ Number}$

$A_1\ B_1 = \frac{A_1 X B_1}{N}$

(ii) Calculate the difference between observed frequencies (fo) and expected frequencies (f1) then squared and used to formula for calculate the value of chi-square (X^2)

$$X^2 = \sum\left[\frac{F_O - F_E}{F_E}\right]^2 \text{ OR } \sum\left[\frac{O-E}{E}\right]^2$$

(iii) Degree of freedom :-

(a) In contingency table d.f = (c-1) (r-1)

Excepting a contingency table d.f = (n-1)

n= number of items

(iv) Interpretation: - If calculated value of chi-square is more than the table value at 1% or 5% level of significance in different degree of freedom then null hypothesis is rejected and alternative hypothesis is accepted.

Q-4 Write short notes or the following:

Ans.

(a) Yate's Correction

(b) Degree of freedom

(a) Yate's Correction :- The chi-square (X^2) is a continuous distribution based on the assumption that the ultimate class frequency is not less than 5 otherwise the value of X^2 null be overestimated and may cause too many rejections of null hypothesis. Therefore to maintain the continuity of X^2 & to draw the correct inference in 2x2 table we apply a correction given by F. Yates known as "Yates" correction for continuity.

In 2x2 table the cell relating to which the frequency is less than 5 beat being least also 0.5 is added in such a way that the marginal total of remaining three cells not changed. By this procedure the difference between observed frequencies and expected frequencies is reduced by 0.5 and the value of X^2 approaches to the actual one and correct inferences can be drawn.

(b) Degree of freedom:- The determination of degree of freedom play an important role in X^2 to test a hypothesis. The value of chi depends on the number of degree of freedom. The dates that are given in the form of a series of variable are in rows or columns. The number of frequencies which can be calculated independently, are known as the degree of freedom

In case of data that are given in the form of a series of variables in rows or columns, the degree of freedom will be the number of item is the total series less one (v=n-1)

In case the number of frequencies are put in cells in a contingency table, the degree of freedom will be equal to the product of the number of rows less one and the number of columns less one V=(R-I) (C-I)

Q-5 Discuss the area of chi-square application?

Ans. Area of chi-square application are as following:-

(i) Test of Independence: - Chi-square test is test of independence of 2 or more attributes are associated or dissociated. By applying the technique of X^2 it can be ascertained whether vaccination is effective in preventing the cholera or not., literacy and employment are associated or not, smoking and cancer are independent or associated etc.

(ii) Test of Homogeneity: The chi-square test is also a test of homogeneity which is used to determine whether two or more independent random samples have been drawn from the same population not.

(iii) Test of Goodness of Fit: - The chi-square test is designed to test the difference between observed frequency distribution and theoretical frequency distribution. By testing this difference it can be found out whether the observed frequency distribution is approximating the theoretical frequency distribution or not the difference between them is significant or insignificant.

Multiple Choice Question

Q-1 When 2 data are compared for particle purpose there, Difference stated through the technique in called

(a) Variance (b) Analysis of variance

(c) Decision theory (d) Business forecasting

Q-2 The formula of calculating correction factor is

(a) $\frac{T}{N}$ (b) $\frac{T^2}{N}$

(c) $\frac{T^2-1}{4}$ (d) $\frac{T^2}{2}$

Q-3 When the source of variance is between the samples what the degree of freedom is.

(a) k-1 (b) N-K

(a) N-1 (d) $N^2 - 1$

Q-4 The technique of analysis of variance was first originated by :

(a) Fisher (b) davies

(c) Connor (d) Spearmean

Q-5 The value of F is mostly obtained according to level of significance

(a) 5 or 1% (b) 7 or 8 %

(c) 6 or 4 % (d) 2 or 3%

Q-6 If the value of $T^2=0$, value of N=12 then the value of C.F. is

(a) 15 (b) 12

(c) 0 (d) 1

Q-7 When the source of variance is between the rows the degree of freedom is:

(a) C-1 (b) r-1

(c) (C-1)(r-1) (d) N-1

Q-8 If the value of T^2 is 1600 and N=16 then the value of C.F.is

(a) 100 (b) 200

(c) 102 (d) 105

Answer Key :

1-B	2-B	3-A	4-A	5-A	6-C	7-D	8-A

Chapter 12
Analysis of Variance

Q-1 Explain the assumption of Analysis of variance.

Ans. Analysis of variance is the classification and cross –classification of statistical data with the view of testing whether the means of specific classifications differ significantly or they are homogenous.

Assumption:

Following are the assumption of analysis of variance.

(i) The sum of the variances of different components is equal to the total variance.

(ii) The units of the sample are selected independently and randomly, otherwise the utility of variance ratio is lessened because of the presence of correlation

(iii) The population from which samples have been selected is normally distributed.

Q-2 Discuss the utility of analysis of variance?

Ans. Following are the utility of analysis of variance:

(i) The analysis of variance can be used to test the significance of linearity of regression, simple and multiple corrections.

(ii) Homogeneity of the data can also be studied in analysis of variance because the data are classified into different parts on two bases.

(iii) The F-Co-efficient or variance ratio is used to find whether the difference between the variance of different samples is significant.

(iv) The analysis of variance is also used to test the hypothesis of equality amongst several means.

Q-3 Explain briefly the technique of analysis of variance.

Ans. The technique of analysis variance can be divided into two parts:

(i) One way classifications (2) Two-way classification

(i) One way classification: - The one way classification refers to the data classified on the basis of one factor only such as effect of fertilizers on four types of plots. It can be computed by following methods.

Direct Method :- Following steps one followed to calculate the value of F

(I) Hypothesis $=H_o= u_1=u_2=u_3$

$H_1=u_1\neq u_2\neq u_3$

(II) Variance between samples

Calculate = $\bar{X}_1$, $\bar{X}_2$, $\bar{X}_3$

Calculate grand mean $\frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3}{n_1+n_2+n_3}$

calculate SSB (Sum of squares between samples)

SSB = $n_1(\bar{X}_1 - \bar{\bar{X}})^2 + n_2(\bar{X}_2 - \bar{\bar{X}})^2 + n_3(\bar{X}_3 - \bar{\bar{X}})^2 + n_K(\bar{X}_K - \bar{\bar{X}})^2$

Find degree of freedom

Df_1 = K-1 (K = No of samples)

(ii) variance within samples :-

calculate mean of each sample $\bar{X}_1, X_2$ ---- X_n

SSW= $\sum(\bar{X}_1 - X_1)^2 + \sum(\bar{X}_2 - X_2)^2 + \sum(\bar{X}_3 - X_3)^2 \ldots\ldots \sum(X_K - X_K)^2$

Find degree of freedom

Df_2 = N-K

(iii) Analysis of variance of Table (ANOVA Table)

Source of variance	**Degree of freedom**	**squares(SSB)**	**Variance (MSB)**	**Variance ratio OR F**
Between samples	**K-1**	$\sum[n(\bar{X} - \bar{\bar{X}})]^2$	**SSB÷(K-1)**	$\frac{MSB}{MSW}$
Within samples	**N-K**	$\sum(X - \bar{X}_K)^2$ **(SSW)**	**SSW÷N-K=MSW**	
Total	**N-1**	$\sum(X - \bar{\bar{X}})^2$ **(SST)**	**SST÷(N-1)**	

N= Total number of items

K= No. of samples

(iv) Calculation of variance ratio (f)

F= $\frac{Variance\ between\ samples}{Variance\ within\ sample}$

F= $\frac{Grater\ variance}{Smaller\ variance}$

(v) The table value of F is obtained from the table according to the degree of freedom. The degree of freedom of greater variance is known as V_1 and that of smaller variance of V_2

(vi) Interpretation of F: The computed value of F is compared with the table value. If the calculated value of is more than the table value the null hypothesis is, rejected and difference is regarded significant and vice versa.

(vii) Level of significance: The table value of F is obtained according to different level of significance but practically 5% or 1% level of significance is used.

Indirect Method

Indirect method is followed as follow:-

The total of all sample units are obtained

$\sum X_1 + \sum X_2 + \sum X_3$ ----------

Each and every unit of sample is squared and totaled.

$\sum X^2 + \sum X_2^2 + \sum X_3^2 +$----------

(iv) Correction Factor:-

C= $\frac{[\sum X_1 + \sum X_2 + \sum X_3 \pm ----]^2}{N}$ OR $\frac{T^2}{N}$

T= Total of the values of all the samples

(v) Total sum of squares (TSS):-

$$TSS=\left[\sum x_1^2 + \sum x_2^2 + \sum x_3^2 - - \sum x_k^2\right] - \frac{T^2}{N}$$

(vi) Sum of square between samples (SSB)

$$SSB = \left[\frac{(\sum x_1)^2}{n_1} + \frac{(\sum x_2)^2}{n_2} + \frac{(\sum x_3)^2}{n_3} + \frac{(\sum x_k)^2}{n_k}\right] - \frac{T^2}{N}$$

(vii) Sum of squares between samples (SSB) :-

SSW= $SST - SSB$

(viii) ANOVA Table is prepared

(ix) Variance Ratio (F) is calculated as follow $F = \frac{variance\ between\ the\ samples}{variance\ neither\ the\ samples}$

(x) Inference is made by comparing the computed value of F with table value of f.

3-Coding Method:

The variance ratio (f) remains unaffected if all the values of the samples are divided or multiplied by a common factor or a common factor is deducted from or added to every value of the samples

The coding method is use full in the case when the values are large and calculations become lengthy.

2-Two-way classification:

Following steps should be followed to calculated variance ratio (f) in the case of two way classification

(i) Coding method can be used to simplify the calculations
(ii) Calculations of correction factor:-
C.F. $=\frac{T^2}{N}$
(iii) Total sum of square (TSS)
TSS= $\sum x_1^2 + \sum x_2^2 + \sum x_3^2 - - \sum x_4^2 - \frac{T^2}{N}$
(iv) Sum of Squares between columns:
$= \left[\frac{(\sum x_c)^2}{n_c}\right] - \frac{T^2}{N}$
$= \sum x_c^2$ = Total of squared values in each columns
N_c= Number of items in each columns

(v) Sum of squares between Rows(SSR)
SSR=$\sum \left[\frac{\sum Xr^2}{n_r}\right] - \frac{T^2}{N}$ $\sum X_r^2$= sum of squared value of each row
N_r = No of items in each row

(vi) Sum of squares of the Residual (SSE)
SSE= TSS-(SSC+SSR)

(vii) Number of degree of freedom:
It is calculated as follows:
No. of degree of freedom between column =(C-1)
No. of degree of freedom between rows =(r-1)
No. of degree of freedom for residual (c-1) (r-1)
Total No of degree of freedom = (N-1) OR (Cr-1)

VIII) -Prepare ANOVA Table – Two way classification

Source of variance	Sum of Square	Degree of Freedom	Mean sum of square	F Ratio
Between Columns	**SSC**	**C-1**	**MSC=SSC/(C-1)**	**F=$\frac{MSC}{MSE}$**
Between Rows	**SSR**	**(r-1)**	**MSR=SSR/(R-1)**	**F=$\frac{MSR}{MSE}$**
Residual	**SSE**	**(C-1)(R-1)**	**SSE/(C-1)X(R-1)= MSE**	

(IX) Interpretation: The calculated value of F is compared with the table value and if the calculated or computed value of F is greater than the table value at specific levels of significances the null hypothesis is rejected and concluded that the difference is significant otherwise vice-versa.

Multiple Choice Question

Q-1 When 2 data are compared for particle purpose there, Difference stated through the technique in called

(b) Variance (b) Analysis of variance

(c) Decision theory (d) Business forecasting

Q-2 The formula of calculating correction factor is

(b) $\frac{T}{N}$ (b) $\frac{T^2}{N}$

(c) $\frac{T^2-1}{4}$ (d) $\frac{T^2}{2}$

Q-3 When the source of variance is between the samples what the degree of freedom is.

(b) k-1 (b) N-K

(b) N-1 (d) $N^2 - 1$

Q-4 The technique of analysis of variance was first originated by :

(b) Fisher (b) davies

(c) Connor (d) Spearmean

Q-5 The value of F is mostly obtained according to level of significance

(b) 5 or 1% (b) 7 or 8 %

(c) 6 or 4 % (d) 2 or 3%

Q-6 If the value of $T^2=0$, value of N=12 then the value of C.F. is

(b) 15 (b) 12

(c) 0 (d) 1

Q-7 When the source of variance is between the rows the degree of freedom is:

(b) C-1 (b) r-1

(c) (C-1)(r-1) (d) N-1

Q-8 If the value of T^2 is 1600 and N=16 then the value of C.F.is

(b) 100 (b) 200

(c) 102 (d) 105

Answer Key :

1-B	2-B	3-A	4-A	5-A	6-C	7-D	8-A

Chapter-13
Decision Theory

Q.1 What is Decision Theory?

Ans. In every business organization a manager has to take a number of decisions. The job of a manager is related with planning, organizing, monitoring and controlling. Decision making is an integral part of all the functions. Decisions are based on the criteria decided by the organizational objectives of business. Problem of decision making will arise in only those situations where there are a number of ways in which a job can be performed.

Q.2 Discuss the terminology used in Decision making.

Ans. (i) **Acts (Alternatives):-**
Acts are the alternative courses of action that are available to the decision maker. The decision making procedure involves selection of one act from a set of alternative acts. Acts are denoted by a1, a2, a3 …etc.

(ii) **Events:-**
Events are also called 'states of nature' or 'outcomes'. Events identify the occurrences which are beyond the control of the decision maker and which determine the level of success for a given act. The various events are symbolized by E1, E2, E3 …. Etc.

(iii) **Pay off:-**
Each combination of a course of action and an event is associated with a pay off Thus pay-off measures the net benefit to the decision maker that results from a given combination of decision alternatives and events.

(iv) Pay-off table:- Various pay off can be put in the shape of a table to form a pay off table. A pay off table shows the relation between all possible states of nature, all possible acts and the values associated with the consequences.

(v) **Opportunity loss:-**
The opportunity loss' or 'regret' is the difference between pay-off realized and the maximum pay-off which could have been realized if another strategy was chosen. This loss is incurred because of failure to adopt the best possible outcome.

Q.3 Write the methods used for decision making under uncertainty.

Ans. For decision making under uncertainty following methods can be used

i. Maxi Max criterion
ii. Maxi-Min criterion
iii. Mini-max Regret criterion
iv. Coefficient of optimism criterion – Hurnicy criterion
v. Laplace criterion

i. Maxi-Max Criterion:-

Under this criterion the decision maker chooses the strategy that makes the best of the best. The decision maker is optimistic by nature and he would always think that the state of nature would be the best from his point of view. In this method first the maximum outcome within every alternative strategy is located and then the alternative with the maximum pay-off is selected.

ii. Maxi-Mini Criterion:-

Under this criterion the decision-maker is assumed to be pessimist by nature and he would always think that the worst would happen and would like to safeguard himself against this condition. In this method first the minimum pay-off for each act is located and then the alternative which maximizes the minimum pay-off is selected.

iii. Mini-Max Regret Criterion:-

In this criterion the decision-maker would select the strategy in which the maximum regret or cost is the lowest. The regret or the opportunity loss is simply the difference between pay-off realized.

iv. Coefficient of optimism criterion:-

In this criterion degree of optimism is represented by $\propto$(alpha), the coefficient of optimism varying between 0 and 1. If $\propto = 0$ it denotes total pessimism and if $\propto=1$, it denotes total optimism. After

determining $\propto$ the value of Hurwitz is found out by using the following formula:

Hai = Maximum Pay-off from alternative ai x $\propto$

+ Minimum Pay off from alternative ai x (1-$\propto$)

Hai = Value of Harwictz

v. Laplace Criterion:-

When no definite information about the probability of occurrence of various states of nature laplace can be used. This involves three steps:-

i) Assigning equal opportunity to each state of nature.
ii) Calculating the expected pay-offs.
iii) Selecting the strategy whose expected pay-off is maximum

Q.4 Discuss various Methods of Decision making under risk?

Ans. Under risk decision-maker forces several events and he can't predict the outcome of an event. The decision-making is probabilistic under Risk the following decision-criterion are used for evaluating the alternative strategies-

(i) Expected Monetary Value (EMV) Criterion:-

First a Pay-off table is constructed and then probabilities are assigned to the various events. It is to be noted that the events exhaust all possible alternative; the probabilities add to one EMV is calculated for each act by multiplying the pay-off values by the assigned probabilities and add the resulting values to obtain EMV of the act, Under Risk the decision. Maker chooses that strategy which has highest EMV.

(ii) Expected opportunity loss (EOL) Criterion:-

Under Risk criterion the expected opportunity loss or expected value of regret is minimized. Expected opportunity loss for any course of action is calculated by multiplying the conditional opportunity loss by its probability of occurrence. In EOL criterion the decision-maker would choose the strategy with the minimum expected opportunity loss.

(iii) Expected Value of Perfect Information (EVPI):-

In this it is assumed that the decision-maker has authentic and perfect information available. This improves the quality of decision.

EVPI = EPPI – EMV

EVPI = Minimum EOL

Where

EPPI = Expected Profit with Perfect Information

EMV= Expected Monetary Value of optimal Act.

Multiple Choice Question

Q-1 Decision making process involves:

(a) Decision (b) management

(c) Control (d) planning

Q-2 ----------are the alternative courses of action that are available to the decision maker:

(a) acts (b) events
(c) decision (d) data

Q-3 Events are denoted by what symbol:
(a) E_1 (b) K
(b) $\overline{X}$ (d) w

Q-4 The decision making is :
(a) certain (b) probabilistic
(c) doubtful (d) None of the above

Q-5 When the stages are shown by tree formation is called:
(a) Tree diagram (b) pie diagram
(c) Bar diagram (d) picto gram

Q-6 If EPPI =71 and EMV=54 then the value of EVPI is
(a) 20 (b) 17
(c) 24 (d) 16

ANSWER KEY:

1-A	**2-A**	**3-A**	**4-B**	**5-A**	**6-B**	

Chapter-14

Correlation

Q-1 What is the meaning of Correlation?

Ans. When two data series are connected in such a way that change in one cause change in the other accordingly, or increase or decrease in one leads to change in the same or opposite direction in the other, they are called correlated. Mathematical scale of side by side change in tendency of two series of facts is called correlation.

Q-2 How many types of correlation?

Ans. Correlation may be the following types:-

Types of Correlation

On the basis of change

Basis on No.In Directionin

1. Positive correlation
2. Negative correlation
3. Partial

Basis on change

Ratio of variable

1. Linear correlation Simplic
2. Non-linear correlation Multiple

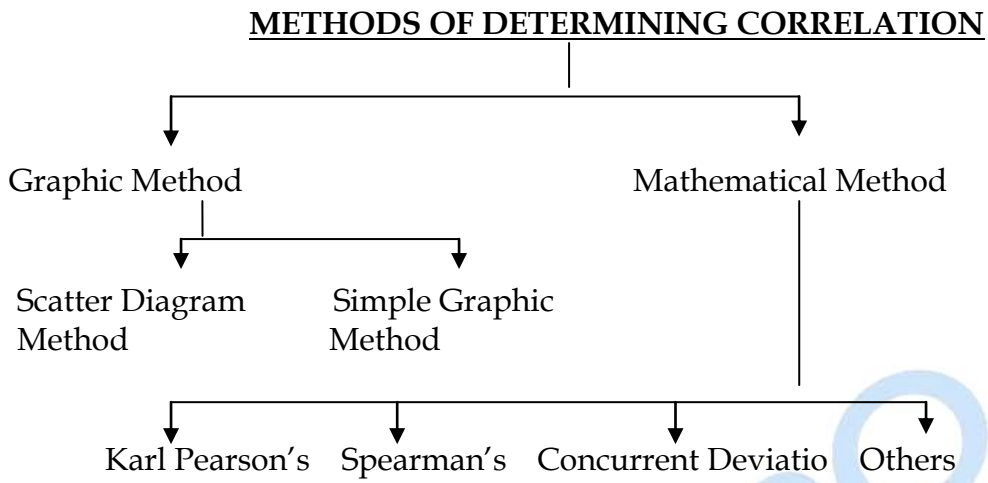
Q-3. Define Degree of Correlation.

Ans. Change in two connected data series may be either in equal ratio or changing ratio. This situation can be measured through coefficient of correlation. This change may be either positive or negative. This coefficient may be in the following degrees:-

Degree of Correlation	Positive	Negative
1. Perfect correlation	+1	-1
2. High degree correlation	+.75 to + 1	-.75 to -1
3. Moderate degree correlation	+.25 to .75	-.25 to -.75
4. Low degree correlation	+ 0 to 0.25	- 0 to –0.25
5. Absence	Zero (0)	Zero (0)

Q.4 Discuss the different Method of Correlation.

Ans. Various Methods of determining correlation between two groups of series can easily be understood with the help of the following chart-



GRAPHIC METHOD

1. Scatter Diagram:-
It is a graphic Method for finding correlation between two groups of variables. One group of variables, an independent variable is shown on X-axis and the other group of variable a dependent one is shown on Y-axis. The movement of the pairs of these variables is plotted on the graph paper by 'dots' on it. The greater the scatteredness of the plotted points on the graph paper, the lesser is the correlation ship between two variables. If the points plotted come closely in the shape of a straight line, the higher is the degree of correlation ship.

2. Simple Graphic Method:-
Under this method individual values are plotted on the graph paper for both the variables. If both the curve drawn are closer to each other, one may infer that these exist a high degree of correlation.

MATHEMATICAL METHOD

1. Karl Pearson's:-

Direct Method:- $r = \frac{\sum dxdy}{N\sigma_x\sigma_y}$

Where:-

r = Correlation coefficient

$\sum dx = (x - \bar{x})$

$\sum dy = (y - \bar{y})$

$\sum dxdy$ = Product of corresponding deviation of x and y variables

σ_x = Standard deviation of x or $\sqrt{\frac{\sum d^2x}{N}}$

σ_y = Standard deviation of x or $\sqrt{\frac{\sum d^2y}{N}}$

Short Cut Method

$$r = \frac{\sum dxdy.N - (\sum dx.\sum dy)}{\sqrt{\sum d^2x.N - (\sum dx)^2}\sqrt{\sum d^2y.N - (\sum dy)^2}}$$

$\sum dx$= sum of deviation from assumed mean of x variable

$\sum dy$= sum of deviation from assumed mean of y variable

2. Spearman's Rank Difference Method:-

$$r_{(R)} = 1 - \frac{6\sum D^2}{N(N^2-1)}$$

$\sum d^2$ = Total of square of rank Difference

N = Number of pairs of observation.

Note: -In the formula we add the factor $\frac{m(m^2-1)}{12}$ to the value $\sum d^2$. Here m means the number of items an item has repeated. The correlation factor is to be added for each repeated value.

3. Concurrent Deviation Method

$$r_c = \pm\sqrt{\pm\frac{2(C-n)}{n}}$$

r_c = Coefficient of correlation by concurrent deviation method

C = Number of concurrent deviation

n = Number of pairs of observation compared

Q.5 What do you understand by multiple & partial correlation?

Ans. When we study of two or more dependent variables on one variable is known as multiple correlations.

Ex:- The combined effect of rainfall and fertilizers on the crop of wheat will be known as multiple correlation. It is devoted by R 1. 23

If we keep the effect of fertilizer constant and study only the effect of rainfall on the crop of wheat then it will be known as partial correlation. It is denoted by $r_{12.3}$

Multiple correlation coefficients

$$R_{1.23} = \sqrt{\frac{r^2 12 + r^2 13 - 2r\ 12\ r_{23}\ r\ 13}{1 - r^2 23}}$$

$$R_{2.13} = \sqrt{\frac{r^2 1 + r^2 23 - 2r\ 12\ r\ 13\ r\ 23}{1 - r^2 13}}$$

$$R_{3.12} = \sqrt{\frac{r^2 31 + r^2 32 - 2r\ 12\ r\ 13\ r\ 23}{1 - r^2 12}}$$

Partial Correlation Coefficient:-

$$R_{12.3} = \frac{r12 - r13.r23}{\sqrt{(1-r^2 12)(1-r^2 23)}}$$

$$r_{13.2} = \frac{r13 - r12.r23}{\sqrt{(1-r^2 12)(1-r^2 23)}}$$

$$r_{23.1} = \frac{r23 - r12.r13}{\sqrt{(1-r^2 12)(1-r^2 13)}}$$

Multiple Choice Question

Q-1 _____is analysis of co variation between two or more variables:
Ans: (a) Variation (b) Correlation
(c) Standard deviation (d) None

Q-2 The correlation is _____measures
Ans: (a) absolute (b) relative
(c) both (d) none

Q-3 The correlation is measure of _____relationship between two variables
Ans: (a) qualitative (b) quantitative
(c) both (d) none

Q-4 The degree to which the variables are interrelated to measured by
Ans: (a) Coefficient of correlation (b) Coefficient of variation
(b) Coefficient of regression (d) none

Q-5 The value of correlation coefficient lies between
Ans: (a) 0&1 (b)-1&0
(c) -0&8 (d)-1&1

Q-6 The correlation is perfect & positive if
(a) r=0 (b) r=-1
(c) r= 1 (d) none

Q-7 The correlation is perfect & negative if?
Ans (a) r=0 (b)r=-1
(c) r=1 (d) none

Q-8 There is no correlation if
Ans (a) r=0 (b) r=-1
(c) r=1 (d) none

Q-9 The variables are said to be independent if r=
(a) 1 (b)0
(c) -1 (d) 2

Q-10 The unit of correlation coefficient is :
Ans: (a) unit of x (b) unit of y
(c) unit of both (d) no unit

ANSWERS KEY :-

1-A	**2-B**	**3-B**	**4-A**	**5-D**	**6-C**	**7-B**	**8-A**	**9-B**	**10-D**

Chapter-15
Regression Analysis

Q.1 What do you understand by Regression?

Ans. The meaning of regression is just reversing the meaning of progression. Progression means "to move forward" while regression means "to move backward" or in statistical terms "the return to the mean value." Regression Analysis is a different technique that estimates the value of a concerned dependent variable. On the basis of the known value of the other independent variable.

Q.2 What is difference between correlation and Regression?

Ans.

Base	Correlation	Regression
1. Relationship	Correlation means the relationship between two or more variables	Regression expresses average relationship between two or more variables.
2. Cause and Effect	Correlation is unable to tell which series is the cause and which is the effect	In regression the given independent variable is the cause and dependent variable is the effect.
3. Effect of change in scale	Coefficient of correlation is in effected by the change of origin and scale.	Regression coefficient is effected by change in scale but not by change of origin
4. Symmetric	It is mutually symmetric r xy = ryx	Regression coefficient are not symmetric in x and y i.e. bxy ≠ byx
5. Limits	It is pure number lying between -1 and +1	bxy or byx either of may be more than -1 or +1
6. Application	It is confined only to the study of linear relationship between the variables.	It studies linear and non-linear relationship between the variables and therefore has much wider applications.

Q.3 Discuss the types of Regression Analysis.

Ans. Regression Analysis is of two types-

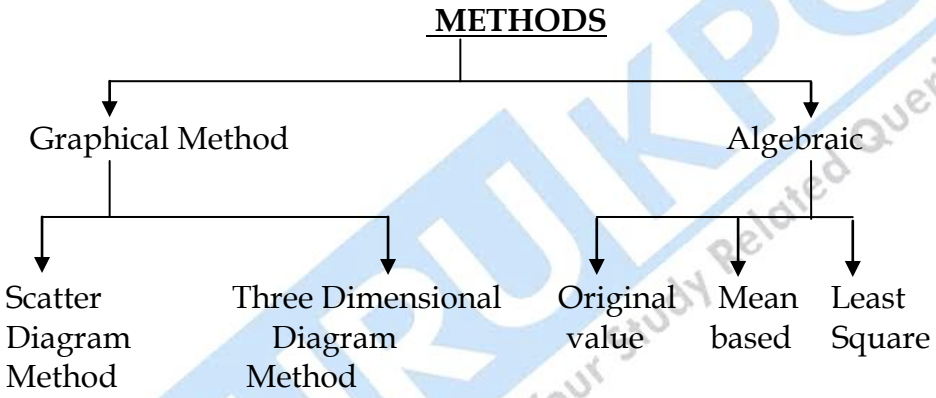
1. Linear and curvilinear regression,

2. 2. Simple and multiple regressions.

1. Linear and Curvilinear regression:- When variables of two related series x-y are marked on graph, it forms a scatter diagram. If two best fit lines are drawn passing through the marked dots, these very lines are called regression lines. If these lines are straight, it is called linear regression. If these lines are in the form of a smooth curve, it is called curvilinear regression.
2. Simple and Multiple regression:- When we study average relationship of only two related variables, it is called simple regression, when regression analysis method is used to examine relationship of more than two variables, it is called multiple regression.

Q.4 What are methods of constructing Regression Lines.?

Ans. Two Regression lines may be obtained in the following ways:-



1. Graphic Method:- In this method points in respect of concerned variables are arranged on graph paper and these points give a picture of scatter diagram. A regression line may be drawn in between these points by free hand or by scale rule.
2. Algebraic Method:- Regression equation –

x on y : $(x - \bar{x}) = bxy\,(y-\bar{y})$

y on x : $(y-\bar{y}) = byx\,(x-\bar{x})$

bxy = Regression coefficient of x on y

byx = Regression coefficient of y on x

(i) with the help of original value

$$bxy = \frac{\sum XY.N - \sum X \sum Y}{\sum Y^2.N - (\sum Y)^2}$$

$$byx = \frac{\sum XY.N - \sum X \sum Y}{\sum X^2.N - (\sum X)^2}$$

(ii) When deviation are taken from actual mean

$$bxy = \frac{\sum XY.}{\sum Y^2.} \qquad byx = \frac{\sum XY}{\sum X^2}$$

$$x^2 = (x-\bar{x})^2 \qquad y^2 = (y-\bar{y})^2$$

(iii) When deviations are taken from assumed mean

$$bxy = \frac{\sum dxdy.N - (\sum dx)(\sum dy)}{\sum dy^2.N - (\sum dy)^2}$$

$$byx = \frac{\sum dxdy.N - (\sum dx)(\sum dy)}{\sum dx^2.N - (\sum dx)^2}$$

(iv) When standard deviation and coefficient of correlation is given

$$Bxy = r\frac{\sigma x}{\sigma y} \qquad byx = r\frac{\sigma y}{\sigma x}$$

(v) When standard deviation and product of deviation from actual mean are given.

$$bxy = \frac{\sum xy}{N\sigma^2 y} \qquad byx = \frac{\sum xy}{N\sigma^2 x}$$

Least Square Method:-

When the data represent a sample from a large population, the least square line is a 'best estimate' of the population regression line.

$$\sum y = Na + b\sum x$$

$$\sum xy = a\sum x + b\sum x^2$$

Where:

$\sum y$ = Total of y series

$\sum y$ = Total of x series

$\sum xy$ = Total of product of x and y series

N = Number of observed pairs of value

'a' and 'b' are calculated by using are following equation:-

X on y	Y on x
$x = a + by$	$Y = a + b\,x$
$\sum x = Na + b\sum y$	$\sum y = Na + b\sum x$
$\sum xy = a\sum y + b\sum y^2$	$\sum xy = a\sum x + b\sum x^2$

Q.5 What is the concept of standard error of estimate in Regression?

Ans. Best estimate can be done in the dependent variable for the given value of independent variable with the help of regression lines. To examine how near to reality such estimates are 'Standard error of estimate' is used. Standard error of estimate is the mid scale of deviations of real values and computed or trend values of dependent series.

This is calculated as follow:-

X on Y

$$- Sxy = \sqrt{\frac{\sum(x-xc)^2}{N}}$$

X on Y

$$- Syx = \sqrt{\frac{\sum(y-yc)^2}{N}}$$

Or

$$Sxy = \sqrt{\frac{\sum x^2 - a\sum x - b\sum xy}{N}}$$

$$Syx = \sqrt{\frac{\sum y^2 - a\sum y - b\sum xy}{N}}$$

Multiple Choice Question

Q-1 _____ is a mathematical measure of relationship between two of more variables.

(a) Correlation (b) determination

(c) Alienation (d) regression

Q-2 Regression analysis confined to the study of only

(a) Simple (b) complex

(c) Multiple (d) none of these

Q-3 The regression analysis for studying more than two variables is known as regression.

(a) Simple (b) complex

(c) Multiple (d) none

Q-4 If the regression curve is a straight line at is called as _____regression

(a) Simple (b) linear

(c) Complex (d) non linear

Q-5 Number of regression lines is

(a) One (b) two

(c) Three (d) none

Q-6 If the regression is not straight line then it is called as _____regression

(a) Linear (b) non linear

(c) Simple (d) multiple

Q-7 The degree of linear regression equation is

(a) 1 (b) 0

(c) 2 (d) none

Q-8 _____is the line which gives the best estimate of one variable with other

(a) Simple line (b) correlation line

(c) normal line (d) regression line

Q-9 The two lines of regression are

(a) X on Y (b) Y on X

(c) both (d) none

Q-10 The regression lines which gives best estimate of X for a given value of Y is line of

(a) X on Y (b) Y on X

(c) both (d) none

ANSWERS:

1-d	**2-a**	**3-c**	**4-b**	**5-b**	**6-b**	**7-a**	**8-d**	**9-c**	**10-a**

Chapter-16
Time Series Analysis

Q.1 Define time series?

Ans. When data are arranged in the under of their occurrence, the resulting statistical series is called a 'time series'. In a time series independent variable represents measurement of time and dependent variable represents the effect on the data due to change in time.

Q.2 Name the components of time series?

Ans. The components of time series are as:-

1) Secular Trend or long term trend
2) Seasonal Variations
3) Cyclical Variations
4) Irregular or Random Variations

Q.3 What are the various methods of measurement of Secular Trend?

Ans. The following are the important methods of measuring Secular Trend-

1) Free Hand Curve Method
2) Semi Average Method
3) Moving Averages Method
4) Method of Least square

1) Free Hand Curve Method;-

This is the simplest method of all the methods. In this methods one plots the given time series on a graph paper and then drawn a freehand curve or line which represents the trend of that time series.

2) Semi Average Method:-

In this method the time series is divided into two equal parts, with same number of years. The value of both the parts will be added and then their arithmetic mean will be calculated. We thus get two points. Each point is plotted at the midpoint of the

class interval covered by the respective part. After this the two points are joined by a straight line. This line will give us the required trend line.

3) The Method of Moving Averages:-

In this method it is important to select period for moving averages. Like 4 yearly moving averages, 5 yearly moving averages 6 yearly moving averages. The periodicity of moving averages is an important factor and this depends on the length of the cycle and nature of data. This period can be even or odd.

4) Method of Least Square:

This Method is more accurate and precise and can be used even for forecasting. We can fit a straight line by this method. For fitting a straight line trend; we use the following basic equation.

$Yc = a + bx$

For solving 'a' and 'b' we use the following two normal equations.

$\sum y = Na + b\sum x$

$\sum xy = a\sum x + b\sum x^2$

Multiple Choice Questions

Q-1 "A set of data depending on the time is called time series." It is stated by-

(a) Patterson (b) Croxton and cowden

(c) Kenny & Keeping (d) Ya-lun-chu

Q-2 Time series analysis is based on –

(a) Data (b) Water

(c) Time (d) All of the above

Q-3 A time series is affected by -

(a) Long term trend (b) Seasonal variation

(c) Cyclical Variation (d)All of the above

Q-4 In time series analysis, O represent-

(a) Original series (b) Optional Solution

(c) Original Solution (d) Optimal Data

Q-5 A time series model is represented by-

(a) O=T+S+C+I (b) O=T+S+C-I

(c) O=T-S+C-I (d)O= $\frac{I+S}{C+I}$

Q-6 The weighted moving average is calculated by-

(a) $\frac{\sum WY}{\sum W}$ (b) $\frac{\sum W}{\sum WY}$

(c) $\sum W + \sum WY$ (d) $\sum W = \sum W Y$

Q-7 A Straight line trend is represented by-

(a) Yc=a+bx (b) Yc=a-bx

(c) $Yc = \frac{a}{bx}$ (d) Yx=$\frac{a-bx}{x}$

Q-8 When data are arranged in the order of their occurrence, the resulting statistical series is called a-

(a) Time Series
(b) Average
(c) Data Series
(d) Water Series

Q-9 What is the formula of multiplicative Model:

(a) O=TXSXCXI=TSCI
(b) O=T-XS+CxI=T-SCI
(c) $O=\frac{TXS}{CXI}=\frac{TS}{CI}$
(d) Water Series

Q-10 What are the formula of general average:

(a) $O=\frac{number\ of\ seasons}{sum\ of\ Average}$
(b) $\frac{Number\ of\ seasons}{sum\ of\ average}x100$
(c) $\frac{100-numbers\ of\ seasons}{sum\ of\ average}$
(d) $\frac{sum\ of\ average}{number\ of\ season}x100$

Answer Key

1-a	**2-c**	**3-d**	**4-a**	**5-a**	**6-a**	**7-a**	**8-a**	**9-a**	**10-d**

Chapter 17
Business Forecasting

Q-1 What is business forecasting?

Ans: Business forecasting is the process of estimating future business conditions analyzing the past business conditions adjusted with current changes and innovation on the basis of statistical and mathematical methods.

Q-2 Discuss the characteristic of Business forecasting?

Ans: Main characteristic of business forecasting are as following:-

(a) Based on Past and Present Conditions:-It is based on past and present economic conditions of the business. For forecast the future of a business, various past and present data and information are analyzed.

(b) Estimation of future: - The business forecasting is to forecasting the future regarding probable economic conditions.

(c) Based on Mathematical and statistical methods: By using statistical and mathematical methods the actual trend of the economic trend which may take place in future can be forecasted.

(d) Scope: The forecasting can be physical as well as financial.

(e) Period: The forecasting can be made for long term, short term, and medium term any specific term.

Q-3 Explain the theories of Business Forecasting.

Ans: In the present time the business forecasting in made on the basis of some specific theories and the estimates based on these theories prove to be more accurate, pure and scientific. Some important theories are as follows:

(i) Time lag or sequence theory: This theory is based on the assumption that the change in business is not simultaneous but successive. This theory explains that except for irregular and seasonal variation in time services, other variations do not occur together but they occur after a regular interval in the form of a sequence.

(ii) Specific Historical Analogy Theory: It is based on the assumption that history repeats itself. Under this theory a specific period in past is selected where the business conditions are found similar to the present conditions and conclusion is made that what happened in past will also happen in present condition also because the present condition are similar to past conditions.

(iii) Action-Reaction Theory:- This theory is based on Newton's Third law of motion which say that for every action there is a reaction is opposite direction. There are two assumption of this theory :

(i) For every action there is a reaction and

(ii) The reaction is influenced by the interval of due original action.

Under this theory the reaction is always in opposite direction and a certain level of business activity is normal. If the actual data have gone upward from the normal level them it will come down from the normal level. On the basis of this theory there can be four phases of a business activity:

(a) Boom or prosperity

(b) Adversity

(c) Depression

(d) Improvement.

(iv) Cross-cut Economic Analysis Theory:- This theory assumed that two business conditions cannot be similar causes can produce the same results. Under this theory, forecasting is made on the basis of analysis and interpretation of present condition because the past events have no relevance with present conditions.

(v) Economic Rhythm Theory: The basic assumption of this theory is that history repeats itself and hence assumes that all economic and business events behave in a rhythmic order. This theory assumes that the speed and time of all business cycle are more or less same and by using mathematical and statistical method trend is obtained which will represent a long term tendency of growth or decline.

Q-4 Explain utility of Business forecasting.

Ans: On the basis of the forecasting the business man can forecast the demand of the product, price of the product condition of the market. The business decision can also be received on the basis of business forecasting. Main advantages of business forecasting are as follows:-

(i) Useful in controlling the Business cycles.

(ii) Helpful in increasing profit and reduction in losses.

(iii) Useful to Administration

(iv) Utility to Society

(v) Facility Control

(vi) Management Decision

(vii) Basic for capital market

(viii) To Achieve the Goals.

Q-5 Describe methods of Business Forecasting.

Ans: Business forecasting is being used by everyone to handle the increasing varieties of forecasting various methods have been developed in past years. Each method has its own use because the forecasting problem differs from one situation to another.

Following are the main methods of business forecasting:

(i) Business Barometers: With the help of Business Barometers the trend and fluctuations in business conditions are made known and by forecasting a decision can be taken relating to the problem. The business barometer in the term used as an indicator of future conditions. The construction of business barometer consists of gross national product, whole sale price consumer prices, employment, industrial production, consumer credit, stock prices, volume of bank deposits and currency outstanding bond yields etc.

(ii) Time series analysis: Time series analysis is also used for the purpose of making business forecasting. Through time series business forecasting is possible only when the business data of various years are available which reflects a definite trend and seasonal variation. By time series analysis the long term trend,

secular trend, seasonal and cyclical variations are asctained, analyzed and separated from the data of various years.

(iii) Extrapolation or Mathematical Projection: Extrapolation or mathematical projection is the simplest methods of business forecasting. This method is based on the assumption that the data do not show a sudden change and there is a continuity and uniformity in the given series of data. The extrapolation relies on the relative consistency in the pattern of past movement and the variable will follow its established path of growth.

(iv) Regression Analysis: Regression analysis used to forecast the most appropriate value of dependent variables on the basis of corresponding known value of independent variables. When two variables are inter-related then possible value for a dependent value can be forecasted on the basis of the value of independent variable along with the nature and quantity of the relation of the value of two variables.

(v) Modern Econometric Method: In this method binomial equations are formulated on the basis of values of variables in respect of various section of economy. Econometric models or growth models take the form of a set of simultaneous equations. On the basis of the models the future activities

of the business or economy are forecasted which reflects the inter-relationship of different business and economic activities.

(vi) Exponential Smoothing Method: It is the best method of business forecasting. This method is used to maintain the smoothing function of moving averages in time series. The exponential smoothing method is a weighted moving average of a time series in which small weights are provided to past data and present data are provided higher weights. The effect of past data is lessened by providing higher weights to the closely current data.

Multiple Choice Question

Q-1 _____is the calculation of reasonable possibilities about the future based on analysis of all the latest relevant information by tested on logically.

(c) Statistical average (b) demand forecasting

(c) Probability (d) Business forecasting

Q-2 Forecasting based on:

(c) Future (b) past

(c) Present (d) past and present

Q-3 The success and failure of business depend upon:

(c) Forecasting (b) management

(d) Controlling (d) leading

Q-4 According to economic rhythm theory which in given are more weighted.

(a) Past condition (b) present conditions

(b) Future conditions (d) None of the above

Q-5 Business indices are the indicator of:

(a) Loss (b) Project

(c) Future condition (d) Present condition

Q-6 The mathematical expression of the arithmetic curve is:

(a) Y=a+bx (b) y-a-bx

(c) y=axbx (d) y= a% bx

Answer key

1-D	**2-D**	**3-A**	**4-A**	**5-C**	**6-A**

Chapter 18

Skewness

Q-1 What is skevenss?

Ans: SKeveness refers to lack of symmetry. If a distribution is normal there would be no skeveness in it. The curve drawn from the normal distribution would be symmetrical. In case of skewed distribution the curve drawn would be tilled either to the left or towards the right.

Q-2 Discuss the Difference between Dispersion and Skewness .

Ans:

Basic Dispersion	**Skewness .**
Nature	These measures depicts Measure of
The scatterdness or	Skewness . show
Spread of values from	whether the
A measure of central	series is symmetrical
Tendency	or not Base It depends upon the It depends
upon the	Averages of second order averages
	of first and Second order
Variability	It do not reveal on which
It indicates the	
Side of the central value direction of variations	
Have more variations.	
Conclusion	All the measures of Coefficient of Skewness
Dispersion are positive	can be positive or –ve. Presentation Dispersion
can't be presented	Skewness . can easily be
By means of diagrams	presented by diagrams

Q-3 How will the Skewness . be tested is a distribution?

Ans. For test to ascertain whether the distribution is normal or skewed the following facts be noticed:

1- Relationship between Averages-

If is a distribution the values of mean, median and mode are not identical, it is Skewness . distribution.

2- Distance of pain for quantities from median:-

If in a distribution the values of Q_3 and Q_1 one equidistant from

median value, it is symmetrical distribution. If they are not equi distant, it is a skewed distribution.

3- Frequencies on either sides of Mode:- If the total frequencies on both sides of the model value are not equal, it is a skewed distribution.
4- Total of Deviations: If the sum of positive deviation from the value of median on mode is equal to sum of negative deviations, there is no Skewness . in the distribution.
5- The Curve: When the data of a distribution are plotted on a graph paper and if the curve is not bell-shaped (Normal) it is a skewed distribution.

Q-4 What are the measures of Skewness .?

Ans: Measures of Skewness . are meant to give an idea about the direction and degree of asymmetry in a variable. This measure can be absolute or relative. Absolute measures of Skewness . till us the degree of asymmetry and whether it is positive or negative in values. It is not suitable for comparative study. To compare the degree of Skewness . between two or more distribution relative measures of Skewness . are computed.

There are four measures of Skewness .:

1. Karl Pearson's Measure
2. Bowley's Measure
3. Kelly's Measure
4. Moments Measure

1- Karl Pearson's Measure:

Absolute measure

$$SK = MEAN\,(X) - Mode\,(Z)$$

If mode is ill- defined
SK= 3 [Mean (x) – median (M)]

(v) Relative Measure

(a) $J=\frac{X-Z}{\sigma}$

(b) $J=\frac{3(X-M)}{\sigma}$

2- Bowley's Measure:-

1- Absolute Measure

$Sk = Q_3 + Q_1 - 2M$

2- Relative Measure

$J_Q = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$

3- Kelly's Measure:-

(i) Absolute Measure:

$SK = P_{90} + P_{10} - 2P_{50}$

(ii) Relative Measure:

$\frac{P_{90}P_{10}\ P_{10} - 2P_{50}}{P_{90} - P_{10}}$

4- Moments measure of skeveness:

This measure of Skewness is based on moments about dispersion. Its computation is very hard and complicated. It is because of lack of simplicity, that its use is restricted.

Multiple Choice Question

Skewness

Q-1 _____ is a characteristic of any frequency distribution which is a measure to test the symmetry or asymmetry of a series.

(a) Correlation (b) Statistical Averages

(b) Dispersion (d) Skewness

Q-2 Symmetrical distribution the values of mean , median and made are :

(a) Unequal (b) Positive

(c) Negative (d) equal

Q-3 What is the order of averages is positive skewness :

(a) X < M <≠ (b) X > M> ≠

(c) X = M= ≠ (d) None of these

Q-4 In a symmetrical distribution, sum of two quartiles (Q_3 and Q_1) is equal to median:

(a) False (b) True

(c) Pertly True (d) Partly False

Q-5 Formation or scattering of the various terms in the series is shown by :

(a) Skewenees (b) Dispersion

(c) Quartiles (d) Both a and b

Q-6 What are the limits of coefficient of skewness when the Karl Pearson's formula is used:

(a) 1+ and -1 (b) ± 3

(c) Both of above (d) None of these

Q-7 The formula for measure of skewness based upon percentiles or deciles is suggested by

(a) Karl Pearson's (b) Spearman's

(c) Kelly (d) Prof. king

Q-8 The curve which appears in symmetrical distribution is a :

(a) Bell- shaped (b) skewed to left

(c) skewed to right (d) None of these

Q-9 If coefficient of skewness of a distribution is 0.4 and its standard deviation is 8 and mean is 30, find its made:

(a) 22.4 (b) 20

(c) 12.8 (d) 26.8

Q-10 Difference between the two quartiles in a frequency distribution is 20 their sum is 50 and median is 26. Find the coefficient of skewness.

(a) 0.1 (b) -0.1

(c) -10 (d) 10

Q-11 $Sk=Q_3+Q_1-2M$, This formula is of which method :

(a) Karl pearson's (b) Bowely's

(c) Kelly (d) None of these

Q-12 When we go on adding up the frequencies of subsequent groups in order we get:

(a) Frequency (b) relative frequency distribution

(c) Cumulative frequency (d) None of these

ANSWER KEY:

1 -D	**2-D**	**3-B**	**4-A**	**5-B**	**6-B**
7-C	**8-A**	**9-D**	**10-B**	**11-B**	**12-C**

Chapter 19

Moments and Kurtosis

Q-1 Define the meaning of Moments.

Ans: In statistics moments are the averages of 'deviations' (d') , 'Squares of deviations '(d^2) cubes of deviations' (d^3) and fourth squares of deviations (d^4) of different values from the arithmetic mean.

If deviations are taken from actual arithmetic mean, the sum of these deviations will always be zero. If these are squared and averages them we can find variance, if these are cubed them averages will be skives, whereas if these are raised to fourth power and average them me find kurtosis.

Q-2 Describe the various methods of calculating moments?

Ans: Moments calculated from actual arithmetic mean are called as central moments, or moments about the mean. The central moments are devoted by u (we)

μ, Stands for first moment about mean,

μ_2 Stand for second moment about mean.

μ_3 Stands for third moment about mean.

μ_4 Stands for fourth moments about mean.

There are three methods of calculating central movements-

1- Direct Methods
2- Short-Cut Method
3- Step Deviation Method

1- Direct Methods:

Individual Series

$$\mu_1 = \frac{\sum(x-\bar{x})}{N} = \frac{\sum d}{N}$$

$$\mu_2 = \frac{\sum(x-\bar{x})^2}{N} = \frac{\sum d^2}{N}$$

$$\mu_3 = \frac{\sum(x-\bar{x})^3}{N} = \frac{\sum d^3}{N}$$

$$\mu_4 = \frac{\sum(x-\bar{x})^3}{N} = \frac{\sum d^4}{N}$$

Frequency Series

$$\mu_1 = \frac{\sum fd}{N} = 0$$

$$\mu_2 = \frac{\sum fd^2}{N} =$$

$$\mu_3 = \frac{\sum fd^3}{N}$$

$$\mu_4 = \frac{\sum d^4}{N}$$

Short cut methods :-

Frequency Series

$$\nu_1 = \frac{\sum dX}{N} \quad \frac{\sum (X-A)}{N} \qquad \nu_1 = \frac{\sum f dX}{N}$$

$$\nu_2 = \frac{\sum dX^2}{N} \quad \frac{\sum (X-A)^2}{N} \qquad \nu_2 = \frac{\sum f dx^2}{N}$$

$$\nu_3 = \frac{\sum dX^3}{N} \quad \frac{\sum (X-A)^3}{N} \qquad \nu_3 = \frac{\sum f dx^3}{N}$$

$$\nu_4 = \frac{\sum dX^4}{N} \quad \frac{\sum (X-A)^4}{N} \qquad \nu_4 = \frac{\sum f dx^4}{N}$$

Q-3 Step deviation Method:-

Ans. When there are equal class intervals this method can be used. In this method deviations are divided by a common factor and the values of d'x, d'x^2, d' x^3, d'x^4 are calculated. In ca≠≠se of frequency series the values of fd'x, fd'x^2, fd'x^3, fd'x^4 are ascertained. On the basis of above values V_1', V_2',V_3' & V_4' are computed. Thereafter the following **formula is used:-**

$$\mu_1 = (\nu_1' - \nu_1')X\, i = 0$$

$$\mu_2 = (\nu_2' - \bar{\nu}_1^2)X\, i^2 \neq$$

$$\mu_2 = (\nu_3' - \bar{3}\nu_2'\, \nu_1' + 2\nu_1'3)X\, i^3$$

$$\mu_2 = (\nu_4' - \bar{4}\nu_3'\, \nu_1' + 6\nu_2'\, \nu_1'^2 - 3\nu_1'4)\, i^4$$

Q-4 What is the meaning of Kurtosis?

Ans: Kurtosis is a statistical Measure which tells about the degree of flatner

Or

Peakedness in the region of the mode of a frequency curve.

Q-5 State the various type of Kurtosis?

Ans: Kurtosis denotes to the clustering of frequencies whether these are in the middle part of the distribution or at the end.

There are following three types of kurtosis.

1- LEPTOKURTIC --→ Peaked curve

2- PLATYKURTIC--→ Flat topped curve

3- MESOKURTIC-→ Normal curve

Q-6 How to Measure kurtosis?

Ans: The measurement of kurtosis is based on forum and second moment. By Karl Pearson's formula:

$$\beta_2 = \frac{u_4}{u_2}$$

(i) If β_2= 3, the Curve is said to be normal

(ii) If β_2>3, the Curve is peak-topped

(iii) If β_3< 3, the Curve is flat topped.

Multiple Choice Question

Q-1 Central problem are denoted by:

(a) $\bar{}$ X (b)M

(c) γ (d) μ

Q-2 If μ4=329.49 and μ2=84.41 then co-efficient of skewness based on moments is

(a) 0.180 (b)0.120

(c) 0.40 (d) 0.75

Q-3 Give the formula of V_1

(a) $\frac{\sum DX}{N} = \{\frac{\sum (X-A)}{N}\}$

(b) $\frac{\sum D4X}{N} = \frac{(X-A)^4}{N}$

(c) $\frac{\sum d^2 n}{N} = \frac{\sum (X-A)^2}{N}$ (d) $\frac{\sum d^3 x}{N} = \frac{\sum (X-A)^3}{N}$

ANSWER KEY:

1-D	**2-A**	**3-A**

Chapter 20

Quadratic Equation

A quadratic equation is a second-order polynomial equation in a single variable x

$$a x^2 + b x + c = 0, \tag{1}$$

with $a \neq 0$. Because it is a second-order polynomial equation, the fundamental theorem of algebra guarantees that it has two solutions. These solutions may be both real, or both complex.

The roots x can be found by completing the square,

$$x^2 + \frac{b}{a} x = -\frac{c}{a} \tag{2}$$

$$\left(x + \frac{b}{2a}\right)^2 = -\frac{c}{a} + \frac{b^2}{4a^2} = \frac{b^2 - 4ac}{4a^2} \tag{3}$$

$$x + \frac{b}{2a} = \frac{\pm\sqrt{b^2 - 4ac}}{2a}. \tag{4}$$

Solving for x then gives

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{5}$$

This equation is known as the quadratic formula.

Q.1 **Solve $x^2 + 3x - 4 = 0$**

Ans Using $a = 1$, $b = 3$, and $c = -4$, my solution looks like this

$$x = \frac{-(3) \pm \sqrt{(3)^2 - 4(1)(-4)}}{2(1)}$$

$$= \frac{-3 \pm \sqrt{9 + 16}}{2} = \frac{-3 \pm \sqrt{25}}{2}$$

$$= \frac{-3 \pm 5}{2} = \frac{-3 - 5}{2}, \frac{-3 + 5}{2}$$

$$= \frac{-8}{2}, \frac{2}{2} = -4, 1 = x$$

Then, as expected, the solution is $x = -4$, $x = 1$.

Chapter 2

Logarithms

A logarithm is just an exponent.

To be specific, the logarithm of a number x to a base b is just the exponent you put onto b to make the result equal x. For instance, since $5^2 = 25$, we know that 2 (the power) is the logarithm of 25 to base 5. Symbolically, $\log_5(25) = 2$.

More generically, if $x = b^y$, then we say that y is "the logarithm of x to the base b" or "the base-b logarithm of x". In symbols, $y = \log_b(x)$. Every exponential equation can be rewritten as a logarithmic equation, and vice versa, just by interchanging the x and y in this way.

Another way to look at it is that the $\log_b x$ function is defined as the inverse of the b^x function. These two statements express that inverse relationship, showing how an exponential equation is equivalent to a logarithmic equation:

$x = b^y$ is the same as $y = \log_b x$

Logarithmic Rules:

Product rule: $\log_b AC = \log_b A + \log_b C$

Ex: $\log_4 64 = \log_4 4 + \log_4 16 = \log_4(4 \cdot 16)$

Quotient rule: $\log_b(A/C) = \log_b A - \log_b C$

Ex: $log3(27/9)=log3(27)−log3(9)=3−2=1$

Power rule: $\log_b A^C = C(\log_b A)$

Ex: $\log_3 9^2 = 2\log_3 9$

Chapter 3

Progressions

Arithmetic Progressions (A. P.)

The sequence a , (a +d), (a +2d), (a +3d), (a +4d), . . . Is called an arithmetic progression.

' a ' is the first term and ' d ' is the common difference of the A.P.

The nth term of the A.P. is a + (n-1)d

The sum of n terms of the A.P. is $S_n = \frac{n}{2}$ [2a + (n - 1) d]

Or $S_n = \frac{n}{2}$ [T_1 + T_n] where T_1 and T_n are the first and nth terms of the A.P.

Three numbers a , b, c are in A.P. if 2b = a +c.

Arithmetic mean= $\frac{a+b}{2}$

Q.1 If the sum of n terms of an AP is 2n2+3n, what is the kth term?

Ans Sn=2n2+3n

=>Sn+1=2(n+1)2+3(n+1)=2n2+7n+5

Clearly, an+1=Sn+1-Sn=4n+5=4(n+1)+5

Thus, ak=4k+1.

Q.2 If sum of n terms of an AP is 3n2+5n then which of its terms is 164?

Ans Sn=3n2+5n

=>S1=a=8

S2=a+a+d=22=>d=6

Now, ak=164

=>8+(k-1)6=164

=>k-1=26

=>k=27.

Geometric Progression (G.P.)

The sequence a , $a r$, ar^2 , ar^3 , . . . Is called a geometric progression.

' a' is the first term and ' r ' is the common ratio of the geometric progression.

The nth term of the G.P. is $T_n = ar^{(n-1)}$.

The sum of n terms of the G.P. is $S_n = \frac{[a(r^n-1)]}{(r-1)}$

Or $S_n = \frac{[a(1-r^n)]}{(1-r)}$.

Three numbers a , b, c are in G.P. if $b^2 = ac$.

The sum of infinite number of terms of the G.P is $S_\infty = \frac{a}{(1-r)}$, provided $|r| < 1$.

Geometric Mean= $\sqrt{ab}$

Q.3 Find a GP for which the sum of first two terms is -4 and the fifth term is 4 times the third term.

Ans a1+a2=-4. a5=4a3.

a+ar=-4 and ar4=4ar2

or, a(1+r)=-4 and r2=4

Thus, r=2 or -2. (Since a and r are non zero)

If r=2, a=-4/3

If r=-2, a=4

Thus, GP is -4/3,-8/3,-16/3,…

or, 4,-8,16,…

Harmonic Progression (H.P.)

The sequence c_1, c_2, c_3 , . . . Are said to be in Harmonic Progression if their reciprocals $\frac{1}{c_1}, \frac{1}{c_2}, \frac{1}{c_3}$,. . . Are in Arithmetic Progression.

Harmonic Mean= $\frac{2ab}{a+b}$

Sum of the first n natural numbers, $1+2+3+ \ldots + n = \frac{n(n+1)}{2}$

Sum of squares of the first n natural numbers, $1^2+2^2+3^2+ \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$

Sum of the cubes of the first n natural numbers, $1^3+2^3+3^3+ \ldots + n^3 = \frac{n^2(n+1)^2}{4}$

Q.4 The 7th term of an H.P. is 1/10 and 12th term is 1/25 Find the 20th term, and the nth term.

Ans. Let the H.P. be 1/a + 1/(a + d) + 1(a + 2d) + ...
The 7th term = 1/(a + 6d) = 1/10 => a +6 d = 10

The 12th term = 1/(a + 11d) = 1/25 => a +11 d = 25

Solving these two equations, a = -8, d = 3
Hence 20th term = 1/(a+19d) = 1/[-8 + 9(3)] = 1/49

nd nth term = 1/[a +(n -1)d] = 1/[-8 +(n -1) 3] = 1/[3n - 11]

Relationship between Arithmetic, Harmonic and Geometric Means

NOTE: Let a,b be two positive real numbers. Then
AM * HM=GM^2.
Also, AM>GM>HM

Chapter 4

Matrices and determinants

In mathematics, a **matrix** (plural **matrices**) is a rectangular array of numbers, symbols, or expressions, arranged in *rows* and *columns*. The individual items in a matrix are called its *elements* or *entries*. An example of a matrix with 2 rows and 3 columns is

$$\begin{bmatrix} 1 & 9 & -13 \\ 20 & 5 & -6 \end{bmatrix}.$$

A *matrix* is a rectangular arrangement of mathematical expressions that can be simply numbers

Commonly the *m* components of the matrix are written in a rectangular arrangement in the form of a column of *m* rows:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

For example,

$$\mathbf{A} = \begin{bmatrix} 9 & 13 & 5 \\ 1 & 11 & 7 \\ 3 & 7 & 2 \\ 6 & 0 & 7 \end{bmatrix}.$$

Addition

The *sum* **A**+**B** of two *m*-by-*n* matrices **A** and **B** is calculated entrywise:
$(\mathbf{A} + \mathbf{B})_{i,j} = \mathbf{A}_{i,j} + \mathbf{B}_{i,j}$, where $1 \le i \le m$ and $1 \le j \le n$.

Example: $\begin{bmatrix} 1 & 3 & 1 \\ 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 5 \\ 7 & 5 & 0 \end{bmatrix} = \begin{bmatrix} 1+0 & 3+0 & 1+5 \\ 1+7 & 0+5 & 0+0 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 6 \\ 8 & 5 & 0 \end{bmatrix}$

Scalar multiplication

The *scalar multiplication* $c\mathbf{A}$ of a matrix $\mathbf{A}$ and a number c (also called a scalar in the parlance of abstract algebra) is given by multiplying every entry of $\mathbf{A}$ by c:
$(c\mathbf{A})_{i,j} = c \cdot \mathbf{A}_{i,j}$.

Example: $$2 \cdot \begin{bmatrix} 1 & 8 & -3 \\ 4 & -2 & 5 \end{bmatrix} = \begin{bmatrix} 2 \cdot 1 & 2 \cdot 8 & 2 \cdot -3 \\ 2 \cdot 4 & 2 \cdot -2 & 2 \cdot 5 \end{bmatrix} = \begin{bmatrix} 2 & 16 & -6 \\ 8 & -4 & 10 \end{bmatrix}$$

Transpose

The *transpose* of an m-by-n matrix $\mathbf{A}$ is the n-by-m matrix $\mathbf{A}^{\mathrm{T}}$ (also denoted $\mathbf{A}^{\mathrm{tr}}$ or ${}^{\mathrm{t}}\mathbf{A}$) formed by turning rows into columns and vice versa:
$(\mathbf{A}^{\mathrm{T}})_{i,j} = \mathbf{A}_{j,i}$.

Example: $$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -6 & 7 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 2 & -6 \\ 3 & 7 \end{bmatrix}$$

Multiplication

Multiplication of two matrices is defined only if the number of columns of the left matrix is the same as the number of rows of the right matrix. If $\mathbf{A}$ is an m-by-n matrix and $\mathbf{B}$ is an n-by-p matrix, then their *matrix product* $\mathbf{AB}$ is the m-by-p matrix whose entries are given by dot product of the corresponding row of $\mathbf{A}$ and the corresponding column of $\mathbf{B}$:

$$[\mathbf{AB}]_{i,j} = A_{i,1}B_{1,j} + A_{i,2}B_{2,j} + \cdots + A_{i,n}B_{n,j} = \sum_{r=1}^{n} A_{i,r}B_{r,j},$$

Where $1 \le i \le m$ and $1 \le j \le p$. For example, the underlined entry 2340 in the product is calculated as $(2 \times 1000) + (3 \times 100) + (4 \times 10) = 2340$:

$$\begin{bmatrix} \underline{2} & \underline{3} & \underline{4} \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \underline{1000} \\ 1 & \underline{100} \\ 0 & \underline{10} \end{bmatrix} = \begin{bmatrix} 3 & \underline{2340} \\ 0 & 1000 \end{bmatrix}.$$

Determinants

Multiplication

Determinants are mathematical objects that are very useful in the analysis and solution of systems of linear equations. As shown by Cramer's rule, a nonhomogeneous system of linear equations has a unique solution iff the determinant of the system's matrix is nonzero (i.e., the matrix is nonsingular). For example, eliminating x, y, and z from the equations

$$a_1 x + a_2 y + a_3 z = 0 \quad (1)$$
$$b_1 x + b_2 y + b_3 z = 0 \quad (2)$$
$$c_1 x + c_2 y + c_3 z = 0 \quad (3)$$

gives the expression

$$a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1 = 0, \quad (4)$$

which is called the determinant for this system of equation. Determinants are defined only for square matrices.

If the determinant of a matrix is 0, the matrix is said to be singular, and if the determinant is 1, the matrix is said to be uni modular.

The determinant of a matrix $\mathbf{A}$,

$$\begin{vmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \\ \vdots & \vdots & \ddots & \vdots \\ z_1 & z_2 & \cdots & z_n \end{vmatrix} \quad (5)$$

Is commonly denoted $\det(\mathbf{A})$, $|\mathbf{A}|$, or in component notation as $\sum(\pm a_1 b_2 c_3 \cdots)$, $D(a_1 b_2 c_3 \cdots)$, or $|a_1 b_2 c_3 \ldots|$ Note that the notation $\det(\mathbf{A})$ may be more convenient when indicating the absolute value of a determinant, i.e., $|\det(\mathbf{A})|$ instead of $||\mathbf{A}||$. The determinant is implemented in *Mathematical* as `Det`[*m*].

A 2×2 determinant is defined to be

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{vmatrix} a & b \\ c & d \end{vmatrix} \equiv a d - b c. \quad (6)$$

A $k\times k$ determinant can be expanded "by minors" to obtain

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1k} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kk} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k2} & a_{k3} & \cdots & a_{kk} \end{vmatrix} \quad (7)$$
$$-a_{12} \begin{vmatrix} a_{21} & a_{23} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k3} & \cdots & a_{kk} \end{vmatrix} + \ldots \pm a_{1k} \begin{vmatrix} a_{21} & a_{22} & \cdots & a_{2(k-1)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{k(k-1)} \end{vmatrix}.$$

A general determinant for a matrix $\mathbf{A}$ has a value

$$|\mathbf{A}| = \sum_{i=1}^{k} a_{ij} C_{ij}, \quad (8)$$

with no implied summation over j and where C_{ij} (also denoted a^{ij}) is the cofactor of a_{ij} defined by

$$C_{ij} \equiv (-1)^{i+j} M_{ij}. \tag{9}$$

and M_{ij} is the minor of matrix $\mathbf{A}$ formed by eliminating row i and column j from $\mathbf{A}$. This process is called determinant expansion by minors (or "Laplacian expansion by minors," sometimes further shortened to simply "Laplacian expansion").

A determinant can also be computed by writing down all permutations of $\{1, \ldots, n\}$, taking each permutation as the subscripts of the letters a, b, ..., and summing with signs determined by $\epsilon_p = (-1)^{i(p)}$, where $i(p)$ is the number of permutation inversions in permutation p (Muir 1960, p. 16), and $\epsilon_{n_1 n_2 \cdots}$ is the permutation symbol. For example, with $n = 3$, the permutations and the number of inversions they contain are 123 (0), 132 (1), 213 (1), 231 (2), 312 (2), and 321 (3), so the determinant is given by

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 b_2 c_3 - a_1 b_3 c_2 - a_2 b_1 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_3 b_2 c_1. \tag{10}$$

If a is a constant and $\mathbf{A}$ an $n \times n$ square matrix, then

$$|a\,\mathbf{A}| = a^n |\mathbf{A}|. \tag{11}$$

Given an $n \times n$ determinant, the additive inverse is

$$|-\mathbf{A}| = (-1)^n |\mathbf{A}|. \tag{12}$$

Determinants are also distributive, so

$$|\mathbf{A}\,\mathbf{B}| = |\mathbf{A}|\,|\mathbf{B}|. \tag{13}$$

This means that the determinant of a matrix inverse can be found as follows:

$$|\mathbf{I}| = \left|\mathbf{A}\,\mathbf{A}^{-1}\right| = |\mathbf{A}|\left|\mathbf{A}^{-1}\right| = 1, \tag{14}$$

where $\mathbf{I}$ is the identity matrix, so

$$|\mathbf{A}| = \frac{1}{\left|\mathbf{A}^{-1}\right|}. \tag{15}$$

Determinants are multi linear in rows and columns, since

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{vmatrix} = \begin{vmatrix} a_1 & 0 & 0 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{vmatrix} + \begin{vmatrix} 0 & a_2 & 0 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{vmatrix} \tag{16}$$

and

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ 0 & a_5 & a_6 \\ 0 & a_8 & a_9 \end{vmatrix} + \begin{vmatrix} 0 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & a_8 & a_9 \end{vmatrix} + \begin{vmatrix} 0 & a_2 & a_3 \\ 0 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{vmatrix}. \quad (17)$$

The determinant of the similarity transformation of a matrix is equal to the determinant of the original matrix

$$\begin{aligned} \left|\mathsf{B}\,\mathsf{A}\,\mathsf{B}^{-1}\right| &= |\mathsf{B}|\,|\mathsf{A}|\left|\mathsf{B}^{-1}\right| && (18)\\ &= |\mathsf{B}|\,|\mathsf{A}|\,\frac{1}{|\mathsf{B}|} && (19)\\ &= |\mathsf{A}|. && (20) \end{aligned}$$

The determinant of a similarity transformation minus a multiple of the unit matrix is given by

$$\begin{aligned} \left|\mathsf{B}^{-1}\,\mathsf{A}\,\mathsf{B} - \lambda\,\mathsf{I}\right| &= \left|\mathsf{B}^{-1}\,\mathsf{A}\,\mathsf{B} - \mathsf{B}^{-1}\,\lambda\,\mathsf{I}\,\mathsf{B}\right| && (21)\\ &= \left|\mathsf{B}^{-1}\,(\mathsf{A} - \lambda\,\mathsf{I})\,\mathsf{B}\right| && (22)\\ &= \left|\mathsf{B}^{-1}\right|\,|\mathsf{A} - \lambda\,\mathsf{I}|\,|\mathsf{B}| && (23)\\ &= |\mathsf{A} - \lambda\,\mathsf{I}|. && (24) \end{aligned}$$

The determinant of a transpose equals the determinant of the original matrix,

$$|\mathsf{A}| = \left|\mathsf{A}^{\mathrm{T}}\right|, \quad (25)$$

and the determinant of a complex conjugate is equal to the complex conjugate of the determinant

$$|\overline{\mathsf{A}}| = \overline{|\mathsf{A}|}. \quad (26)$$

Let ϵ be a small number. Then

$$|\mathsf{I} + \epsilon\,\mathsf{A}| = 1 + \epsilon\,\mathrm{Tr}\,(\mathsf{A}) + O\left(\epsilon^2\right), \quad (27)$$

where $\mathrm{Tr}\,(\mathsf{A})$ is the matrix trace of A. The determinant takes on a particularly simple form for a triangular matrix

$$\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{k\,1} \\ 0 & a_{22} & \cdots & a_{k\,2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{k\,k} \end{vmatrix} = \prod_{n=1}^{k} a_{n\,n}. \quad (28)$$

Important properties of the determinant include the following, which include invariance under elementary row and column operations.

1. Switching two rows or columns changes the sign.

2. Scalars can be factored out from rows and columns.

3. Multiples of rows and columns can be added together without changing the determinant's value.

4. Scalar multiplication of a row by a constant c multiplies the determinant by c.

5. A determinant with a row or column of zeros has value 0.

6. Any determinant with two rows or columns equal has value 0.

Property 1 can be established by induction. For a 2×2 matrix, the determinant is

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - b_1 a_2 \quad (29)$$

$$= -(b_1 a_2 - a_1 b_2) \quad (30)$$

$$= -\begin{vmatrix} b_1 & a_1 \\ b_2 & a_2 \end{vmatrix} \quad (31)$$

For a 3×3 matrix, the determinant is

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

$$= -\left(a_1 \begin{vmatrix} c_2 & b_2 \\ c_3 & b_3 \end{vmatrix} + b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} - c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \right) = -\begin{vmatrix} a_1 & c_1 & b_1 \\ a_2 & c_2 & b_2 \\ a_3 & c_3 & b_3 \end{vmatrix}$$

$$= -\left(-a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} + b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} b_2 & a_2 \\ b_3 & a_3 \end{vmatrix} \right) = -\begin{vmatrix} b_1 & a_1 & c_1 \\ b_2 & a_2 & c_2 \\ b_3 & a_3 & c_3 \end{vmatrix}$$

$$= -\left(a_1 \begin{vmatrix} c_2 & b_2 \\ c_3 & b_3 \end{vmatrix} - b_1 \begin{vmatrix} c_2 & a_2 \\ c_3 & a_3 \end{vmatrix} + c_1 \begin{vmatrix} b_2 & a_2 \\ b_3 & a_3 \end{vmatrix} \right) = -\begin{vmatrix} c_1 & b_1 & a_1 \\ c_2 & b_2 & a_2 \\ c_3 & b_3 & a_3 \end{vmatrix}.$$

Example: Find the determinant of the matrix using diagonals.

$$\begin{bmatrix} 14 & 5 & -2 \\ -3 & 5 & 4 \\ 6 & 9 & 3 \end{bmatrix}$$

We will find the determinant of the matrix using diagonals.

Repeat the first two columns to the right of the determinant:

$$\begin{vmatrix} 14 & 5 & -2 & 14 & 5 \\ -3 & 5 & 4 & -3 & 5 \\ 6 & 9 & 3 & 6 & 9 \end{vmatrix}$$

Find the product of each diagonal entry with the first row entries, 14, 5 and –2.

$$\begin{vmatrix} 14 & 5 & -2 \\ -3 & 5 & 4 \\ 6 & 9 & 3 \end{vmatrix} \begin{matrix} 14 & 5 \\ -3 & 5 \\ 6 & 9 \end{matrix} = (14)(5)(3)+(5)(4)(6)+(-2)(-3)(9)$$

= 210 + 120 + 54

= 384

Find the product of each diagonal entry with the last row entries, 6, 9 and 3.

$$\begin{vmatrix} 14 & 5 & -2 \\ -3 & 5 & 4 \\ 6 & 9 & 3 \end{vmatrix} \begin{matrix} 14 & 5 \\ -3 & 5 \\ 6 & 9 \end{matrix} = (6)(5)(-2)+(9)(4)(14)+(3)(-3)(5)$$

= –60 + 504 – 45

= 399

Subtract the second product from the first product:

384 – 399 = –15

The determinant is –15.

Cramer's Rule

Given a system of linear equations, Cramer's Rule is a handy way to solve for just one of the variables without having to solve the whole system of equations. Let's use the following system of equations:

$$\begin{aligned} 2x + y + z &= 3 \\ x - y - z &= 0 \\ x + 2y + z &= 0 \end{aligned}$$

We have the left-hand side of the system with the variables (the "coefficient matrix") and the right-hand side with the answer values. Let D be the determinant of the coefficient matrix of the above system, and let D_x be the determinant formed by replacing the x-column values with the answer-column values:

system of equations	coefficient matrix's determinant	answer column	D_x: coefficient determinant with answer-column values in x-column
$2x + 1y + 1z = 3$ $1x - 1y - 1z = 0$ $1x + 2y + 1z = 0$	$D = \begin{vmatrix} 2 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 1 \end{vmatrix}$	$\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$	$D_x = \begin{vmatrix} 3 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 2 & 1 \end{vmatrix}$

Similarly, D_y and D_z would then be:

$$D_y = \begin{vmatrix} 2 & 3 & 1 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{vmatrix} \qquad D_z = \begin{vmatrix} 2 & 1 & 3 \\ 1 & -1 & 0 \\ 1 & 2 & 0 \end{vmatrix}$$

Evaluating each determinant, we get:

$$D = \begin{vmatrix} 2 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 1 \end{vmatrix} = (-2)+(-1)+(2) \\ -(-1)-(-4)-(1) = 3$$

$$D_x = \begin{vmatrix} 3 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 2 & 1 \end{vmatrix} = (-3)+(0)+(0) \\ -(0)-(-6)-(0) = -3+6 = 3$$

$$D_y = \begin{vmatrix} 2 & 3 & 1 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{vmatrix} = (0)+(-3)+(0) \\ -(0)-(0)-(3) = -3-3 = -6$$

$$D_z = \begin{vmatrix} 2 & 1 & 3 \\ 1 & -1 & 0 \\ 1 & 2 & 0 \end{vmatrix} = (0)+(0)+(6) \\ -(-3)-(0)-(0) = 6+3 = 9$$

Cramer's Rule says that $x = D_x \div D$, $y = D_y \div D$, and $z = D_z \div D$. That is:

$x = {}^{3}/_{3} = 1$, $y = {}^{-6}/_{3} = -2$, and $z = {}^{9}/_{3} = 3$

That's all there is to Cramer's Rule. To find whichever variable you want (call it "ß" or "beta"), just evaluate the determinant quotient $D_ß \div D$. (Please don't ask me to explain why this works. Just trust me that determinants can work many kinds of magic.)

Q.1 Given the following system of equations, find the value of *z*.

$$\begin{aligned} 2x + y + z &= 1 \\ x - y + 4z &= 0 \\ x + 2y - 2z &= 3 \end{aligned}$$

Ans To solve only for *z*, I first find the coefficient determinant.

$$D = \begin{vmatrix} 2 & 1 & 1 \\ 1 & -1 & 4 \\ 1 & 2 & -2 \end{vmatrix} = (4)+(4)+(2) \\ -(-1)-(16)-(-2) = 10-13 = -3$$

Then I form D_z by replacing the third column of values with the answer column:

$$D_z = \begin{vmatrix} 2 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 2 & 3 \end{vmatrix} = (-6)+(0)+(2) \\ -(-1)-(0)-(3) = -4-2 = -6$$

Then I form the quotient and simplify: $\frac{D_z}{D} = \frac{-6}{-3} = 2$

$z = 2$

The point of Cramer's Rule is that you don't have to solve the whole system to get the one value you need.

Almost.

What if the coefficient determinant is zero? You can't divide by zero, so what does this mean? "$D = 0$" means that the system of equations has no unique solution. The system may be inconsistent (no solution at all) or dependent (an infinite solution, which may be expressed as a parametric solution such as "(a, $a + 3$, $a – 4$)"). In terms of Cramer's Rule, "$D = 0$" means that you'll have to use some other method.

Chapter 5
Differentiation

These allow us to find an expression for the derivative of any function we can write down algebraically explicitly or implicitly.

Rules of differentiation

Differentiation of a constant term

$\frac{dc}{dx} = 0$, for any constant c
Example:
$\frac{d3}{dx} = 0$

The Identity rule
$\frac{dx}{dx} = 1$,

The Sum rule
$\frac{d(u+v)}{dx} = \frac{du}{dx} + \frac{dv}{dx}$, where u and v are functions of x

Example: $\frac{d(8x^3+7x^2+2x+1)}{dx} = \frac{d8x^3}{dx} + \frac{d7x^2}{dx} + \frac{d2x}{dx} + \frac{d1}{dx}$

Ans= $24x^2 + 7x + 2$

The product rule

$\frac{d(uv)}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}$, where u and v are functions of x

Example: $\frac{d(8x^3 * 7x^2)}{dx} = 8x^3\frac{d7x^2}{dx} + 7x^2\frac{d8x^3}{dx}$

Ans: $(8x^3 * 14x) + (7x^2 * 24x^2) = 112x^4 + 168x^4 = 280x^4$

The chain rule
y = f(u), u = g(x), f and g differentiable. Then, $\frac{dy}{dx} = \frac{dy}{du} * \frac{du}{dx}$

Example : We calculate the derivative of $1+\sqrt[3]{1+x^2}$:

$$u = 1+x^2 \Rightarrow \frac{du}{dx} = 2x$$

$$y = 1+\sqrt[3]{1+x^2} = 1+\sqrt[3]{u} = 1+u^{1/3}$$

By the power rule

$$\frac{dy}{du} = \frac{1}{3}u^{-2/3} = \frac{1}{3}\left(1+x^2\right)^{-2/3}$$

By the chain rule

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \frac{1}{3}\left(1+x^2\right)^{-2/3} 2x$$

We calculate the derivative of: $\sqrt{1+\sqrt[3]{1+x^2}}$

$$u = 1+\sqrt[3]{1+x^2} \Rightarrow \frac{du}{dx} = \frac{2x}{3}\left(1+x^2\right)^{-2/3}$$

$$y = \sqrt{1+\sqrt[3]{1+x^2}} = \sqrt{u} = u^{1/2}$$

By the power rule

$$\frac{dy}{du} = \frac{1}{2}u^{-1/2} = \frac{1}{2\sqrt{1+\sqrt[3]{1+x^2}}}$$

By the chain rule

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \frac{1}{2\sqrt{1+\sqrt[3]{1+x^2}}}\frac{2x}{3}\left(1+x^2\right)^{-2/3}$$

The quotient rule

$\frac{d\left(\frac{u}{v}\right)}{dx} = \frac{v\frac{du}{dx}-u\frac{dv}{dx}}{v^2}$, where $v \neq 0$

Example: Calculate the derivative: $y = \frac{x+2}{3x+1}$

u = x + 2

v = 3x +1

$$\frac{d}{dx}\left(\frac{x+2}{3x+1}\right) = \frac{\frac{d(x+2)}{dx}(3x+1) - \frac{d(3x+1)}{dx}(x+2)}{(3x+1)^2}$$
$$= \frac{1(3x+1) - 3(x+2)}{(3x+1)^2}$$
$$= \frac{-5}{(3x+1)^2}$$

The power rule

$\frac{dx^n}{dx} = nx^{n-1}$, for any power n, integer, rational or irrational

Calculate the derivative: y = (3x2 - x)4

$$u = 3x^2 - x \implies \frac{du}{dx} = 6x - 1$$
$$\frac{d}{dx}(3x^2 - x)^4 = 4(3x^2 - x)^3(6x - 1)$$

Differentiation of implicit function or Partial differentiation

Suppose the function f(x) is defined by an equation: g(f(x),x)=0, rather than by an explicit formula. Then g is a function of two variables, x and f. Thus g may change if f changes and x does not, or if x changes and f does not.

Let the change in g arising from a change, df, in f and none in x be a(f,x)df, and let the change in g from a change, dx, in x and none in f be b(f,x).

The total change in g must vanish since g is a constant, (0), which gives us

a(f,x)df + b (f,x)dx = 0

$\frac{df}{dx} = -\frac{b(f,x)}{a(f,x)}$, where b(f,x) and a(f,x) are partial differentiation with respect to x

Suppose $g(x, f) = f^3(x) - x = 0$

Then $0 = dg = 3f^2df - dx$ that is

$$\frac{\partial g}{\partial f} = 3f^2, \quad \frac{\partial g}{\partial x} = -1$$

We get

$$\frac{df}{dx} = -\frac{\partial g}{\partial x} / \frac{\partial g}{\partial f} = \frac{f^{-2}(x)}{3}$$

2. g(x, f) = xf(x) - 1 = 0

We get

$$\frac{\partial g}{\partial f} = x, \quad \frac{\partial g}{\partial x} = f$$

and

$$\frac{df}{dx} = -\frac{\partial g}{\partial x} / \frac{\partial g}{\partial f} = -\frac{f(x)}{x}$$

Since f(x) = 1 / x, we get (1 / x)' = - (1 / x^2)

Key Terms

Statistics: - Collection, classification, presentation, analysis, and interpretation of comparable numerical facts for pre-determined objects is called statistics.

Plural sense:- As statistical Data, numerical statement of facts.

Singular Sense:- As statistical methods , complete body of the principtes and techniques used in collecting and analyzing such data.

Data: Data means collection of facts in numerical form, like data of births and death in India.

Classification: The process of dividing the collected data into various groups on the basis of their resemblance and similarities is known as classification.

Individual Series: When Original value of each individual item or unit is written separately.

Discrete Series:- Variable which can be measured in definite values and the values are in whole number and there, is appreciable definite difference between two successive values.

Continuous Series:- When units cannot be measured with mathematical accuracy, and the units are measured using approximation and are placed within the limits of the class.

Exclusive series: A series of class intervals in which the values equal to the upper limit of a class are not included in that class, 20 is not included in the class 10-20

Inclusive Series: In inclusive series both the limits of a class are included in that very class. 11-20,21-30 ,11 &20 both are included in the class 11-20.

Tabulation: - Systematic arrangement of data in columns and rows is known as tabulation.

General Purpose Table: The primary and usually the sole purpose reference table is to present the Data in such a manner the individual items may be formed readily by a reader.

Special purpose table: Such table is also called as summary table or analytical table. Such tables are designed to highlight some particular set of facts in simple form.

Original Table: Original table is one which contains data in the form in which they were originally collected. Such tables are also called as classification table.

Derivative Table: Derivative table is that which presents totals, averages, percentages, rates, ratios, coefficients etc., derived from the original data.

Simple or single table: A simple table is one in which collected data are presented according to one characteristics only.

Complex table: In such tables more than one characteristic of data are presented.

Central tendency: Average value lies in between the minimum and maximum values in the series, mostly in the center.

Arithmetic mean: When the sum of all the terms in a series is divided by the number of items, the quotient obtained is called arithmetic mean.

Median:- The central figure , which is obtained when the variables in a series are arranged in ascending or in descending order, is called median of that series.

Mode:- The term 'Mode' has been derived from the French word. 'La Mode' which means "the most fashionable"

Geometric mean: Geometric mean of a series is that root of the product of all its terms as the number of terms in the series.

Harmonic mean: Harmonic means is the reciprocal of the arithmetic mean of the reciprocals of the terms in a series or group.

Dispersion: - Average distance of the various terms from the averages value is known as dispersion.

Range:- Difference between the maximum value and minimum value in a series is called its range.

Percentile Range: Difference between the value of 90^{th} percentile and 10^{th} percentile is range [80%]

Quartile Deviation: Half of the difference between the third quartile and first quartile is known as quartile deviation.

Standard Deviation: Standard deviation is square root of the mean of sequence of deviation from the arithmetic mean.
Variance: - Square of the standard deviation is known as variance. It is also known as second moment of dispersion.

Mean Deviation: Mean deviation is the measure of dispersion which is based upon all the items in a variable.

Probability:- The ratio of the favorable events to the total number of equally likely events is called probability.

Permutation: Arrangement or placement of some given items in various possible orders is known as permutation.

Combination: A selection or group without any consideration of order is called combination.

Compound Events: Two or more events happening together are known as compound events.

Probability theory: In the general language one use the term probability in the sense of happening or not happening of events.

Addition Theorem: When probability of happening of any one of the two or more mutually exclusive events is to be found addition theorem is used P(AUB)=P(A)+P(B)

Multiplication Theorem: Probability of two or more independent events happening together is found using the multiplication theorem.

Theoretical Frequency Distributions: If an expected frequency is deduced on the basis of theoretical consideration is known as theoretical frequency distribution.

Binomial Distribution: The result of nay trial can be classified only under two mutual exclusive categories called success and failure.

Poisson distribution: The poison distribution is applicable where the successful events in the total event are few.
Normal Distribution: Normal distribution is the most important distribution used in statistical practices. It acts like a directive instrument or appliance in physical and social science medicine, agriculture & engineering etc.

Estimator:- Any sample statistic which is used to estimate a population parameter is called an estimator.
Estimate: An estimate is a specific observed value of statistic. Estimate is based on future.
Point Estimate: A point estimate is a single number which is used to estimate an unknown population parameter. The sample means (X) is the best estimator of the population mean (u)

Internal Estimate: - A range of values within which a population parameter is likely to lie is called internal estimate.

Census Method:- This method of data collection is normally used where the size of the universe is not big there is a need for accurate results.

Sampling Method: - A Sample, as the name applies, is a smaller representation of a larger whole.

Sampling Distribution:- If random samples of certain size are taken from the universe and statistics for each sample are computed, vir. mean, median, standard deviation etc.

Standard Error: The difference between sampling results &population sampling results is called sampling error.

Degrees of Freedom: - The degrees of freedom refer to the number of values in a sample we can specify freely, once we know something about the sample.

Student Distribution: 'Student' & published his research findings under that name which was later on renowned as student's t-distribution.

Fisher's z –test: To test the significance of correlation coefficient in small sample prof. Ronald Fisher originated a specific device according to which the coefficient of correlation is transformed into z statistic.

Variance Ratio Test: F-test : The F-test is used to test the significance of difference between two variances.

Test of Significance: The test of significance for small samples is based on the assumption that the population where from the sample is selected is of the nature of normality.

Chi-Square Test :- The chi-Square test is a non-parametric test where no assumption is made about the parameters of population.

Null Hypothesis: Null hypothesis is adopted and it is assumed that both the attributes are independent and the difference between observed frequencies and expected frequencies is nil i.e.H_O: $F_O = F_E$

Testing the Hypothesis: If calculated value is more than the table value, we reject the null hypothesis and it is said that both the attributes are not independent and they are associated.

Yate's Correction: - To maintain the continuity of x^2 & to draw the correct inference in 2x2 table we apply a correction given by f. Yates known as 'yate's correction for continuity.
Degrees of Freedom: The numbers of frequencies which can be calculated independently are known as the degree of freedom

Analysis of Variance: - Analysis of variance is the technique used to estimate the separation of variance ascribable to one group of cases from the variance ascribable to other groups.
One way classification: One way classification the data are classification according to one factor or the influence of one factor on different sample group.

Two way classifications: Two way classifications refer to the classification of data according to two factor or influence of two factor on sample groups.

Latin Square: - Latin square is a technique of controlling two sources of errors.

Decision:- Process results in the selection from a set of alternative course of action, that courses of action which is regarded to meet the objective of the decision problems more satisfactorily as compared to others.
Acts: Acts are the alternative courses of action that are available to the decision maker.

Alternatives: - The decision making procedure involves selection of one act from a set of alternative acts.

Events: Events are also called "States of nature" or outcomes.

Decision-making under risk: The risk involved in taking a particular decision is large and this amount of risk is reflected in deciding possible occurrence of a state of outcome.
Pay off Table: Various pay-off can be put in the shape of a table to from a pay-off table.

Expected value of perfect information: - Under this it is assumed that the decision maker has automatic and perfect information available.

Expected opportunity loss criterion: The EOL criterion the decision maker would choose the strategy with the minimum expected opportunity loss.

Correlation: - when changes in one series are dependent on the changes in the other degree of such inter dependence is called correlation.

Correlation & Causation: - If change in one series is cause and change in the other series is 'effect' it is called correlation & causation.

Positive Correlation: If the direction of independent variables of two series is same, the correlation is positive.

Negative Correlation: If increase in value of one variable leads to decrease in value of other variable, i.e. hangs in both move in reverse direction.

Linear Correlation: - If the distance between changes of values of two variables is constant, there exists linear correlation between such variables.

Multiple Correlations: When correlation between more than two variables is calculated, it is called multiple correlations.

Perfect Positive: If the variations in two variables are in constant ratio in same direction, the correlation is perfect positive.

Perfect correlation: - If the variations in two variables are in a constant ratio, the correlation is said to be perfect.

Absence of correlation: - When there is no concurrent tendency or reverse tendency between two variables i.e there is complete absence of interdependence between them. In such a case the correlation will be here.

Regression:- Measurement of average relationship between two or more variables in the form of original units is called regression.

Linear regression: If a straight line is derived from the marked points in scatter diagram, it will be linear regression.

Curvilinear regression: If it is in the form of a curve, it will be called curvilinear regression.

Dependent Variable: Dependent variable is one whose values are influenced by the other variable while independent variable is one which influences the value of other variable.

Dependent variable: Dependent variable is one whose value is influenced by the other variable while independent variable is one which influences the value of other variable.

Independent Variable: Independent variable is also known as regress or predictor or explainer while the dependent variable is also termed as regressed, predicted or explained variable.

Time series: Series based on values of any time period is called time series.

Secular trend: - Increasing or decreasing ordinary basic trend of a time series is called lone-term trend.

Seasonal Variable: The changes caused by special festivals, climate, seasons etc. are called seasonal fluctuations or variations. Their time period is a year or less than that.

Irregular or Random variations: Fluctuations caused by unexpected events are called irregular fluctuations.

Semi-average method: This method means average of the first half of the data and average of the second half of the Data separately. Long term trend is calculated it.

Moving average method: In it trend is calculated by finding moving average of Original Data.
Least Squares Method: Sum of squares of Deviations of different items of time series is the least. An algebraic equation is used in it : $Y_c=a+bx$

Business forecasting: Business forecast is just an estimate of future conditions as a systematic basic. It requires balanced consideration of factors, influencing future development.

Trend: The trend reveals the direction of business in past as well as the trend of these directions in future.

Cyclical Oscillations: The cyclical oscillations denote the current rate and changes is acceleration on deceleration of business activities in future.

Exponential Smoothing Method: This method is used to maintain the smoothing function of moving averages in time series.

Time log or Sequence Theory: This theory is based on the assumption that the changes in business are not simultaneous but successive.

Skewness: Skewness is a characteristic of any frequency distribution which is a measure to test the symmetry or asymmetry of a series.
Symmetrical Distribution: When the frequency is a distribution first increase in a regular order and ofter reaching a maximum point decrease in the same regular order, the distribution is said to be symmetrical.

Positive skewness: Mean is the largest, median in between and mode is the smallest (i.e.g:X>M>Z)

Negative skewness: Mean is the smallest, median in between and mode is the largest, (i.e.X<M<Z)

Measure of skewness: Measure of skewness are the devices to find out the direction and the extent of asymmetry in a statistical series.

Kurtosis: Kurtosis is the property of a distribution which expresses its relative peakedness.

Moment: Moment is the arithmetic mean of the various powers of deviations taken from mean.

Sheppard's Correction: In case of continuous series, due to the assumption that values are concentrated at the midpoint, errors in the even moments cheep in to correct them sheppard's correction.

Beta Coefficient: Beta Coefficients are calculated on the basis of mutual relations as ratios between different moments from these co-efficient we know about skewenss & kurtosis.

Central Moments: Moments calculated from actual arithmetic mean are called as central moments or moments about them mean(u)

Bibliography

1. Hoy, M., J. Livernois, C. McKenna, R. Rees and T. Stengos (2001) Mathematics for Economics, MIT Press.
2. Pemberton, M. and N. Rau (2001) Mathematics for Economists: An Introductory Textbook, Manchester University Press.
3. G. S. Maddala, Econometrics, McGraw-Hill, New York, 1978.
4. Jacob Schwartz, Mathematics as a tool for economic understanding.
5. Eduard Stiefel, An Introduction to Numerical Mathematics, Academic Press, New York.
6. Takayama, Mathematical Economics, Holt, Rinehart, and Winston, 1974.
7. Paul Samuelson, Economics, 9th ed., McGraw-Hill, New York 1973.
8. Joel Franklin, Methods of Mathematical Economics, Springer-Verlag, New York, 1980.
9. Victor Klee and George Minty, How good is the simplex algorithm? Inequalities, 3 (1972), Academic Press, New York.

Important Websites

- ***www.economics.utoronto.ca/osborne/MathTutorial/***
- ***wps.aw.com/aw_klein_mathecon_2/***
- ***economics.about.com/cs/studentresources/f/math.htm***
- ***www.sosmath.com/algebra/logs/log1/log1.html***
- ***www.physics.uoguelph.ca/tutorials/LOG/***