KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University Established Under Section 3 of UGC Act 1956)
Coimbatore – 641 021.

LECTURE PLAN
DEPARTMENT OF BIOCHEMISTRY

STAFF NAME : Dr.K.POORNIMA
SUBJECT NAME: Biostatistics and Research Methodology SUB.CODE:17BCP305A
SEMESTER : III CLASS: II M.Sc (BC)

Sl. No	Duration of Period	Topics to be Covered	Support material
UNIT I			
1	1	Definitions, scope of Biostatistics	T1: 3-9
2	1	Variables in biology	T3:22-25
3	1	Collection of data – Primary and secondary data	T1:10-12 T2:14-38
4	1	Classification of data – Characteristics and types	T1:15-18 T2:39-42
5	1	Tabulation of data – General rules and types of tables	T1:18-23 T2:42-63
6	1	Graphical and diagrammatic representation of data	T1:23-48 T2:71-100
7	1	Continuation of graphical and diagrammatic representation of data	T1:23-48 T2:71-100
8	1	Measures of central tendency – Arithmetic mean	T1:50-58
9	1	Median and mode	T1:61-72
10	1	Measures of dispersion- Range	T1:85-87
11	1	Standard deviation, Coefficient of variation	T1:97-109
12	1	Revision and possible questions discussion of unit I	
		Total no of hours planned for UNIT I = 12	
Unit II			
1	1	Correlation: Meaning and definition	T1:139-141
2	1	Scatter diagram	T1:141-144
3	1	Karl Pearson's correlation coefficient	T1:146-156
4	1	Rank correlation	T1:156-159
5	1	Correlation problems	T1:172-176
6	1	Regression: Regression in two variables	T1:177-186
7	1	Regression coefficient problems	T1:191-200
8	1	Uses of regression	T1:106

9	1	Revision and possible questions discussion of Unit II	
		Total no of hours planned for UNIT II = 09	
Unit III			
1	1	Test of significance-Introduction and procedures	T3:218-225
2	1	Tests based on mean only – Large samples	T1:245-251
3	1	Small samples – Student's – t - distribution	T3:225-237
4	1	F - test	T1:272-276
5	1	Chi square test – Goodness of fit	T1:294-318
6	1	Analysis of variance – One way classification	T1:319-328
7	1	Analysis of variance – Two way classification	T1:329-333
8	1	CRD, RBD Designs	R1:462-466
9	1	Revision and possible question discussion of unit 3	
		Total no of hours planned for UNIT III = 09	
Unit- IV			
1	1	Research scope and significance	T4:1-7
2	1	Research process	T4:10-20
3	1	Types of research	T4:2-4
4	1	Characteristics of good research	T4:21-22
5	1	Problems in research – Identifying research problems	T4:24-30
6	1	Research design - Introduction	T4:31-32
7	1	Features of good design	T4:33-35
8	1	Revision and possible questions discussion of Unit IV	
		Total no of hours planned for UNIT IV = 08	
UNIT V			
1	1	Sampling design – Meaning and concepts	T4:55-56
2	1	Steps in sampling	T4:56-57
3	1	Criteria for good sample design	T4:57-58
4	1	Scaling measurements - Introduction	T4:78-79
5	1	Scaling measurements - Techniques	T4:79-81
6	1	Types of scales	T4:76-77
7	1	Revision and possible questions discussion of Unit V	
		Total no of hours planned for UNIT V = 07	
Discussion on previous year end semester examination question paper discussion			
1	1	Discussion on previous year ESE question paper - 1	
2	1	Discussion on previous year ESE question paper - 2	
3	1	Discussion on previous year ESE question paper - 3	
Total number of hours planned to complete this course: 48			

Support Materials

T1: S.Palanichamy and M. Manoharan (2008). Statistical methods for biologists; 3rd edition; Palani Paramount Publications.

T2: P. Ramakrishnan (1995). Biostatistics; 1st edition; Saras Publication.

T3: P.N. Arora and P.K. Malhan (2005). Biostatistics; 1st edition, Himalaya Publishers house, Mumbai.

T4: C.R. Kothari (2009). Research Methodology-Methods and Techniques, 3rd edition; New Age International Pvt Ltd, New Delhi.

R1: Gupta. S.P. (2007). Statistical Methods; Sultan and Chand Company, New Delhi.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)
(Established Under Section 3 of UGC Act 1956)
Coimbatore - 641021.
(For the candidates admitted from 2017 onwards)

DEPARTMENT OF BIOCHEMISTRY

SUBJECT : BIOSTATISTICS

SEMESTER : III

SUBJECT CODE: 17BCP305A **CLASS : II M.Sc. BC**

SCOPE

This course has been intended to provide the student insights into helpful areas of statistics and research methodology.

OBJECTIVE

Students get an idea about collection, interpretation and presentation of statistical data and identify the research problem**.**

UNIT I

Definitions-Scope of Biostatistics- Variables in biology, collection, classification and tabulation of data- Graphical and diagrammatic representation.

Measures of central tendency – Arithmetic mean, median and mode. Measures of dispersion- Range, standard deviation, Coefficient of variation.

UNIT II

Correlation: Meaning and definition - Scatter diagram –Karl Pearson's correlation coefficient. Rank correlation.

Regression: Regression in two variables – Regression coefficient problems – uses of regression.

UNIT III

Test of significance: Tests based on Means only-Both Large sample and Small sample tests – Student's t test, F-test, Chi square test - goodness of fit. Analysis of variance – one way and two way classification. CRD, RBD Designs.

UNIT IV

Research: Scope and significance – Types of Research – Research Process – Characteristics of good research – Problems in Research – Identifying research problems. Research Designs – Features of good designs.

UNIT V

Sampling Design : Meaning – Concepts – Steps in sampling – Criteria for good sample design. Scaling measurements – Techniques – Types of scale.

SUGGESTED READINGS

Gupta, S.P., (2007). Statistical Methods, Sultan Chand & Co, New Delhi.

Kothari, C.R., (2009). Research Methodology – Methods and Techniques, 3rd edition, New Age International Pvt. Ltd, New Delhi.

Sundar Rao, P.S.S., and Richard, J., (2006). Introduction to Biostatistics and ResearchMethods, PHI Publication, New Delhi.

Sandhu, T., (1990). Research Techniques in Biological Sciences, Anmol Publishers, New Delhi.

UNIT-I
SYLLABUS

Definitions-Scope of Biostatistics- Variables in biology, collection, classification and tabulation of data- Graphical and diagrammatic representation.
Measures of central tendency – Arithmetic mean, median and mode. Measures of dispersion-Range, standard deviation, Coefficient of variation.

Introduction

Statistical tools are found useful in progressively increasing of disciplines. In ancient times the statistics or the data regarding the human force and wealth available in their land had been collected by the rulers. Now-a-days the fundamental concepts of statistics are considered by many to be essential part of their knowledge.

Origin and Growth

The origin of the word 'statistics' has been traced to the Latin word 'status', the Italian word 'statista' , the French word 'statistique' and the German word 'statistik'. All these words mean political state.

Meaning

The word 'statistics' is used in two different meanings. As a plural word it means data or numerical statements. As a singular word it means the science of statistics and statistical methods. The word 'statistics' is also used currently as singular to mean data.

Definitions

Statistics is " the science of collection, organization, presentation, analysis and interpretation of numerical data". – Dr S.P.Gupta.

"Statistics are numerical statement of facts in any department of enquiry, placed in relation to each other". – Dr.A.L.Bowley.

Functions

The following are the important functions of statistics.

* Collection
* Numerical Presentation
* Diagrammatic Presentation
* Condensation
* Comparison
* Forecasting
* Policy Making

* Effect Measuring
* Estimation
* Tests of significance.

Characteristics

* Statistics is a Quantitative Science.
* It never considers a single item.
* The values should be different.
* Inductive logic is applied.
* Statistical results are true on the average.
* Statistics is liable to be misused.

Collection of data

Data constitutes the base. The findings of an investigation depend on correctness and completeness of the relevant data. Sources of data are of two kinds- primary source and secondary source. The term source means origin or place from which data comes or got. A primary source is one that itself collects the data; a secondary source is one that makes available data which were collected by some other agency. Based on source, data are classified under two categories- Primary data and secondary data.

Primary data

The data which is collected by actual observation or measurement or count is called primary data.

Secondary Data

The data which are compiled from the records of others is called secondary data.

Methods of collection of primary Data

Primary Data is collected in any one of the following methods:

* Direct personal interviews
* Indirect oral interviews
* Information from correspondence
* Mailed questionnaire method.
* Schedules sent through enumerators.

Sources of secondary data

Secondary data can be compiled either from published sources or from unpublished sources.

Classification

Classification is the process of arranging data into groups or classes according to the common characteristics possessed by the individual items.

Basis

Data can be classified on the basis of one or more of the following:

i) Geographical Classification or Spatial Classification

Some data can be classified area-wise such as states, towns etc.

ii) Chronological or Temporal or Historical Classification

Some data can be classified on the basis of time and arranged chronologically or historically.

iii) Qualitative Classification

Some data can be classified on the basis of attributes or characteristics.

iv) Quantitative Classification

Some data can be classified in terms of magnitudes.

Tabulation

Tabulation is the process of arranging data systematically in rows and columns of a table.

There are two methods or modes in which data can be presented. They are

i) Statistical Tables

ii) Diagrams or Graphs

Parts of a table

A good table has the following parts or components:

* Identification number
* Title
* Prefatory Note or Head note
* Stubs
* Captions
* Body of the table
* Foot note
* Source

Frequency Distribution

The easiest method of organizing data is a frequency distribution, which converts raw data into a meaningful pattern for statistical analysis.

The following are the *steps* of constructing a frequency distribution:
1. Specify the number of class intervals. A class is a group (category) of interest. No totally accepted rule tells us how many intervals are to be used. Between 5 and 15 class intervals are generally recommended. Note that the classes must be both *mutually exclusive and all-inclusive.* Mutually exclusive means that classes must be selected such that an item can't fall into two classes, and all-inclusive classes are classes that together contain all the data.

2. When all intervals are to be the same width, the following rule may be used to find the required class interval width:
W = (L - S) / K
where:
W= class width, **L=** the largest data, **S=** the smallest data, **K=** number of classes

Example
Suppose the age of a sample of 10 students are:
20.9, 18.1, 18.5, 21.3, 19.4, 25.3, 22.0, 23.1, 23.9, and 22.5
We select K=4 and W=(25.3 - 18.1)/4 = 1.8 which is rounded-up to 2. The frequency table is as follows:

Class Interval	Class Frequency	Relative Frequency
18-20	3	30 %
20-22	2	20 %
22-24	4	40 %
24- 26	1	10 %

Cumulative Frequency Distribution
When the observations are numerical, cumulative frequency is used. It shows the total number of observations which lie above or below certain key values. Cumulative Frequency for a population = frequency of each class interval + frequencies of preceding intervals. For example, the cumulative frequency for the above problem is: 3, 5, 9, and 10.

Diagrams and Graphs
Diagrams
Diagrams are various geometrical shapes such as bars, circles etc . Diagrams are based on scales but are not confirmed to points or lines. They are more attractive and easier to understand than graphs and are widely used in advertisement and publicity.

Rules for construction

* Title
* Proportion between width and height
* Size
* scale
* Index
* Suitable Diagram
* Simplicity
* Neatness
* Foot-Note and source
* Identification numbers.

Types of Diagram

The frequently used diagrams are divided into the following four heads:

1. One Dimensional diagram- Bar Diagram
2. Two Dimensional diagram – Pie Diagram, Rectangle, squares and circles
3. Three Dimensional diagram – Cubes
4. Pictograms and Cartograms.

Histograms are used to graph absolute, relative, and cumulative frequencies.

Ogive is also used to graph cumulative frequency. An ogive is constructed by placing a point corresponding to the *upper end of each class* at a height equal to the cumulative frequency of the class. These points then are connected. An ogive also shows the relative cumulative frequency distribution on the right side axis.

A less-than ogive shows how many items in the distribution have a value less than the upper limit of each class.

A more-than ogive shows how many items in the distribution have a value greater than or equal to the lower limit of each class.

A less-than cumulative frequency polygon is constructed by using the upper true limits and the cumulative frequencies.

A more-than cumulative frequency polygon is constructed by using the lower true limits and the cumulative frequencies.

Pie chart is often used in newspapers and magazines to depict budgets and other economic information. A complete circle (the pie) represents the total number of measurements. The size of a slice is proportional to the relative frequency of a particular category. For example, since a complete circle is equal to 360 degrees, if the relative frequency for a category is 0.40, the slice assigned to that category is 40% of 360 or (0.40)(360)= 144 degrees.

Measures of Central Tendency and Dispersion

INTRODUCTION

In this chapter we are going to deal with Measures of central tendency and about the measures of dispersion. The measures of central tendency concentrate about the values in the central part of the distribution. Plainly speaking an average of a statistical series is the value of the variable which is the representative of the entire distribution. If we know the average alone we cannot form a complete idea about the distribution so for the completeness of the idea we use Measures of dispersion.

Measures of Central Tendency

According to Professor Bowley the measures of central tendency are "statistical constants which enable us to comprehend in a single effort the significance of the whole "

The following are the three measures of central tendency in this chapter we deal with

- Arithmetic Mean or simply Mean
- Median
- Mode

Arithmetic Mean or simply Mean

Arithmetic Mean or simply Mean is the total values of the item divided by

their number of the items. It is usually denoted by $\bar{X}$.

Individual series

$\bar{X} = \Sigma X / N$

Example:

The expenditure of ten families are given below .Calculate arithmetic mean.

30 70 10 75 500 8 42 250 40 36

Solution

Here N=10

$\Sigma X = 30 + 70 + 10 + 75 + 500 + 8 + 42 + 250 + 40 + 36 = 1061$

$\bar{X} = 1061 / 10 = 106.1$

Discrete series

$\bar{X} = \Sigma f X / \Sigma f$

Example

Calculate the mean number of person per house.

No. of person : 2 3 4 5 6

No. of house :10 25 30 25 10

Solution

X	f	f X
2	10	20
3	25	75
4	30	120
5	25	125
6	10	60
	$\Sigma f = 100$	$\Sigma f X = 400$

$\bar{X} = 400 / 100 = 4$.

Continuous series

$\bar{X} = \Sigma f m / \Sigma f$ where m represents the mid value .

Mid-value = (upper boundary + lower boundary) / 2.

Example

Calculate the mean for the following.

Marks	20-30	30-40	40-50	50-60	60-70	70-80
No. of student	5	8	12	15	6	4

Solution:

C.I	**f**	**m**	**f m**
20-30	5	25	125
30-40	8	35	280
40-50	12	45	540

50-60	15	55	825
60-70	6	65	390
70-80	4	75	300
	Σ f = 50		Σ f m= 2460

$\bar{X}$ = 2460 / 50 = 49.2.

Median

The median is the value for the middle most items when all the items are in the order of magnitude. It is denoted by M or Me.

Individual series

For odd number of item

Position of the median = (N+1) / 2

For even number of item

Position of the median = [(N / 2)+((N/2)+1)] / 2

Example

Calculate median for the following.

22 10 6 7 12 8 5

Solution

Here N =7

Arrange in ascending order or descending order.

5 6 7 8 10 12 22

(N+1) / 2= (7+1) /2

= 4 th item = 8

Discrete series

Position of the median = (N+1) / 2 th item.

Example

Find the median for the following.

X : 10 15 17 18 21

F: 4 16 12 5 3

Solution

X f c.f

10	4	4
15	16	20
17	12	32
18	5	37
21	3	40
	N= 40	

(N+1) /2 = (40+1) / 2 = 20.5^{th} item

= (20^{th} item +21^{st} item) /2 =(15+17) /2

= 16.

Continuous series

M = L+[((N/2) –c.f) x i] / f.

Where L- lower boundary, f-frequency, i-size of class interval,
c.f- cumulative frequency.

Example

Calculate the median height given below.

Height :	145-150	150-155	155-160	160-165	165-170	170-175
No. of student:	2	5	10	8	4	1

Solution :

Height	No. of student	c.f
145-150	2	2
150-155	5	7
155-160	10	17
160-165	8	25
165-170	4	29
170-175	1	30
	Σ f = 30	

Position of the median = N/2 th item = 30 / 2 =15.

M = L+[((N/2) –c.f) x i] / f.

= 155+ [(15-7)x5] / 10 = 155+(40/10) = 159.

Mode :

Mode is the value which has the greatest frequency density. Mode is usually denoted by Z.

Individual series

The value which occur more times are identified as mode.

Example

Determine the mode

32, 35,42, 32, 42,32.

Solution:

Unimode = 32.

Discrete series

Determine the mode

Size of dress	**No. of set**
18	55
20	120
22	108
24	45

here mode represents highest frequency .

Mode = 20

Continuous series

$Z = L + [\, i(f_1 - f_0) / (2f_1 - f_0 - f_2)]$

Where L- lower boundary , f_1-frequency of the modal class, f_0 – frequency of the preceding modal class, f_2- frequency of the succeeding modal class, i-size of class interval , c.f- cumulative frequency.

Example

Determine the mode

Marks	0-10	10-20	20-30	30-40	40-50
No.of student	5	20	35	20	12

Solution

Marks	No. of student
0-10	5
10-20	20
20-30	35
30-40	20
40-50	12

$Z = L + [\ i(f_1 - f_0) / (2f_1 - f_0 - f_2)]$

$= 20+[10(35-20)/(2(35)-20-20)] = 20+5$

$= 25.$

Empirical relation

- Mode = 3 median -2 mean.

Measures of Dispersion

Measure of dispersion deals mainly with the following three measures

- Range
- Standard deviation
- Coefficient of variation

Range

Range is the difference between the greatest and the smallest value.

- Range = L – S , where L-largest value & S-Smallest value
- Coefficient of range = (L-S) /(L+S)

Individual series

Example

Find the value of range and its coefficient of range for the following data.

8 ,10, 5, 9,12,11

Solution

Range = L – S

= 12- 5 =7

Coefficient of range = (L-S) /(L+S)

= (12-5) / (12+5)

= 7 /17 = 0.4118

Continuous series

Range = L – S, where L-Mid-value of largest boundary & S-Mid-value of smallest boundary

Calculate the range.

Marks	20-30	30-40	40-50	50-60	60-70	70-80
No.of student	5	8	12	15	6	4

Solution

C.I	**f**	**m**
20-30	5	25
30-40	8	35
40-50	12	45
50-60	15	55
60-70	6	65
70-80	4	75

Here L=75 & S=25

Range = L – S = 75-25 = 50

Standard deviation

The standard deviation is the root mean square deviation of the values from the arithmetic mean .It is a positive square root of variants. It is also called root mean square deviation. This is usually denoted by σ .

Individual series

$$\sigma = \sqrt{(\Sigma x^2 / N) - (\Sigma x / N)^2}$$

Example

Calculate standard deviation for the following data.

40,41,45,49,50,51,55,59,60,60.

Solution

X	X^2
40	1600
41	1681
45	2025
49	2401
50	2500
51	2601
55	3025

59	3481
60	3600
<u>60</u>	<u>3600</u>
510	$\Sigma x^2 = 26504$

$$\sigma = \sqrt{(\Sigma x^2 / N) - (\Sigma x / N)^2}$$

$$= \sqrt{(26514/10) - (510/10)^2}$$

$$= 7.09$$

Discrete series

$$\sigma = \sqrt{(\Sigma fx^2 / \Sigma f) - (\Sigma fx / \Sigma f)^2}$$

Example

Calculate standard deviation for the following data.

X :	0	1	2	3	4	5
F :	1	2	4	3	0	2

Solution

X	**f**	**fx**	**x^2**	**fx^2**
0	1	0	0	0
1	2	2	1	2
2	4	8	4	16
3	3	9	9	27
4	0	0	16	0
5	<u>2</u>	<u>10</u>	25	<u>50</u>
	$\Sigma f = 12$	$\Sigma fx = 29$		$\Sigma fx^2 = 95$

$$\sigma = \sqrt{(\Sigma fx^2 / \Sigma f) - (\Sigma fx / \Sigma f)^2}$$

$$= \sqrt{(95/12) - (29/12)^2}$$

$$= 1.44$$

Continuous series

$$\sigma = \sqrt{(\Sigma fm^2 / \Sigma f) - (\Sigma fm / \Sigma f)^2}$$

Example

C.I :	0-10	10-20	20-30	30-40	40-50
F :	2	5	9	3	1

Solution

C.I	f	m	fm	m^2	fm^2
0-10	2	5	10	25	50
10-20	5	15	75	225	1125
20-30	9	25	225	625	5625
30-40	3	35	105	1225	3675
40-50	1	45	45	2025	2025
	20		460		12500

$$\sigma = \sqrt{(\Sigma fm^2 / \Sigma f) - (\Sigma fm / \Sigma f)^2}$$

$$= \sqrt{(12500/20) - (460/20)^2}$$

$$= 9.79$$

Coefficient of variation

Coefficient of variation = [standard deviation / arithmetic mean] x100

Example

Calculate the coefficient of variation.

Mean= 51, standard deviation = 7.09

Solution

Coefficient of variation = [standard deviation / arithmetic mean] x100

= (7.09 /51) x100

= 13.9

POSSIBLE QUESTIONS
UNIT I

PART A (20 x 1 = 20 Marks)
Question number 1 – 20 online examinations

PART B (5 x 2= 20Marks)

1. Define Statistics.
2. Write the formula to calculate the angle in Pie Diagram.
3. Define Classification.
4. Define Mid-Value and also find the Mid-Value of 100 – 110.
5. Define Frequency Distribution.
6. What do you mean by size of the Class Interval?
7. Write a formula to calculate Percentage Bar Diagram.
8. Define Histogram.
9. What are the types of classification?
10. Write about Geographical Classification with example.
11. Define Frequency Distribution.
12. Write any two functions of Statistics.
13. What is Bi-variate data?
14. Define Frequency Curve.
15. Calculate the Mean for the following.

X	20	30	35	15	10
f	2	3	4	3	2

16. Define Median and give Example.
17. Calculate the Range and its Coefficient for the following data.

X	:	12	14	16	18	20
f	:	1	3	5	3	1

18. What is mean by Bimodal?
19. Calculate the Median for the following data.

    80 100 50 90 120 110

20. Write the relation between Standard Deviation and Variance.
21. Calculate the Average number of students per class for the following data.

    26 46 33 25 36 27 34 29

22. Find Median and Mode for the following data.

    13 16 17 15 18 14 19 15 12

23. Define Range.
24. Find the Arithmetic Mean for the following data.
    70 60 75 50 42 95 46
25. Calculate the Range and its Coefficient for the following data.
    17 10 56 19 12 11 18 14
26. Find the median for 57, 58, 61, 42, 38, 65, 72, and 66
27. Write the empirical relation for Mode.

PART C (5 X 6 = 30 Marks)

1. Explain about the Classification of data.
2. Draw a suitable Pie Diagram to represent the following submitted as a part of the budget proposal of the govt. of India for the year 1995 – 96.

Item of Expenditure	Percentage
i) Interest	25
ii) Defense	15
iii) Other non plan expenditure	20
iv) States share of taxes and duties	15
v) State and UT plan assistance	10
vi) Central plan	15
Total	100

3. You are given the average expenditure of a family for a month.

Item	Average Expenditure per month (Rs)
Food	2,400
Clothing	200
Rent	800
Medicare	150
Entertainment	450

   Draw a Pie Diagram for the above data.
4. The following table shows the total sale in July 2005 of major brands of cars in India. Represent the following data by using Simple Bar Diagram.

Brand	July 2005 Sales (in Rs. '000)
Maruti	81
Hyundai Santro	39
Honda city	14
Matiz	20

Opel Astra	10
Fait Uno	12
Ford	37
Mitsubishi Lancer	11
Mercedes	3

5. Draw the less than Ogive and hence find the Median of the following data.

Marks	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 - 79	80 - 89	90 - 99
No. of students	7	11	24	32	9	14	2	1

6. Draw Percentage Bar Diagram for the following data.

   Food Rs.200
   Clothing Rs.60
   Education Rs.70
   Rent Rs.130
   Miscellaneous Rs.40

7. Draw a Multiple Bar Diagram for the following data.

Year	**Sales (000 Rs)**	**Gross Profit (000 Rs.)**	**Net Profits (000 Rs)**
1974	100	30	10
1975	120	40	15
1976	130	45	25
1977	150	50	25

8. Nixon Corporation manufactures computers. The following data are the numbers of computers produced at the company for sample of 25 days.

   24 32 27 23 33 33 29 25 23 28 21
   26 31 22 27 33 27 23 28 29 31 35
   34 22 26

   Construct frequency distribution using classes 21 - 23, 24 - 26, 27 - 29, 30 - 32 and 33 - 35. And draw a Histogram to the frequency distribution.

9. The frequency distribution representing the number of days annually the employees at the Voltas Ltd. who were absent due to illness is

Number of days absent	0-2	3-5	6-8	9-11	12-14	Total
Frequency	5	12	20	10	3	50

Draw a Frequency Polygon to the above Frequency Distribution.

10. Calculate the Mode for the following Continuous Frequency Distribution.

Salary (in Rs. 1000s) :	0 – 19	20 - 39	40 - 59	60 – 79	80 - 99
No. of Employees:	5	20	35	20	12

11. Find the Mean and the Standard Deviation for the given below data set.

10	14	20	12	21	16	19	17	14	25

12. Calculate the Standard Deviation and Coefficient of Variance (CV) for the following data.

X	0 – 10	10 - 20	20 - 30	30 – 40	40 - 50
f	2	5	9	3	1

13. Calculate the Median for the following Continuous Frequency Distribution.

Wages (in Rs.) :	0 - 19	20 - 39	40 - 59	60 – 79	80 - 99
No. of Workers:	5	20	35	20	12

14. Calculate the Coefficient of Variation for the following data.

X	6	9	12	15	18
f	7	12	13	10	8

15. Calculate the Median for the following.

Hourly Wages (in Rs.)	40 - 50	50 – 60	60 - 70	70 - 80	80 - 90	90 - 100
Number of Employees	10	20	15	30	15	10

16. The following data give the details about salaries (in thousands of rupees) of seven employees randomly selected from a Pharmaceutical Company.

Serial No.	1	2	3	4	5	6	7
Salary per Annum ('000)	89	57	104	73	26	121	81

Calculate the Standard Deviation and Coefficient of variance of the given data.

17. Calculate the Arithmetic Mean for the following data.

Height (cms):	160	161	162	163	164	165	166
No. of Persons :	27	36	43	78	65	48	28

18. Calculate the Coefficient of Variance for the following data.

77 73 75 70 72 76 75 72 74 76

Karpagam Academy of Higher Education
Department of Biochemistry
Biostatistics and Research Methodology
UNIT I

1	1	The most suitable form of presentation for publicity and Propaganda is	**diagram**	numerals	graph	map	**diagram**
2	1	Mailed enquiry method can be adopted if the respondents are	illiterate	**literate**	blind	physically challenged	**literate**
3	1	Pie-chart represents the components of a factor by	percentages	angles	**sectors**	deciles	**sectors**
4	1	The number of questions of a questionnaire should be	5	20	25	**as small as possible keeping in view the purpose of the survey**	**as small as possible keeping in view the purpose of the survey**
5	1	Primary data are	**always more reliable compared to secondary data**	less reliable compared to secondary data	depends on the care with which data have been selected	depends on the agency collecting the data	**always more reliable compared to secondary data**
6	1	The census data published for state wise population in India will be known as	Quantitative classification	Two-way classification	**Geographical classification**	Quanlitative classification	**Geographical classification**
7	1	Statistics implies	both data and science	data only	**data, science and measure in samples**	samples	**data, science and measure in samples**
8	1	Data taken from the publication, 'Agricultural Situation in India' will be considered as	primary data	**secondary data**	primary and secondary data	published data	**secondary data**
9	1	Year wise recording of data of food production will be called as:	Geographical classification	**Chronological classification**	Quantitative classification	Quanlitative classification	**Chronological classification**
10	1	Who is the father of Biostatistics?	R.A.Fisher	W.Gosset	**Sir Francis Galton**	S.C.Gupta	**Sir Francis Galton**
11	1	Number of source of data is	**2**	3	4	1	**2**
12	1	Compared with primary data secondary data are	more reliable	**less reliable**	equally reliable	not reliable	**less reliable**
13	1	In quantitative classification data are classified on the basis of	attributes	time	location	**magnitudes**	**magnitudes**
14	1	Classification according to class-intervals would yield	raw data	discrete data	qualitative data	**grouped data**	**grouped data**
15	1	In qualitative classification data are classified on the basis of	**attributes**	time	location	class intervals	**attributes**
16	1	In geographical classification data are classified on the basis of	**area**	attributes	time	location	**area**
17	1	Which source is one that itself collects the data?	**primary data**	secondary data	published data	non published data	**primary data**
18	1	Individual observations are called	**raw data**	grouped data	ungrouped data	simple data	**raw data**
19	1	Which one is Geographical classification?	1990-91	**North**	Male	442	**North**
20	1	In discrete frequency distribution values are given as	class intervals	grouped data	**ungrouped data**	equal frequency data	**ungrouped data**
21	1	In continuous frequency distribution values are given as	**class intervals**	grouped data	ungrouped data	equal frequency data	**class intervals**
22	1	Which of the following is the one dimensional diagram?	**square diagram**	multiple bar diagram	rectangular diagram	Pie-Chart	**square diagram**
23	1	In a bar diagram, the base line is	Horizontal	**Vertical**	False baseline	true baseline	**Vertical**
24	1	Pictograms are shown by	Dots	Lines	Circles	**Pictures**	**Pictures**
25	1	Histogram is suitable for the data presented as	**continuous grouped frequency distribution**	discrete grouped frequency distribution	individual series	discrete series	**continuous grouped frequency distribution**
26	1	Numerical data presented in descriptive form are called	classified presentation	tabular presentation	graphical presentation	**textual presentation**	**textual presentation**
27	1	A simple table represents	**only one factor or variable**	always two factors or variables	two or more number of factors or variables	only four factors or variables	**only one factor or variable**
28	1	The row headings of a table are known as	sub-titles	**stubs**	reference notes	captions	**stubs**
29	1	The origin of the word statistics has been traced from the Latin word	statista	**status**	statistic	statistique	**status**
30	1	Relative error is always	Positive	Negative	**Positive or Negative**	Zero	**Positive or Negative**
31	1	The column headings of a table are known as	sub-titles	stubs	reference notes	**captions**	**captions**
32	1	Statistical data are collected for,	Collecting data without any purpose	**a given purpose**	any purpose	for all purpose	**a given purpose**
33	1	In grouped data, the number of classes preferred are	Minimum possible	**adequate**	Maximum possible	any arbitrarily chosen number	**adequate**
34	1	Class interval is measured as	The sum of the upper and lower limit	Half of the sum of the upper and lower limit	**The difference between the upper and lower limit**	Upper limit + lower limit	**The difference between the upper and lower limit**
35	1	The shape of the trilinear charts is that of a	Cone	Cube	**Equilateral triangle**	Pyramid	**Equilateral triangle**
36	1	Mean is a measure of	**central value**	dispersion	correlation	regression	**central value**
37	1	Mode is that value in a frequency distribution which possesses	minimum frequency	frequency one	**maximum frequency**	last frequency	**maximum frequency**
38	1	The most stable measure of central tendency is	**the mean**	the median	the mode	range	**the mean**
39	1	Sum of the deviations about mean is:	minimum	**zero**	maximum	one	**zero**
40	1	The formula used to calculate arithmetic mean for individual series by direct method is	**∑X/N**	∑Fx/N	∑Fm/N	N/ ∑Fx	**∑X/N**
41	1	The formula used to calculate arithmetic mean for individual series by short cut method is	**A+ ∑d/N**	A+ ∑Fd/N	A+∑Fm/N	A+ N/ ∑d	**A+ ∑d/N**
42	1	In continuous series the formula for A.M is	X/N	N/∑X	∑X/N	**∑fm/N**	**∑fm/N**
43	1	A.M = 8, N = 12 then ∑X =	76	80	86	**96**	**96**
44	1	12, 34, 56, 34, 45, 11 in this series the mode is	12	56	**34**	11	**34**
45	1	The data given as 5, 12, 16, 24, 35, 44 will be called as	a continuous series	a discrete series	**an individual series**	time series	**an individual series**
46	1	Which of the following divides the series into two equal parts	Mean	Mode	**Median**	Range	**Median**
47	1	Mode of the following data 3, 6, 5, 7, 8, 4, 9	3	7	**no mode**	5	**no mode**
48	1	Which of the following is not a measure of dispersion?	Range	Quartile deviation	Standard deviation	**Median**	**Median**
49	1	Range of the given values is given by	**L-S**	L+S	S-L	LS	**L-S**
50	1	Which one of the following is relative measure of dispersion?	range	Q.D	S.D	**Coefficient of variation**	**Coefficient of variation**
51	1	Range of a set of values is 65 and maximum value in the series is 83. The minimum value of the series is	74	9	**18**	65	**18**
52	1	If standard deviation is 5, then the variance is	5	**25**	625	2.23068	**25**
53	1	If the value of mode and mean is 60 and 66 then the value of median is	**64**	46	54	44	**64**
54	1	Standard deviation is also called	**Root mean square deviation**	Mean square deviation	Root deviation	Root median square deviation	**Root mean square deviation**
55	1	X: 10-15 15-20 20-25 25-30 30-35 35-40	8	**30**	33	12	**30**
56	1	f: 12 4 10 6 11 8 , range is					
57	1	If the S.D and the C.V of a series are 5 and 25, then the mean	**500**	200	250	100	**500**
58	1	Which one of the following is a measure of central tendency	**median**	range	variation	correlation	**median**
59	1	Mean of the following values is 5 15 20 10 20	5	**14**	41	20	**14**
60	1	The position of the median for an individual series is taken as	**(n+1)/2**	(n+2)/2	n/2	n/4	**(n+1)/2**

UNIT-II
SYLLABUS

Correlation: Meaning and definition - Scatter diagram –Karl Pearson's correlation coefficient. Rank correlation.
Regression: Regression in two variables – Regression coefficient problems – uses of regression.

Simple Linear Correlation

The term Correlation refers to the relationship between the variables. Simple correlation refers to the relationship between two variables. Various types of correlation are considered.

Positive or Negative when the values of two variables change in the same direction, their positive correlation between the two variables.

Example :

X	50	60	70	95	100	105
Y	23	32	37	41	46	50

Example :

X	34	25	18	10	7
Y	51	49	42	33	19

Simple or Partial or Multiple

When only two variables are considered as under positive or negative correlation above the correlation between them is called Simple correlation. When more than two variables as considered the correlation between two of them when all other variables are held constant, i.e., when the linear effects of all other variables on them are removed is called partial correlation. When more than two variables are considered the correlation between one of them and its estimate based on the group consisting of the other variables is called multiple correlation.

Methods

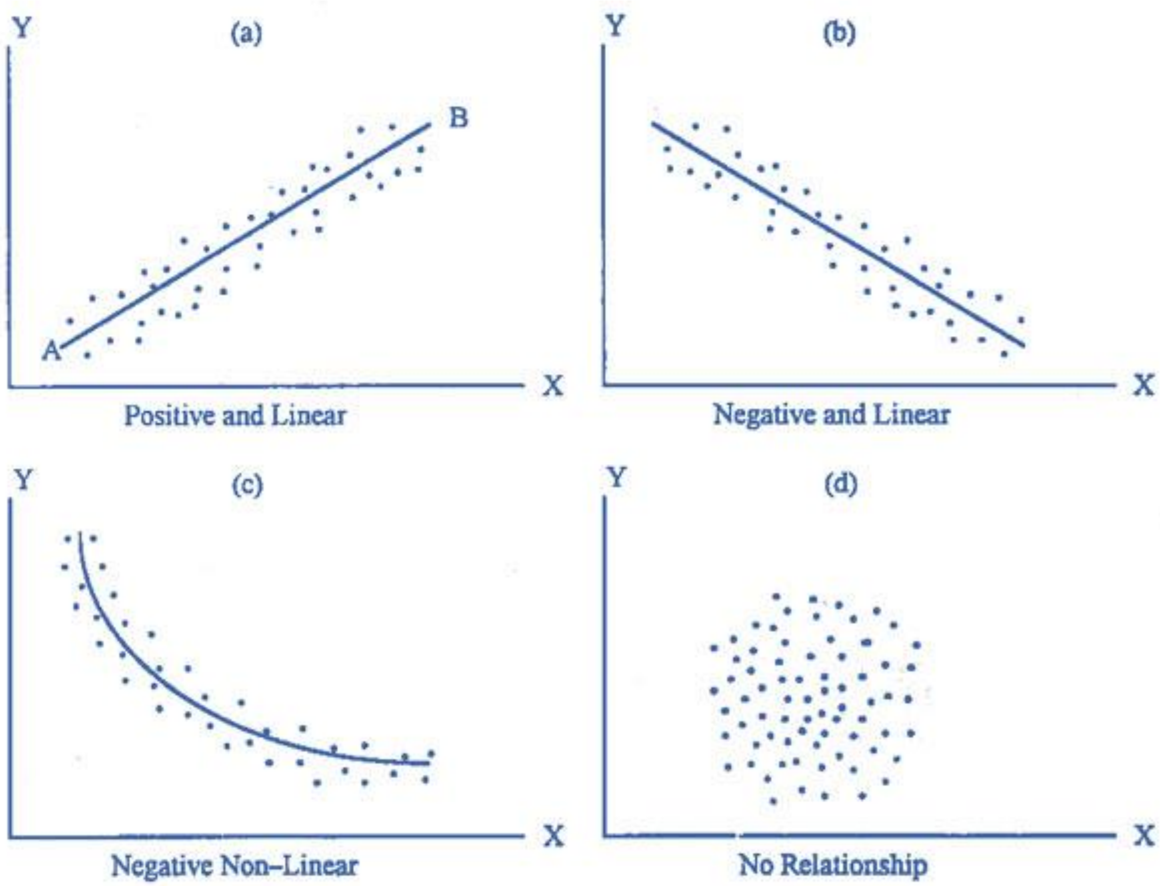
The following four methods are available under simple linear correlation and among them; product moment method is the best one.

- Scatter Diagram
- Karl Pearson's correlation coefficient or product moment correlation coefficient (r)
- Spearman's rank correlation coefficient (ρ)

- Correlation coefficient by concurrent deviation method (r_c).

Scatter Diagram

Scatter diagram is a graphic picture of the sample data. Suppose a random sample of n pairs of observations has the values $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \ldots, (X_n, Y_n)$. These points are plotted on a rectangular co-ordinate system taking independent variable on *X*-axis and the dependent variable on *Y*-axis. Whatever be the name of the independent variable, it is to be taken on *X*-axis. Suppose the plotted points are as shown in figure (a). Such a diagram is called scatter diagram. In this figure, we see that when *X* has a small value, *Y* is also small and when *X* takes a large value, *Y* also takes a large value. This is called direct or positive relationship between *X* and *Y*. The plotted points cluster around a straight line. It appears that if a straight line is drawn passing through the points, the line will be a good approximation for representing the original data. Suppose we draw a line ***AB*** to represent the scattered points. The line ***AB*** rises from left to the right and has positive slope. This line can be used to establish an approximate relation between the random variable *Y* and the independent variable *X*. It is nonmathematical method in the sense that different persons may draw different lines. This line is called the regression line obtained by inspection or judgment.



Making a scatter diagram and drawing a line or curve is the primary investigation to assess the type of relationship between the variables. The knowledge gained from the scatter diagram can be used for further analysis of the data. In most of the cases the diagrams are not as simple as in figure (a). There are quite complicated diagrams and it is difficult to choose a proper mathematical model for representing the original data. The scatter diagram gives an indication of the appropriate model which should be used for further analysis with the help of method of least squares. Figure (b) shows that the points in the scatter diagram are falling from the top left corner to the right. This is a relation called inverse or indirect. The points are in the neighborhood of a certain line called the regression line.

As long as the scattered points show closeness to a straight line of some direction, we draw a straight line to represent the sample data. But when the points do not lie around a straight line, we do not draw the regression line. Figure (c) shows that the plotted points have a tendency to fall from left to right in the form of a curve. This is a relation called non-linear or curvilinear. Figure (d) shows the points which apparently do not follow any pattern. If X takes a small value, Y may take a small or large value. There seems to be no sympathy between X and Y. Such a diagram suggests that there is no relationship between the two variables.

Karl Pearson's Coefficient

Karl Pearson's Product-Moment Correlation Coefficient or simply Pearson's Correlation Coefficient for short, is one of the important methods used in Statistics to measure Correlation between two variables.

A few words about Karl Pearson. Karl Pearson was a British mathematician, statistician, lawyer and a eugenicist. He established the discipline of mathematical statistics. He founded the world's first statistics department In the University of London in the year 1911. He along with his colleagues Weldon and Galton founded the journal "Biometrika" whose object was the development of statistical theory.

The Correlation between two variables X and Y, which are measured using Pearson's Coefficient, give the values between +1 and -1. When measured in population the Pearson's Coefficient is designated the value of Greek letter rho (ρ). But, when studying a sample, it is designated the letter r. It is therefore sometimes called Pearson's r. Pearson's coefficient reflects the linear relationship between two variables. As mentioned above if the correlation coefficient is +1 then there is a perfect positive linear relationship between variables, and if it is -1 then there is a perfect negative linear relationship between the variables. And 0 denotes that there is no relationship between the two variables.

The degrees -1, +1 and 0 are theoretical results and are not generally found in normal circumstances. That means the results cannot be more than -1, +1. These are the upper and the lower limits.

Pearson's Coefficient computational formula

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{N}\right)\left(\sum Y^2 - \frac{(\sum Y)^2}{N}\right)}}$$

Sample question: compute the value of the correlation coefficient from the following table:

Subject	**Age x**	**Weight Level y**
1	43	99
2	21	65
3	25	79
4	42	75
5	57	87
6	59	81

Step 1: Make a chart. Use the given data, and add three more columns: xy, x2, and y2.

Subject	Age x	Weight Level y	xy	x^2	y^2
1	43	99			
2	21	65			
3	25	79			
4	42	75			
5	57	87			
6	59	81			

Step 2: Multiply x and y together to fill the xy column. For example, row 1 would be 43 × 99 = 4,257

Step 3: Take the square of the numbers in the x column, and put the result in the x^2 column.

Subject	Age x	Weight Level y	xy	x^2	y^2
1	43	99	4257	1849	
2	21	65	1365	441	
3	25	79	1975	625	
4	42	75	3150	1764	
5	57	87	4959	3249	
6	59	81	4779	3481	

Step 4: Take the square of the numbers in the y column, and put the result in the y^2 column.

Step 5: Add up all of the numbers in the columns and put the result at the bottom.2 column. The Greek letter sigma (Σ) is a short way of saying "sum of."

Subject	Age x	Weight Level y	xy	x^2	y^2
1	43	99	4257	1849	9801
2	21	65	1365	441	4225
3	25	79	1975	625	6241
4	42	75	3150	1764	5625
5	57	87	4959	3249	7569
6	59	81	4779	3481	6561
Σ	247	486	20485	11409	40022

Step 6: Use the following formula to work out the correlation coefficient. The answer is: 1.3787 × 10-4
the range of the correlation coefficient is from -1 to 1. Since our result is 1.3787 × 10-4, a tiny

positive amount, we can't draw any conclusions one way or another.

Spearman's Rank Correlation Coefficient

The Spearman correlation coefficient is often thought of as being the Pearson correlation coefficient between the ranked variables. In practice, however, a simpler procedure is normally used to calculate ρ. The n raw scores X_i, Y_i are converted to ranks x_i, y_i, and the differences $d_i = x_i - y_i$ between the ranks of each observation on the two variables are calculated.

If there are no tied ranks, then ρ is given by

$$\rho = 1 - \left(\frac{6 \sum d^2}{N(N^2 - 1)}\right)$$

If tied ranks exist, Pearson's correlation coefficients between ranks should be used for the calculation:

One has to assign the same rank to each of the equal values. It is an average of their positions in the ascending order of the values.

Example

X : 21 36 42 37 25

Y : 47 40 37 42 43. For the data given above , calculate the rank correlation coefficient.

Solution

		RANK			
X	Y	X	Y	d	D^2
21	47	5	1	4	16
36	40	3	4	-1	1
42	37	1	5	-4	16
37	42	2	3	-1	1
25	43	4	2	2	4
Total				$\sum d = 0$	$\sum d^2 = 38$

$$\rho = 1 - \left(\frac{6 \sum d^2}{N(N^2 - 1)}\right)$$

$$= 1 - \left(\frac{6 \times 38}{5(5^2 - 1)}\right)$$

= 1- 1.9 = - 0.9

Tied Ranks

When one or more values are repeated the two aspects- ranks of the repeated values and changes in the formula are to be considered.

Example

Find the rank correlation coefficient for the percentage of marks secured by a group of 8 students in Economics and Statistics.

Marks in Economics:	50	60	65	70	75	40	70	80
Marks in Statistics:	80	71	60	75	90	82	70	50

Solution

Let X be Marks in Economics and Y be Marks in Statistics

		RANK			
X	Y	X	Y	d	D^2
50	80	7	3	4	16
60	71	6	5	1	1
65	60	5	7	-2	4
70	75	3.5	4	-0.5	0.25
75	90	2	1	1	1
40	82	8	2	6	36
70	70	3.5	6	-2.5	6.25
80	50	1	8	-7	49
		Total		$\sum d = 0$	$\sum d^2 = 113.5$

$$\rho = 1 - \left(\frac{6\{\sum d^2 + m(m^2-1)/12\}}{N(N^2-1)} \right)$$

When m=2 , $m(m^2-1)/12 = 0.5$

Therefore $\rho = 1 - \left(6\{113.5+0.5\}/8(8^2-1)\}\right)$

$= 1- 1.3571 = -0.3571$

Simple Linear Regression

The line which gives the average relationship between the two variables is known as the regression equation. The regression equation is also called estimating equation.

Uses

1. Regression analysis is used in statistics and other displines.
2. Regression analysis is of practical use in determining demand curve, supply curve, consumption function, etc from market survey.
3. In Economics and Business, there are many groups of interrelated variables.
4. In social resarch, the relation between variables may not known; the relation may differ from place to place.
5. The value of dependent variable is estimated corresponding to any value of the independent variable using the appropriate regression equation.

Method of Least Squares

from a scatter diagram, there is virtually no limit as to the number of lines that can be drawn to make a linear relationship between the 2 variables

- the objective is to create a BEST FIT line to the data concerned
- the criterion is the called the method of least squares
- i.e. the *sum of squares* of the *vertical deviations* from the points to the line be a minimum (based on the fact that the dependent variable is drawn on the vertical axis)
- the linear relationship between the dependent variable (Y) and the independent variable(x) can be written as Y = a + bX , where a and b are parameters describing the vertical intercept and the slope of the regression.
- Similarly the linear relationship between the dependent variable (XY) and the independent variable(Y) can be written as X = a' + b'Y , where a and b are parameters describing the vertical intercept and the slope of the regression.
-

Calculating a and b:

The values of a and b for the given pairs of values of (xi,yi) i=1,2,3…..are determined,
Using the normal equations as ,

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

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

Similarly, the values of a' and b' for the given pairs of values of (xi,yi) i=1,2,3…..are determined,
Using the normal equations as ,
$\sum x = Na' + b'\sum y$

$\sum xy = a'\sum y + b'\sum y^2$

Methods of forming the regression equations

- Regression equations on the basis of normal equations.
- Regression equations on the basis of X and Y and b_{YX} and b_{XY}.

Problem

From the following data, obtain the two regression equations.

X 6 2 10 4 8

Y 9 11 5 8 7 use normal equations.

Solution

X	**Y**	**XY**	**X^2**	**Y^2**
6	9	54	36	81
2	11	22	4	121
10	5	50	100	25
4	8	32	16	64
8	7	56	64	49
$\sum x=0$	$\sum y=0$	$\sum xy=214$	$\sum x^2=220$	$\sum y^2=340$

Let the regression equation Y on X is Y = a + bX

The normal equations are ,

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

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

By substituting the values from the table, we get

5a+30b = 40 -------1

30a + 180b = 214 --------2

Solving these two equations we get,

a=11.90 and b= -0.65

Therefore the regression Y on X is Y = 11.90-0.65X.

Let the regression equation X on Y is X = a' + b'Y

The normal equations are,

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

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

By substituting the values from the table, we get
5a'+40'b = 30 -------3
40a' + 340b' = 214 --------4
Solving these two equations we get,
a' = 16.40 and b= -1.30
Therefore the regression equation X on Y is X = 16.40-1.30Y

Example From the data given below, find

(i) the two regression equations
(ii) The correlation coefficient between the variables X and y
(iii) The value of Y when X= 30

X : 25 28 35 32 31 36 29 38 34 32
Y : 43 46 49 41 36 32 31 30 33 39

Solution

X	Y	x= X- X`	Y= Y-Y`	xy	x^2	y^2
25	43	-7	5	-35	49	25
28	46	-4	8	-32	16	64
35	49	3	11	33	9	121
32	41	0	3	0	0	9
31	36	-1	-2	2	1	4
36	32	4	-6	-24	16	36
29	31	-3	-7	21	9	49
38	30	6	-8	-48	36	64
34	33	2	-5	-10	4	25
32	39	0	1	0	0	1
320	380	0	0	-93	140	398

X` = 32, Y`= 38, $b_{xy} = \sum xy / \sum y^2 = -0.2337$, $b_{yx} = \sum xy / x^2 = -0.6643$

iv) Regression equation of Y on X ,(Y - Y` )= $b_{yx}$ (X-X`)

(Y – 38) = -0.6643(X-32) ⇒ Y = 59.26-0.6643X

(ii) Regression equation of X on Y , (X - X` )= $b_{xy}$ (Y-Y`)

(X – 32) = -0.2337Y +8.88 ⇒ X = 40.88 - 0.233 Y

(iii) $r = + \sqrt{b_{yx} b_{xy}} = -0.3940$

(iv) Y = 59.26-0.6643x30= 39

Properties of Regression coefficients

1. The two regression equations are generally different and are not to be interchanged in their usage.
2. The two regression lines intersect at (X, Y).
3. Correlation coefficient is the geometric mean of two regression coefficients.
4. The two regression coefficients and the correlation coefficient have the same sign.
5. Both the regression coefficients and the correlation coefficient cannot be greater than one numerically and simultaneously.
6. Regression coefficients are independent of change of origin but are affected by the change of scale.
7. Each regression coefficient is in the unit of the measurement of the dependent variable.
8. Each regression coefficient indicates the quantum of change in the dependent variable corresponding to unit increase in the independent variable.

POSSIBLE QUESTIONS
UNIT II

PART A (20 x 1 = 20 Marks)
Question number 1 – 20 online examinations

PART B (5 x 2= 20Marks)

1) What are the types of Correlation?
2) Write any two properties of Correlation.
3) What is the range of Correlation Coefficient?
4) Define Positive Correlation.
5) What is meant by Regression?
6) What are the formulae for Regression co-efficients?
7) Distinguish between Correlation and Regression.
8) Write the formula for Rank Correlation, when more than one rank is repeated.
9) If $b_{xy} = -0.2337$ and $b_{yx} = -0.6643$ then find the Correlation Coefficient.

10) What is Negative Correlation? Give an example?
11) Write down the formula for Karl Pearson's Coefficient of Correlation.
12) Define Scatter Diagram.
13) What is Simple Correlation?
14) Define Regression Equation.
15) When X = 40, Y = 60, $\sigma_X = 10$, $\sigma_Y = 15$ and r = 0.7 find the Regression Equation of Y on X.

PART C (5 X 6 = 30 Marks)

1) Calculate the Correlation Coefficient from the following variables.

Sales in ('0000)	57	58	59	59	60	61	62	64
Advertisement Expenditure ('000)	17	16	15	18	12	14	19	11

2) Marks obtained by 8 students in Accountancy (X) and Statistics (Y) are given below. Compute Rank Correlation Coefficient.

X	25	20	28	22	40	60	20
Y	40	30	50	30	20	10	30

3) Calculate the two Regression Equations from the following data.

X	10	12	13	12	16	15
Y	40	38	43	45	37	43

4) Calculate Karl Pearson's Coefficient of Correlation from the following data.

Wages	100	101	102	102	100	99	97	98
Cost of Living	98	99	99	97	95	92	95	94

5) From the data given below find the two Regression Equations.

X	10	12	13	12	16	15
Y	20	28	23	25	27	30.

i) Estimate Y when X = 20. ii) Estimate X when Y = 35.

6) A comparison of the undergraduate Grade Point Averages of 10 corporate employees with their scores in a managerial trainee examination produced the results shown *in* the following table.

Exam Score	89	83	79	91	95	82	69	66	75	80
GPA	2.4	3.1	2.5	3.5	3.6	2.5	2.0	2.2	2.6	2.7

Measure the Correlation Coefficient between Exam scores and GPA by using Rank Method and also interpret the data given with the help of Scatter Diagram.

7) Develop the Regression Equation that best fit the data given below using annual income as an independent variable and amount of life insurance as dependent variable.

Annual Income (Rs. in 000's)	62	78	41	53	85	34

Amount of Life Insurance (Rs. in 00's)	25	30	10	15	50	7

8) The ranks of ten students in Economics and Statistics subjects are as follows.

Economics	3	5	8	4	7	10	2	1	6	9
Statistics	6	4	9	8	1	2	3	10	5	7

Calculate Spearman's Rank Correlation Coefficient.

9) You are given the following data:

	X	Y
Arithmetic Mean	36	85
Standard Deviation	11	8
Correlation coefficient between X and Y	= 0.66	

Find the two Regression Equations. And also find Correlation Coefficient.

UNIT=II

No.	Marks	Question	Option 1	Option 2	Option 3	Option 4	Answer
1	2	The study of regression is considerably used by	**Economists and Businessmen**	Bioscience researcher	Mathematician	Astronaut	**Economists and Businessmen**
2	2	The average relationship existing between X and Y variables is described by	Regression graph	Regression diagram	**Regression line**	Regression bar	**Regression line**
3	2	The regression equation of Y on X is expressed as	X=a+bY	X=b+aY	**Y=a+bX**	Y=b +aX	**Y=a+bX**
4	2	The regression coefficient of X on Y is	**$r * (\sigma_x / \sigma_y)$**	$r * (\sigma_y / \sigma_x)$	σ_x / σ_y	σ_y / σ_x	**$r * (\sigma_x / \sigma_y)$**
5	2	The under root of the product of two regression coefficients is equal to	coefficient of variation	coefficient of regression	**correlation coefficient**	correlation zero	**correlation coefficient**
6	2	Scatter diagram method is a	**Graphic**	Mathematical	Numerical	Algebraical	**Graphic**
7	2	If every change in X produces a corresponding decrease in Y then X and Y are said to be	uncorrelated	independent	positively correlated	**negatively correlated**	**negatively correlated**
8	2	Rank correlation was discovered by	R.A.Fisher	Sir Francis Galton	Karl Pearson	**Spearman**	**Spearman**
9	2	Correlation is used to measure	**closeness of relationship between variables**	one variable from another	nature of the distribution	central value	**closeness of relationship between variables**
10	2	Coefficient of correlation lies between	**1 and –1**	0 and 1	0 and ∞	0 and –1	**1 and –1**
11	2	While drawing a scatter diagram if all points appear to form a straight linegetting downward from left to right, then it is inferred that there is	a perfect positive correlation	simple positive correlation	**a perfect negative correlation**	no correlation	**a perfect negative correlation**
12	2	The range of the rank correlation coefficient is	0 to 1	**-1 to 1**	0 to ∞	-∞ to +∞	**-1 to 1**
13	2	The technique used in measuring the closeness of the relationship between theVariables is referred to as	Range	standard deviation	median	**correlation analysis**	**correlation analysis**
14	2	If both the variables are varying in the same direction is known as	**positive correlation**	negative correlation	linear correlation	non linear correlation	**positive correlation**
15	2	X : 10 12 15 18 20 Y : 15 20 22 25 37 These two variables are	**Positively correlated**	negatively correlated	linearly correlated	non linearly correlated	**Positively correlated**
16	2	X : 20 30 40 60 80 Y : 40 30 22 15 10 These two variables are	Positively correlated	**negatively correlated**	linearly correlated	non linearly correlated	**negatively correlated**
17	2	If the two variables are varying in opposite direction the correlation is said to be	positive	**negative**	simple	partial	**negative**
18	2	When we study the relationship between the yield of rice per acre and both the amount of rainfall and the amount of fertilizers used, it is the problem of	**multiple correlation**	Simple correlation	non linear correlation	linear correlation	**multiple correlation**
19	2	Study of three or more variables simultaneously is known as	Simple correlation	**multiple correlation**	linear correlation	non linear correlation	**multiple correlation**
20	2	If the amount of change in one variable tends to bear constant ratio to the amount ofchange in the other variable then the correlation is said to be	non linear	**linear**	multiple	simple	**linear**
21	2	X : 10 20 30 40 50 Y : 70 140 210 280 350 The above said two variables are	non linear	**linear**	multiple	simple	**linear**
22	2	If the amount of change in one variable does not bear a constant ratio to the amount of change in the other variable is known as	**non linear**	linear	multiple	simple	**non linear**
23	2	In scatter diagram if the points are lying in a straight line from lower left hand corner to the upper right hand corner the correlation is said to be	perfectly negative	**perfectly positive**	positive	negative	**perfectly positive**
24	2	In scatter diagram if the points are lying in a straight line from upper left hand corner to the lower right hand corner the correlation is said to be	**perfectly negative**	perfectly positive	positive	negative	**perfectly negative**
25 26	2 2	A simple and non mathematical method of studying correlation between the variables is	Karl pearson's coefficient of correlation	**scatter diagram method**	concurrent deviation method	rank correlation	**scatter diagram method**
27	2	Correlation is the	Relationship between two values	**Relationship between two variables**	Significance between two variables	Significant level between two values	**Relationship between two variables**
28	2	If r =+1, the given two variables are	**perfectly positive**	perfectly negative	no correlation	Negative	**perfectly positive**
29	2	If r =-1, the given two variables are	perfectly positive	**perfectly negative**	positive	negative	**perfectly negative**
30	2	If r =0, the given two variables are having	perfect positive correlation	perfect negative correlation	**no correlation**	slight correlation	**no correlation**
31	2	Coefficient of correlation lies between	**1 and –1**	0 and 1	0 and ∞	0 and –1	**1 and –1**
32	2	The range of the rank correlation coefficient is	0 to 1	**-1 to 1**	0 to ∞	-∞ to +∞	**-1 to 1**
33	2	Formula for rank correlation is					
34	2	The formula for computing Pearsonian r is	**$\Sigma xy / N\sigma_x \sigma_y$**	$\Sigma x / N\sigma_x \sigma_y$	$\Sigma y / N\sigma_x \sigma_y$	$\Sigma xy / \sigma_x \sigma_y$	**$\Sigma xy / N\sigma_x \sigma_y$**
35	2	In rank correlation the sum of the differences of ranks between two variables shall be	**zero**	more than 1	less than 1	10	**zero**
36	2	Estimation of the value of one variable from the given value of another variable is done by	correlation analysis	**regression analysis**	chi square test	student's t test	**regression analysis**
37	2	If advertising and sales are correlated the expected amount of sales for a give Advertising expenditure is calculated by	correlation analysis	**regression analysis**	chi square test	student's t test	**regression analysis**
38	2	If advertising and sales are correlated the required amount of expenditure for attaininga given amount of sales is calculated by	correlation analysis	**regression analysis**	chi square test	student's t test	**regression analysis**
39	2	If yield of rice and rainfall are correlated the amount of rain required to achieve a certain production figure is calculated by	correlation analysis	**regression analysis**	chi square test	student's t test	**regression analysis**
40	2	The dictionary meaning of the term 'regression' is	**act of returning**	removing	reacting	relating	**act of returning**
41	2	The term 'Regreeion' was first used by	R.A.Fisher	**Sir Francis Galton**	Karl Pearson	Spearman	**Sir Francis Galton**
42	2	In 1877, the relationship between the height of the fathers and sons was studied by	R.A.Fisher	**Sir Francis Galton**	Karl Pearson	Spearman	**Sir Francis Galton**
43	2	The regression equation of X on Y is expressed as	**X=a+bY**	X=b+aY	Y=a+bX	Y=b +Ax	**X=a+bY**
44	2	The measure of the average relationship between two or more variables in terms ofthe original units of the data is referred to as	correlation	**regression**	standard deviation	correlation coefficient	**regression**
45	2	One of the most frequently used techniques in Economics and Business research is	correlation	**regression**	standard deviation	correlation coefficient	**regression**
46	2	The variable which is used to predict the variable of interest is called the	dependent variable	**independent variable**	common variable	multidependent variable	**independent variable**
47	2	The variable we are trying to predict is called	**dependent variable**	independent variable	common variable	multidependent variable	**dependent variable**
48	2	In correlation analysis r_{xy} and r_{yx} are	**symmetric**	not symmetric	multiples of two	multiples of five	**symmetric**
49	2	In regression analysis the regression coefficients b_{xy} and b_{yx} are	symmetric	**not symmetric**	multiples of two	multiples of five	**not symmetric**
50	2	In the regression equation Y = a+bX, Y is a	**dependent variable**	independent variable	common variable	multidependent variable	**dependent variable**
51	2	In the regression equation Y = a+bX, X is a	dependent variable	**independent variable**	common variable	multidependent variable	**independent variable**
52	2	The regression coefficient of Y on X is	$r * (\sigma_x / \sigma_y)$	**$r * (\sigma_y / \sigma_x)$**	σ_x / σ_y	σ_y / σ_x	**$r * (\sigma_y / \sigma_x)$**
53	2	In the regression equation Y = a+bX, b is the	X intercept	Y intercept	**slope**	dependent variable	**slope**
54	2	In the regression equation Y = a+bX, a is the	X intercept	**Y intercept**	slope	dependent variable	**Y intercept**
55	2	In rank correlation the value of D is	R_1+R_2	**R_1-R_2**	R_1xR_2	R_1/R_2	**R_1-R_2**
56	2	If three ranks are equal at 5^{th} place, the rank given to them is	5	**6**	7	5.5	**6**
57	2	When equal ranks are assigned the adjustments made by adding	**1/12(m3-m)**	1/6(m2-m)	1/5(m2+m)	1/12(m2-m)	**1/12(m3-m)**
58	2	The only method used with ranks not the actual value is	Karl pearson's coefficient of correlation	**rank correlation**	scatter diagram method	concurrent deviation method	**rank correlation**
59	2	If the data are of a qualitative nature like honesty,efficiency and intelligence etc the method used is	Karl pearson's coefficient of correlation	**rank correlation**	scatter diagram method	concurrent deviation method	**rank correlation**
60	2	If the dots in scatter diagram is too scattered we can say that	r = +1	r = -1	**r = 0**	r = 5	**r = 0**

UNIT-III
SYLLABUS

Test of significance: Tests based on Means only-Both Large sample and Small sample tests – Student's t test, F-test, Chi square test - goodness of fit. Analysis of variance – one way and two way classification. CRD, RBD Designs.

Hypothesis:

A statistical hypothesis is an assumption that we make about a population parameter, which may or may not be true concerning one or more variables.

According to Prof. Morris Hamburg "A hypothesis in statistics is simply a quantitative statement about a population".

Hypothesis testing:

Hypothesis testing is to test some hypothesis about parent population from which the sample is drawn.

Example:

A coin may be tossed 200 times and we may get heads 80 times and tails 120 times, we may now be interested in testing the hypothesis that the coin is unbiased.

To take another example we may study the average weight of the 100 students of a particular college and may get the result as 110lb. We may now be interested in testing the hypothesis that the sample has been drawn from a population with average weight 115lb.

Hypotheses are two types

1. Null Hypothesis
2. Alternative hypothesis

Null hypothesis:

The hypothesis under verification is known as *null hypothesis* and is denoted by H_0 and is always set up for possible rejection under the assumption that it is true.

For example, if we want to find out whether extra coaching has benefited the students or not, we shall set up a null hypothesis that *"extra coaching has not benefited the students".* Similarly, if we want to find out whether a particular drug is effective in curing malaria we will take the null hypothesis that *"the drug is not effective in curing malaria*".

Alternative hypothesis:

The rival hypothesis or hypothesis which is likely to be accepted in the event of rejection of the null hypothesis H_0 is called alternative hypothesis and is denoted by H_1 or H_a.

For example, if a psychologist who wishes to test whether or not a certain class of people have a mean I.Q. 100, then the following null and alternative hypothesis can be established.

The null hypothesis would be

$$H_0 : \mu = 100$$

Then the alternative hypothesis could be any one of the statements.

$$H_1 : \mu \neq 100$$

$$(or)\, H_1 : \mu > 100$$

$$(or)\, H_1 : \mu < 100$$

Errors in testing of hypothesis:

After applying a test, a decision is taken about the acceptance or rejection of null hypothesis against an alternative hypothesis. The decisions may be four types.

1) The hypothesis is true but our test rejects it.(type-I error)
2) The hypothesis is false but our test accepts it. .(type-II error)
3) The hypothesis is true and our test accepts it.(correct)
4) The hypothesis is false and our test rejects it.(correct)

The first two decisions are called errors in testing of hypothesis.

i.e.1) Type-I error

2) Type-II error

1) Type-I error: The type-I error is said to be committed if the null hypothesis (H_0) is true but our test rejects it.

2) Type-II error: The type-II error is said to be committed if the null hypothesis (H_0) is false but our test accepts it.

Level of significance:

The maximum probability of committing type-I error is called level of significance and is denoted by α.

α = P (Committing Type-I error)

= P (H_0 is rejected when it is true)

This can be measured in terms of percentage i.e. 5%, 1%, 10% etc…….

Power of the test:

The probability of rejecting a false hypothesis is called power of the test and is denoted by $1-\beta$.

Power of the test =P (H_0 is rejected when it is false)
= 1- P (H_0 is accepted when it is false)
= 1- P (Committing Type-II error)
= $1-\beta$

- A test for which both α and β are small and kept at minimum level is considered desirable.
- The only way to reduce both α and β simultaneously is by increasing sample size.
- The type-II error is more dangerous than type-I error.

Critical region:

A statistic is used to test the hypothesis H_0. The test statistic follows a known distribution. In a test, the area under the probability density curve is divided into two regions i.e. the region of acceptance and the region of rejection. The region of rejection is the region in which H_0 is rejected. It indicates that if the value of test statistic lies in this region, H_0 will be rejected. This region is called critical region. The area of the critical region is equal to the level of significance α. The critical region is always on the tail of the distribution curve. It may be on both sides or on one side depending upon the alternative hypothesis.

One tailed and two tailed tests:

A test with the null hypothesis $H_0:\theta=\theta_0$ against the alternative hypothesis $H_1:\theta\neq\theta_0$, it is called a two tailed test. In this case the critical region is located on both the tails of the distribution.

A test with the null hypothesis $H_0:\theta=\theta_0$ against the alternative hypothesis $H_1:\theta>\theta_0$ (right tailed alternative) or $H_1:\theta<\theta_0$ (left tailed alternative) is called one tailed test. In this case the critical region is located on one tail of the distribution.

$H_0:\theta=\theta_0$ against $H_1:\theta>\theta_0$ ------- right tailed test
$H_0:\theta=\theta_0$ against $H_1:\theta<\theta_0$ ------- left tailed test

Sampling distribution:

Suppose we have a population of size 'N' and we are interested to draw a sample of size 'n' from the population. In different time if we draw the sample of size n, we get different samples of different observations i.e. we can get ${}^{N}c_n$ possible samples. If we calculate some particular statistic from each of the ${}^{N}c_n$ samples, the distribution of sample statistic is called sampling distribution of the statistic. For example if we consider the mean as the statistic, then the distribution of all possible means of the samples is a distribution of the sample mean and it is called sampling distribution of the mean.

Standard error:

Standard deviation of the sampling distribution of the statistic t is called standard error of t.

i.e. S.E (t)= $\sqrt{Var(t)}$

Utility of standard error:

1. It is a useful instrument in the testing of hypothesis. If we are testing a hypothesis at 5% l.o.s and if the test statistic i.e. $|Z| = \left|\frac{t-E(t)}{S.E(t)}\right| > 1.96$ then the null hypothesis is rejected at 5% l.o.s otherwise it is accepted.
2. With the help of the S.E we can determine the limits with in which the parameter value expected to lie.
3. S.E provides an idea about the precision of the sample. If S.E increases the precision decreases and vice-versa. The reciprocal of the S.E i.e. $\frac{1}{S.E}$ is a measure of precision of a sample.
4. It is used to determine the size of the sample.

Test statistic:

The test statistic is defined as the difference between the sample statistic value and the hypothetical value, divided by the standard error of the statistic.

i.e. test statistic $Z = \frac{t-E(t)}{S.E(t)}$

Procedure for testing of hypothesis:

1. Set up a null hypothesis i.e. $H_0 : \theta = \theta_0$.
2. Set up a alternative hypothesis i.e. $H_1 : \theta \neq \theta_0$ or $H_1 : \theta > \theta_0$ or $H_1 : \theta < \theta_0$
3. Choose the level of significance i.e. α.
4. Select appropriate test statistic Z.
5. Select a random sample and compute the test statistic.
6. Calculate the tabulated value of Z at α% l.o.s i.e. Z_α.
7. Compare the test statistic value with the tabulated value at α% l.o.s. and make a decision whether to accept or to reject the null hypothesis.

Large sample tests:

The sample size which is greater than or equal to 30 is called as large sample and the test depending on large sample is called large sample test.

The assumption made while dealing with the problems relating to large samples are

Assumption-1: The random sampling distribution of the statistic is approximately normal.

Assumption-2: Values given by the sample are sufficiently closed to the population value and can be used on its place for calculating the standard error of the statistic.

Large sample test for single mean (or) test for significance of single mean:

For this test

The null hypothesis is $H_0 : \mu = \mu_0$

against the two sided alternative $H_1 : \mu \neq \mu_0$

where μ is population mean

μ_0 is the value of μ

Let $x_1, x_2, x_3, ..., x_n$ be a random sample from a normal population with mean μ and variance σ^2

i.e. if $X \sim N(\mu, \sigma^2)$ then $\bar{x} \sim N\left(\mu, \sigma^2/n\right)$, Where $\bar{x}$ be the sample mean

Now the test statistic $Z = \frac{t - E(t)}{S.E(t)} \sim N(0,1)$

$$= \frac{\bar{x} - E(\bar{x})}{S.E(\bar{x})} \sim N(0,1)$$

$$\Rightarrow Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Now calculate $|Z|$

Find out the tabulated value of Z at α% l.o.s i.e. Z_α

If $|Z| > Z_\alpha$, reject the null hypothesis H_0

If $|Z| < Z_\alpha$, accept the null hypothesis H_0

Note: if the population standard deviation is unknown then we can use its estimate s, which will be calculated from the sample. $s = \sqrt{\frac{1}{n-1}\sum(x-\bar{x})^2}$.

Large sample test for difference between two means:

If two random samples of size n_1 and n_2 are drawn from two normal populations with means μ_1 and μ_2, variances σ_1^2 and σ_2^2 respectively

Let $\bar{x}_1$ and $\bar{x}_2$ be the sample means for the first and second populations respectively

Then $\bar{x}_1 \sim N\left(\mu_1, \sigma_1^2/n_1\right)$ and $\bar{x}_2 \sim N\left(\mu_2, \sigma_2^2/n_2\right)$

Therefore $\bar{x}_1 - \bar{x}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$

For this test

The null hypothesis is $H_0 : \mu_1 = \mu_2 \Rightarrow \mu_1 - \mu_2 = 0$

against the two sided alternative $H_1 : \mu_1 \neq \mu_2$

Now the test statistic $Z = \frac{t - E(t)}{S.E(t)} \sim N(0,1)$

$$= \frac{(\bar{x}_1 - \bar{x}_2) - E(\bar{x}_1 - \bar{x}_2)}{S.E(\bar{x}_1 - \bar{x}_2)} \sim N(0,1)$$

$$\Rightarrow Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S.E(\bar{x}_1 - \bar{x}_2)} \sim N(0,1)$$

$$\Rightarrow Z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$ [since $\mu_1 - \mu_2$ =0 from H$_0$]

Now calculate $|Z|$

Find out the tabulated value of Z at α% l.o.s i.e. Z_α

If $|Z| > Z_\alpha$, reject the null hypothesis H$_0$

If $|Z| < Z_\alpha$, accept the null hypothesis H$_0$

Note: If σ_1^2 and σ_2^2 are unknown then we can consider S_1^2 and S_2^2 as the estimate value of σ_1^2 and σ_2^2 respectively..

Large sample test for single standard deviation (or) test for significance of standard deviation:

Let $x_1, x_2, x_3, \ldots\ldots\ldots\ldots\ldots, x_n$ be a random sample of size n drawn from a normal population with mean μ and variance σ^2,

for large sample, sample standard deviation s follows a normal distribution with mean σ and variance $\sigma^2/2n$ i.e. $s \sim N\left(\sigma, \sigma^2/2n\right)$

For this test

The null hypothesis is $H_0 : \sigma = \sigma_0$

against the two sided alternative $H_1 : \sigma \neq \sigma_0$

Now the test statistic $Z = \dfrac{t - E(t)}{S.E(t)} \sim N(0,1)$

$$= \frac{s - E(s)}{S.E(s)} \sim N(0,1)$$

$$\Rightarrow Z = \frac{s - \sigma}{\sigma/\sqrt{2n}} \sim N(0,1)$$

Now calculate $|Z|$

Find out the tabulated value of Z at α% l.o.s i.e. Z_α

If $|Z| > Z_\alpha$, reject the null hypothesis H_0

If $|Z| < Z_\alpha$, accept the null hypothesis H_0

Large sample test for difference between two standard deviations:

If two random samples of size n_1 and n_2 are drawn from two normal populations with means μ_1 and μ_2, variances $\sigma_1{}^2$ and $\sigma_2{}^2$ respectively

Let s_1 and s_2 be the sample standard deviations for the first and second populations respectively

Then $s_1 \sim N\left(\sigma_1, \sigma_1{}^2/2n_1\right)$ and $\bar{x}_2 \sim N\left(\sigma_2, \sigma_2{}^2/2n_2\right)$

Therefore $s_1 - s_2 \sim N\left(\sigma_1 - \sigma_2, \dfrac{\sigma_1{}^2}{2n_1} + \dfrac{\sigma_2{}^2}{2n_2}\right)$

For this test

The null hypothesis is $H_0 : \sigma_1 = \sigma_2 \Rightarrow \sigma_1 - \sigma_2 = 0$

against the two sided alternative $H_1 : \sigma_1 \neq \sigma_2$

Now the test statistic $Z = \frac{t - E(t)}{S.E(t)} \sim N(0,1)$

$$= \frac{(s_1 - s_2) - E(s_1 - s_2)}{S.E(s_1 - s_2)} \sim N(0,1)$$

$$\Rightarrow Z = \frac{(s_1 - s_2) - (\sigma_1 - \sigma_2)}{S.E(s_1 - s_2)} \sim N(0,1)$$

$$\Rightarrow Z = \frac{(s_1 - s_2)}{\sqrt{\frac{\sigma_1^{\;2}}{2n_1} + \frac{\sigma_2^{\;2}}{2n_2}}} \sim N(0,1)$$ [since $\sigma_1 - \sigma_2$=0 from H_0]

Now calculate $|Z|$

Find out the tabulated value of Z at α % l.o.s i.e. Z_α

If $|Z| > Z_\alpha$, reject the null hypothesis H_0

If $|Z| < Z_\alpha$, accept the null hypothesis H_0

Large sample test for single proportion (or) test for significance of proportion:

Let x is number of success in n independent trails with constant probability p, then x follows a binomial distribution with mean np and variance npq.
In a sample of size n let x be the number of persons processing a given attribute then the sample proportion is given by $\hat{p} = \frac{x}{n}$

Then $E(\hat{p}) = E\left(\frac{x}{n}\right) = \frac{1}{n}E(x) = \frac{1}{n}np = p$

And $V(\hat{p}) = V\left(\frac{x}{n}\right) = \frac{1}{n^2}V(x) = \frac{1}{n^2}npq = \frac{pq}{n}$

$S.E(\hat{p}) = \sqrt{\frac{pq}{n}}$

For this test

The null hypothesis is $H_0 : p = p_0$

against the two sided alternative $H_1 : p \neq p_0$

Now the test statistic $Z = \frac{t - E(t)}{S.E(t)} \sim N(0,1)$

$$= \frac{\hat{p} - E(\hat{p})}{S.E(\hat{p})} \sim N(0,1)$$

$$\Rightarrow Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \sim N(0,1)$$

Now calculate $|Z|$

Find out the tabulated value of Z at α % l.o.s i.e. Z_α

If $|Z| > Z_\alpha$, reject the null hypothesis H_0

If $|Z| < Z_\alpha$, accept the null hypothesis H_0

Large sample test for single proportion (or) test for significance of proportion:

let x_1 and x_2 be the number of persons processing a given attribute in a random sample of size n_1 and n_2 then the sample proportions are given by $\hat{p}_1 = \frac{x_1}{n_1}$ and $\hat{p}_2 = \frac{x_2}{n_2}$

Then $E(\hat{p}_1) = p_1$ and $E(\hat{p}_2) = p_2 \Rightarrow E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$

And $V(\hat{p}_1) = \frac{p_1 q_1}{n_1}$ and $V(\hat{p}_2) = \frac{p_2 q_2}{n_2} \Rightarrow V(\hat{p}_1 - \hat{p}_2) = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$

$S.E(\hat{p}_1) = \sqrt{\frac{p_1 q_1}{n_1}}$ and $S.E(\hat{p}_2) = \sqrt{\frac{p_2 q_2}{n_2}} \Rightarrow S.E(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$

For this test

The null hypothesis is $H_0 : p_1 = p_2$

against the two sided alternative $H_1 : p_1 \neq p_2$

Now the test statistic $Z = \frac{t - E(t)}{S.E(t)} \sim N(0,1)$

$$= \frac{\hat{p}_1 - \hat{p}_2 - E(\hat{p}_1 - \hat{p}_2)}{S.E(\hat{p}_1 - \hat{p}_2)} \sim N(0,1)$$

$$\Rightarrow Z = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{S.E(\hat{p}_1 - \hat{p}_2)} \sim N(0,1)$$

$$\Rightarrow Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \sim N(0,1)$$

$$\Rightarrow Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{pq}{n_1} + \frac{pq}{n_2}}} \sim N(0,1) \quad \text{Since } p_1 = p_2 \text{ from } H_0$$

$$\Rightarrow Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

When p is not known p can be calculated by $p = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$ and $q = 1 - p$

Now calculate $|Z|$

Find out the tabulated value of Z at α % l.o.s i.e. Z_α

If $|Z| > Z_\alpha$, reject the null hypothesis H_0

If $|Z| < Z_\alpha$, accept the null hypothesis H_0

- **As σ is unknown,**

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} = \left[\bar{x} - z_{\alpha/2} \frac{s}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \frac{s}{\sqrt{n}} \right]$$

Step 2: If μ_0 falls into the above confidence intervals, then do *not* reject H_0. Otherwise, reject H_0.

Example 1:

The average starting salary of a college graduate is $19000 according to government's report. The average salary of a random sample of 100 graduates is $18800. The standard error is 800.

(a) Is the government's report reliable as the level of significance is 0.05.

(b) Find the p-value and test the hypothesis in (a) with the level of significance $\alpha = 0.01$.

(c) The other report by some institute indicates that the average salary is $18900. Construct a 95% confidence interval and test if this report is reliable.

[solutions:]

(a)

$$H_0 : \mu = \mu_0 = 19000 \text{ vs. } H_a : \mu \neq \mu_0 = 19000,$$
$$n = 100, \bar{x} = 18800, s = 800, \alpha = 0.05$$

Then,

$$|z| = \left| \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \right| = \left| \frac{18800 - 19000}{800/\sqrt{100}} \right| = |-2.5| = 2.5 > z_{\alpha/2} = z_{0.025} = 1.96.$$

Therefore, reject H_0.

(b)

$$\text{p-value} = P(|Z| > |z|) = P(|Z| > 2.5) = 2 \cdot P(Z > 2.5) = 0.0124 > 0.01$$

Therefore, ***not*** reject H_0.

(c)

$$H_0 : \mu = \mu_0 = 18900 \text{ vs } H_a : \mu \neq \mu_0 = 18900,$$

A 95% confidence interval is

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} = 18800 \pm z_{0.025} \frac{800}{\sqrt{100}} = 18800 \pm 1.96 \cdot 80 = [18643.2, 18956.8].$$

Since $\mu_0 = 18900 \in [18643.2, 18956.8]$, Therefore, ***not*** reject H_0.

Example 2:
A sample of 49 provides a sample mean of 38 and a sample standard deviation of 7. Let $\alpha = 0.05$. Please test the hypothesis

$$H_0 : u = 40 \ vs. \ H_a : u \neq 40.$$

based on
(a) classical hypothesis test
(b) p-value
(c) confidence interval.
[solution:]

$$\bar{x} = 38, \ s = 7, \ u_0 = 40, \ n = 49, \ z = \frac{\bar{x} - u_0}{s/\sqrt{n}} = \frac{38 - 40}{7/\sqrt{49}} = -2.$$

(a)

$$|z| = 2 > 1.96 = z_{0.025}$$

we reject H_0.
(b)

$$p - value = P(|Z| > |z|) = P(|Z| > 2) = 2 * (1 - 0.9772) = 0.0456 < 0.05 = \alpha$$

we reject H_0.
(c)
$100 \times (1-\alpha)\% = 95\%$ confidence interval is

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} = 38 \pm z_{0.025} \frac{7}{\sqrt{49}} = 38 \pm 1.96 = [36.04, 39.96].$$

Since $40 \notin [36.04, 39.96]$, we reject H_0.

Hypothesis Testing for the Mean (Small Samples)

For samples of size less than 30 and when σ is unknown, if the population has a normal, or nearly normal, distribution, the *t*-distribution is used to test for the mean μ.

Using the t-Test for a Mean μ when the sample is small		
Procedure	**Equations**	**Example 4**

State the claim mathematically and verbally. Identify the null and alternative hypotheses	State H_0 and H_a	$H_0 : \mu \geq 16500$ $H_a : \mu < 16500$ $n = 14, \bar{x} = 15700, s = 1250$
Specify the level of significance	Specify α	$\alpha = 0.05$
Identify the degrees of freedom and sketch the sampling distribution	$d.f = n - 1$	$d.f. = 13$
Determine any critical values. If test is left tailed, use One tail, α column with a negative sign. If test is right tailed, use One tail, α column with a positive sign. If test is two tailed, use Two tails, α column with a negative and positive sign.	Table 5 (*t*-distribution) in appendix B	The test is left-tailed. Since test is left tailed and $d.f = 13$, the critical value is $t_0 = -1.771$
Determine the rejection regions.	The rejection region is $t < t_0$	The rejection region is $t < -1.771$
Find the standardized test statistic	$t = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \approx \frac{\bar{x} - \mu}{s/\sqrt{n}}$	$t = \frac{15700 - 16500}{1250/\sqrt{14}} \approx -2.39$
Make a decision to reject or fail to reject the null hypothesis	If t is in the rejection region, reject H_0, Otherwise do not reject H_0	Since $-2.39 < -1.771$, reject H_0
Interpret the decision in the context of the original claim.		Reject claim that mean is at least 16500.

Chi-Square Tests and the F-Distribution

Goodness of Fit

DEFINITION A **chi-square goodness-of-fit test is** used to test whether a frequency distribution fits an expected distribution.

The test is used in a multinomial experiment to determine whether the number of results in each category fits the null hypothesis:

H_0: The distribution fits the proposed proportions

H_1: The distribution differs from the claimed distribution.

To calculate the test statistic for the chi-square goodness-of-fit test, you can use observed frequencies and expected frequencies.

DEFINITION The **observed frequency O** of a category is the frequency for the category observed in the sample data.

The **expected frequency E** of a category is the calculated frequency for the category. Expected frequencies are obtained assuming the specified (or hypothesized) distribution. The expected frequency for the *i*th category is

$E_i = np_i$

where *n* is the number of trials (the sample size) and p_i is the assumed probability of the *i*th category.

The Chi-square Goodness of Fit Test: The sampling distribution for the goodness-of-fit test is a chi-square distribution with $k-1$ degrees of freedom where k is the number of categories. The test statistic is

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

where *O* represents the observed frequency of each category and *E* represents the expected frequency of each category. To use the chi-square goodness of fit test, *the following must be true*

.

1. The observed frequencies must be obtained using a random sample.
2. The expected frequencies must be ≥ 5.

Performing the Chi-Square Goodness-of-Fit Test (p 496)		
Procedure	Equations	Example **(p 497)**
Identify the claim. State the null and alternative hypothesis.	State H_0 and H_1	H_0: Classical 4%

		Country 36% Gospel 11% Oldies 2% Pop 18% Rock 29%
Specify the significance level	Specify α	$\alpha = 0.01$
Determine the degrees of freedom	d.f. = #categories - 1	$d.f. = 6 - 1 = 5$
Find the critical value	χ^2_α: Obtain from Table 6 Appendix B	$\varphi^2_{0.01}(d.f = 5) = 15.086$
Identify the rejection region	$\chi^2 \geq \chi^2_\alpha$	$\chi^2 \geq 15.086$
Calculate the test statistic	$\chi^2 = \sum \frac{(O-E)^2}{E}$	Survey results, n = 500 Classical O= 8 E = .04*500 = 20 Country O = 210 E = .36*500 = 180 Gospel O = 7 E = .11*500 = 55 Oldies O = 10 E = .02*500 = 10 Pop O = 75 E = .18*500 = 90 Rock O= 125 E = .29*500 = 145 Substituting $\chi^2 = 22.713$
Make the decision to reject or fail to reject the null hypothesis	Reject if χ^2 is in the rejection region Equivalently, we reject if the P-value (the probability of getting as extreme a value or more extreme) is $\leq \alpha$	Since 22.713 > 15.086 we reject the null hypothesis Equivalently $P(X \geq 22.713) < 0.01$ so reject the null hypothesis. (Note Table 6 of Appendix B doesn't have a value less than 0.005.)
Interpret the decision in the context of the original claim		Music preferences differ from the radio station's claim.

Using Minitab to perform the Chi-Square Goodness-of-Fit Test (Manual p 237)

The data from the example above (Example 2 p 497) will be used.

Enter Three columns: Music Type: Classical, etc, Observed: 8 etc, Distribution 0.04, etc. (Note the names of the columns 'Music Types', 'Observed' and 'Distribution' are entered in the gray row at the top.)

Select **Calc->Calculator**, **Store the results in** C4, and calculate the **Expression** C3*500, click **OK**, Name C4 'Expected' since it now contains the expected frequencies

Music Type	Observed	Distribution	Expected
Classical	8	0.04	20
Country	210	0.36	180
Gospel	72	0.11	55
Oldies	10	0.02	10
Pop	75	0.18	90
Rock	125	0.29	145

Next calculate the chi-square statistic, $(O-E)^2/E$ as follows: Click **Calc->Calculator**. **Store the results in** C5 and calculate the **Expression** (C2-C4)**2/C4. Click on **OK** and C5 should contain the calculated values.

7.2000
5.0000
41.8909
0.0000
2.5000
2.7586

Next add up the values in C5 and the sum is the test statistic as follows: Click on **Calc->Column Statistics**. Select **Sum** and enter C5 for the **Input Variable**. Click OK. The chi-square statistic is displayed in the session window as follows:

Sum of C5

```
Sum of C5 = 22.7132
```

Next calculate the P-value: Click on **Calc->Probability Distributions->Chi-square**. Select **Cumulative Probability** and enter 5 **Degrees of Freedom** Enter the value of the test statistic

22.7132 for the **Input Constant**. Click **OK**.

The following is displayed on the Session Window.

Cumulative Distribution Function

```
Chi-Square with 5 DF
      x  P( X <= x )
22.7132     0.999617
```

P(X≤22.7132) = 0.999617 So the P-value = 1 – 0.999617 = 0.000383. This is less than α = 0.01 so we reject the null hypothesis.

Instead of calculating the P-value, we could have found the critical value from the Chi-Square table (Table 6 Appendix B) for 5 degrees of freedom as we did above. The value is 15.086, and since our test statistic is 22.7132, we reject the null hypothesis.

Chi-Square with M&M's

H_0: Brown: 13%, Yellow: 14%, Red: 13%, Orange: 20%, Green 16%, Blue 24%
Significance level: α = 0.05
Degrees of freedom: number of categories – 1 = 5
Critical Value: $\chi^2{}_{0.05}(d.f.=5)=11.071$
Rejection Region: $\chi^2 \geq 11.071$
Test Statistic: $\chi^2 = \sum \frac{(O-E)^2}{E}$, where O is the actual number of M&M's of each color in the bag and E is the proportions specified under H_0 times the total number.
Reject H_0 if the test statistic is greater than the critical value (1.145)

Section 10.2 Independence

This section describes the chi-square test for independence which tests whether two random variables are independent of each other.

DEFINTION An ***r x c* contingency table** shows the observed frequencies for the two variables. The observed frequencies are arranged in r rows and c columns. The intersection of a row and a column is called a **cell.**

The following is a contingency table for two variables A and B where f_{ij} is the frequency that A equals A_i and B equals B_j.

	A1	**A2**	**A3**	**A4**	**A**
B1	f_{11}	f_{12}	f_{13}	f_{14}	$f_{1.}$
B2	f_{21}	f_{22}	f_{23}	f_{24}	$f_{2.}$
B3	f_{31}	f_{32}	f_{33}	f_{34}	$f_{3.}$
B	$f_{.1}$	$f_{.2}$	$f_{.3}$	$f_{.4}$	f

If A and B are independent, we'd expect

$$f_{ij} = prob(A = A_i) * prob(B = B_j) * f = \left(\frac{f_{i.}}{f}\right)\left(\frac{f_{.j}}{f}\right) f = \frac{(f_{i.})(f_{.j})}{f}$$

$$\frac{(sum of\ row\ i) * (sum of\ column\ j)}{sample\ size}$$ (

Example 1 Determining the expected frequencies of CEO's ages as a function of company size under the assumption that age is independent of company size.

	<= 39	40 - 49	50 - 59	60 - 69	>= 70	Total
Small/midsize	42	69	108	60	21	300
Large	5	18	85	120	22	250
Total	47	87	193	180	43	550

	<= 39	40 - 49	50 - 59	60 - 69	>= 70	Total

Small/midsize	$\frac{300*47}{550} \approx 25.64$	$\frac{300*87}{550} \approx 47.45$	$\frac{300*193}{550} \approx 105.27$	$\frac{300*180}{550} \approx 98.18$	$\frac{300*43}{550} \approx 23.45$	300
Large	$\frac{250*47}{550} \approx 21.36$	$\frac{250*87}{550} \approx 39.55$	$\frac{250*193}{550} \approx 87.73$	$\frac{250*180}{550} \approx 81.82$	$\frac{250*43}{550} \approx 19.55$	250
Total	47	87	193	180	43	550

After finding the expected frequencies under the assumption that the variables are independent, you can test whether they are independent using the chi-square independence test.

DEFINITION A **chi-square independence test** is used to test the independence of two random variables. Using a chi-square test, you can determine whether the occurrence of one variable affects the probability of occurrence of the other variable.

To use the test,

1. The observed frequencies must be obtained from a random sample
2. Each expected frequency must be ≥ 5

The sampling distribution for the test is a chi-square distribution with

$$(r-1)(c-1)$$

degrees of freedom, where r and c are the number of rows and columns, respectively, of the contingency table. The test statistic for the chi-square independence test is

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

where O represents the observed frequencies and E represents the expected frequencies.

To begin the test we state the null hypothesis that the variables are independent and the alternative hypothesis that they are dependent.

Performing a Chi-Square Test for Independence (p 507)		
Procedure	Equations	Example2 **(p 507)**
Identify the claim. State the	State H_0 and H_1	H_0: CEO's ages are

null and alternative hypotheses.		independent of company size H_1: CEO's ages are dependent on company size.
Specify the level of significance	Specify α	$\alpha = 0.01$
Determine the degrees of freedom	$d.f. = (r-1)(c-1)$	$d.f. = (2-1)(5-1) = 4$
Find the critical value.	χ^2_α :Obtain from Table 6, Appendix B	$\chi^2_\alpha \geq 13.277$
Identify the rejection region	$\chi^2 \geq \chi^2_\alpha$	$\chi^2 \geq 13.277$
Calculate the test statistic	$\chi^2 = \sum \frac{(O-E)^2}{E}$	$\sum \frac{(O-E)^2}{E} \approx 77.9$ Note that O is in the table of actual CEO's ages above, and E is in the table of Expected CEO's ages (if independent of size) above
Make a decision to reject or fail to reject the null hypothesis	Reject if χ^2 is in the rejection region. Equivalently, we reject if the P-value (the probability of getting as extreme a value or more extreme) is $\leq \alpha$	Since 77.9 > 13.277 we reject the null hypothesis Equivalently $P(X \geq 77.0) < \alpha$ so reject the null hypothesis. (Note Table 6 of Appendix B doesn't have a value less than 0.005.)
Interpret the decision in the context of the original claim		CEO's ages and company size are dependent.

The test statistic (77.887) is greater than the critical value obtained from Table 6, Appendix B (13.277) so the null hypothesis is rejected. (Alternatively the P-Value (0.000) is less than the level of significance, α (0.01) so the null hypothesis is rejected.)

3. An urban geographer randomly samples 20 new residents of a neighborhood to determine their ratings of local bus service. The scale used is as follows: 0–very dissatisfied, 1–dissatisfied, 2– neutral, 3–satisfied, 4–very satisfied. The 20 responses are 0,4,3, 2,2,1,1,2,1,0,01,2,1,3,4,2,0,4,1. Use the sign test to see whether the population median is 2.

   Solution:

   There are 5 observations above the hypothesized median. Because the sample size is larger than 10, we test using the sample proportion $p = 5/20 = 0.25$. Using the PROB-VALUE method the steps in this test are:

   1) H_0: $\pi = 0.5$ and H_A: $\pi \neq 0.5$
   2) We will use the *Z*-distribution
   3) We will use the 5%-level, thus $\alpha = 0.05$
   4) The test statistic is $z = (0.25 - 0.5)/\sqrt{0.25/20} = -2.24$
   5) Table A-4 shows that $P(|Z| > 2.24) \approx 0.025$.
   6) Because PROB-VALUE < α, we reject H_0. We conclude π is different than 0.5, and thus the median is different than 2.

4. A course in statistical methods was team-taught by two instructors, Professor Jovita Fontanez and Professor Clarence Old. Professor Fontanzez used many active learning techniques, whereas Old employed traditional methods. As part of the course evaluation, students were asked to indicate their instructor preference. There was reason to think students would prefer Fontanez, and the sample obtained was consistent with that idea: of the 12 students surveyed, 8 preferred Professor Fontanez and 2 preferred Professor Old. The remaining students were unable to express a preference. Test the hypothesis that the students prefer Fontanez. (*Hint:* Use the sign test.)

   Solution:

   Although the sample is large enough for a normal approximation, we will use the binomial distribution to illustrate its application. Of the 12 observations, 8 preferred Prof. Fontanez, thus we need the probability of observing 8 or more successes in 12 trials of a Bernoulli process with the probability of success equal to 0.5. From Table A-1, we get

$$P(X \geq 8) = P(X=8) + P(X=9) + P(X=10) + P(X=11) + P(X=12)$$

$$P(X \geq 8) = 0.1208 \quad + 0.0537 \quad + 0.0161 \quad + 0.0029 \quad + 0.002 \quad = 0.1937$$

Adopting the 5% uncertainty level, we see that PROB-VALUE > α. Thus we fail to reject H_0. We cannot conclude students prefer Fontanez.

5. Use the data in Table 10-8 to perform two Mann–Whitney tests: (a) compare uncontrolled intersections and intersections with yield signs, and (b) compare uncontrolled intersections and intersections with stop signs.

   Solution:

   (a) The rank sums are 119.5 and 90.5 for the yield-signed and uncontrolled intersections respectively. Given the small sample size, we use an exact test rather than the normal approximation. The associated PROB-VALUE is 0.272. Adopting a 5% level of uncertainty, we fail to reject the hypothesis of no difference. We cannot conclude the samples were drawn from different populations.

   (b) The rank sums are 130.5 and 59.5 for the stop-signed and uncontrolled intersections respectively. Given the small sample size, we use an exact test rather than the normal approximation. The associated PROB-VALUE is 0.013. Adopting a 5% level of uncertainty, we reject the hypothesis of no difference. We conclude the samples were drawn from different populations.

6. Solid-waste generation rates measured in metric tons per household per year are collected randomly in selected areas of a township. The areas are classified as high-density, low density, or sparsely settled. It is thought that generation rates probably differ because of differences in waste collection and opportunities for on-site storage. Do the following data support this hypothesis?

High Density	Low Density	Sparsely Settled
1.84	2.04	1.07

3.06	2.28	2.31
3.62	4.01	0.91
4.91	1.86	3.28
3.49	1.42	1.31

Solution:

We will use the multi-sample Kruskal-Wallis test with an uncertainly level α = 0.1. The null hypothesis is that all samples have been drawn from the same population. The rank sums are 55, 39 and 26 for the high density, low density, and sparsely settled samples respectively. The Kruskal-Wallis statistic is

$$H = \frac{12}{15(15+1)}\left(\frac{55^2}{5} + \frac{39^2}{5} + \frac{26^2}{5}\right) - 3(15+1) = 4.22$$

Using the $\chi 2$ distribution with $3 - 1 = 2$ degrees of freedom, the associated PROB-VALUE is 0.121. We fail to reject the null hypothesis. The sample does not support the hypothesis of differing waste generation rates.

7. The distances travelled to work by a random sample of 12 people to their places of work in 1996 and again in 2006 are shown in the following table.

	Distance (km)			Distance (km)	
Person	1996	2006	Person	1996	2006
1	8.6	8.8	7	7.7	6.5
2	7.7	7.1	8	9.1	9
3	7.7	7.6	9	8	7.1
4	6.8	6.4	10	8.1	8.8
5	9.6	9.1	11	8.7	7.2
6	7.2	7.2	12	7.3	6.4

Has the length of the journey to work changed over the decade?

Solution:

The sample can be considered as twelve paired observations. By taking differences between paired values, we get measures of the change for each individual. If the median change for the population is zero, we expect a sample to have a median difference near zero. Thus we will do a sign test for the median difference with a hypothesized value of zero. In other words, the hypotheses are $H_0 : \eta = 0$ and $H_A: \eta \neq 0$. We denote samples values whose distance decreased with a minus sign. Sample values with a positive difference get a plus sign. The sample becomes

S = {-,+,+,+,+,0,+,-,+,-,+,+}

Ignoring the tie, this is a sample of size 11 with 8 values above the hypothesized median. We are using Format (C) of Table 10-1, thus the PROB-VALUE is $2P(X \geq 8)$ where X is a binomial variable with $\pi = 0.5$. From the equation for the binomial, the PROB-VALUE is found to be 0.113. At the α = 10% level, we fail the reject the null hypothesis. We cannot conclude there has been a change in distance.

8. One hundred randomly sampled residents of a city subject to periodic flooding are classified according to whether they are on the floodplain of the major river bisecting the city or off the floodplain. These households are then surveyed to determine whether they currently have flood insurance of any kind. The survey results are as follows:

	On the Floodplain	Off the Floodplain
Insured	50	10
No Insurance	15	25

Test a relevant hypothesis.

Solution:

We will do a χ^2 test for a relationship between insurance and house location. The null hypothesis is no relationship (independence). Augmenting the data with expected frequencies, we have:

	On the Floodplain	Off the Floodplain
Insured	50 (39)	10 (21)
No Insurance	15 (26)	25 (14)

The corresponding χ^2 value is 22.16. Table A-8 shows that with 1 degree of freedom, $P(\chi^2 > 20)$ is zero to 3 decimal places. Thus for any reasonable level of uncertainty (any $\alpha < 0.0005$), we can reject the null hypothesis.

9. The occurrence of sunshine over a 30-day period was calculated as the percentage of time the sun was visible (i.e., not obscured by clouds). The daily percentages were:

Day	Percentage of sunshine	Day	Percentage of sunshine	Day	Percentage of sunshine
1	75	11	21	21	77
2	95	12	96	22	100
3	89	13	90	23	90
4	80	14	10	24	98
5	7	15	100	25	60
6	84	16	90	26	90
7	90	17	6	27	100
8	18	18	0	28	90
9	90	19	22	29	58
10	100	20	44	30	0

If we define a sunny day as one with over 50% sunshine, determine whether the pattern of occurrence of sunny days is random.

Solution:

For this we can use the number-of-runs test. Rather than calculate runs across two samples, here we will simply note if a day has 50% or more sunshine. The sample becomes

S={+,+,+,+,-,+,+,-,+,+,-,+,+,-,+,+,-,-,-,-,+,+,+,+,+,+,+,+,-}

We see that the sample consists of 12 runs. There are $n_x = 21$ sunny days, and $n_y = 9$ cloudy days. Because $n_x < 20$, we cannot use the normal approximation given in Table 10-5. Instead the probability is computed using combinatorial rules, and is approximately 0.4. This is far too large for rejection of the randomness hypothesis. We cannot conclude the pattern is non-random.

10. Test the normality of the DO data (a) using the Kolmogorov–Smirnov test with the ungrouped data of Table 2-4 and (b) using the χ^2 test with $k = 6$ classes of Table 2-6.

Solution:

(a) We will take the mean and standard deviation as known rather than estimated from the sample. Doing so results in calculated PROB-VALUES that are smaller than the true values (i.e., we are more likely to reject the null hypothesis). For the DO data the mean and standard deviation are 5.58 and 0.39 respectively. We sort the data, and then find the differences between the observed and expected cumulative distributions. The table below shows the results for a few of the 50 observations:

x_i	$S(x_i)$	$F(x_i)$	$\lvert S(x_i)-F(x_i)\rvert$
4.2	0.020	0.015	0.005
4.3	0.040	0.023	0.017
4.4	0.060	0.032	0.028
…	…	…	…
5.9	0.780	0.692	0.088

…	…	...	…
6.7	0.960	0.960	0.000
6.8	0.980	0.972	0.008
6.9	1.000	0.981	0.019

The maximum difference is 0.088. Table A-9 shows that with 50 degrees of freedom, the corresponding PROB-VALUE is about 0.6. We obviously cannot reject the hypothesis of normality.

(b) Here we will take the mean and standard deviation as unknown, to be estimated from the sample. In other words, we estimate two parameters from the sample. In building the χ^2 table, we combine the first two and the last two categories in Table 2-6 to ensure at least 2 expected frequencies per cell. This reduces the number of categories 4, as seen in the table below:

Group	Minimum	Maximum	O_j	E_j	$(O_j-E_j)^2/E_j$
1	4.000	4.990	9	3.3	10.13
2	5.000	5.490	10	17.0	2.89
3	5.500	5.990	20	21.7	0.14
4	6.000	6.990	11	7.0	2.24

The observed Chi-square value is 15.4. With $k - p - 1 = 4 - 2 - 1 = 1$ degrees of freedom, Table A-8 shows that the PROB-VALUE is less than 0.0005. We therefore reject the null hypothesis.

Note that with only 4 classes, we can obtain only a rough idea of the distribution of DO. The 4 classes given in Table 2-6 do not yield a distribution that is at all similar to the normal distribution. In practice one would need many classes (and observations) for the χ^2 test to be reliable.

Analysis of Variance (ANOVA)

I. Introduction

In Regression, the decomposition of the total sum of squares (SST) into the "explained" sum of squares (SSR) and the "unexplained" sum of squares (SSE) took place in the Analysis of Variance or ANOVA table. However, ANOVA also refers to a statistical technique used to test for diffferences between the means for several populations. While the procedure is related to regression, in ANOVA the independent variable(s) are qualitative rather than quantitative. In both regression and ANOVA the dependent variable is quantitative.

Example 1: As city manager, one of your responsibilities is purchasing. The city is looking to buy lightbulbs for the city's streetlights. Aware that some brands' lightbulbs might outlive other brands' lightbulbs, you decide to conduct an experiment. Seven lightbulbs each are purchased from four brands (GE, Dot, West, and a generic) and placed in streetlights. The lifetime of each of the 28 lightbulbs is then recorded in the file "**Lightbulbs**."

In this example, the lifetime of a lightbulb, in thousands of hours, is the quantitative dependent variable of interest. The company marketing the lightbulb, i.e., the brand-name, is the qualitative independent variable. The variable "brand name" has four possible values (or four "levels" in the terminology of ANOVA). The letter "*k*" will be used for the number of "levels" of the independent variable or "factor". Here, k = 4 for the four brands being tested. We say, "the factor *brand-name* has four levels: GE, Dot, West, and generic."

The "populations" referred to in these notes are simply the different levels of the factor. So that, in this example, we are interested in whether the mean lifetimes for the four populations of lightbulbs differ. Since we cannot know with certainty, however, the true mean lifetime of all lightbulbs carrying a certain brand-name, we rely upon statistics to determine if the differences observed *between* samples drawn from the four brands are statistically significant. (Non-significant differences are those that can plausibly be attributed to chance, i.e., sample-to-sample, variation alone.)

II. The (one-way) ANOVA Model

In order to perform tests of statistical significance, a model is assumed. The model used in ANOVA is similar in many respects to the model employed in regression. In fact, You may find it useful in these notes to make analogies between the model and formulas in ANOVA and the the corresponding model and formulas in regression. In the model below, recall that the dependent (or **response**) variable is quantitative as in regression, but the independent (or **factor**) variable is now qualitative. We begin with a model in which a single independent variable is used to describe the dependent variable. This **One-Way** ANOVA is analogous to simple regression.

Terminolgy

Although regression and analysis of variance are closely related, historically they developed separately. As a result they each adopted their own terminolgy. Unfortunately, this often means that similar things are referred to differently in the two procedures. Below is a list of some of the names used in ANOVA and what they refer to.

- Response: the dependent variable
- Factor(s): the independent variable(s)
- Levels: the possible values of a factor
- Treatments: another name for levels in one-way ANOVA, but there will be a distinction between levels and treatments when we discuss two-way ANOVA later. The term treatments derives from medicine, where the different treatments were the drugs or procedures being tested on patients, and agriculture, where the treatments were the different fertilizers or pesticides being tested on crops.
- The μ_i are called the "factor-level means" or the "treatment means" in one-way ANOVA and represent the true mean value of the response variable for the i^{th} population of treatments.

Example 1 (continued): For the lightbulb problem,

- the response is the lifetime of a particular lightbulb (in thousands of hours)
- the factor is the brand-name
- there are four levels or treatments: GE, Dot, West, and generic
- μ_{GE} is the mean lifetime of all GE bulbs, μ_{Dot} is the mean lifetime of all Dot bulbs, etc.

IV. Hypothesis Test

As usual, we rely on a hypothesis test to determine if the sample means for the k samples drawn (one from each population) differ enough for the difference to be statistically significant (more than would likely occur due to random chance alone).

Example 1 (continued): It is important that the student understand why probability is important here. It is not unusual for one manufacturer to source a product marketed under many brand-names. For example, there are only a handfull of companies manufacturing denim jeans, but there are dozens of brand-name jeans available to the consumer. Similarly, not all lightbulbs are manufactured by the companies marketing them. It is not inconceivable, therefore, that all four brands of lightbulbs being tested by the city come off of the same assembly line. Yet, when tested, they would still yield four *different* sample means simply because of sample-to-sample variation. As city manager, you might be more than a little embarassed to discover that the brand that you've touted as superior to all others is actually different in name only! Lawsuits have been lost for far less.

Hypotheses:

- **H_0:**, i.e., all population means are equal. This is equivalent to saying that the *k* treatments have no differential effect upon the value of the response.
- **H_A:** At least two of the means differ. This says that different treatments produce different values of the response variable, on average.

Test Statistic:

F = $\frac{MSR}{MSE}$ **,** where ***MSR*** = the **<u>M</u>ean <u>S</u>quare** for **<u>T</u>reatments**,
and ***MSE*** = the **<u>M</u>ean <u>S</u>quare** for **<u>E</u>rror**

Note: What I'm calling ***MSR*** is often called *MST* in the literature. I've chosen to continue the use of *MSR* to highlight the similarity between regression and analysis of variance. MSE remains the same for both regression and analysis of variance. Formulas for the mean squares are given later in the notes.

Logic:

The analysis of variance uses the ratio of two ***variances***, *MSR* and *MSE*, to determine whether population ***means*** differ; hence the name "analysis of variance." Recall that one of the assumptions of the model is that the variance σ^2 is the same for all populations. ***MSE*** provides an unbiased estimate of σ^2 in ANOVA just as it does in regression (see regression notes). If the population means are all equal, which is the null hypothesis, it can be shown that *MSR also* provides an unbiased estimate of σ^2. If all of the population means are equal, therefore, we would expect **F** to be nearly equal to **1** since *MSR* and *MSE* should yield similar estimates of the variance σ^2.

If some population means differ from others, however, *MSR* will tend to be bigger than *MSE* resulting in an **F - Ratio** substantially larger than **1**. Thus we reject **H_0** for large values of **F**, just as in regression.

V. The ANOVA Table: Sums of Squares and Degrees of Freedom

A. Introduction

At the heart of any analysis of variance is the ANOVA Table. The formulas for the sums of squares in ANOVA are simplified if the *k* samples are all of the same size n_S. In the interests of simplicity, therefore, the following discussion assumes that all *k* samples contain the same number of observations n_S.

B. Notation

- The index i represents the i[th] population or treatment, where i ranges from 1 to k
- The index j represents the j[th] obsevation within a sample, where j ranges from 1 to n_S
- n is the total number of observations from all samples
- y_{ij} is the value of the j[th] observation in the i[th] sample
- $\bar{y}_i$ is the mean of the i[th] sample
- $\bar{\bar{y}}$ (read "y double-bar") is the mean of all n observations, $\bar{\bar{y}} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_S} y_{ij}$, or the mean of the sample means (hence the "double-bar" in the name), $\bar{\bar{y}} = \frac{\bar{y}_1 + \bar{y}_2 + \cdots + \bar{y}_k}{k}$

C. Sums of Squares

<u>S</u>um of <u>S</u>quares for <u>T</u>reatments, $SSR = n_S \sum_{i=1}^{k}\left(\bar{y}_i - \bar{\bar{y}}\right)^2$ is the "Between Group" variation, where the k "groups" or populations are represented by their sample means. If the sample means differ substantially then SST will be large.

<u>S</u>um of <u>S</u>quares for <u>E</u>rror, $SSE = \sum_{i=1}^{k}\sum_{j=1}^{n_S}\left(y_{ij} - \bar{y}_i\right)^2$ is the "Within Group" variation and represents the random or sample-to-sample variation

<u>Total</u> <u>S</u>um of <u>S</u>quares, $SST = \sum_{i=1}^{k}\sum_{j=1}^{n_S}\left(y_{ij} - \bar{\bar{y}}\right)^2$ is the total variation in the values of the response variables over all k samples. (Note: SST is the same as in regression)

D. Degrees of Freedom

Degrees of freedom for treatments, $\mathbf{df_{SSR} = k-1}$. Rather than memorizing this formula, just imagine the number of dummy variables that you would have to create to conduct the equivalent analysis in regression. Since you always leave one possibility out in regression, you would need to create k – 1 dummy variables. Since the resulting regression model would have k – 1 independent variables, SSR (SST here) would have k – 1 degrees of freedom.

Degrees of freedom for error, $\mathbf{df_{SSE} = \mathit{n} - \mathit{k}}$.

Total degrees of freedom, $\mathbf{df_{SST} = \mathit{n} - 1}$. This is the same result obtained in regression.

Note: The two component degrees of freedom sum to the total degrees of freedom, just as in regression.

E. Mean Squares

<u>M</u>ean <u>S</u>quare for <u>T</u>reatments, $MSR = \dfrac{SSR}{k - 1}$ is equivalent to MSR in regression

<u>M</u>ean <u>S</u>quare for <u>E</u>rror, $MSE = \dfrac{SSE}{n - k}$ is the same as MSE in regression. As in regression, MSE is an unbiased estimator of the common population variance □$^{□}$.

F. F – Ratio

The statistic used to test the null hypothesis □$_{□□}$□□□□$_{□□}$□□□□□□□□□□$_{□k}$ is $F = \dfrac{MSR}{MSE}$. As mentioned earlier, if the null hypothesis is correct then this ratio should be close to one. If some of the sample means differ substantially, however, the ratio will be much larger. Large values of F therefore correspond to strong evidence for rejecting H_0. Statgraphics reports a *P*-value for the test.

G. Summary

The ANOVA Table below summarizes some of the information in this section

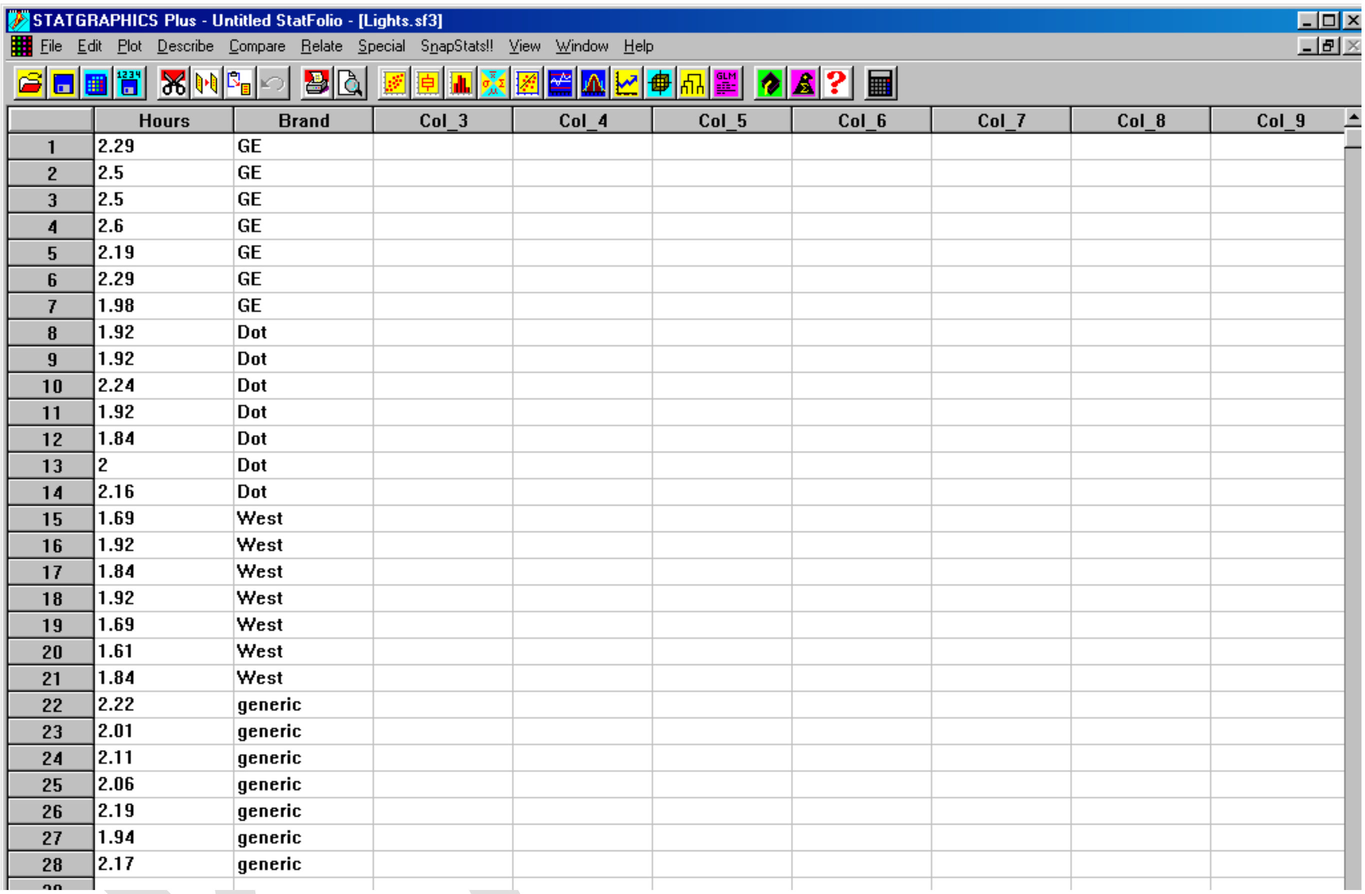
ANOVA Table for One-Way Analysis of Variance

Source	*Sum of Squares*	*Df*	*Mean Square*	*F-Ratio*	*P-Value*
Between groups	***SSR***	*k* - 1	***MSR* = *SSR*/(***k*-1)	***F* = *MSR/MSE***	
Within groups	***SSE***	*n* - *k*	***MSE* = *SSE***/(n-k)		
Total (Corr.)	***SST***	*n* - 1			

VI. Using Statgraphics

To perform a one-way analysis of variance in Statgraphics, follow *<u>C</u>ompare* > *<u>A</u>nalysis of Variance* > *<u>O</u>ne-Way ANOVA* and enter the response and factor into the dependent variable and factor fields, respectively.

Example 1 (continued): For the lightbulb problem, the spreadsheet might look like the one below. Notice that the qualitative factor Brand doesn't need to be numeric. Statgraphics will treat the factor in ANOVA as qualitative, so there is no need to recode it as a numeric variable. For the same reason there is no need to create dummy variables as in regression.



	Hours	Brand	Col_3	Col_4	Col_5	Col_6	Col_7	Col_8	Col_9
1	2.29	GE							
2	2.5	GE							
3	2.5	GE							
4	2.6	GE							
5	2.19	GE							
6	2.29	GE							
7	1.98	GE							
8	1.92	Dot							
9	1.92	Dot							
10	2.24	Dot							
11	1.92	Dot							
12	1.84	Dot							
13	2	Dot							
14	2.16	Dot							
15	1.69	West							
16	1.92	West							
17	1.84	West							
18	1.92	West							
19	1.69	West							
20	1.61	West							
21	1.84	West							
22	2.22	generic							
23	2.01	generic							
24	2.11	generic							
25	2.06	generic							
26	2.19	generic							
27	1.94	generic							
28	2.17	generic							

This leads to the ANOVA Table below. Looking at the *P*-value for the *F*-test, we conclude that there is strong evidence that at least two of the mean lifetimes differ.

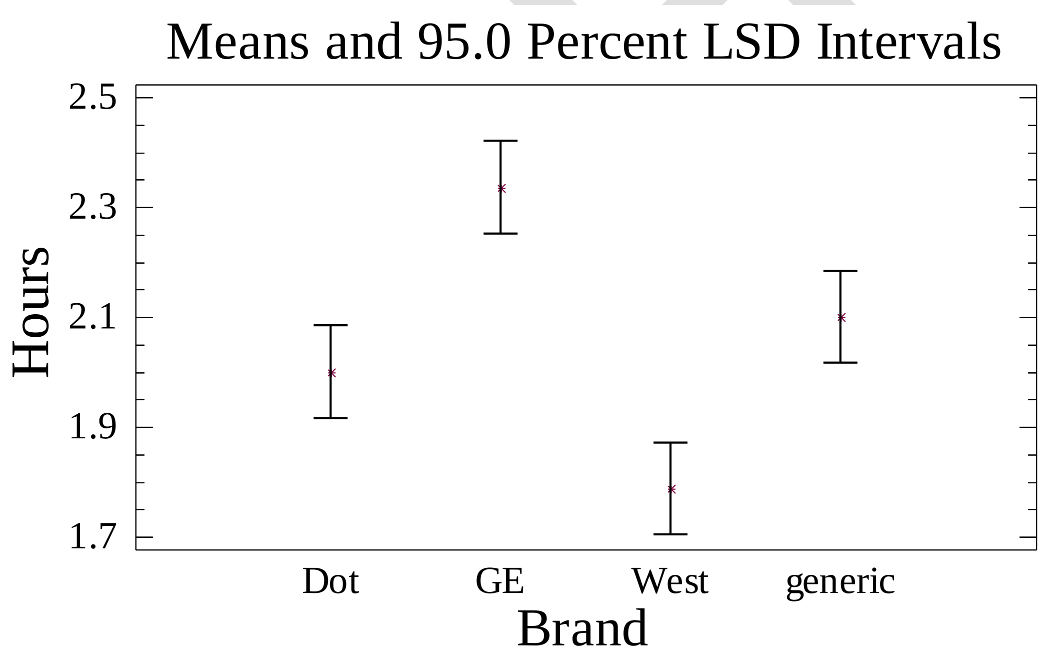
ANOVA Table for Hours by Brand

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio
Between groups	1.08917	3	0.363057	15.62
Within groups	0.557714	24	0.0232381	
Total (Corr.)	1.64689	27		

Once the city manager has detected a difference in mean lifetimes, he/she would naturally wish to determine which brand's lightbulbs are superior. Statgraphics has a graphical option called a "Means Plot" which graphs 95% confidence intervals for the mean lifetimes of the four brands. If the 95% confidence intervals for two brands don't overlap then the city manager may conclude, at the 5% level of significance, that the true mean lifetimes for the two brands differ. If, on the other hand, the intervals *do* overlap the manager cannot draw a statistically significant conlusion at the 5% level of significance. (Remember, it's quite possible that the two brands' bulbs come off of the same assembly line, so don't try to force conclusions that can't be supported statistically!)

Below is the *Means Plot*. There is clearly evidence, at the 5% level of significance, that the GE bulbs last longer, on average, than bulbs from the other brands. Similarly, there is evidence, at the 5% level, that the West bulbs fail sooner, on average, than bulbs from the other brands. The sample differences between the Dot and generic bulbs, however, may be due to chance alone. (We don't actually *know* that Dot and generic bulbs are interchangeable, but the sample doesn't provide strong enough evidence to discount the possibilty.)



VII. Two-Way ANOVA

When the effects of two qualitative factors upon a quantitative response variable are investigated, the procedure is called two-way ANOVA. Although a model exists for two-way analysis of variance, similar to the multiple regression model, it will not be covered in this class. Neither will we cover the details of the ANOVA Table. Nevertheless, there are some new considerations in two-way ANOVA stemming from the presence of the second factor in the model.

Example 2: The EPA (Environmental Protection Agency) tests public bodies of water for the presence of *coliform* bacteria. Aside from being potentially harmful to people in its own right, this bacteria tend to proliferate in polluted water, making the presence of *coliform* bacteria a surrogate for polution. Water samples are collected off public beaches, and the number of *coliform* bacterial per cc is determined. (See the file "**Bacteria**."

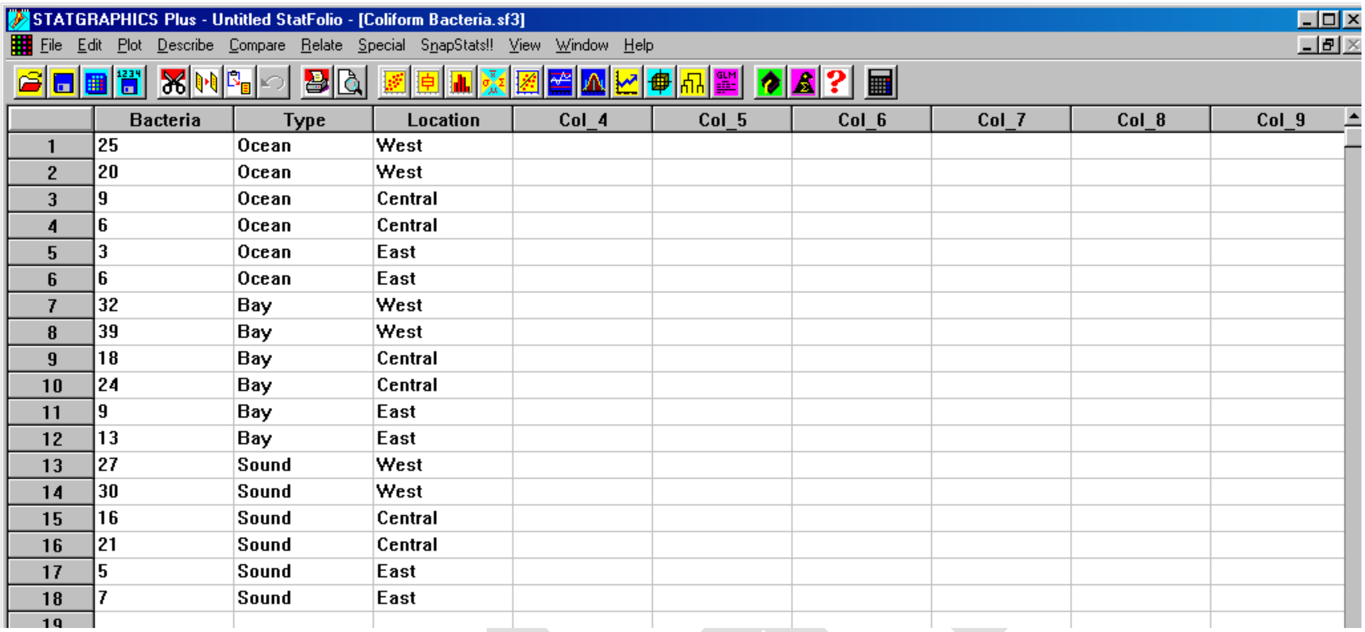
The EPA is interested in determining the factors that affect *coliform* bacterial formation in a particular county. The county has beaches adjacent to the ocean, a bay, and a sound. The EPA beleives that the amount of "flushing" a beach gets may affect the ability of polution to accumulate in the waters off the beach. The EPA also believes that the geographical location of the beach may be significant. (There could be several reasons for this: the climate may be different in different parts of the county, or the land-use may vary across the county, etc.)

As luck would have it, there is at least one beach for each combination of type (ocean, bay, sound) and location (west, central, east) within the county. Because of this, the EPA decides to sample a beach at each of the 9 possible combinations of type and location and conduct a two-way analysis of variance for *coliform* bacterial count. Two independent samples are taken at each beach to allow for an estimation of the natural variation in *coliform* bacterial count (this "repetition" is needed for the computation of MSE, which estimates the sample-to-sample variance in bacterial counts).

VIII. Two-Way ANOVA Using Statgraphics

To perform a two-way analysis of variance in Statgraphics, follow *Compare > Analysis of Variance > Multifactor ANOVA* and enter the response and factors into the dependent variable and factor fields, respectively.

Example 2 (continued): Since data from such a study often appears in the form of a two-way table, with one factor as the row variable, the second as the column variable, and the observations as values in the row-by-column cells, it is important to remember that each variable must have its own column in the spreadsheet as in the example below. (This may require that you re-format the original spreadsheet prior to beginning the analysis.)



	Bacteria	Type	Location	Col_4	Col_5	Col_6	Col_7	Col_8	Col_9
1	25	Ocean	West						
2	20	Ocean	West						
3	9	Ocean	Central						
4	6	Ocean	Central						
5	3	Ocean	East						
6	6	Ocean	East						
7	32	Bay	West						
8	39	Bay	West						
9	18	Bay	Central						
10	24	Bay	Central						
11	9	Bay	East						
12	13	Bay	East						
13	27	Sound	West						
14	30	Sound	West						
15	16	Sound	Central						
16	21	Sound	Central						
17	5	Sound	East						
18	7	Sound	East						
19									

The default ANOVA Table below has separate rows for the factors Type (called factor A) and Location (called factor B). A test of the significance of each factor is performed and the corresponding p-value displayed. It appears that both the type of beach and its location affect *coliform* bacterial count.

But does the effect of the beach type on bacteria count depend upon its location within the county? If the particular pairings of factor levels are important, the factors are said to "interact."

Source	Sum of Squares	Df	Mean Square	F-Rat
MAIN EFFECTS				
A:Type	364.778	2	182.389	16.
B:Location	1430.11	2	715.056	62.
RESIDUAL	148.222	13	11.4017	
TOTAL (CORRECTED)	1943.11	17		

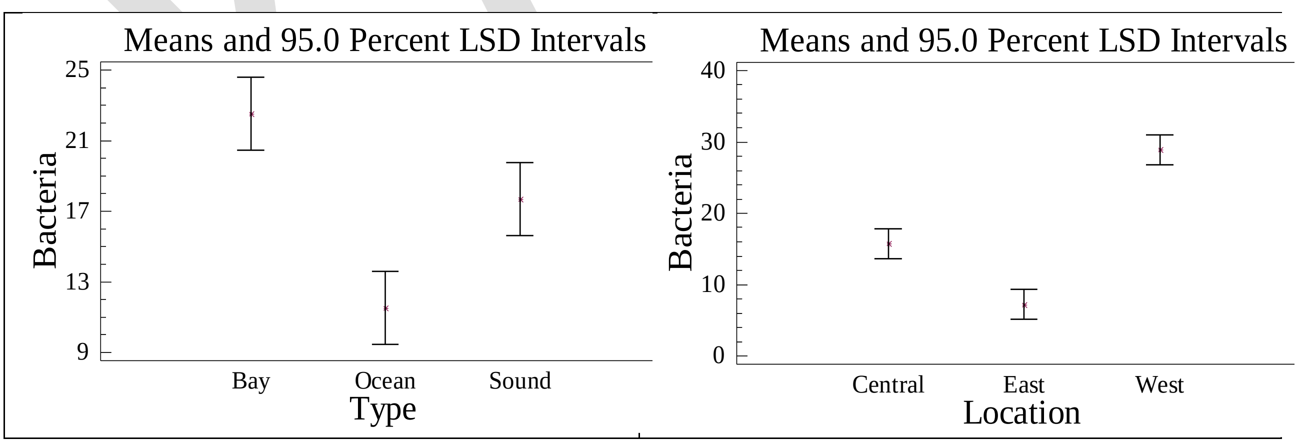
Before interpreting the results in the ANOVA table above, we should consider the role that interaction plays. If the effect of beach type on bacteria formation depends on the location of the beach then it is better to investigate the *combinations* of the levels of the factors type and location for their affect on bacteria. It will come as no surprise to you that there is a hypothesis test for interactions.

H$_0$: The factors Type and Location do *not* interact.
H$_A$: The factors Typ and Location *do* interact

To check for interaction, use the right mouse button and *Analysis Options* and enter “2” for the *Maximum Order Interaction*. The resulting output for our example below shows a *P*-value of 0**.**3047 for the test for interactions. Thus the evidence for interaction is not particularly strong. The practical effect of discounting interaction is that we are able to return to the previous output (the one without interactions) and interpret the *P*-values for the factors Type and Location separately. Since the *P*-values for both factors are significant, we conclude that factors affect bacteria growth.

Source	Sum of Squares	Df	Mean Square	F-Rat
MAIN EFFECTS				
A:Type	364.778	2	182.389	18.
B:Location	1430.11	2	715.056	70.
INTERACTIONS				
AB	57.2222	4	14.3056	1.
RESIDUAL	91.0	9	10.1111	
TOTAL (CORRECTED)	1943.11	17		

Having determined that the type of beach and the beach’s location are both significant, we next investigate the nature of the relationship between these factors and bacteria count. Once again we turn to the means plots under *Graphical Options*. Statgraphics dfaults to a means plot for the factor Type because this was the first factor entered in the *Input Dialog Box*. To get a means plot for the factor Location, use *Pane Options* to select it. The two means plots appear below.



Individually, these means plots are interpreted as in one-way ANOVA. There is evidence, at the 5% level of significance, that the mean bacteria count at ocean beaches is less than for other types, and that the mean count is highest at bay beaches. Similarly, the mean count is lowest in the east and greatest in the west, with all differences being statistically significant at the 5% level of significance. Furthermore, ***because interactions were judged not-significant***, we can add the main effects together and say that the least polluted beaches tend to be located in the east on the ocean, while the most polluted tend to be in the west on bays. We could not have added the separate (or main) effects in this way if there had been significant interact, for in that case the effect upon bacteria count at a particular type of beach (ocean, for example) may be very different for different locations.

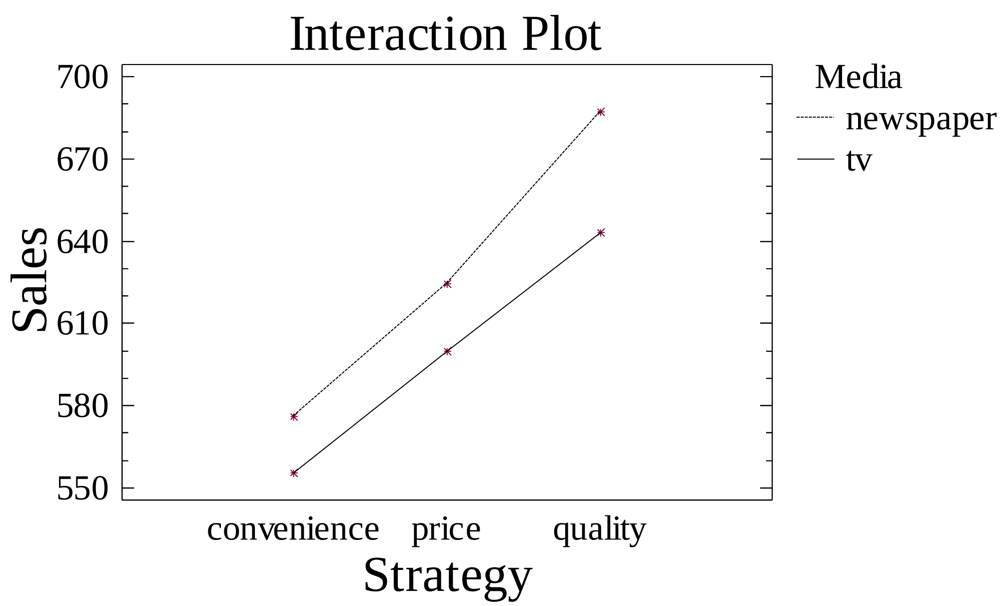
Example 3: The last two examples are based on a marketing study. A new apple juice product was entering the marketplace. It had three distinct advantages relative to existing apple juices. First, it was not a concentrate and was therefore considered to be of higher "quality" than many similar products. Second, as one of the first juices packaged in cartons, it was cheaper than competing products. Third, partly because of the packaging, it was more convenient. The director of marketing for the company would like to know which advantage should be emphasized in advertisements. The director would also like to know whether local television or newspapers are better for sales.

Consequently, six cities with similar demographics are chosen, and a different combination of "Marketing Strategy " and "Media" is tried in each. The unit sales of apple juice for the ten weeks immediately following the start of the ad campaigns are recorded for each city in the file **Apple Juice (two-way)**. The two-way table below describes the city assignments for the six possible combinations of levels for the two factors. Below the assignment table is the ANOVA Table for interactions.

	Convenience	Quality	Price
Local Television	City 1	City 3	City 5
Newspaper	City 2	City 4	City 6

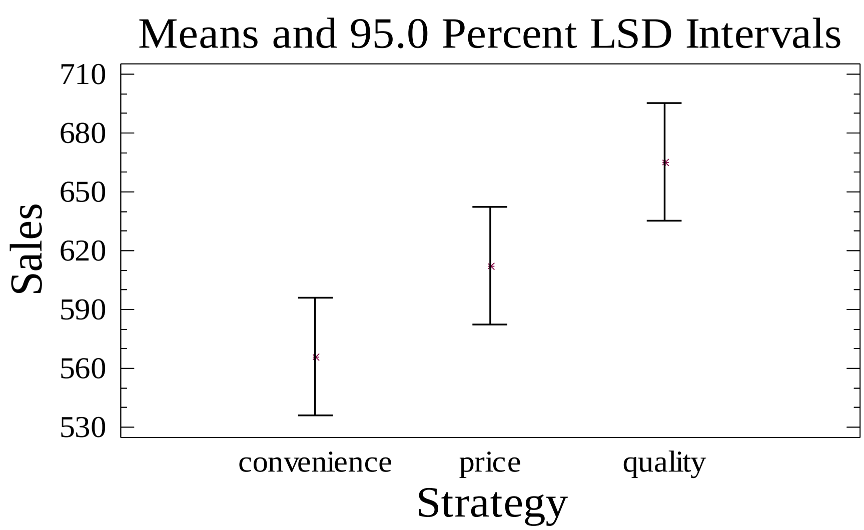
Source	Sum of Squares	Df	Mean Square	F-Rat
MAIN EFFECTS				
A:Strategy	98838.6	2	49419.3	5.
B:Media	13172.0	1	13172.0	1.
INTERACTIONS				
AB	1609.63	2	804.817	0.
RESIDUAL	501137.0	54	9280.31	
TOTAL (CORRECTED)	614757.0	59		

Interactions are not significant to the model (p-value equals 0.9171), a fact which is reinforced by looking at the *Interaction Plot* under *Graphical Options*. Note that the two curves are almost parallel, a sign that interactions are not significant.



Removing interactions, we obtain the ANOVA Table below, from which we conclude that the marketing strategy is significant, but the media used probably isn't. Since only marketing strategy apppears to affect sales, we'll restrict ourselves to the means plot for the factor Strategy below. Only the difference in mean sales when emphasizing quality versus emphasizing convenience is statistically significant at the 5% level of significance.

Source	Sum of Squares	Df	Mean Square	F-Rat
MAIN EFFECTS				
A:Strategy	98838.6	2	49419.3	5.
B:Media	13172.0	1	13172.0	1.
RESIDUAL	502746.0	56	8977.61	
TOTAL (CORRECTED)	614757.0	59		



Example 4: This is just the apple juice problem revisited (see file "**Apple Juice – Remix**"). By a judicious rearrangement of sales figures, I've created a marketing study in which interactions are significant. (See the two-way table below for the new assignments.) The comparison of the interaction plots for this example and example 3 should help to clarify the role of interactions in the interpretation of ANOVA output. The small *P*-value of 0**.**0474 for the hypothesis test of interactions implies that certain combinations of marketing strategy and media are important to sales.

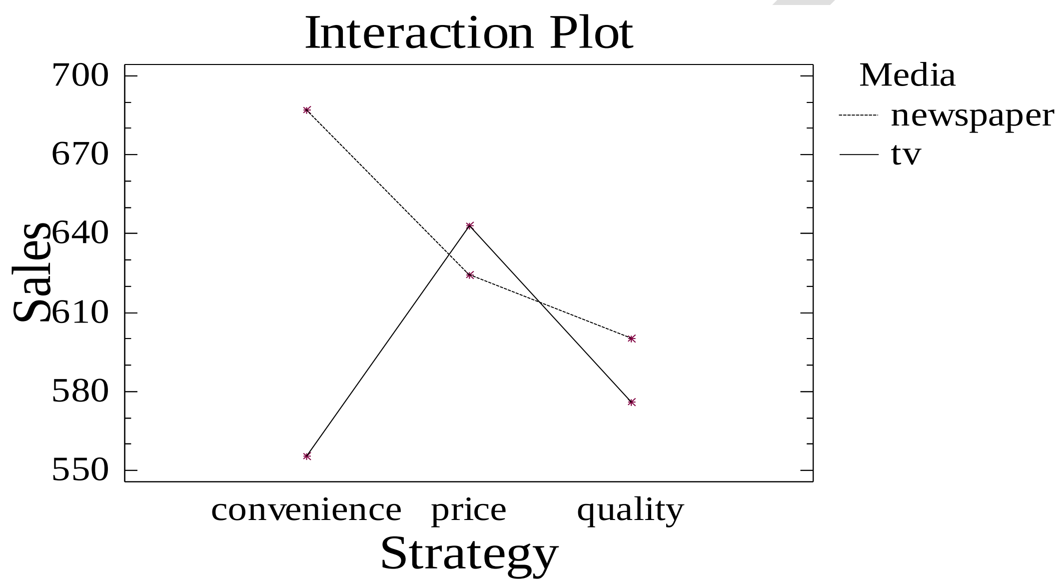
	Convenience	Quality	Price
Local Television	City 1	City 2	City 3
Newspaper	City 4	City 5	City 6

```
Source                    Sum of Squares     Df    Mean Square     F-Rat
------------------------------------------------------------------------
MAIN EFFECTS
 A:Strategy                      22393.6      2        11196.8        1.
 B:Media                         31327.4      1        31327.4        3.

INTERACTIONS
 AB                              59899.3      2        29949.6        3.

RESIDUAL                        501137.0     54        9280.31
------------------------------------------------------------------------
TOTAL (CORRECTED)               614757.0     59
```

Looking at he interaction plot, notice that emphasizing convenience lead to both the lowest and highest mean sales, depending upon whether local television or newspapers were used. Thus, it wouldn't make sense to talk about the effect of emphasizing convenience without consideration of the media used, i.e., we should only interpret levels of the two factors taken together (the combinations). Therefore, we will not investigate the means plots for Strategy and Media. From the interaction plot, it appears that the most effective campaign would emphasize convenience in newspapers. The least effective combination is to emphasize convenince on local television. (Note: Since the interaction plot doesn't display confidence intervals for the six possible combinations, we cannot attach a particular significance level to our conclusions as we could with the means plots.)



Hypothesis testing

Hypothesis testing
Up till now, we have dealt mainly with descriptive statistics, but as we've mentioned before, we also have inferential statistics at our disposal. These are statistics that allow us to make statements about one or more population based upon one or more samples that we've taken from the population. This allows us to test various hypotheses. There are many ways in which to do this, and we will cover only a few in this course.
In this chapter, we will go through the rationale behind hypothesis testing, and how we go about determining whether to reject, or fail to reject a null hypothesis, and in the process, decide whether our sample statistic is significantly different from some other measurement important to our experimental objectives.

Null vs alternate hypothesis

Hypothesis testing is a systematic model to summarise the evidence in order to decide between possible hypotheses.

Inferential stats are based upon the idea of a null hypothesis and an alternative hypothesis. The null hypothesis (written as HO:) is a statement written in such a way that there is no difference between two items. When we test the null hypothesis, we will determine a P value, which provides a numerical value for the likelihood that the null hypothesis is true. If it is unlikely that the null hypothesis is true, then we will reject our null hypothesis in favour of an alternate hypothesis (written as HA), and this states that the two items are not equal.

One can use the analogy of the criminal justice system. If one is arrested, the null hypothesis is that one is innocent of the crime. The state has the burden of showing that this null hypothesis is not likely to be true. If the state does that, then the judge rejects the null hypothesis and accepts the alternate hypothesis that you are guilty. Thus the state has to show that you are not innocent, in order to reject the null hypothesis. It is similar in statistics – your statistical test must show that the two items are different.

In the trial, if the null hypothesis is not rejected, your innocence has not been proven. It is just that the state failed to support your guilt. You are never proven innocent in terms of the trial: the state simply failed to show your guilt. The same is true in stats: if you fail to show that the two items are different, then you fail to reject the null hypothesis, but you have not proven that the null hypothesis is true. In other words, you can reject the null hypothesis, but it is incorrect to say you have "accepted" the null hypothesis – this implies that you have shown the null hypothesis to be correct, and that is not the case. Philosophically, it is important to understand these concepts.

Remember that we always state our hypotheses in terms of population parameters.

Example

When analysing data collected from a sample, we want to use that data to answer a biological question. We want to use our sample estimates to make some inferences about the biological population under study. Lets use an example the average height of students at UWC versus the average height of students at Wits. We want to use the estimates of our samples of these two populations in order to determine if there is a difference in height between students attending the two different universities.

For our null hypothesis, we will state that with respect to the parameter of the average height, there is no difference between the two groups.

When we examine our sample estimates (our form of descriptive statistics), we see that the two groups are different. But does this mean that the populations are different? There are two possibilities: the first is that the populations are different, and this is why their estimates are different. In other words, our null hypothesis is wrong and should be rejected. The second possibility is that the populations are the same, and the difference seen in the samples is just due to random error. In other words, our initial assumption (the null hypothesis) is correct, and so we fail to reject the null hypothesis. And so we have to decide which of the two possibilities is the correct one.

Sample difference

We must now ask: how much difference is there in our sample?

We need to quantify the difference. Inferential statistics are numbers that quantify differences.

We will ask questions like: is the difference big or small? A small difference could happen just by random sampling error, so if the difference is small we will assume the second of the possibilities we mentioned (ie that the populations are the same). A big difference is unlikely to occur just by chance, so if the difference is big, then we will assume the first of the possibilities (ie that the populations are different).

Alpha possibility

To determine if we have a small or big difference, we ask what is the probability of obtaining this much difference just by chance if we have sampled populations that are not different (ie, if our null hypothesis is correct)? This probability is called “alpha (α) probability”.

A small difference has a large probability (>0.05) of occurring. A big difference has a small probability (≤0.05) of occurring.

Since the inferential statistic quantifies the difference, we must determine the probability of finding the particular value of the statistic. The sampling distribution of the statistic allows us to determine probability. If the alpha (α) probability of the statistic is > 0.05, then we fail to reject the null hypothesis. If the alpha (α) probability of the statistic is ≤0.05, then we reject the null hypothesis.

Type I, II errors

When we reject or fail to reject a null hypothesis, we hope we are making the right decision. However, there is always some probability of us being wrong.

There are two possible ways in which we could be wrong:

We might reject a null hypothesis that we should have rejected – in other words, we concluded that there is a difference when there really isn’t. Statisticians call this a Type I (one) error.

We might also fail to reject a null hypothesis that we should have rejected – in other words, we failed to find a difference that actually does exist. Statisticians call this a Type II (two) error. When we fail to reject a null hypothesis, the probability that we have committed a Type II error is called the beta (β) probability. The ability of a statistical test to avoid making a Type II error is called the power of a test. Power therefore refers to how well a test can detect a difference when it actually exists. A powerful test is one that can detect small differences.

When we make scientific conclusions, we want to have both α and β as small as possible. They are inversely related – in other words, as the one goes up, the other goes down. Statisticians have shown both theoretically and empirically, that you can minimize both α and β by using a value of 0.05. If you use a smaller α, the β goes up too high. This is why statisticians generally recommend that null hypotheses be rejected at the α = 0.05 value. Although it might seem like an arbitrary value to us biologists, there are actually good mathematical reasons for using 0.05.

The only way to simultaneously decrease both α and β is to increase your sample size.

Reasoning of hypothesis testing

These are the steps we generally follow when hypothesis testing. We make a statement in both the null and sometimes also alternative hypothesis form, about a parameter. Then we select the experimental units that comprise our sample, and gather data that allow us to calculate a sample parameter. We use these parameters in order to decide whether or not to reject the null hypothesis. In order to do that though, we need to know what probability level we want to use, and we then use our chosen probability level in order to decide whether or not to reject the null hypothesis.

Hypothesis stating

When we state our hypotheses, we need to decide whether we are going to be applying a one-sided or two-sided test, and then state our hypotheses accordingly.

Setting criterion

Next we set our criterion, in other words, we chose an appropriate probability level. As already discussed, we generally use an alpha level of 0.05. This will determine the regions of the distribution of our parameter (in this instance, we are looking at sample means) in which our sample mean will fall, and whether it means we reject or fail to reject the null hypothesis.

Z scores

You will already have encountered the idea of Z scores and how they are used in order to determine the probability that your parameter falls within either the expected, or the unexpected range of possibilities. We will quickly review what we need to know:

We are able to convert our sample mean for example, into a score that fits in somewhere on the standard distribution graph, and we call this the Z score. Z scores are a special application of the transformation rules. The z score for an item, indicates how far and in what direction, that item deviates from its distribution's mean, expressed in units of its distribution's standard deviation. The mathematics of the z score transformation are such that if every item in a distribution is converted to its z score, the transformed scores will necessarily have a mean of zero and a standard deviation of one. So we are able to calculate a Z score we call Z-critical, and this is the Z score that defines the boundary of the region you will use in order to reject, or fail to reject, the null hypothesis. The Z-test score is the value you will have calculated from your sample value.

Test statistics

We have more than one type of test statistic that we can use, other than the Z-test score. We also have the T-test score for example, and we will discuss this at greater length in the next chapter. As we've already seen, these test scores allow us to convert our original measurement from our data set, into units that feature on the standard distribution, and this allows us to look up the probability of our score occurring randomly for example, in the table.

Setting a criterion

Z scores are especially informative when the distribution to which they refer is normal. In every normal distribution, the distance between the mean and a given Z score cuts off a fixed proportion of the total area under the curve. Statisticians have provided us with tables such as table B2 in your textbook by Zar, indicating the value of these proportions for each possible Z. If your Z-test value falls beyond either of the two areas cut off by Z-critical, then you can reject the null hypothesis in favour of your alternate hypothesis. Should your Z-test value fall within the area under the peak of the curve, then you have to conclude that your sample (or samples) have yielded a statistic that is not among those cases that would only occur alpha proportion of the time, if the hypothesis tested is true. You then fail to reject the null hypotheis.

Making a decision
In this particular instance, we have selected alpha to equal 0.05. In other words, we want to know whether our sample mean lies in the null distribution region of 95%, or does it fall in a more extreme part of the distribution, where there is a greater than 95% chance of the mean being drawn from the population. In order to answer the question, we convert our sample mean to a Z-score, and observe where it falls in the Z, or standard, distribution. If it falls outside the region of Z = 1.65 (the critical region), then we are able to reject the null hypothesis.

One-tailed tests
When rejecting the null hypothesis in favour of the alternative hypothesis. We have more than one type of alternative hypothesis to select, depending on our particular experiment. We have both non-directional alternative hypotheses which we call two-tailed tests and we will discuss these later on in this chapter, and we have directional hypotheses, or one-tailed tests. In a one-tailed test, the direction of deviation from the null value is clearly specified. We place all of alpha in the one tail in a one-tailed test.
One-tailed tests should be approached with caution though: they should only be used in the light of strong previous research, theoretical or logical considerations. You need to have a very good reason beforehand that the outcome will lie in a certain direction.
One application in which one-tailed tests are used is in industrial quality control settings. For example, a company making statistical checks on the quality of its medical products is only interested in whether their product has fallen significantly below an acceptable standard. They are not usually interested in whether the product is better than average, this is obviously good, but in terms of legal liabilities and general consumer confidence in their product, they have to watch that their product quality is not worse than it should be. Hence one-tailed tests are often used.

Right-tailed tests
If using a one-tailed test, it can be either a right-tailed or a left-tailed test, and this just refers to the expected direction of your result. If you expect your sample mean to fall in the region beyond Z critical, then we refer to it as a right tailed test. If the sample mean is indeed larger than Z critical, then we can reject the null hypothesis. If it is less, then we may not.

Left-tailed tests

With a left-tailed test, we are expecting our sample mean to be less than Z critical, and we want to know whether it differs significantly from the population mean of 100 in the example we have here. If the sample mean is indeed less than Z critical, we are able to reject the null hypothesis.

Two-tailed tests
A two-tailed test requires us to consider both sides of the Ho distribution, so we split alpha, and place half in each tail. With a two-tailed test, we want to know whether our sample mean is significantly bigger or smaller than the population mean.

Two-tailed hypothesis testing
If our chosen alpha is 0.05 therefore, we will divide that into half, and use that figure to calculate our Z critical score, which will indicate the position on the distribution curve where, should our Z test score be observed to be either larger than, or smaller than, the Z critical, we can reject the null hypothesis.

One- and two-tail comparison
Here we have a comparison between a one-tail, and two-tail test, using the same alpha value.

POSSIBLE QUESTIONS
UNIT III

PART A (20 X 1 = 20 Marks)
Question number 1 – 20 Online examination

PART-B (5 x 2 = 10 Marks)

1. Three processes A, B and C are tested to see whether their outputs are equivalent. The following observations of outputs are made:

A	10	12	13	11	10	14	15	13
B	9	11	10	12	13			
C	11	10	15	14	12	13		

Given, table value of F for (2,16) d.f at 5% level of significance is 3.63. Carry out the analysis of variance and state your conclusion.
2. Explain one-way classification in ANOVA.
3. What are the criterions for a uniformly most powerful test?
4. Describe power functions and OC functions
5. Three processes A, B and C are tested to see whether their outputs are equivalent. The following observations of outputs are made:

A	10	12	13	11	10	14	15	13
B	9	11	10	12	13			

C	11	10	15	14	12	13

Given, table value of F for (2,16) d.f at 5% level of significance is 3.63. Carry out the analysis of variance and state your conclusion.

6. Describe power functions and OC functions

PART-C (5 x 6 = 30 Marks)

1. Suppose the National Transportation Safety Board (NTSB) wants to examine the safety of compact cars, midsize cars, and full-size cars. It collects a sample of three for each of the treatments (cars types). Using the hypothetical data provided below, test whether the mean pressure applied to the driver's head during a crash test is equal for each types of car. Use α = 5%.

	Compact cars	**Midsize cars**	**Full-size cars**
	643	469	484
	655	427	456
	702	525	402
$\bar{X}$	666.67	473.67	447.33
S	31.18	49.17	41.68

UNIT III

1	3	The hypothesis under test is	simple hypothesis	alternative hypothesis	**null hypothesis**	complex hypothesis	**null hypothesis**
2	3	A test based on a test statistic is classified as	**randomized test**	non-randomized test	sequential test	Bayes test	**randomized test**
3	3	Student's t test is applicable in case of	**small samples**	for sample of size between 5 and 30	large samples	sample size more than 100	**small samples**
4	3	Reject H_0 when it is false is known as	Type I error	Type II error	**Correct decision**	wrong decision	**Correct decision**
5	3	Accept H_0 when it is true is known as	Type I error	Type II error	**Correct decision**	wrong decision	**Correct decision**
6	3	Accept H_0 when it is false is known as	Type I error	**Type II error**	Correct decision	wrong decision	**Type II error**
7	3	In a sample of 10 items the degree of freedom for student's t test is	10	**9**	12	11	**9**
8	3	If the sample size is less than 30 then those samples may be regarded as	large samples	**small samples**	parameter	attitude	**small samples**
9	3	The range of statistic-t is	-1 to +1	**$-\infty$ to $+\infty$**	0 to ∞	0 to 1	**$-\infty$ to $+\infty$**
10	3	The distribution used to test goodness of fit is	F distribution	**χ^2 distribution**	t distribution	Z distribution	**χ^2 distribution**
11	3	Large sample theory is applicable when	**n>30**	n<30	n=30	n=10	**n>30**
12	3	95% of fiducial limits of population mean are	**A.M ± 1.96 S.E**	A.M ± 3.96 S.E	A.M ± 2.58 S.E	A.M ± 3.58 S.E	**A.M ± 1.96 S.E**
13	3	99% of fiducial limits of population mean are	A.M ± 1.96 S.E	A.M ± 3.96 S.E	**A.M ± 2.58 S.E**	A.M ± 3.58 S.E	**A.M ± 2.58 S.E**
14	3	A part of the population selected for study is called as	statistic	**sample**	parameter	event	**sample**
15	3	Test statistic Z =	**(X -μ)/S.E(X)**	(X +μ)/S.E(X)	Xμ/S.E(X)	(X* μ)/S.E(X)	**(X -μ)/S.E(X)**
16	3	In students t test standard error of single mean =	**S/√n-1**	S/√n-2	S/√n+1	S/√n+2	**S/√n-1**
17	3	Null hypothesis is denoted by	H_1	**H_0**	H_2	H	**H_0**
18	3	Alternative hypothesis is denoted by	H_0	**H_1**	H_2	H	**H_1**
19	3	The value of 5% level of significance is	2.58	**1.96**	1.64	2.33	**1.96**
20	3	The value of 1% level of significance is	**2.58**	1.64	1.96	2.33	**2.58**
21	3	If the sample size n=2, the students t-distribution reduces to	Normal distribution	F- distribution	**Cauchy distribution**	z – distribution	**Cauchy distribution**
22	3	If n, the sample size is larger than 30, the students t-distribution reduces to	**Normal distribution**	F- distribution	Cauchy distribution	Z- distribution	**Normal distribution**
23	3	The d.f. for students t based on a random sample of size n is	**n-1**	n	n-2	(n-2)/2	**n-1**
24	3	Degrees of freedom for statistic-chi-square in case of contingency table of order(3x3)is	3	**4**	2	1	4
25	3	Students 't' test is defined by the statistic	$t = x-\mu/\sqrt{s}/n$	$\sqrt{t} = x-\mu/\sqrt{s}/n$	**$t = x-\mu/s/\sqrt{n}$**	$t = x-\mu/s/n$	**$t = x-\mu/s/\sqrt{n}$**
26	3	Chi-square variate has _______ degree of freedom	2	**1**	5	7	**1**
27	3	Degree of freedom is denoted by	**Φ**	**v**	χ^2	fi	**v**
28	3	Students t- distribution, the name was given by	Arthur Henshalme	Benjamin	**William S.Cosset**	Karl Pearson	**William S.Cosset**
29	3	Chi-Square test was first used in testing statistical hypothesis by	**Karl Pearson**	Robert mayer	A.Fischer	Spearman	**Karl Pearson**
30	3	While testing significance of the difference of two sample means in case of small samples, the degree of freedom is calculated by	$V=n_1+n_2$	$V=n_1+n_2-1$	**$V=n_1+n_2-2$**	$V=n_1+n_2+2$	**$V=n_1+n_2-2$**
31	3	Analysis of variance utilizes	**F- test**	Chi-square test	Z-test	t-test	**F- test**
32	3	Analysis of variance technique originated in	**Agricultural research**	Industrial research	Biological research	Business research	**Agricultural research**
33	3	Analysis of variance technique was developed by	Gosset	**R.A.Fisher**	Laplace	Karl Pearson	**R.A.Fisher**
34	3	Design of experiment was introduced by	**R.A.Fisher**	Spearman	Karl Pearson	Lorenz	**R.A.Fisher**
35	3	The variation within samples is known as	error	**block**	degrees of freedom	local control	**block**
36	3	CRD's statistical analysis is same as	two way classification	randomized block design	**One way classification**	completely randomized design	**One way classification**
37	3	If in a RBD having 5 treatments and 4 replications, a treatment is added, the increase in error difference will be	1	2	**3**	4	**3**
38	3	In a RBD with 4 blocks and 5 treatments having one missing value, the error difference will be	12	**11**	10	9	**11**
39	3	The ratio of the number of replications required in CRD and RBD for the same amount of information is	6:04	**10:06**	10:08	6:10	**10:06**
40	3	Error sum of squares in RBD as compared to CRD using the same material is:	more	**less**	equal	not comparable	**less**
41	3	In the analysis of data of a RBD with b block and v treatments, the error d.f. are:	b(v-1)	v(b-1)	**(b-1) (v-1)**	A=0	**(b-1) (v-1)**
42	3	A randomized block design (RBD) has:	**two way classification**	one way classification	three way classification	no classification	**two way classification**
43	3	A CRD is also known as:	unsystematic design	**non-restrictional design**	single block design	restrictional design	**non-restrictional design**
44	3	CRD's are mostly used in	**field experiments**	experiments on animals	pot experiments	mathematical calculation	**field experiments**
45	3	In a CRD with 't' treatments and 'n' experimental units, error d.f. is equal to	**n-t**	n-t-1	n-t+1	t-n	**n-t**
46	3	The formula for obtaining a missing value in RBD by minimizing the error mean square was given by	W.G.Cochran	T.Wishart	**F.Yates**	J.W.Tukey	**F.Yates**
47	3	The idea of preliminary test of significance was propounded by	W.G.Cochran	**A.E.Paull**	F.Yates	J.W.Tukey	**A.E.Paull**
48	3	The range of F-variate is	**$-\infty$ to ∞**	0 to 1	0 to ∞	$-\infty$ to 0	**$-\infty$ to ∞**
49	3	A business firm's sales data classified on the basis of different salesman and sales in different region is example of	One way ANOVA	**Two way ANOVA**	T-test	Chi square test	**Two way ANOVA**
50	3	The ratio F is equal to	**S_1^2/S_2^2**	S_1/S_2	S_2/S_1	S_2^2/S_1^2	**S_1^2/S_2^2**
51	3	If the calculated value of F is greater than the table value, the difference in sample means is	**Significant**	Not significant	Less significant	Insignificant	**Significant**
52	3	Coding method should be used specially	When given figures are simple or convenient	**When given figures are big or inconvenient**	When given figures are simple or inconvenient	When given figures are big or convenient	**When given figures are big or inconvenient**
53	3	If the calculated value of F is lesser than the table value, the difference in sample mean is	Significant	**Not significant**	Less significant	Insignificant	**Not significant**
54	3	In one way ANOVA the degrees of freedom is equal to	**n-1**	1-n	v-n	X	**n-1**
55	3	We can determine one way ANOVA if there are	**Differences within one factor**	Differences between two factors	Similarities within one factor	Similarities between two factors	**Differences within one factor**
56	3	We can determine two way ANOVA if there are	Differences within one factor	**Differences between two factors**	Similarities within one factor	Similarities between two factors	**Differences between two factors**
57	3	In two way ANOVA, sum of squares due to error is	SSC-SSR	**SST-(SSC+SSR)**	SSC+(SST-SSR)	SST-SSR	**SST-(SSC+SSR)**
58	3	Mean square of residual variance is	SSC/c-1	**SSE/(r-1)(c-1)**	SSR/(r-1)(c-1)	SSR/(r-1)	**SSE/(r-1)(c-1)**
59	3	The value of F in one way classification is	**between column variable/ within column variable**	within column variable/ between column variable	within row variable/ between column variable	between row variable/ within row variable	**between column variable/ within column variable**
60	3	In randomized block design the units are called as	**experimental units**	blocks	plots	block units	**experimental units**

UNIT V

1	5	All items in any field of inquiry constitute a --------------------	**Population**	sample	total items	infinite samples	**Population**
2	5	Population census conducted once in a -----------------	century	**decade**	fortnight	two decade	**decade**
3	5	A definite plan for obtaining a sample from a given population is known as	Research design	**sample design**	clear plan	accurate plan	**sample design**
4	5	---------------- is determined before data are collected	Research design	**sample design**	clear plan	accurate plan	**sample design**
5	5	The following is an example of finite universe	Number of stars in the sky	listeners of a specific radio programme	Throwing of a dice	**the population of a city**	**the population of a city**
6	5	The following are infinite universe except	Number of stars in the sky	listeners of a specific radio programme	Throwing of a dice	**the population of a city**	**the population of a city**
7	5	One of the following is the example of infinite universe	the population of a city	the number of workers in a factory	the number of students in a college	**listeners of a specific radio programme**	**listeners of a specific radio programme**
8	5	A --------------- results from errors in the sampling procedures	**systematic bias**	sampling error	designing error	research bias	**systematic bias**
9	5	When the size of the sample increases sampling error ---------	**decreases**	increases	no change	zero	**decreases**
10	5	To study the economic status of a town or village the sampling design used is	probability sampling	**non-probability sampling**	based on each item	based on items at random	**non-probability sampling**
11	5	Probability sampling is also known as	quota sampling	judgement sampling	**random sampling**	purposive sampling	**random sampling**
12	5	One of the following is not a non-probability sampling	deliberate sampling	purposive sampling	judgement sampling	**chance sampling**	**chance sampling**
13	5	The probability of selecting sample of size 3 from a finite population of 6 elements is	**1/20**	3/6	2/6	1/6	**1/20**
14	5	Selecting every i^{th} item on a list is known as	**systematic sampling**	area sampling	multi-stage sampling	cluster sampling	**systematic sampling**
15	5	If a 4% sample is desired, the sample selection will be	one in every 10^{th} item	**one in every 25^{th} item**	one in every 50^{th} item	one in every 4^{th} item	**one in every 25^{th} item**
16	5	If a population from which a sample is to be drawn doesnot constitute a homogenous group the technique is known as	**Stratified sampling**	area sampling	multi-stage sampling	cluster sampling	**Stratified sampling**
17	5	In stratified sampling the different sub populations are called as	sub datas	**strata**	substitutes	random samples	**strata**
18	5	The most efficient and an optimal design of complex random sampling is	**Stratified sampling**	area sampling	multi-stage sampling	cluster sampling	**Stratified sampling**
19	5	Using proportional allocation, the sample sizes for different strata are 15,9 and 6 respectively which is in proportion to the sizes of the strat	3000:4000:1000	**4000:2400:1600**	2400:1600:4000	1800:2200:4000	**4000:2400:1600**
20	5	If clusters happens to be some geographic subdivisions it is known as	Stratified sampling	**area sampling**	multi-stage sampling	cluster sampling	**area sampling**
21	5	__________ is a process of mapping aspects of a domain onto other aspects of a range according to some rule of correspondence	strata	**measurement**	sub strata	clusters	**measurement**
22	5	The data which are numerical in name only and donot share any properties of the numbers we deal in ordinary arithmetic is	**nominal data**	ordinal data	interval data	ratio data	**nominal data**
23	5	In those situations,when we can't do anything except set up inequalities, we refer to the data as	nominal data	**ordinal data**	interval data	ratio data	**ordinal data**
24	5	On Mohs' scale number 5&2 refers to apatite and gypsum.In this 5 > 2 refers to	**apatite is harder than gypsum**	apatite is softer than gypsum	gypsum is softer than apatite	both are equally hard	**apatite is harder than gypsum**
25	5	On Mohs' scale the numbers 6&9 refers to feldspar and sapphire respectively in this 6 < 9 refers to	feldspar is the most hardest	sapphire and feldspar are equally hard	**feldspar is softer than sapphire**	sapphire is softer than feldspar	**feldspar is softer than sapphire**
26	5	is simply a system of assigning number symbols to events in order to label them	**nominal scale**	ordinal scale	interval scale	ratio scale	**nominal scale**
27	5	The assignment of number of basketball players in order to identify them is the usual example of	**nominal scale**	ordinal scale	interval scale	ratio scale	**nominal scale**
28	5	In nominal scale the measure of central tendency used is	mean	median	**mode**	quartiles	**mode**
29	5	Generally used measure of dispersion for nominal scales is	standard deviation	mean deviation	quartile deviation	**no measure of dispersion**	**no measure of dispersion**
30	5	The most common test of statistical significance that can be utilized in nominal scales is	F-test	t-test	Z-test	**chi-square test**	**chi-square test**
31	5	For the measure of correlation ----------- can be worked out for nominal scales	**contingency coefficient**	scatter diagram	Karl-pearson's coefficient of correlation	graphs	**contingency coefficient**
32	5	is the least powerful level of measurement	**nominal scale**	ordinal scale	interval scale	ratio scale	**nominal scale**
33	5	The lowest level of the ordered scale that is commonly used is the -----------------	nominal scale	**ordinal scale**	interval scale	ratio scale	**ordinal scale**
34	5	A students rank in his graduation class involves the use of an -------------	nominal scale	**ordinal scale**	interval scale	ratio scale	**ordinal scale**
35	5	Multiplication and division can only be used with this scale but not with other scales.	nominal scale	ordinal scale	interval scale	**ratio scale**	**ratio scale**
36	5	Measures of central tendency used in ratio scale is	median	mode	arithmetic mean	**geometric and harmonic mean**	**geometric and harmonic mean**
37	5	Researchers in physical sciences have the advantage to describe variables in	nominal scale	ordinal scale	interval scale	**ratio scale**	**ratio scale**
38	5	Researchers in behavioural sciences have the advantage to describe variables in	nominal scale	ordinal scale	**interval scale**	ratio scale	**interval scale**
39	5	The most precise type of scale is	nominal scale	ordinal scale	interval scale	**ratio scale**	**ratio scale**
40	5	refers to the extent to which a test measures what we actually wish to measure	**Validity**	Reliability	Practicality	speed	**Validity**
41	5	has to do with the accuracy and precision of a measurement procedure	validity	**reliability**	practicality	speed	**reliability**
42	5	is concerned with a wide range of factors of economy,convenience and interpretability	validity	reliability	**practicality**	speed	**practicality**
43	5	If the instrument contains a representative sample of the universe,the --------- is good	Criterion related validity	**content validity**	construct validity	obstruct validity	**content validity**
44	5	enables the researcher to study the perceptual structure of a set of stimuli and the cognitive processes underlying the developm	rational scaling	nominal scaling	**multi-dimentional scaling**	interval scale	**multi-dimentional scaling**
45	5	_________ describes the procedures of assigning numbers to various degrees of opinion,attitude and other concepts	**scaling**	sampling	validity	reliability	**scaling**
46	5	Categorical scales are also known as	nominal scale	ordinal scale	interval scale	**rating scale**	**rating scale**
47	5	Comparative scales are also known as	nominal scale	**ranking scale**	interval scale	ratio scale	**ranking scale**
48	5	Classification without indicating order, distance or unique origin is	**nominal scale**	ordinal scale	interval scale	ratio scale	**nominal scale**
49	5	________ indicates magnitude relationship of 'more than' or 'less than' but indicate no distance or unique origin	nominal scale	**ordinal scale**	interval scale	ratio scale	**ordinal scale**
50	5	Which scales have both order and distance values but no unique origin	nominal scale	ordinal scale	**interval scale**	ratio scale	**interval scale**
51	5	The most widely used scale construction technique developed on ad hoc basis is	**arbitrary approach**	item analysis approach	consensus approach	factor scales	**arbitrary approach**
52	5	Qualitative description of a limited number of aspects of a thing or of traits of a person is	nominal scale	ordinal scale	interval scale	**rating scale**	**rating scale**
53	5	The scaling technique in the form of 'always-often-occasionally-rarely-never' is	nominal scale	ordinal scale	interval scale	**rating scale**	**rating scale**
54	5	Greater sensitivity of measurement is achieved in rating scale if there is	**more points on a scale**	less points on a scale	only two points	zero points	**more points on a scale**
55	5	If the respondents are either easy raters or hard raters, the type of error occurs is	error of central tendency	**error of leniency**	error of hallo effect	error of reliability	**error of leniency**
56	5	When raters are reluctant to give extreme judgments, the result is the	**error of central tendency**	error of leniency	error of hallo effect	error of reliability	**error of central tendency**
57	5	Composite standard method of paired comparision is given by	**J.P.Guilford**	E.L.Thorndike	M.D.Thurston	R.A.Fisher	**J.P.Guilford**
58	5	It is easy to construct Likert-type scale in comparison to Thurstone type because	only written examination is done	**can be performed without a panel of judges**	only oral questions are asked	It is performed with a panel of judges	**can be performed without a panel of judges**
59	5	An attempt to measure the psychological meanings of an object to an individual is	**semantic differential scale**	ordinal scale	interval scale	ratio scale	**semantic differential scale**
60	5	In which section of dissertation ,we can give our suggestions?	Introduction	Result	**Discussion**	Su mmary	**Discussion**

SYLLABUS

UNIT IV

Research: Scope and significance – Types of Research – Research Process – Characteristics of good research – Problems in Research – Identifying research problems. Research Designs – Features of good designs.

OBJECTIVES OF RESEARCH

The purpose of research is to discover answers to questions through the application of scientific procedures. The main aim of research is to find out the truth which is hidden and which has no discovered as yet. Though each research study has its own specific purpose, we may think of research objectives as falling into a number of following broad groupings:

- To gain familiarity with a phenomenon or to achieve new insights into it (studies with this object in view are termed as *exploratory* or *formulative* research studies);
- To portray accurately the characteristics of a particular individual, situation or a (studies with this object in view are known as *descriptive* research studies);
- To determine the frequency with which something occurs or with which it as so with something else (studies with this object in view are known as *diagnostic* re: studies);
- To test a hypothesis of a causal relationship between variables (such studies are *hypothesis-testing* research studies).

MOTIVATION IN RESEARCH

What makes people to undertake research? This is a question of fundamental importance

possible motives for doing research may be either one or more of the following:

1. Desire to get a research degree along with its consequential benefits;

2. Desire to face the challenge in solving the unsolved problems, i.e., concern over practical problems initiates research;
3. Desire to get intellectual joy of doing some creative work;
4. Desire to be of service to society;
5. Desire to get respectability.

However, this is not an exhaustive list of factors motivating people to undertake research st

Many more factors such as directives of government, employment conditions, curiosity abot things, desire to understand causal relationships, social thinking and awakening, and the like n

well motivate (or at times compel) people to perform research operations.

TYPES OF RESEARCH

The basic types of research are as follows:

i. Descriptive vs. Analytical: Descriptive research includes surveys and fact-finding enquiries of different kinds. The major purpose of description of the state of affairs as it exists at present. In social science and business research we quite often use the term Ex post facto research for descriptive research studies. The main characteristic of this method is that the researcher has no control over the variables; he can only report what has happened or what is happening. Most ex post facto research projects are used for descriptive studies in which the researcher seeks to measure such items as, for example, frequency of shopping, preferences of people, or similar data. Ex post facto studies also include attempts by researchers to discover causes even when they cannot control the variables. The methods of research utilized in descriptive research are survey methods of all kinds, including comparative and correlation methods. In analytical research, on the other hand; the researcher has to use facts or information already available, and analyze these to make a critical evaluation of the material.

ii. Applied vs. Fundamental: Research can either be applied (or action) research or fundamental (to basic or pure) research. Applied research aims at finding a solution for an immediate problem facing a society or an industrial business organization, whereas fundamental research is mainly concerned with generalizations and with the formulation of a theory. "Gathering knowledge for knowledge's sake is termed 'pure' or 'basic' research." Research concerning some natural phenomenon or relating to pure mathematics are examples of fundamental research. Similarly, research studies, concerning human behavior carried on with a view to make generalizations about human behavior, are also examples of fundamental research, but research aimed at certain conclusions (say, a solution) facing a concrete social or business problem is an example of applied research. Research to identify social, economic or political trends that. May affect a particular institution or the copy research (research to find out whether certain cornmunications will be read and understood) or the marketing research or evaluation research are examples of applied research. Thus, the central aim of applied research is to discover a solution for some pressing practical problem, whereas basic research is directed towards finding information that has a broad base of applications and thus, adds to the already existing organized body of scientific knowledge.

iii. Quantitative vs. Qualitative: Quantitative research is based on the measurement of quantity or amount. It is applicable to phenomena that can be expressed in terms of quantity. Qualitative research, on the other hand, is concerned with qualitative phenomenon, i.e., phenomena relating to or involving quality or kind. For instance, when we are interested "in investigating the reasons for human behavior (i.e., why people think or do certain things), we quite often talk of 'Motivation Research', an important type of qualitative research. This type of research aims at discovering the underlying motives and desires, using in depth interviews for the purpose. Other techniques of such research are word association tests, sentence completion tests, story completion tests and similar other projective techniques. Attitude r opinion research i.e., research\h designed to find out how people feel or what they think about a particular subject or institution is also qualitative research. Qualitative research is specially important in the behavioral sciences where the aim is to discover the underlying motives of human behavior. Through such research we can analyze the various factors which motivate people to behave in a particular manner or which make people like or dislike a particular thing. It may be stated, however, that to apply qualitative research in Practice is relatively a difficult job and therefore, while doing such research, one should seek guidance from experimental psychologists.

iv. Conceptual vs. Empirical: Conceptual research is that related to some abstract idea(s) or
theory. It is generally used by philosophers and thinkers. to develop new concepts or to
reinterpret existing ones. On the other hand, empirical research relies on experience or
observation alone, often without due regard for system and theory. It is data-based research, coming up with conclusions which are capable of being verified by observation or experiment. We can also call it as experimental type of research. In such a research it is necessary to get at facts firsthand, at their source, and actively to go about doing certain things to stimulate the production of desired information. In such a research, the researcher must first provide himself with a working hypothesis or guess as to the probable results. He then works to get enough facts (data) to prove or disprove his hypothesis. He then sets up experimental designs which he thinks will manipulate the persons or the materials concerned so as to bring forth the desired information. Such research is thus

characterized by the experimenter's control over the variables under study and his deliberate manipulation of one of them to study its effects. Empirical research is appropriate when proof is sought that certain variables affect other variables in some way. Evidence gathered through experiments or empirical studies is today considered to be the most powerful support possible for a given hypothesis.

v. Some Other Types of Research: All other types of research are variations of one or more
of the above stated approaches, based on either the purpose of research, or the time required to accomplish research, on the environment in which research is done, or on the
basis of some other-similar factor. Form the point of view of time; we can think of research
either a one-time research O n autodial research. In the former case the research is' confine to a single time-period, whereas in the latter case the research is carried on over
several time-periods. Research can be field-setting research or Laboratory research or
simulation research, depending upon the environment in which it is to be carried out. Research can as well be understood as clinical or diagnostic research such research follow case-study methods or in depth approaches to reach the basic causal relations. Such studies usually go deep into the causes of things or events that interest us using very small samples and very deep probing data gathering devices. The research may be exploratory or it may be formalized. The objective of exploratory research is the development of hypotheses rather than their testing, whereas formalized research studies are those with substantial structure and with specific hypotheses to be tested. Historical research1 is that which utilizes historical sources like documents, remains, etc. to study events or ideas of the past; including the philosophy of persons and

groups at any remote point of time Research can also be classified a conclusion-oriented and decision-oriented. While doing conclusion oriented research, a researcher is free to pick up a problem, redesign the enquiry as he proceeds and is prepared to conceptualize as he wishes. Decision-oriented research is always for the need of a decision maker and the researcher in this case is not free to embark upon research according to his own inclination. Operations research is an example

of decision oriented research since it is a scientific method of providing executive departments. With a quantitative basis for decisions regarding operations under their control.

Importance of Knowing How Research is done

The study of research methodology gives the student the necessary training in gathering material and arranging or card-indexing them, participation in the fieldwork when required, and also training in techniques for the collection of data appropriate to particular problems, in the use of statistics, questionnaires and controlled experimentation and in recording evidence, sorting it out and interpreting it. In fact, importance of knowing the methodology or how research is done stems from the following considerations: -

i. For one who is preparing himself for a career of carrying out research, the importance of knowing research methodology and research techniques is obvious since the same constitute the tools of his trade. The knowledge of methodology provides good training specially to the new research worker and enables him to do better research. It helps him to develop disciplined thinking or a 'bent of mind' to observe the field objectively. Hence, those aspiring for careerism in research must develop the skill of using research techniques and must thoroughly understand the logic behind them.
ii. Knowledge of how to do research will inculcate the ability to evaluate and use research results with reasonable confidence. In other' words, we can state that the knowledge of research methodology is helpful in various fields such as government

or business administration, community development and social work where persons are increasingly called upon to evaluate and use research results for action.

The chart indicates that the research process consists of a number of closely related activities, as shown through I to VII. But such activities overlap continuously rather than following a strictly prescribed sequence. At times, the first step determines the nature of the last step to be undertaken. If subsequent procedures have not been taken into account in the early stages, serious difficulties may arise which may even prevent the completion of the study. One should remember that the various steps involved in a research process are not mutually exclusive; nor they are separate and distinct. They do not necessarily follow each other in any specific order and the researcher has to be constantly anticipating' at each step in the research process the requirements of the subsequent steps. However, the following order concerning various provides a useful procedural guideline regarding the research process: (1) formulating the research problem; (2) extensive literature survey; (3) developing the hypothesis; (4) preparing the research design; (5) determining sample design; (6) collecting the data; (7) execution of the project; (8) analysis of data; (9) hypothesis testing; (10) generalisations and interpretation, and (11) preparation of the report or presentation of the results, i.e., formal write-up of conclusions reached.

A brief description of the above stated steps will be helpful.

1. Formulating the research problem: There are two types of research problems, viz., those which relate to stales of nature and those which relate to relationships between variables. At the very outset the researcher must single out the problem he wants to study, i.e., he must decide the general area of interest or aspect of a subject-matter that he would like to inquire into. Initially the problem may be stated in a broad general way and then the ambiguities, if any, relating to the problem be resolved. Then, the feasibility of a particular solution has to be considered before a working formulation of the problem can be set up. The formulation of a general topic into a specific research problem, thus, constitutes the first step in a scientific enquiry. Essentially two steps are involved in formulating the research problem, viz., understanding the problem thoroughly, and rephrasing the same into meaningful terms from an analytical point of vie

The best way of understanding the problem is to discuss it with one's own colleagues or with those having some expertise in the matter. In an academic institution the researcher can seek the help from a guide who is usually an experienced man and has several research problems in mind. Often, the guide puts forth the problem in general terms and it is up to the researcher to narrow it down and phrase the problem in

operational terms. n private business, units or in governmental organisations, the problem is usually earmarked by the administrative agencies with whom the researcher can discuss as to how the problem originally came about and what considerations are involved in its possible solutions.

The researcher must at the same time examine all available literature to get himself acquainted with the selected problem. He may review two types of literature-the conceptual literature concerning the concepts and theories, and the empirical literature consisting of studies made earlier which are similar to the one propose 1'The basic outcome of this review will be the know ledge as to what data and other materials are available for operational purposes which will enable the researcher to specify his own research problem in a meaningful context. After this the researcher rephrases the problem into analytical or operational terms i.e., to put the problem in as specific terms as possible. This task of formulating, or defining, a research problem is a step of greatest importance in the entire research process. The problem to be investigated must be defined unambiguously for that will help discriminating relevant data from irrelevant ones. Care must, however, be taken to-verify the objectivity and validity of the background facts concerning the problem. Professor W.A. Neiswanger correctly states that the statement of the objective is of basic importance because it determines the data which are to be collected, the characteristics of the data which are relevant, relations which are to be explored, the choice of techniques to be used in these explorations and the form of the final report. If there are certain pertinent terms, the same should be clearly defined along with the task of formulating the problem. In fact, formulation of the problem often follows a sequential pattern where a number of formulations are set up, each formulation more specific than the preceeding one, each one phrased in more analytical terms, and each more realistic in terms of the available data and resources

.

2. Extensive literature survey: Once the problem is formulated, a brief summary of it should be written down. It is compulsory for a research worker writing a thesis for a Ph.D. degree to write a synopsis of the topic and submit it to the necessary Committee or the Research

Board for approval. At this juncture the researcher should undertake extensive literature survey connected with the problem. For this purpose, the abstracting and indexing journals and published or unpublished bibliographies are the first place to go to. Academic journals, conference proceedings, government reports, books etc., must be tapped depending on the nature of the problem. In this process, it should be remembered that one source will lead to another. The earlier studies, if any, which are similar to the study in hand, should be carefully studied. A good library will be a great help to the researcher at this stage

3. Development of working hypotheses: After extensive literature survey, researcher should
state in clear terms the working hypothesis or hypotheses. Working hypothesis is tentative assumption made in order to draw out and test its logical or empirical consequences. As such the manner in which research hypotheses are developed is particularly important since they provide the focal point for research. They also affect the manner in which tests must be conducted in the analysis of data and indirectly the quality of data which is required for the analysis In most types of research, the development of working hypothesis plays' an important role. Hypothesis should be very specific and limited to the piece of research in hand because it has to be tested. The role of the hypothesis is to guide the researcher by delimiting the area of research and to keep him on the right track. It sharpens his thinking and focuses attention on the more important facets of the problem. It also indicates the type of data required and the type of methods of data analysis to be used. How does one go about developing working hypotheses? The answer is by using the following approach:

- Discussions with colleagues and experts about the problem, its origin and the objectives in seeking a solution;
- Examination of data and records, if available, concerning the problem for possible trends, peculiarities and other clues;
- Review of similar studies in the area or of the studies on similar problems; and
- Exploratory personal investigation which involves original field interviews on a limited scale with interested parties and individuals with a view to secure greater insight into the practical aspects of the problem.

Thus, working hypotheses arise as a result of a-priori thinking about the subject, examination of the available data and material including related studies and the counsel of experts and interested parties. Working hypotheses are more useful when stated in precise and clearly defined terms. It may as well be remembered that occasionally we may encounter a problem where we do not need working hypotheses, specially in the case of exploratory or formulative researches which do not aim at testing the hypothesis. But as a general rule, specification of working hypotheses in another basic step of the research process in most research problems.

4. **Preparing the research design:** The research problem having been formulated in clear cut terms, the researcher will be required to prepare a research design, i.e., he will have to state the conceptual structure within which research would be conducted. The preparation of such a design facilitates research to be as efficient as possible yielding maximal information. In other words, the function of research design is to provide for the collection of relevant evidence with minimal expenditure of effort, time and money. But how all these can be achieved depends mainly on the research purpose. Research purposes may be grouped into four categories, viz., (i) Exploration, (ii) Description, (iii) Diagnosis, and (iv) Experimentation. A flexible research design which provides opportunity for considering many different aspects of a problem is considered appropriate if the purpose of the research study is that of exploration. But when the purpose happens to be an accurate description of a situation or of an association between variables, the suitable design will be one that minimises bias and maximises the reliability of the data collected and analysed.

There are several research designs, 'such as, experimental and non-experimental hypothesis testing. Experimental designs can be either informal designs (such as before-and-after without control after-only with control, before-and-after with control) or formal designs (such as completely randomized design, randomized block design, Latin square design, simple and complex factorial designs), out of which the researcher must select one for his own project.

The preparation of the research design, appropriate for a particular research problem, involves usually the consideration of the following:

- The means of obtaining the information;
- The availability and skills of the researcher and his staff (if any)

- Explanation of the way in which selected means of obtaining information will be organised and the reasoning leading to the selection;
- The time available for research; and
- The cost factor relating to research, i.e., the finance available for the purpose.

5. Determining sample design: All the items under consideration in any field of inquiry constitute a 'universe' or 'population' complete enumeration of all the items in the 'population' is known as a census inquiry. It can be presumed that in such an inquiry when all the items are' covered no element of chance is left and highest accuracy is obtained. But in practice this may not be true. Even the slightest element of bias in such an inquiry will get larger and larger as the number of observations increases. Moreover, there is no way of checking the element of bias or its extent except through a resurvey or use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy. Not only this, census inquiry is not possible in practice under many circumstances. For instance, blood testing is done only on sample basis. Hence, quite often we select only a few items from the universe for our study purposes. The items so selected constitute what is technically called a sampl0 '

The researcher must decide the way of selecting a sample or what is popularly known as the sample design other words, a sample design is a definite plan determined before any data are actually collected for obtaining a sample from a given population. Thus, the plan to select 12 of a city's 200 drugstores in a certain we constitutes a sample design. Samples can be either probability samples or non-probability samples. With probability sample's each element has a known probability of being included in the sample but e non-probability samples do not allow the researcher to determine this probability. Probability samples are those based on. simple random sampling, systematic sampling, stratified sampling, cluster/area sampling whereas non-probability samples are those based on convenience sampling, judgement sampling and quota sampling techniques. A brief mention of the important sample designs is as follows:

(i) ***Deliberate sampling:*** Deliberate sampling is also known as purposive or non-probability sampling. This sampling method involves purposive or deliberate selection of particular units of the universe for constituting a sample which represents the universe. When population elements are selected for inclusion in the sample based on the ease of access, it can be called *convenience sampling.* If a researcher wishes to secure data from, say, gasoline buyers, he may select a fixed number of petrol stations and may conduct interviews at these stations. This would be an example of convenience sample of gasoline buyers. At times such a procedure may give very biased results particularly when the population is not homogeneous. On the other hand, in *judgement sampling* the researcher's judgement is used for selecting items which he considers as representative of the population. For example, a judgement sample of college students might be taken to secure reactions to a new method of teaching. Judgement sampling is used quite frequently in qualitative research where the desire happens to be to develop hypotheses rather than to generalise to larger populations.

(ii) ***Simple random sampling:*** This type of sampling is also known as chance sampling or probability sampling where each and every item in the population has an equal chance of inclusion in the sample and each one of the possible samples, in case of finite universe, has the same probability of being selected. For example, if we have to select a sample of 300 items from a universe of 15,000 items, then we can put the names or numbers of all the 15,000 items on slips of paper and conduct a lottery. Using the random number tables is another method of random sampling. To select the sample, each item is assigned a number from 1 to 15,000. Then, 300 five digit random numbers are selected from the table. To dothis we select some random starting point and then a systematic pattern is used in proceeding through the table. We might start in the 4th row, second column and proceed down the column to the bottom of the table and then move to the top of the next column to the right. When a number exceeds the limit of the numbers in the frame, in our case over 15,000, it is simply passed over and the next number selected that does fall within the relevant range. Since the numbers were placed in the table in a completely random fashion, the resulting sample is random. This procedure gives each item an equal probability of being selected. Incase of infinite population, the selection of each item in a random sample is controlled by the same probability and that successive selections are independent of one another.

(iii) ***Systematic sampling:*** In some instances the most practical way of sampling is to select every 15th name on a list, every 10th house on one side of a street and so on. Sampling of this type is known as systematic sampling. An element of randomness is usually introduced into this kind of sampling by using random numbers to pick up the unit with which to start.This procedure is useful when sampling frame is available in tire form of a list. In such a design the selection process starts by picking some random point in the list and then every *n*th element is selected until the desired number is secured.

(iv)***Stratified sampling:*** If the population from which a sample is to be drawn does not constitute a homogeneous group, then stratified sampling technique is applied so as to obtain a representative sample. In this technique, the population is stratified into a number of non-overlapping subpopulations or strata and sample items are selected from each stratum. Ifthe items selected from each stratum is based on simple random sampling the entire procedure first stratification and then simple random sampling, is known as *stratified random sampling.*

(iv) ***Quota sampling:*** In stratified sampling the cost of taking random samples from individual strata is of tell so expensive that interviewers are simply given quota to be filled from different strata, the actual selection of items for sample being left to the interviewer's judgement. This is called quota sampling. The size of the quota for each stratum is generally proportionate to the size of that stratum in the population. Quota sampling is thus an important form of non-probability sampling. Quota samples generally happen to be judgement samples rather than random samples.

(vi) Cluster sampling and area sampling: Cluster sampling involves grouping the population and then selecting the groups or the clusters rather than individual elements for inclusion in the sample. Suppose some departmental store wishes to sample its credit card holders. It has issued its cards to 15,000 customers. The sample size is to be kept say 450. For cluster sampling this list of 15,000 card holders could be formed into 100 clusters of 150 card holders each. Three clusters might then be selected for the sample randomly. The sample size must often be larger than the simple random/sample to ensure the same level of accuracy because is cluster sampling procedural potential for order bias and other source of error is usually accentuated. The clustering approach can, however, make the sampling procedure relatively easier and increase the efficiency of field work, specially: in the case of personal interviews.

Area sampling is quite close to cluster sampling and is often talked about when the total geographical area of interest happens to be big one. Under area sampling we first divide the total area into a number of smaller non-overlapping areas, generally called geographical clusters, then a number of these smaller areas are randomly selected, and all units in these small areas are included in the sample. Area sampling is specially helpful where we do not have the list of the population concerned. It also makes the field interviewing more efficient since interviewer can do many interviews at each location.

(vii) Multi-stage sampling: This is a further development of the idea of cluster sampling. This technique is meant for big inquiries extending to a considerably large geographical area like an entire country. Under multi-stage sampling the first stage may be to select large primarysampling units such as states, then districts, then towns and finally certain families within towns. If the technique of random-sampling is applied at all stages, the sampling procedure is described as multi-stage random sampling.

(viii) Sequential sampling: This is somewhat a complex sample design where the ultimate size of the sample is not fixed in advance but is determined according to mathematical decisions on the basis of information yielded as survey progresses. This design is usually' adopted under acceptance sampling plan in the context of statistical quality control.

In practice, several of the methods of sampling described above may well be used in the same study in which case it can be called mixed sampling. It may be pointed out here that normally one should resort to random sampling so that bias can be eliminated and sampling error can be estimated. But purposive sampling is considered desirable when the universe happens to be small and a known characteristic of it is to be studied intensively. Also, there are conditions under which sample designs other than random sampling may be considered better for reasons like convenience and low costs. The sample design to be used must be decided by the researcher taking into consideration the nature of the inquiry and other related factors.

6. Collecting the data: In dealing with any real life problem it is often found that data at hand are inadequate, and hence, it becomes necessary to c6ilect data that are appropriate) There are several ways of collecting the appropriate data which differ considerably in context of money costs, time and other resources at the disposal of the researcher.

Primary data can be collected either through experiment or through survey. If the researcher conducts an experiment, he observes some quantitative measurements, or the data, with the help of which he examines the truth contained in his hypothesis. But in the case of a survey, data can be collected by anyone or more of the following ways:

(i) **By observation:** This method implies the collection of information by way of investigator's own observation, without interviewing the respondents. The information obtained relates to what is currently happening and is not complicated by either the past behaviour or future intentions or attitudes of respondents. This method is no doubt an expensive method and the information provided by this method is also very limited. As such this method is not suitable in inquiries where large samples are concerned.

(ii) **Through personal interview:** The investigator follows a rigid procedure and seeks answers to a set of pre-conceived questions through personal interviews. This method of collecting data is usually carried out in a structured way where output depends upon the ability of the interviewer to a large extent. .

(iii) **Through telephone interviews:** This method of collecting information involves contacting the respondents on telephone itself. This is not a very widely used method but it plays an important role in industrial surveys in developed regions, particularly, when the survey has to be accomplished in a very limited time.

(iv) **By mailing of questionnaires:** The researcher and the respondents do come in contact with each other if this method of survey is adopted. Questionnaires are mailed to the respondents with a request to return after completing the same. It is the most extensively used method in various economic and business surveys. Before applying this method, usually a Pilot Study for testing the questionnaire is conduced which reveals the weaknesses, if any, of the questionnaire. Questionnaire to be used must be prepared very carefully so that it may prove to be effective in collecting the relevant information.

(v) Through schedules: Under this method the enumerators are appointed and given training. They are provided with schedules containing relevant questions. These enumerators go to respondents with these schedules. Data are collected by filling up the schedules by enumerators on the basis of replies given by respondents. Much depends upon the capability of enumerators so far as this method is concerned. Some occasional field checks on the work of the enumerators may ensure sincere work.

The researcher should select one of these methods of collecting the data taking into consideration the nature of investigation, objective and scope of the inquiry, financial resources,available time and the desired degree of accuracy. Though he should pay attention to all these factors but much depends upon the ability and experience of the researcher. In this context *Dr A.L. Bowley* very aptly remarks that in collection of statistical data commonsense is the chief requisite and experience the chief teacher.

1. **Execution of the project:** Execution of the project is a very important step in the research process. If the execution of the project proceeds on correct lines, the data to be collected would be adequate and dependable. The researcher should see that the project is executed in a systematic manner and in time. If the survey is to be conducted by means of structured questionnaires, data can be readily machine-processed. In such a situation, questions as well as the possible answers may be coded. If the data are to be collected through interviewers, arrangements should be made for proper selection and training of the interviewers. The training may be given with the help of instruction manuals which explain clearly the job of the interviewers at each step. Occasional field checks should be made to ensure that the interviewers are doing their assigned job sincerely and efficiently. A careful watch should be kept for unanticipated factors in order to keep the survey as much realistic as possible. This, in other words, means that steps should be taken to ensure that the survey is under statistical control so that the collected information is in accordance with the pre-defined standard of accuracy. If some of the respondents do not cooperate, some suitable methods should be designed to tackle till's problem. One method of dealing with the non-response problem is to make a list of the non-respondents and take a small sub-sample of them, and then with the help of experts vigorous efforts can be made for securing response.

8. Analysis of data: After the data have been collected, the researcher turns to the task of analysing them. The analysis of data requires a number of closely related operations such as establishment of categories, the application of these categories to raw data through coding, tabulation and then drawing statistical inferences. The unwieldy data should necessarily be condensed into a few manageable groups and tables for further analysis. Thus, researcher should classify the raw data into some purposeful and usable categories. *Coding* operation is usually done at this stage through which the categories of data are transformed into symbols that may be tabulated and counted. *Editing* is the procedure that improves the quality of the data for coding. With coding the stage is ready for tabulation. *Tabulation* is a part of the technical procedure wherein the classified data are put in the form of tables. The mechanical devices can be made use of at this juncture. A great deal of data, specially in large inquiries, is tabulated by computers. Computers not only save time but also make it possible to study large number of variables affecting a problem simultaneously.

Analysis work after tabulation is generally based on the computation of various percentages, coefficients, etc., by applying various well defined statistical formulae. In the process of analysis, relationships or differences supporting or conflicting with original or new hypotheses should be subjected to tests of significance to determine with what validity data can be said to indicate any conclusion(s}. For instance, if there are two samples of weekly wages, each sample being drawn from factories in different parts of the same city, giving two different mean values, then our problem may be whether the two mean values are significantly different or the difference is just a matter of chance. Through the use of statistical tests we can establish whether such a difference is a real one or is the result of random fluctuations. If the difference happens to be real, the inference will be that the two samples

3. Good research is empirical: It implies that research is related basically to one or more aspects of a real situation and deals with concrete data that provides a basis for external validity to research results.

4. Good research is, replicable: This characteristic allows research results to be verified by replicating the study and thereby building a sound basis for decisions.

Problems Encountered by Researchers in India

Researchers in India, particularly those engaged in empirical research, are facing several problems.

Some of the important problems are as follows:

1. *The lack of a scientific training in the methodology of research* is a great impediment for researchers in our country. There is paucity of competent researchers. Many researchers take a leap in the dark without knowing research methods. Most of the work, which goes in the name of research, is not methodologically sound. Research to many researchers and even to their guides, is mostly a scissor and paste job without any insight shed on the collated materials. The consequence is obvious, viz., the research results, quite often, do not reflect the reality or realities. Thus, a systematic study of research methodology is an urgent necessity. Before undertaking research projects, researchers should be well equipped with all the, methodological aspects. As such, *efforts should be made to provide short- duration intensive courses for meeting this requirement.*
2. There is *insufficient interaction* between the university research departments on one side and business establishments, government departments and research institutions on the other side. A great deal of primary data of non-confidential nature remains untouched / untreated by the researchers for want of proper contacts. *Efforts should be made to develop satisfactory liaison among all concerned for better and realistic researches.* There is need for developing some mechanisms of a university-industry interaction programme so that academics can get ideas from practitioners on what needs to be researched and practitioners can apply the research done by the academics.
3. Most of the business units in our country do not have the confidence that the material supplied by them to researchers will not be misused and as such they are often reluctant in supplying the needed information to researchers, The concept of secrecy seems to be sacrosanct to business organisations in the country so much so that it proves an impermeable barrier to researchers. Thus, *there is the need for generating the confidence that the information/data obtained from a business unit will not be misused.*
4. *Research studies overlapping one another are undertaken quite often for want of adequate information.* This results in duplication and fritters away resources. This problem can be solved by proper compilation and revision, at regular intervals, of a list of subjects on which and the places where the research is going on. Due attention should be given toward identification of research problems in various disciplines of applied science which are of

immediate concern to the industries.

5. *There does not exist a code of conduct for researchers* and inter-university and inter- departmental rivalries are also quite common. Hence, there is need for developing a code of conduct for researchers which, if adhered sincerely, can win over this problem.
6. Many researchers in our country also face *the difficulty of adequate and timely secretarial assistance*, including computerial assistance. This causes unnecessary delays in the completion of research studies. All possible efforts be made in this direction so that efficient secretarial assistance is made available to researchers and that too well in time. University Grants Commission must play a dynamic role in solving this difficulty.
7. *Library management and functioning is not satisfactory at many places* and much of the time and energy of researchers are spent in tracing out the books, journals, reports, etc., rather than in tracing out relevant material from them. .
8. *There is also the problem that many of our libraries are not able to get copies of old and new Acts / Rules, reports and other government publications in time.* This problem is felt more in libraries which are away in places from Delhi and/or the state capitals. Thus efforts should be made for the regular and speedy supply of all governmental publications to reach our libraries.
9. *There is also the difficulty of timely availability of published data* from various government and other agencies doing this job in our country. Researcher also faces the problem on account of the fact that the published data vary quite significantly because of differences in coverage by the concerning agencies.
10. There may, at times, take place *the problem of conceptualization* and also problems relating to the process of data collection and related things.

DEFINING THE RESEARCH PROBLEM

In research process, the first and foremost step happens to be that of selecting and properly defining a research problem. A researcher must find the problem and formulate it so that it becomes susceptible to research. Like a medical doctor, a researcher must examine all the symptoms (presented to him or observed by him) concerning a problem before he can diagnose correctly. To define a problem correctly, a researcher must know: what a problem is?

WHAT IS A RESEARCH PROBLEM?

A research problem, in general, refers to some difficulty which a researcher experiences in the context of either a theoretical or practical situation and wants to obtain a solution for the same. Usually we say that a research problem does exist if the following conditions are met with:

a. There must be an individual (or a group or an organisation), let us call it *'I'*, to whom the problem can be attributed. The individual or the organisation, as the case may be, occupies an environment, say *'N'*, which is defined by values of the uncontrolled variables, Y_j.

b. There must be at least two courses of action, say C_1 and C_2, to be pursued. A course of action is defined by one or more values of the controlled variables. For example, the number of items purchased at a specified time is said to be one course of action.

c. There must be at least two possible outcomes, say O_1 and O_2 of the course of action, of which one should be preferable to the other. In other words, this means that there must be at least one outcome that the researcher wants, i.e., an objective.

d. The courses of action available must provides some chance of obtaining the objective, but they cannot provide the same chance, otherwise the choice would not matter. Thus, if p ($O_j \setminus I, C_1, N$) represents the probability that an outcome O_j will occur, if *I* select C in *N*, then P ($O_1 \setminus I, C_1, N$) ;# P ($O_1 \setminus I, 2, N$) . In simple words, we can say that the choices must have unequal efficiencies for the desired outcomes.

Over and above these conditions, the individual or the organisation can be said to have the problem only if 'I' does not know what course of action is best, i.e., '1', must be in doubt about the solution. Thus, an individual or a group of persons can be said to have a problem which can be technically described as a research problem, if they (individual or the group), having one or more desired outcomes, are confronted with two or more courses of action that have some but not equal efficiency for the desired objective(s) and are in doubt about which course of action is best. We can, thus, state the components [1] of a research problem as under:

i. There must be an individual or a group which has some difficulty or the problem.

ii. There must be some objective(s) to be attained at. If one wants nothing, one cannot have a problem.

iii. There must be alternative means (or the courses of action) for obtaining the objective(s) one wishes to attain. This means that there must be *at least two means* available to a researcher for if he has no choice of means, he cannot have a problem.

iv. There must remain some doubt in the mind of a researcher with regard to the selection of alternatives. This means that research must answer the question concerning the relative efficiency of the possible alternatives.

v. There must be some environment(s) to which the difficulty pertains.

Thus, a research problem is one which requires a researcher to find out the best solution for the given problem, i.e., to find out by which course of action the objective 'can be attained optimally in the context of a given environment. There are several factors which may result in making the problem complicated. For instance, the environment may change affecting the efficiencies of the courses of action or the values of the outcomes; the number of alternative courses of action may be very large; persons not involved in making the decision may be affected by it and react to it favourably or unfavourably, and similar other factors. All such elements (or at least the important ones) may be thought of in context of a research problem.

SELECTING THE PROBLEM

The research problem undertaken for study must be carefully selected. The task is a difficult one, although it may not appear to be so. Help may be taken from a research guide in this connection. Nevertheless, every researcher must find out his own salvation for research problems cannot be borrowed. A problem must spring from the researcher's mind like a plant springing from its own seed. If our eyes need glasses, it is not the optician alone who decides about the number of the lens we require. We have to see ourselves and enable him to prescribe for us the right number by cooperating with him. Thus, a research guide can at the most only help a researcher choose a subject. However, the following points may be observed by a researcher in selecting a research problem or a subject for research:

i. Subject which is overdone should not be normally chosen, for it will be a difficult task to throw any new light in such a case.

ii. Controversial subject should not become the choice of an average researcher

iii. Too narrow or too vague problems should be avoided.

iv. The subject selected for research should be familiar and feasible so that the related research material or sources of research are within one's reach. Even then it is quite difficult to supply definitive ideas concerning how a researcher should obtain ideas for his research. For this purpose, a researcher should contact an expert or a professor in the University who is already engaged in research. He may as well read articles published in current literature available on the subject and may think how the techniques and ideas discussed therein might be applied to the solution of other problems. He may discuss with others what he has in mind concerning a problem. In this way he should make all possible efforts in selecting a problem.

v. The importance of the subject, the qualifications and the training of a researcher, the costs involved, the time factor are' few other criteria that must also be considered in selecting a problem-In other words, before the final selection of a problem is done, a researcher must ask himself the following questions:
    a. Whether he is well equipped in terms of his background to carry out the research?
    b. Whether the study falls within the budget he can afford?
    c. Whether the necessary cooperation can be obtained from those who must

participate in research as subjects?

If the answers to all these questions are in the affirmative, one may become sure so far as the practicability of the study is concerned.

vi. The selection of a problem must be preceded by a preliminary study. This may not be necessary when the problem requires the conduct of a research closely similar to one that has already been done. But when the field of inquiry is relatively new and does not have available a set of well developed techniques, a brief feasibility study must always be undertaken.

If the subject for research is selected properly by observing the above mentioned points, the research will not be a boring drudgery; rather it will be love's labour. In fact, zest for work is a must. The subject or the problem selected must involve the researcher and must have an upper most place in his mind so that he may undertake all pains needed for the study

NECESSITY OF DEFINING THE PROBLEM

Quite often we all hear that a problem clearly stated is a problem half solved. This statement signifies the need for defining a research problem. The problem to be investigated must be defined unambiguously for that will help to discriminate relevant data from the irrelevant ones. A proper definition of research problem will enable the researcher to be on the track whereas an ill-defined problem may create hurdles. Questions like: What data are to be collected? What characteristics of data are relevant and need to be studied? What relations are to be explored. What techniques are to be used for the purpose? and similar other questions crop up in the mind of the researcher who can well plan his strategy and find answers to all such questions only when the research problem has been well defined. Thus, defining a research problem properly is a prerequisite for any study and is a step of the highest importance. In fact, formulation of a problem is often more essential than its solution. It is only on careful detailing the research problem that we can work out the research design and can smoothly carry on all the consequential steps involved while doing research.

TECHNIQUE INVOLVED IN DEFINING A PROBLEM

Let us start with the question: What does one mean when he/she wants to define a research problem? The answer may be that one wants to state the problem along with the bounds within which it is to be studied. In other words defining a problem involves the task of laying down boundaries within which a researcher shall study the problem with a pre-determined objective in view.

How to define a research problem is undoubtedly a herculean task. However, it is a task that must be tackled intelligently to avoid the perplexity encountered in a research operation. The usual approach is that the researcher should himself pose a question (or in case someone else wants the researcher to carry on research, the concerned individual, organisation or an authority should pose the question to the researcher) and set-up techniques and procedures for throwing light on the question concerned for formulating or defining the research problem. But such an approach generally does not produce definitive results because the question phrased in such a fashion is usually in broad general terms and as such may not be in a form suitable for testing.

Defining a research problem properly and clearly is a crucial part of a research study and must in no case be accomplished hurriedly. However, in practice this a frequently overlooked which causes a lot of problems later on. Hence, the research problem should be defined in a systematic manner, giving due weightage to all relating points. The technique for the purpose involves the undertaking of the following steps generally one after the other: (i) statement of the problem in a general way; (ii) understanding the nature of the problem; (iii) surveying the available literature (iv) developing the ideas through discussions; and (v) rephrasing the research problem into a working proposition.

A brief description of all these points will be helpful.

(i) **Statement of the problem in a general way:** First of all the problem should be stated in a broad general way, keeping in view either some practical concern or some scientific or intellectual this purpose, the researcher must immerse himself thoroughly in the subject matter concerning which he wishes to pose a problem. In case of social research, it is considered advisable to do some field observation and as such the researcher may undertake some sort of preliminary survey or what is often called *pilot survey. Then* the researcher can himself state the Problem or he can seek the guidance of the guide or the subject expert in accomplishing this task. Often the guide forth the problem in general terms, and it is then up to the researcher to narrow it down and the problem in operational terms. In case there is some directive from an organisational authority, the problem then can be stated accordingly. The problem stated in a broad general way may contain various ambiguities which must be resolved by cool thinking and rethinking over the problem. At the same time the feasibility of a particular solution has to be considered and the same should be kept in view while starting the problem.

(ii) Understanding the nature of the problem: The next step in defining the problem is to understand its origin and nature clearly. The best way of understanding the problem is to discuss it with those who first raised it in order to find out how the problem originally came about and with what objectives in view. If the researcher has stated the problem himself, he should consider once again all those points that induced him to make a general statement concerning the problem. For a better understanding of the nature of the problem involved, he can enter into discussion with those who have a good knowledge of the problem concerned or similar other problems. The researcher should also keep in view the environment within which the problem is to be studied and understood.

(iii) Surveying the available literature: All available literature concerning the problem at hand must necessarily be surveyed and examined before a definition of the research problem IS given. This means that the researcher must be well-conversant with relevant theories in the field, reports and records as also all other relevant literature. He must devote sufficient time in reviewing of research already undertaken on related problems. This is done to find out what data and other materials, if any, are available for operational purposes. "Knowing what data are available often serves to narrow the problem itself as well as the technique that might be used”. This would also help a researcher to know if there are certain gaps in the theories, or whether the existing theories applicable to the problem under study are inconsistent with each other, or whether the findings of the different studies do not follow a pattern consistent with the theoretical expectations and so on. All this will enable a researcher to take new strides in the field for furtherance of knowledge i.e., he can move up starting from the existing premise. Studies on related problems are useful for indicating the type of difficulties that may be encountered in the present study as also the possible analytical shortcomings. At times such studies may also suggest useful and even new lines of approach to the present problem.

(iv) Developing the ideas through discussions: (Discussion concerning a problem often produces useful information. Various new ideas can be developed through such an exercise. Hence, a researcher must discuss his problem with his colleagues and others who have enough experience in the same area or in working on similar problems. This is quite often known as an *experience survey*. *People* with rich experience are in a position to enlighten the researcher on different aspects of his proposed study and their advice and comments are usually invaluable to the researcher. They help him sharpen his focus of attention on specific aspects within the field. Discussions with such persons should not only be confined to the formulation of the specific problem at hand, but should also be concerned with the general approach to the given problem, techniques that might be used, possible solutions, etc.

(v) Rephrasing the research problem: Finally, the researcher must sit to rephrase the research
problem into a working proposition. Once the nature of the problem has been clearly understood, the environment (within which the problem has got to be studied) has been defined, discussions over the problem have taken place and the available literature has been surveyed and examined, rephrasing the problem into analytical or operational terms is not a difficult task. Through rephrasing, the researcher puts the research problem in as specific terms as possible so that it may become operationally viable and may help in the development of working hypotheses.

a. Technical terms and words or phrases, with special meanings used in the statement of the problem, should be clearly defined.
b. Basic assumptions or postulates (if any) relating to the research problem should be clearly stated.
c. A straight forward statement of the value of the investigation (i.e., the criteria for the selection of the problem) should be provided.
d. The suitability of the time-period and the sources of data available must also be considered by the researcher in defining the problem.
e. The scope of the investigation or the limits within which the problem is to be studied must be mentioned explicitly in defining a research problem

AN ILLUSTRATION

The technique of defining a problem outlined above can be illustrated for better understanding by taking an example as under:

Let us suppose that a research problem in a broad general way is as follows:

"Why is productivity in Japan so much higher than in India"?

In this form the question has a number of ambiguities such as: What sort of productivity is being referred to? With what industries the same is related? With what period of time the productivity is being talked about? In view of all such ambiguities the given statement or the question is much too general to be amenable to analysis. Rethinking and discussions about the problem may result in narrowing down the question to:
"What factors were responsible for the higher labour productivity of Japan's; manufacturing
industries during the decade 1971 to 1980 relative to India's manufacturing industries?" This latter version of the problem is definitely an improvement over its earlier version for the various ambiguities have been removed to the extent possible. Further rethinking and rephrasing might place the problem on a still better operational basis as shown below:
"To what extent did labour productivity in 1971 to 1980 in Japan exceed that of India in respect of 15 selected manufacturing industries? What factors were responsible for the productivity differentials between the two countries by industries?"

With this sort of formulation, the various terms involved such as 'labour productivity', 'productivity differentials', etc. must be explained clearly. The researcher must also see that the necessary data are available. In case the data for one or more industries selected are not available for the concerning time-period, then the said industry or industries will have to be substituted by other industry or industries. The suitability of the time-period must also be examined. Thus, all relevant factors must be considered by a researcher before finally defining a research problem.

CONCLUSION

We may conclude by saying that the task of defining a research problem, very often, follows a sequential pattern-the problem is stated in a general way, the ambiguities are resolved, thinking and rethinking process results in a more specific formulation of the problem so that it may be a realistic one In terms of the available data and resources and is also analytically meaningful. All this results in a well defined research problem that is not only meaningful from an operational point of view, but is equally capable of paving the way for the development of working hypotheses and for means of solving the problem itself.

RESEARCH DESIGN

MEANING OF RESEARCH DESIGN

The formidable problem that follows the task of defining the research problem is the preparation of the design of the research project, popularly known as the "research design". Decisions regarding what, where, when, how much, by what means concerning an inquiry or a research study constitute a research design. "A research design is the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure". In fact, the research design is the conceptual structure within which research is conducted; it constitutes the blueprint for the collection, measurement and analysis of data. As such the design includes an outline of what the researcher will do from writing the hypothesis and its operational implications to the final analysis of da90re explicitly, the desing decisions happen to be in respect of:

(i)What is the study about?
(ii)Why is the study being made?
(iii)Where will the study be carried out?
(iv)What type of data is required?
(v)Where can the required data be found?
(vi)What periods of time will the study include?
(vii) What will be the sample design?
(viii) What techniques of data collection will be used?
(ix) How will the data be analysed?
(x) In what style will the report be prepared?

Keeping in view the above stated design decisions, one may split the overall research design into the following parts:

a. *The sampling design* which deals with the method of selecting items to be observed for the given study.
b. *the observational design* which relates to the conditions under which the

observations are to be made;

c. *the statistical design* which concerns with the question of how many items are to be observed and how the information and data gathered are to be analysed; and
d. *the operational design* which deals with the techniques by which the procedures specified in the sampling, statistical and observational designs can be carried out.

-

From what has been stated above, we can state the important features of a research design as under:

(i) It is a plan that specifies the sources and types of information relevant to the research problem.

(ii) It is a strategy specifying which a roach will be used for gathering and analysing the data.

(iii) It also includes the time and cost bud ets since most studies me done under these constraints.

In brief, research design must, at least, contain-(a) a clear statement of the research problem; (b) procedures and techniques to be used for gathering information; (c) the population to be stud and (d) methods to be used in processing and analysing data.

NEED FOR RESEARCH DESIGN

Research design is needed because it facilitates the smooth sailing of the various research operations thereby making research as efficient as possible yielding maximal information with minimal expenditure of effort, time and money. Just as for better, economical and attractive construction of a house, we need a blueprint (or what is commonly called the map of the house) well thought out and prepared by expert architect. Similarly we need a research design or a plan in advance of data collection and analysis for our research project. Research design stands for advance planning of the methods to be adopted for collecting the relevant data and the techniques to be used in their analysis; keeping in view the objective of the research and the availability of staff, time and money. Preparation of research design should be done with great care as any error in it may upset the entire project. Research design, in fact, has a great bearing on the reliability of the results arrived at and as such constitutes the firm foundation of the entire edifice of the research work.

Even then the need for a well thought out research design is at times not realised by many. The importance which this problem deserves is not given to it. As a result many researches do not serve the purpose for which they are' undertaken. In fact, they may even give misleading conclusion. Thoughtlessness in designing the research project may result in rendering the research exercise futile. It is, therefore, imperative that an efficient and appropriate design must be prepared before starting research operations. The design helps the researcher to organize his ideas in a form whereby it will be possible for him to look for flaws and inadequacies. Such a design can even be given to others for their comments and critical evaluation. In the absence of such a course of action, it will be difficult for the critic to provide a comprehensive review of the proposed study.

FEATURES OF A GOOD DESIGN:

A good design is often characterized by adjectives like flexible, appropriate, efficient, econornical and so on. Generally, the design which minirnises bias and maxirnises the reliability of the data collected and analysed is considered a good design. The design which gives the smallest experimental error is supposed to be the best design in many investigations. Similarly, a design which yields maximal information and provides an opportunity for considering many different aspects of a problem is considered most appropriate and efficient design in respect of many research problems. Thus, the question of good design is related to the purpose or objective of the research problem and also with the nature of the problem to be studied. A design may be quite suitable in one case, but may be found wanting in one respect or the other in the context of some other research problem. One single design cannot serve the purpose of all types of research problems.

A research design appropriate for a particular research pr6blem, usually involves the consideration of the following factors:

(i) the means of obtaining information;
(ii) the availability and skills of the researcher and his staff, if any;
(iii) the objective of the problem to be studied;
(iv) the nature of the problem to be studied; and
(v) the availability of time and money for the research work

If the research study happens to be an exploratory or a formulative one, wherein the major emphasis is on discovery of ideas and insights, the research design most appropriate must be flexible enough to permit the consideration of many different aspects of a phenomenon. But when the purpose of a study is accurate description of a situation or of an association between variables (or in what are called the descriptive studies), accuracy becomes a major consideration and a research design which minimises bias and maxirnises the reliability of the evidence collected is considered a good design. Studies involving the testing of a hypothesis of a causal relationship between variables require a design which will permit inferences about causality in addition to the rninirnisation of bias and maximisation of reliability. But in practice it is the most difficult task to put a particular study in a particular group, for a given research may have in it elements of two or more of the functions of different studies. It is only on the basis of its primary function that a study can be categorised either an exploratory or descriptive or hypothesis-testing study and accordingly the choice of a research deign may be made in case of a particular study. Besides, the availability of time, money, skills of the h staff and the means of obtaining the information must be given due weightage while working the relevant details of the research design such as experimental design, survey design, sample and the like.

IMPORTANT CONCEPTS RELATING TO RESEARCH DESIGN

,~ .

Before describing the different research designs, it will be appropriate to explain the various concepts to designs so that these may be better and easily understood.

1. Dependent and independent variables: A concept which can take on different quantitative is called a variable. As such the concepts like weight, height, income are all examples of variabes, Qualitative phenomena (or the attributes) are also quantified on the basis of the presence or absence of the concerning attribute(s). Phenomena which can take on quantitatively different values even in decimal points are called 'continuous variables': But all variables are not continuous. If they can only be expressed in integer values, they are non-continuous variables or in statistical language 'discrete variables'." Age is an example of continuous variable, but the number of children is an example of non-continuous variable. If one variable depends upon or is a consequence of the other variable, it is termed as a dependent variable, and the variable that is antecedent to the dependent variable is termed as an independent variable. For instance, if we say that height depends upon age, then height is a dependent variable and age is an independent variable. Further, if in addition to being dependent upon age, height also depends upon the individual's sex, then height is a dependent variable and age and sex are independent variables. Similarly, readymade films and lectures are examples of independent variables, whereas behavioural changes, occurring as a result of the environmental manipulations, are examples of dependent variables.

2. Extraneous variable: Independent variables that are not related to the purpose of the study, but may affect the dependent variable are termed as extraneous variables. Suppose the researcher wants to test the hypothesis that there is a relationship between children's gains in social studies achievement and their self-concepts. In this case self-concept is an independent variable and social studies achievement is a dependent variable. Intelligence may as well affect the social studies achievement, but since it is not related to the purpose of the study undertaken by the researcher, it will be termed as an extraneous variable. Whatever effect is noticed on dependent variable as a result of extraneous variable(s) is technically described as an 'experimental error'. A study must always be so designed that *the effect upon the dependent variable is attributed entirely to the independent variable(s), and not to some extraneous variable or variables.*

3. Control: One important characteristic of a good research design is to minimise the influence or effect of extraneous variable(s). The technical term 'control' is used when we design the study minimising the effects of extraneous independent variables. In experimental researches, the term 'control' is used to refer to restrain experimental conditions.

4. Confounded relationship: When the dependent variable is not free from the influence of extraneous variable(s), the relationship between the dependent and independent variables is said to be confounded by an extraneous variable(s).

5. Research hypothesis: When a prediction or a hypothesised relationship is to be tested by scientific methods, it is termed as research hypothesis. The research hypothesis is a predictive statement that relates an independent variable to a dependent variable. Usually a research hypothesis must contain, at least, one independent and one dependent variable. Predictive statements which are not to' be objectively verified or the relationships that are assumed but not to be tested, are not termed research hypotheses.

6. Experimental and non-experimental hypothesis-testing research: When the purpo.se of research is to test a research hypothesis, it is termed as hypothesis-testing research. It can be of the experimental design or of the non-experimental design. Research in which the independent variable is manipulated is termed 'experimental hypothesis-testing research' and a research in which an independent variable is not manipulated is called 'non-experimental hypothesis-testing research'. For instance, suppose a researcher wants to study whether intelligence affects reading ability for a group of students and for this purpose he randomly selects 50 students and tests their intelligence and reading ability by calculating the coefficient of correlation between the two sets of scores. This is an example of non-experimental hypothesis-testing research because herein the independent variable, intelligence, is not manipulated. But now suppose that our researcher randomly selects 50 students from a group of students who are to take a course in statistics and then divides them into two groups by randomly assigning 25 to Group A, the usual studies programme, and 25 to Group B, the special studies programme. At the end of the course, he administers a test to each group in order to judge the effectiveness of the training programme on the student's performance-level. This is an example of experimental hypothesis-testing research because in this case the independent variable, viz., the type of training programme, is manipulated.

7. Experimental and control groups: In an experimental hypothesis-testing research when a group is exposed to usual conditions, it is termed a 'control group', but when the group is exposed to some novel or special condition, it is termed an 'experimental group'. In the above illustration, the Group A can be called a control group and the Group B an experimental group. If both groups A and B are exposed to special studies programmes, then both groups would be termed 'experimental groups.' It is possible to design studies which include only experimental groups or studies which include both experimental and control groups.

8. Treatments: The different conditions under which experimental and control groups are put are usually referred to as 'treatments'. In the illustration taken above, the two treatments are the usual studies programme and the special studies programme. Similarly, if we want to determine through an experiment the comparative impact of three varieties of fertilizers on the yield of wheat, in that case the three varieties of fertilizers will be treated as three treatments.

9. Experiment: The process of examining the truth of a statistical hypothesis, relating to some research problem, is known as an experiment. For example, we can conduct an experiment to examine the usefulness of a certain newly developed drug. Experiments can be of two types, viz., absolute experiment and comparative experiment. If we want to determine the impact of a fertilizer on the yield of a crop, it is a case of absolute experiment; but if we want to determine the impact of one fertilizer as compared to the impact of some other fertilizer, our experiment then will be termed as a comparative experiment. Often, we undertake comparative experiments when we talk of designs of experiments.

10. Experimental unites): The pre-determined plots or the blocks, where different treatments are used, are known as experimental units. Such experimental units must be selected (defined) very carefully.

DIFFERENT RESEARCH DESIGNS

Different research designs can be conveniently described if we categorize them as: (1) research in case of exploratory research studies; (2) research design in case of descriptive and diagnostic research studies, and (3) research design in case of hypothesis-testing research studies.

We take up each category separately.

1. Research design in case of exploratory research studies: Exploratory research studies are also termed as formulative research studies. The main purpose of such studies is that of formulating a problem for more precise investigation or of developing the working hypotheses from an operational point of view. The major emphasis in such studies is on the discovery of ideas and insights. As such the research design appropriate for such studies must be flexible enough to provide opportunity for considering different aspects of a problem under study. Inbuilt flexibility in research design is needed because the research problem, broadly defined initially, is transformed into one with more precise meaning in exploratory studies, which fact may necessitate changes in the research procedure for gathering relevant data. Generally, the following three methods in the context of research design for such studies are talked about: (a) the survey of concerning literature; (b) the experience survey(c) the analysis of 'insight-stimulating' examples.

The survey of concerning literature happens to be the most simple and fruitful method of formulating precisely the research problem or developing hypothesis. Hypotheses stated by earlier workers may be reviewed and their usefulness be evaluated as a basis for further research. It may also .be considered whether the already stated hypotheses suggest new hypothesis. In this way the researcher should review and build upon the work already done by others, but in cases where hypotheses have not yet been formulated, his task is to review the available material for deriving the relevant hypotheses from it.

Besides, the bibliographical survey of studies, already made in one's area of interest may as well as made by the researcher for precisely formulating the problem. He should also make an attempt to apply concepts and theories developed in different research contexts to the area in which he is himself working. Sometimes the works of creative writers also provide a fertile ground for hypothesis - formulation and as such may be looked into by the researcher.

Experience survey means the survey of people who have had practical experience with the problem to be studied. The object of such a survey is to obtain insight into the relationships between variables and new ideas relating to the research problem. For such a survey people who are competent and can contribute new ideas may be carefully selected as respondents to ensure a representation of different types of experience. The respondents so selected may then be interviewed by the investigator. The researcher must prepare an interview schedule for the systematic questioning of informants. But the interview must ensure flexibility in the sense that the respondents should be allowed to raise issues and questions which the investigator has not previously considered. Generally, the experience - collecting interview is likely to be long and may last for few hours. Hence, it is often considered desirable to send a copy of the questions to be discussed to the respondents well in advance. This will also give an opportunity to the respondents for doing some advance thinking over the various issues involved so that, at the time of interview, they may be able to contribute effectively. Thus, an experience survey may enable the researcher to define the problem more concisely and help in the formulation of the research hypothesis. This survey may as well provide information about the practical possibilities for doing different types of research.

Analysis of 'insight-stimulating' examples is also a fruitful method for suggesting hypotheses for research. It is particularly suitable in areas where there is little experience to serve as a guide. This method consists of the Intensive study of selected instances of the phenomenon in which one is interested. For this purpose the existing records, if any, may be examined, the unstructured interviewing may take place, or some other approach may be adopted. Attitude of the investigator, the intensity of the study and the ability of the researcher to draw together diverse information into a unified interpretation are the main features which make this method an appropriate procedure for evoking insights.

Now, what sorts of examples are to be selected and studied? There is no clear cut answer to it. Experience indicates that for particular problems certain types of instances are more appropriate than others. One can mention few examples of 'insight-stimulating' cases such as the reactions of strangers, the reactions of marginal individuals, the study of individuals who are in transition from one stage to another, the reactions of individuals from different social strata and the like. In general, cases that provide sharp contrasts or have striking features are considered relatively more useful while adopting this method of hypotheses formulation.

Thus, in an exploratory of formulative research study which merely leads to insights or hypotheses, whatever method or research design outlined above is adopted, the only thing essential is that it must continue to remain flexible so that many different facets of a problem may be considered as and when they arise and come to the notice of the researcher.

2. Research design in case of descriptive arid diagnostic research studies: Descriptive research studies are those studies which are concerned with describing the characteristics of a particular individual, or of a group, whereas diagnostic research studies determine the frequency with which something occurs or its association with something else. The studies concerning whether certain variables are associated are examples of diagnostic research studies. As against this, studies concerned with specific predictions, with narration of facts and characteristics concerning individual, group or situation are all examples of descriptive research studies. Most of the social research comes under this category. From the point of view of the research design, the descriptive as well as diagnostic studies share common requirements and as such we may group together these two types of research studies. In descriptive as well as in diagnostic studies, the researcher must be able to define clearly, what he wants to measure and must find adequate methods. for measuring it along with a clear cut definition of 'population' he wants to study. Since the aim is to obtain complete and accurate information in the said studies, the procedure to be used must be carefully planned. The research design must make enough provision for protection against bias and must maximise reliability, with due concern for the economical completion of the research study. The design in such studies must be rigid and not flexible and must focus attention on the following:

a) Formulating the objective of the study (what the study is about and why is it being made?)
b) Designing the methods of data collection (what techniques of gathering data will be adopted?)
c) Selecting the sample (how much material will be needed?)

d) Collecting the data (where can the-required data be found and with what time period should the data be related?)
e) Processing and analysing the data.
f) Reporting the findings.

In a descriptive/diagnostic study the first step is to specify the objectives with sufficient precision to ensure that the data collected are relevant. If this is not done carefully, the study may not provide the desired information.

Then comes the question of selecting the methods by which the data are to be obtained. In other words, techniques for collecting the information must be devised. Several methods (viz., observation, questionnaires, interviewing, examination of records, etc.), with their merits and limitations, are available for the purpose and the researcher may user one or more of these methods which have been discussed in detail in later chapters. While designing data-collection procedure, adequate safeguards against bias and unreliability must be ensured. Whichever method is selected, questions must be well examined and be made unambiguous; interviewers must be instructed not to express their own opinion; observersmust be trained so that they uniformly record a given item of behaviour. It is always desirable to pre-test the data collection instruments before they are finally used for the study purposes. In other words, we can say that *"structured instruments"* are used in such studies.

In most of the descriptive/diagnostic studies the researcher takes out sample(s) and then wishes to make statements about the population on the basis of the sample analysis or analyses. More often than not, sample has to be designed. Different sample designs have been discussed in detail in a separate chapter in this book. Here we may only mention that the problem of designing samples should be tackled in such a fashion that the samples may yield accurate information with a minimum amount of research effort. Usually one or more forms of probability sampling, or what is often described as random sampling, are used.

To obtain data free from errors introduced by those responsible for collecting them, it is necessary to supervise closely the staff of field workers as they collect and record information. Checks may be set up to ensure that the data collecting staff perform their duty honestly and without prejudice. "As data are collected, they should be examined for completeness, comprehensibility, consistency and reliability.'?

The data collected must be processed and analysed. This includes steps like coding the interview replies, observations, etc.; tabulating the data; and performing several statistical computations. To the extent possible, the processing and analysing procedure should be planned in detail before actual work is started. This will prove economical in the sense that the researcher may avoid unnecessary labour such as preparing tables for which he later finds he has no use or on the other hand, re-doing some tables because he failed to include relevant data. Coding should be done carefully to avoid error in coding and for this purpose the reliability of coders needs to be checked. Similarly, the accuracy of tabulation may be checked by having a sample of the tables re-done. In case of mechanical tabulation the material (i.e., the collected data or information) must be entered on appropriate cards which is usually done by punching holes corresponding to a given code. The accuracy of punching is to be checked and ensured. Finally, statistical computations are needed and as such averages, percentages and various coefficients must be worked out. Probability and sampling analysis may as well be used. The appropriate statistical operations, along with the use of appropriate tests of significance should be carried out to safeguard the drawing of conclusions concerning the study.

Last of all comes the question of reporting the findings. This is the task of communicating the findings to others and the researcher must do it in an efficient manner. The layout of the report needs to be well planned so that all things relating to the research study may be well presented in simple and effective style.

Thus, the research design in case of descriptive/diagnostic studies is a comparative design throwing light on all points narrated above and must be prepared keeping in view the objective(s) of the study and the resources available. However, it must ensure the minimisation of bias and maximisation of reliability of the evidence collected. The said design can be appropriately referred to as a survey *design* since it takes into account all the steps involved in a survey concerning a phenomenon to studied.

The difference between research designs in respect of the above two types of research studies can be conveniently summarized in tabular form as under:

Table 3.1

Table 3.1

Research Design	Type of study: Exploratory of Formulative	Type of study: Descriptive/Diagnostic
Overall design	Flexible design (design must provide opportunity for considering different aspects of the problem)	Rigid design (design must make enough provision for protection against bias and must maximise reliability)
(i) Sampling design	Non-probability sampling design (purposive or judgement sampling)	Probability sampling design (random sampling)
(ii) Statistical design	No pre-planned design for analysis	Pre-planned design for analysis
(iii) Observational design	Unstructured instruments for collection of data	Structured or well thought out instruments for collection of data
(iv) Operational design	No fixed decisions about the operational procedures	Advanced decisions about operational procedures.

3. Research design in case of hypothesis-testing research studies: Hypothesis-testing research
studies (generally known as experimental studies) are those where the researcher tests the hypotheses of causal relationships between variables. Such studies require procedures that will not only reduce bias and increase reliability, but will permit drawing inferences about causality. Usually experiments meet this requirement. Hence, when we talk of research design in such studies, we often mean the design of experiments.

Professor R.A. Fisher's name is associated with experimental designs. Beginning of such designs was made by him when he was working at Rothamsted Experimental Station

(Centre for Agricultural Research in England). As such the study of experimental designs has its origin in agricultural research. Professor Fisher found that by dividing agricultural fields or plots into different blocks and then by conducting experiments in each of these blocks, whatever information is collected and inferences drawn from them, happens to be more reliable. This fact inspired him to develop certain experimental s for testing hypotheses concerning scientific investigations. Today, the experimental designs are being used in researches relating to phenomena of several disciplines. Since experimental designs originated in the context of agricultural operations, we still use, though in a technical sense, several terms of agriculture (such as treatment, yield, plot, block etc.) in experimental designs.

POSSIBLE QUESTIONS
UNIT IV

PART A (20 x 1 = 20 Marks)
Question number 1 – 20 online examinations

PART B (5 x 2= 20Marks)

1. Explain the steps necessary to carryout research effectively.
2. How do you define a research problem?
3. What are the problems encountered by the researchers in India?
4. What do you mean by research? Explain its significance in modern times.
5. Distinguish between Research methods and Research methodology.
6. Describe the different types of research, clearly pointing out the difference between an experiment and a survey.

PART C (5 X 6 = 30 Marks)

7. Explain in detail the steps involved in research process

8. What are the scope and significance of research? Explain the research process.

9. How will you identify a research problem? Explain the steps involved in execution of research successfully.

10. Explain in detail the research process and the characteristics of good research.

11. Discuss research design and what are the features of good design?

12. Describe fully the techniques of defining a research problem.

13. What is research design? Discuss the basis of stratification to be employed in sampling public opinion on inflation.

14. Give your understanding of a good research design. Is single research design suitable in all research studies? If not, why?

15. Describe some of the important research designs used in experimental hypothesis – testing research study.

16. Write a short note on 'Experience Survey' explaining fully its utility in exploratory research studies.

17. What is research problem? Define the main issues which should receive the attention of the researcher in formulating the research problem.

		UNIT IV					
1	4	Research in common parlance refers to	**A search for knowledge**	a search for attitude	a search for aptitude	a search for anything	**A search for knowledge**
2	4	A scientific and systematic search for pertinent information on a specific topic is known as	knowledge	**research**	information	attitude	**research**
3	4	The meaning of research as " a careful investigation specially through search for new facts in any branch of knowledge is given by	Redman	Mory	Spearman	**The advanced Learner's dictionary of current English**	**The advanced Learner's dictionary of current English**
4	4	Research as a " systematized effort to gain new knowledge" is given by	**Redman and Mory**	Spearman	Karl Pearson	The advanced Learner's dictionary of current English	**Redman and Mory**
5	4	----------------- is the persuit of truth with the help of study	knowledge	**research**	information	attitude	**research**
6	4	In social science and business research we quite often use the term ----------- for descriptive research studies	***Ex post facto research***	*in post facto research*	*Ex pre facto research*	*in pre facto research*	***Ex post facto research***
7	4	The main characteristic of *Ex post facto research* method is that ------------ has no control over the variables	The sample	**the researcher**	the engineer	the citizen	**the researcher**
8	4	Studying of frequency of shopping, preferences of people, or similar data are studied by -----------	Analytical research	applied research	Fundamental research	**descriptive research**	**descriptive research**
9	4	The use of facts already available and analyse these to make a critical evaluation of the material is coming under					
10	4	Finding a solution for an immediate problem facing a society or an industrial/business organization is	Analytical research	**applied research**	Fundamental research	descriptive research	**applied research**
11	4	Which of the following is not a quantitative approach	inferential approach	experimental approach	simulation approach	**qualitative approach**	**qualitative approach**
12	4	The purpose of -------- to research is to form a data base from which to infer the relationship of population	**inferential approach**	experimental approach	simulation approach	qualitative approach	**inferential approach**
13	4	Inferential approach is otherwise called as	**survey research**	deep research	qualitative research	vague research	**survey research**
14	4	Research approach in which artificial environment is created and data generated isknown as	inferential approach	experimental approach	**simulation approach**	qualitative approach	**simulation approach**
15	4	Building models for understanding future conditions is	inferential approach	experimental approach	**simulation approach**	qualitative approach	**simulation approach**
16	4	Research that is concerned with subjective assessment of attitudes,opinions and behaviour is known as	inferential approach	experimental approach	simulation approach	**qualitative approach**	**qualitative approach**
17	4	Increased amount of research make progress ------------	**possible**	impossible	easier	difficult	**possible**
18	4	___________ inculcates scientific and inductive thinking and it promotes the development of logical habits of thinking	knowledge	**research**	attitude	aptitude	**research**
19	4	Government's budgets should be based on	**people's need and availability of resources**	poverty status	business status of the country	economic status	**people's need and availability of resources**
20	4	__________ provides the basis for nearly all Government policies in our economic system	**Research**	information	knowledge	money	**Research**
21	4	Investigating the structure and development of a market for the purpose of formulating efficient policies for purchasing pro	**market research**	business research	operational research	motivational research	**market research**
22	4	Application of mathematical ,logical and analytical techniques to the solution of business problems of cost minimization and	market research	business research	**operational research**	motivational research	**operational research**
23	4	------------ has become an integral tool of business policy these days	business analysis	**market analysis**	operational analysis	motivational analysis	**market analysis**
24	4	Determining why people behave as they do is mainly concerned with market characteristics is	market research	business research	operational research	**motivational research**	**motivational research**
25	4	To those students who are to write a master's or PhD thesis research may means a	source of livelihood	**careerism**	the outlet of new ideas and insights	the generalization of new theories	**careerism**
26	4	To professionals in research methodology ,research may mean a	**source of livelihood**	careerism	the outlet of new ideas and insights	the generalization of new theories	**source of livelihood**
27	4	To philosophers and thinkers,research may means	source of livelihood	careerism	**the outlet of new ideas and insights**	the generalization of new theories	**the outlet of new ideas and insights**
28	4	To analysts and intellectuals,research may mean	source of livelihood	careerism	the outlet of new ideas and insights	**the generalization of new theories**	**the generalization of new theories**
29	4	To literary men and women research may mean	source of livelihood	**the development of new styles and creative work**	the outlet of new ideas and insights	the generalization of new theories	**the development of new styles and creative work**
30	4	______________ is the fountain of knowledge for the sake of knowledge and an important source of providing guidelines	attitude	**research**	aptitude	knowledge	**research**
31	4	Research is a sort of formal ---------- which enables one to understand the new developments in one's field in a better way	**training**	placement	project	case study	**training**
32	4	A step of greatest importance in the entire research process is	development of working hypothesis	extensive literature survey	preparing the research design	**formulating the research problem**	**formulating the research problem**
33	4	The plan to select 12 of a city's 200 drug stores in a certain way constitutes a	**sample design**	executing the research	analytical design	operational research	**sample design**
34	4	Sampling of every 15^{th} name on a list or every 10^{th} house on one side of a street and so on is known as	random sampling	deliberate sampling	**systematic sampling**	stratified sampling	**systematic sampling**
35	4	The most extensively used method of collecting the datas in various economic and business surveys is	by observation	**by mailing of questionnaire**	through schedules	through personal interview	**by mailing of questionnaire**
36	4	In which method of collecting data the enumerators are appointed and given training	by observation	by mailing of questionnaire	**through schedules**	through personal interview	**through schedules**
37	4	At the end of the work there should be .	Summary	Result and discussion	**Bibliography**	Objective.	**Bibliography**
38	4	__________ is normally an early section in the dissertation.	Summary	Result and discussion	**Literature Review**	Objective	**Literature Review**
39	4	Preparatory stage for the literature review is ________.	Collecting data	General survey of related research.	Writing title	Writing contents	General survey of related research.
40	4	Age is an example of	**continuous variable**	non continuous variable	dependent variable	independent variable	**continuous variable**
41	4	The number of children in a family is an example of	continuous variable	**non continuous variable**	dependent variable	independent variable	**non continuous variable**
42	4	Readymade films and lectures are examples of	continuous variable	non continuous variable	dependent variable	**independent variable**	**independent variable**
43	4	Behavioural changes, occurring as a result of the environmental manipulations are	continuous variable	non continuous variable	**dependent variable**	independent variable	**dependent variable**
44	4	Independent variables that are not related to the purpose of the study,but may affect the dependent variable are termed as	continuous variable	non continuous variable	**extraneous variable**	independent variable	**extraneous variable**
45	4	Most of the social research comes under the category of	simple research	**descriptive research**	diagnostic research	exploratory research	**descriptive research**
46	4	One of the following is not an informal experimental design	before and after without control design	after only with control design	before and after with control design	**factorial design**	**factorial design**
47	4	The following are formal experimental designs except	latin square design	CRD	RBD	**after only with control design**	**after only with control design**
48	4	One way ANOVA is used to analyse	latin square design	**CRD**	RBD	after only with control design	**CRD**
49	4	The experimental design frequently used in agricultural research is	**latin square design**	CRD	RBD	after only with control design	**latin square design**
50	4	The latin square design is very similar to	one way ANOVA	**two way ANOVA**	correlation	regression	**two way ANOVA**
51	4	Factorial designs are mainly used in	agricultural phenomena	**economic and social phenomena**	biological research phenomena	physical research phenomena	**economic and social phenomena**
52	4	The statistical accuracy of the experiments is increased by	**principle of replication**	principle of randomization	principle of local control	principle of reliability	**principle of replication**
53	4	Rothamsted experimental station is	**centre for agricultural research in England**	centre for agricultural research in India	centre for agricultural research in America	centre for agricultural research in Canada	**centre for agricultural research in England**
54	4	The following are the basic principles of experimental designs except	principle of replication	principle of randomization	principle of local control	**principle of reliability**	**principle of reliability**
55	4	Studies which determine the frequency with which something occurs or its association with something else is called	descriptive research	**diagnostic research**	exploratory research	hypothesis testing research	**diagnostic research**
56	4	In a hypothesis testing research, if the independent variable intelligence is not manipulated it is called	**non experimental hypothesis testing research**	experimental hypothesis testing research	exploratory research	diagnostic research	**non experimental hypothesis testing research**
57	4	Apredictive statement that relates an independent variable to a dependent variable is called	**research hypothesis**	research plan	research summary	research design	**research hypothesis**
58	4	Summary must be	**Easy to read**	Difficult to read	Easy to understand	Difficult to understand	**Easy to read**
59	4	________ makes a summary attractive to readers.	Figures	Fonts	**Title**	Format	**Title**
60	4	Summary should be written in ________	**Easy terminology**	figures	tables	structures	**Easy terminology**

UNIT-V
SYLLABUS

Sampling Design : Meaning – Concepts – Steps in sampling – Criteria for good sample design. Scaling measurements – Techniques – Types of scale.

MEANING AND CONCEPTS OF SAMPLING DESIGN

In this context one must remember that two costs are involved in a sampling analysis viz., the cost of collecting the data and the cost of an incorrect inference resulting from the data. Researcher must keep in view the two causes -of incorrect inferences viz., systematic bias and sampling error. A *systematic bias* results from errors in the sampling procedures and it cannot be reduced or eliminated by increasing the sample size. At best the causes responsible for these errors can be detected and corrected. Usually a systematic bias is the result of one or more of the following factors:

1. **Inappropriate sampling frame:** If the sampling frame is inappropriate i.e., a biased representation of the universe, it will result in a systematic bias.
2. **Defective measuring device:** If the measuring device is constantly in error, it will result in systematic bias. In survey work, systematic bias can result if the questionnaire or the interviewer is biased. Similarly, if the physical measuring device is defective there will be systematic bias in the data collected through such a measuring device.
3. **Non-respondents:** If we are unable to sample all the individuals initially included in the sample, may arise a systematic bias. The reason is that in such a situation the likelihood of establishing t or receiving a response from an individual is often correlated with the measure of what is to be estimated.
4. **Indeterminancy principle:** Sometimes we find that individuals act differently when kept under observation than what they do when kept in non-observed situations. For instance, if workers are that somebody is observing them in course of a work study ow4'he basis of which the average of time to complete a task will be determined and accordingly the quota will be set for piece they generally tend to work slowly in comparison to the speed with which they work if kept ed. Thus, the indeterminancy principle may also be a cause of a systematic bias.
5. **Natural bias in the reporting of data:** Natural bias of respondents in the reporting of data is often the cause of a systematic bias in many inquiries! There is usually a 'downward bias in the income data collected by government taxation department, whereas we find an upward bias in the income data collected by some social organisation. People in general understate their incomes if asked about it for tax purposes, but they overstate the same if asked for social status or their affluence. Generally in psychological surveys, people tend to give what they think is the' correct' answer rather than revealing their true feelings.

Sampling errors are the random variations in the sample estimates around the true population parameters. Since they occur randomly and are equally likely to be in either direction, their nature happens to be of compensatory type and the expected value of such errors happens to be equal to zero. Sampling error decreases with the increase in the size of the sample, and it happens to be of a smaller magnitude in case of homogeneous population.

Sampling error can be measured for a given sample design and size. The measurement of sampling error is usually called the precision of the sampling plan'. We increase the sample size, the precision can be improved. But increasing the size of the sample has its own limitations viz., large sized sample increases the cost of collecting data and also enhances the systematic bias. Thus the effective way to increase precision is usually to select a better sampling design which has a smaller sampling error for a given sample size at a given cost. In practice, however, people prefer aless precise design because it is easier to adopt the same and also because of the fact that systematic bias can be controlled in a better way in such a design.

In brief, *while selecting a sampling procedure, researcher must ensure that the procedure causes a relatively small sampling error and helps to control the systematic bias in a better way.*

CHARACTERISTICS OF A GOOD SAMPLE DESIGN

From what has been stated above, we can list down the characteristics of a good sample design under:

a) Sample design must result in a truly representative sample.
b) Sample design must be such which results in a small sampling error.
c) Sample design must be viable in the context of funds .available for the research study.
d) Sample design must be such so that systematic bias can be controlled in a better way.
e) Sample should be such that the results of the sample study can be applied, in general, for the universe with a reasonable level of confidence.

DIFFERENT TYPES OF SAMPLE DESIGNS

There are different types of sample designs based on two factors viz., the representation basis and the element selection technique. On the representation basis, the sample may be probability sampling or it may be non-probability sampling. Probability sampling is based on the concept of random selection whereas non-probability sampling is 'non-random' sampling. On element selection basis, the sample may be either unrestricted or restricted. When each sample element is drawn individually from the population at large, then the sample so drawn is known as 'unrestricted sample', whereas all other forms of sampling are covered under the term 'restricted sampling'. The following chart exhibits the sample designs as explained above.

Thus, sample designs are basically of two types viz., non-probability sampling and probability sampling. We take up these two designs separately.

Non-probability sampling: Non-probability sampling is that sampling procedure which does not afford any basis for estimating the probability that each item in the population has of being included in the sample. Non-probability sampling is also known by different names such as deliberate sampling .purposive sampling and judgement sampling. In this type of sampling, items for the sample are selected deliberately by the researcher; his choice concerning the items remains supreme. In other words, under non-probability sampling the organisers of the inquiry purposively choose the particular units of the universe for constituting a sample on the basis that the small mass that they so select out of a huge one will be typical or representative of the whole. For instance, if economic conditions of people living in a state are to be studied, a few towns and villages may be purposively selected for intensive study-on the principle that they can be representative of the entire state. Thus, judgement of the organisers of the study plays an important part in this sampling design.

In such a design, personal element has a great chance of entering into the selection of the sample. The investigator may select a which shall yield results favourable to his point of view if that happens, the entire inquiry may get vitiated. Thus, there is always the danger of bias entering into this type of sampling technique, But in the investigators are impartial, work without bias and have the necessary experience so as to take sound judgement, the results' obtained from an analysis of deliberately selected sample may be tolerably reliable. However, in such a sampling, there is no assurance that every element has some specifiable-chance of being included. Sampling error in of sampling cannot be estimated and the element of bias, great or small, is always there. As such this sampling design in rarely adopted in large inquires of importance. However, in small inquiries and reaches by individuals, this design may be adopted because of the relative advantage of time and money inherent in this method of sampling. *Quota sampling* is also an example of non-probability sampling. Under quota sampling the interviewers are simply given quotas to be filled from the different strata with some restrictions on how they are to be filled. In other words, the actual selection of the sample is left to the interviewer's discretion. This type of sampling is very convenient and is relatively inexpensive. But the samples so selected certainly do not possess the characteristic of random samples. Quota samples are essentially judgement samples and inferences drawn on their bias are not amenable to statistical treatment in a formal way.

Probability sampling: Probability sampling is also known as 'random sampling' or 'chance sampling'. Under this sampling design, every item of the universe has an equal chance of inclusion in the sample. It is, so to say, a lottery method in which individual units are picked up from the whole group not deliberately but by some mechanical process. Here it is blind chance alone that determines whether one item or the other is selected. The results obtained from probability or random sampling can be assured in terms of probability i.e., we can measure the errors of estimation or the significance of results obtained from a random sample, and this fact brings out the superiority of random sampling design over the deliberate sampling design. Random sampling ensures the law of Statistical Regularity which states that if on an average the sample chosen is a random one, the sample will have the same composition and characteristics as the universe. This is the reason why random sampling is considered as the-best technique of selecting a representative sample.

Random sampling from a finite population refers to that method of sample selection which gives each possible sample combination an equal probability of being picked up and each item in the entire population to have an equal chance of being included in the sample. This applies to sampling without replacement i.e., once an item is selected for the sample, it cannot appear in the sample again (Sampling with replacement is used less frequently in which procedure the element selected for the sample is returned to the population before the next element is selected. In such a situation the same element could appear twice in the same sample before the second element is chosen). In brief, the implications of random sampling (or simple random sampling) are:

a. It gives each element in the population an equal probability of getting into the sample; a all choices are independent of one another.
b. It gives each possible sample combination an equal probability of being chosen.

Keeping this in view we can define a simple random sample (or simply a random sample) from a finite population as a sample which is chosen in such a way that each of the ${}^{N}C_{n}$ possible sample has the same probability, $1/{}^{N}C_{n}$, of being selected. To make it more clear we take a certain finite population consisting of six elements (say *a, b,* c, *d, e, f)* i.e., $N = 6$. Suppose that we want to take a sample of size $n = 3$ from it. Then there are ${}^{6}C_{3} = 20$ possible distinct samples of the required size, and they consist of the elements *abc, abd, abe, abf, acd, ace, acf, ade, adf, aef, bcd, bce, bcf, bde, bdf, bef, cde, cdf, cef,* and *def* If we choose one of these samples in such a way that each has the probability 1/20 of being chosen, we will then call this a random sample.

HOW TO SELECT A RANDOM SAMPLE?

With regard to the question of how to take a random sample in actual practice, we could, in simple cases like the. one above, write each of the possible samples on a slip of paper, mix these slips thoroughly in a container and then draw as a lottery either blindfolded or by rotating a drum or by any other similar device. Such a procedure is obviously impractical, if not altogether impossible in complete problems of sampling. In fact, the practical utility of such a method is very much limited.

Fortunately, we can take a random sample in a relatively easier way without taking the trouble of enlisting all possible samples on paper-slips as explained above. Instead of this, we can write the name of each element of a finite population on a slip of paper, put the slips of paper so prepared into a box or a bag and mix them thoroughly and then draw (without looking) the required number of slips for the sample one after the other without replacement. In doing so we must make sure that in successive drawings each of the remaining elements of the population has the same chance of being selected. This procedure will also result in the same probability for each possible sample. We can verify this by taking the above example. Since we have a finite population of 6 elements and we want to select a sample of size 3, the probability of drawing anyone element for our sample in the first draw is *3/6,* the probability of drawing one more element in the second draw is *2/5,* (the first element drawn is not replaced) and similarly the probability of drawing one more element in the third draw is *1/4.* Since these draws are independent, the joint probability of the three elements which constitute our sample is the product of their individual probabilities and this works out to 3/6 x 2/5 x 114 = 1120. This verifies our earlier calculation.

Even this relatively easy method of obtaining a random sample can be simplified in actual practice by the use of random number tables. Various statisticians like Tippett. Yates, Fisher have prepared tables of ran om numbers which can be used for selecting a random sample. Generally, Tippett's random number tables are used for the purpose. Tippett gavel 0400 four figure numbers. He selected 41600 digits from the census reports and combined them ilto fours to give his random numbers which may be used to obtain a random sample.

We can illustrate the procedure by an example. First of all we reproduce the first thirty sets of Tippett's numbers

2952	6641	3992	9792	7979	5911
3170	5624	4167	9525	1545	1396
7203	5356	1300	2693	2370	7483
3408	2769	3563	6107	6913	7691
0560	5246	1112	9025	6008	8126

Suppose we are interested in taking a sample of 10 units from a population of 5000 units, bearing numbers from 3001 to 8000. We shall select 10 such figures from the above random numbers which are not less than 3001 and not greater than 8000. If we randomly decide to read the table numbers left to right, starting from the first row itself, we obtain the following numbers: 6641, 3992, 7979, 5911, 3170, 5624, 4167, 7203, 5356, and 7483.

The units bearing the above serial numbers would then constitute our required random sample.

One may note that it is easy to draw random samples from finite populations with the aid of m number tables only when lists are available and items are readily numbered. But in some ions it is often impossible to proceed in the way we have narrated above. For example, if we to estimate the mean height of trees in a forest, it would not be possible to number the trees, and random numbers to select a random sample. In such situations what we should do is to select trees for the sample haphazardly without aim or purpose, and should treat the sample as a sample for study purposes.

RANDOM FROM AN INFINITE UNIVERSE

So for we have talked about random sampling, keeping in view only the finite populations. But what about random sampling in context of infinite populations? It is relatively difficult to explain the concept m sample from an infinite population. However, a few examples will show the basic characteristic of such a sample. Suppose we consider the 20 throws of a fair dice as a sample from the hypothetically infinite population which consists of the results of all possible throws of the dice. If the probability of getting a particular number, say 1, is the same for each throw and the 20 throws all independent, then we say that the sample is random. Similarly, it would be said to be sampling from an infinite population if we sample with replacement from a finite population and our sample would be considered as a random sample if in each draw all elements of the population have the same probability of being selected and successive draws happen to be independent. In brief, one can say that the selection of each item in a random sample from an infinite population is controlled by the same probabilities and that successive selections are independent of one another.

COMPLEX RANDOM SAMPLING DESIGNS

Probability sampling under restricted sampling techniques, as stated above, may result in comp random sampling designs. Such designs may as well be called 'mixed sampling designs' for many such designs may represent a combination of probability and non-probability sampling procedure selecting a sample. Some of the popular complex random sampling designs are as follows:

(i) Systematic sampling: In some instances, the most practical way of sampling is to select every *i*th item on a list. Sampling of this type is known as systematic sampling. An element of randomness is introduced into this kind of sampling by using random numbers to pick up the unit with which to start. For instance, if a 4 per cent sample is desired, the first item would be selected randomly from the first twenty-five and thereafter every 25th item would automatically be included in the sample. Thus, in systematic sampling only the first unit is selected randomly and the remaining units of the sample are selected at fixed intervals. Although a systematic sample is not a random sample in the strict sense of the term, but it is often considered reasonable to treat systematic sample as if it were a random sample.

Systematic sampling has certain plus points. It can be taken as an improvement over a simple random sample in as much as the systematic sample is spread more evenly over the entire population. It is an easier and less costlier method of sampling and can be conveniently used even in case of large populations. But there are certain dangers too in using this type of sampling. If there is a hidden periodicity in the population, systematic sampling will prove to be an inefficient method of sampling For instance, every 25th item produced by a certain production process is defective. If we are to select a 4% sample of the items of this process in a systematic manner, we would either get all defective items or all good items in our sample depending upon the random starting position. If all elements of the universe are ordered in a manner representative of the total population, i.e., the population list is in random order, systematic sampling is considered equivalent to random sampling. But if this is not so, then the results of such sampling may, at times, not be very reliable. In practice systematic sampling is used when lists of population are available and they are of considerable length.

(ii) Stratified sampling: If a population from which a sample is to be drawn does not constitute homogeneous group, stratified sampling technique is generally applied in order to obtain a representative sample. Under stratified sampling the population is divided into several sub-populations that are individually more homogeneous than the total population (the different sub-populations are c 'strata') and then we select items from each stratum to constitute a sample. Since each stratum is more homogeneous than the total population, we are able to get more precise estimates for each stratum and by estimating more accurately each of the component parts, we get a better estimate the whole. In brief, stratified sampling results in more reliable and detailed information.

The following three questions are highly relevant in the context of stratified sampling:

a) How to form strata?
b) How should items be selected from each stratum?
c) How many items be selected from each stratum or how to allocate the sample size of each stratum?

Regarding the first question, we can say that the strata be formed on the basis of common characteristic(s) of the items to be put in each stratum. This means that various strata be formed in such a way as to ensure elements being most homogeneous within each stratum and most heterogeneous between the different strata. Thus, strata are purposively formed and are usually based on past experience and personal judgement of the researcher. One should always remember that careful consideration of the relationship between the characteristics of the population and the characteristics to be estimated are normally used to define the strata. At times, pilot study may be conducted for determining a more appropriate and efficient stratification plan. We can do so by taking small samples of equal size from each of the proposed strata and then examining the variances within and among the possible stratifications, we can decide an appropriate stratification plan for our inquiry.

In respect of the second question, we can say that the usual method, for selection of items for the sample from each stratum, resorted to is that of simple random sampling. Systematic sampling can be used if it is considered more appropriate in certain situations.

Regarding the third question, we usually follow the method of proportional allocation under which the sizes of the samples from the different strata are kept proportional to the sizes of the strata. That if P_i represents the proportion of population included in stratum i, and n represents the total sample size, the number of elements selected from stratum i is $n. P_i$. To illustrate it, let us suppose that we want a sample of size $n = 30$ to be drawn from a population of size $N = 8000$ which is divided into three strata of size $N_1 = 4000$, $N_2 = 2400$ and $N_3 = 1600$. Adopting proportional allocation, we shall the sample sizes as under for the different strata:

For strata with $N_i = 4000$, we have $P_i = 4000/8000$ hence $n_1 = n. PI = 30\ (4000/8000) = 15$

Similarly, for strata with $N2 = 2400$, we have

$n_2 = n. P_2 = 30\ (2400/8000) = 9$, and

for strata with $N3 = 1600$, we have

$n_3 = n. P_3 = 30\ (1600/8000) = 6$.

Thus, using proportional allocation, the sample sizes for different strata are 15,9 and 6 respectively which is in proportion to the sizes of the strata viz., 4000 : 2400 : 1600. Proportional allocation is considered most efficient and an optimal design when the cost of selecting an item is equal for each stratum, there is no difference in within-stratum variances, and the purpose of sampling happens to be to estimate the population value of some characteristic. But in case the purpose happens to be to compare the differences among the strata, then equal sample selection from each stratum would be more efficient even if the strata differ in sizes. In cases where strata differ not only in size but also in variability and it is considered reasonable to take larger samples from the more variable strata and smaller samples from the less variable strata, we can then account for both (differences in stratum size and differences in stratum variability) by using disproportionate sampling design by requiring:

$$n_1/N_1 \sigma_1 = n_2/N_2 \sigma_2 = \ldots\ldots\ldots = n_k/N_k \sigma_k$$

Where $\sigma_1, \sigma_2, \ldots$ and σ_k denote standard deviations of the k strata, $N_1, N_2, \ldots, N_k$ denote the sizes of the k strata and $n_1, n_2 \ldots, n_k$ denote the sample sizes of k strata. This is called *'optimum allocation'* in the context of disproportionate sampling. The allocation in such a situation results in the following formula for determining the sample sizes different strata:

$$n_i = \frac{n. N_i \sigma_i}{N_1 \sigma_1 + N_2 \sigma_2 + \ldots + N_k \sigma_k} \quad \text{for } i = 1, 2, \ldots \text{ and } k.$$

We may illustrate the use of this by an example.

Illustration 1

A population is divided into three strata so that N_1 = 5000, N_2 = 2000 and N_3 = 3000. Respective

standard deviations are:

a_1 = 15, a_2 = 18 and o , = 5.

How should a sample of size *n* = 84 be allocated to the three strata, if we want optimum allocation using disproportionate sampling design?

Solution: Using the disproportionate sampling design for optimum allocation, the sample sizes for different strata will be determined as under:

Sample size for strata with N_1 = 5000

84(5000) (15)

n_1 =--

(5000) (15) + (2000) (18) + (3000) (5)

= 6300000/126000 = 50

Sample size for strata with N_2 = 2000

84(2000) (15)

n_2 =---

(5000) (15) + (2000) (18) + (3000) (5)

=3024000/126000=24

Sample size for strata with N_3 = 3000

84(3000) (15)

n_3=---

(5000) (15) + (2000) (18) + (3000) (5)

=1260000/126000=10

In addition to differences in stratum size and differences in stratum variability, we ma differences in stratum sampling cost, then we can have cost optimal disproportionate sampling by requiring

$$\frac{n_1}{N_1 \sigma_1 \sqrt{C_1}} = \frac{n_2}{N_2 \sigma_2 \sqrt{C_2}} = \cdots = \frac{n_k}{N_k \sigma_k \sqrt{C_k}}$$

where

C_1 = Cost of sampling in stratum 1

C_2 = Cost of sampling in stratum 2

C_k = Cost of sampling in stratum k

and all other terms remain the same as explained earlier. The allocation in such a situation results in the following formula for determining the sample sizes for different strata:

$$n_i = \frac{n \cdot N_i \sigma_i / \sqrt{C_i}}{N_1 \sigma_1 / \sqrt{C_1} + N_2 \sigma_2 / \sqrt{C_2} + \ldots + N_k \sigma_k / \sqrt{C_k}} \quad \text{for } i = 1, 2, \ldots, k$$

It is not necessary that stratification be done keeping in view a single characteristic. Populations are often stratified according to several characteristics. For example, a system-wide survey designed to determine the attitude of students toward a new teaching plan, a state college system with 20 colleges might stratify the students with respect to class, see and college. Stratification of this type is known as *cross-stratification,* and up to a point such stratification increases the reliability of estimates and is much used in opinion surveys.

From what has been stated above in respect of stratified sampling, we can say that the sample so constituted is the result of successive application of purposive (involved in stratification of items) and random sampling methods. As such it is an example of mixed sampling. The procedure wherein we first have stratification and then simple random sampling is known as stratified random sampling.

(iii) Cluster sampling: If the total area of interest happens to be a big one, a convenient way in which a sample can be taken is to divide the area into a number of smaller non-overlapping areas and then to randomly select a number of these smaller areas (usually called clusters), with the ultimate sample consisting of all (or samples of) units in these small areas or clusters.

Thus in cluster sampling the total population is divided into a number of relatively small subdivisions which are themselves clusters of still smaller units and then some of these clusters are randomly selected for inclusion in the overall sample. Suppose we want to estimate the proportion of machine- parts in an inventory which are defective. Also assume that there are 20000 machine parts in the inventory at a given point of time, stored in 400 cases of 50 each. Now using a cluster sampling, we would consider the 400 cases as clusters and randomly select' *n'* cases and examine all the machine- parts in each randomly selected case.

Cluster sampling, no doubt, reduces cost by concentrating surveys in selected clusters. But certainly it is less precise than random sampling. There is also not as much information in *'n'* observations within a cluster as there happens to be in *'n'* randomly drawn observations. Cluster sampling is used only because of the economic advantage it possesses; estimates based on cluster Samples are usually more reliable per unit cost.

(iv)Area sampling: If clusters happen to be some geographic' subdivisions, in that case cluster sampling is better known as area sampling. In other words, cluster designs, where the primary sampling unit represents a cluster of units based on geographic area, are distinguished as area sampling. The plus and minus points of cluster sampling are also applicable to area sampling.

(v)Multi-stage sampling: Multi-stage sampling is a further development of the principle of cluster sampling. Suppose we want to investigate the working efficiency of nationalised banks in India and we want to take a sample of few banks for this purpose. The first stage is to select large primary sampling unit such as states in a country. Then we may select certain districts and interview all banks in the chosen districts. This would represent a two-stage sampling design with the ultimate sampling units being clusters of districts.

If instead of taking a census of all banks within the selected districts, we select certain towns and interview all banks in the chosen towns. This would represent a three-stage sampling design. If instead of taking a-census of all banks within the selected towns, we randomly sample banks from each selected town, then it is a case of using a four-stage sampling plan. If we select randomly at all stages, we will have what is known as 'multi-stage random sampling design'.

Ordinarily multi-stage sampling is applied in big inquires extending to a considerable large geographical area, say, the entire country. There are two advantages of this sampling design viz., (a) It is easier to administer than most single stage designs mainly because of the fact that sampling frame under multi-stage sampling is developed in partial units. (b) A large number of units can be sampled for a given cost under multistage sampling because of sequential clustering, whereas this is not possible in most of the simple designs.

(vi)Sampling with probability proportional to size: In case the cluster sampling units do not have the same number or approximately the same number of elements, it is considered appropriate to use a random selection process where the probability of each cluster being included in the sample is proportional to the size of the cluster. For this purpose, we have to list the number of elements in each cluster irrespective of the method of ordering the cluster. Then we must sample systematically the appropriate number of elements from the cumulative totals. The actual numbers selected in this way do not refer to individual elements, but indicate which clusters and how many from the cluster are to be selected by simple random sampling or by systematic sampling. The results of this type of sampling are equivalent to those of a simple random sample and the method is less cumbersome and is alsorelatively less expensive. We can illustrate this with the help of an example.

Illustration 2

The following are the number of departmental stores in 15 cities: 35, 17, 10, 32,70,28,26, 19, 26, 66, 37, 44, 33, 29 and 28. If we want to select a sample of 10 stores, using cities as clusters selecting within clusters proportional to size, how many stores from each city should be chose (Use a starting point of 10).

Solution: Let us put the information as under (Table 4.1):

Since in the given problem, we have 500 departmental stores from which we have to select a sample of 10 stores, the appropriate sampling interval is 50. As we have to use the starting point of 10, so we add successively increments of 50 till 10 numbers have been selected. The numbers, thus obtained are: 10,60,110,160,210,260,310,360,410 and 460 which have been shown in the last column of the table (Table 4.1) against the concerning cumulative totals. From this we can say that two stores should be selected randomly from city number five and one each from city number 1,3, 7, 9, 10, 11, 12, and 14. This sample of 10 stores is the sample with probability

UNIT V

MEANING AND CONCEPTS OF SAMPLING DESIGN

In this context one must remember that two costs are involved in a sampling analysis viz., the cost of collecting the data and the cost of an incorrect inference resulting from the data. Researcher must keep in view the two causes -of incorrect inferences viz., systematic bias and sampling error. A *systematic bias* results from errors in the sampling procedures and it cannot be reduced or eliminated by increasing the sample size. At best the causes responsible for these errors can be detected and corrected. Usually a systematic bias is the result of one or more of the following factors:

6. **Inappropriate sampling frame:** If the sampling frame is inappropriate i.e., a biased representation of the universe, it will result in a systematic bias.
7. **Defective measuring device:** If the measuring device is constantly in error, it will result in systematic bias. In survey work, systematic bias can result if the questionnaire or the interviewer is biased. Similarly, if the physical measuring device is defective there will be systematic bias in the data collected through such a measuring device.
8. **Non-respondents:** If we are unable to sample all the individuals initially included in the sample, may arise a systematic bias. The reason is that in such a situation the likelihood of establishing t or receiving a response from an individual is often correlated with the measure of what is to be estimated.
9. **Indeterminancy principle:** Sometimes we find that individuals act differently when kept under observation than what they do when kept in non-observed situations. For instance, if workers are that somebody is observing them in course of a work study ow4'he basis of which the average of time to complete a task will be determined and accordingly the quota will be set for piece they generally tend to work slowly in comparison to the speed with which they work if kept ed. Thus, the indeterminancy principle may also be a cause of a systematic bias.
10. **Natural bias in the reporting of data:** Natural bias of respondents in the reporting of data is often the cause of a systematic bias in many inquiries! There is usually a 'downward bias in the income data collected by government taxation department, whereas we find an upward bias in the income data collected by some social organisation. People in general understate their incomes if asked about it for tax purposes, but they overstate the same if asked for social status or their affluence. Generally in psychological surveys, people tend to give what they think is the' correct' answer rather than revealing their true

feelings.

Sampling errors are the random variations in the sample estimates around the true population parameters. Since they occur randomly and are equally likely to be in either direction, their nature happens to be of compensatory type and the expected value of such errors happens to be equal to zero. Sampling error decreases with the increase in the size of the sample, and it happens to be of a smaller magnitude in case of homogeneous population.

Sampling error can be measured for a given sample design and size. The measurement of sampling error is usually called the precision of the sampling plan'. We increase the sample size, the precision can be improved. But increasing the size of the sample has its own limitations viz., large sized sample increases the cost of collecting data and also enhances the systematic bias. Thus the effective way to increase precision is usually to select a better sampling design which has a smaller sampling error for a given sample size at a given cost. In practice, however, people prefer aless precise design because it is easier to adopt the same and also because of the fact that systematic bias can be controlled in a better way in such a design.

In brief, *while selecting a sampling procedure, researcher must ensure that the procedure causes a relatively small sampling error and helps to control the systematic bias in a better way*.

CHARACTERISTICS OF A GOOD SAMPLE DESIGN

From what has been stated above, we can list down the characteristics of a good sample design under:

f) Sample design must result in a truly representative sample.
g) Sample design must be such which results in a small sampling error.
h) Sample design must be viable in the context of funds .available for the research study.
i) Sample design must be such so that systematic bias can be controlled in a better way.
j) Sample should be such that the results of the sample study can be applied, in general, for the universe with a reasonable level of confidence.

DIFFERENT TYPES OF SAMPLE DESIGNS

There are different types of sample designs based on two factors viz., the representation basis and the element selection technique. On the representation basis, the sample may be probability sampling or it may be non-probability sampling. Probability sampling is based on the concept of random selection whereas non-probability sampling is 'non-random' sampling. On element selection basis, the sample may be either unrestricted or restricted. When each sample element is drawn individually from the population at large, then the sample so drawn is known as 'unrestricted sample', whereas all other forms of sampling are covered under the term 'restricted sampling'. The following chart exhibits the sample designs as explained above.

Thus, sample designs are basically of two types viz., non-probability sampling and probability sampling. We take up these two designs separately.

Non-probability sampling: Non-probability sampling is that sampling procedure which does not afford any basis for estimating the probability that each item in the population has of being included in the sample. Non-probability sampling is also known by different names such as deliberate sampling .purposive sampling and judgement sampling. In this type of sampling, items for the sample are selected deliberately by the researcher; his choice concerning the items remains supreme. In other words, under non-probability sampling the organisers of the inquiry purposively choose the particular units of the universe for constituting a sample on the basis that the small mass that they so select out of a huge one will be typical or representative of the whole. For instance, if economic conditions of people living in a state are to be studied, a few towns and villages may be purposively selected for intensive study-on the principle that they can be representative of the entire state. Thus, judgement of the organisers of the study plays an important part in this sampling design.

In such a design, personal element has a great chance of entering into the selection of the sample. The investigator may select a which shall yield results favourable to his point of view if that happens, the entire inquiry may get vitiated. Thus, there is always the danger of bias entering into this type of sampling technique, But in the investigators are impartial, work without bias and have the necessary experience so as to take sound judgement, the results' obtained from an analysis of deliberately selected sample may be tolerably reliable. However, in such a sampling, there is no assurance that every element has some specifiable-chance of being included. Sampling error in of sampling cannot be estimated and the element of bias, great or small, is always there. As such this sampling design in rarely adopted in large inquires of importance. However, in small inquiries and reaches by individuals, this design may be adopted because of the relative advantage of time and money inherent in this method of sampling. *Quota sampling* is also an example of non-probability sampling. Under quota sampling the interviewers are simply given quotas to be filled from the different strata with some restrictions on how they are to be filled. In other words, the actual selection of the sample is left to the interviewer's discretion. This type of sampling is very convenient and is relatively inexpensive. But the samples so selected certainly do not possess the characteristic of random samples. Quota samples are essentially judgement samples and inferences drawn on their bias are not amenable to statistical treatment in a formal way.

Probability sampling: Probability sampling is also known as 'random sampling' or 'chance sampling'. Under this sampling design, every item of the universe has an equal chance of inclusion in the sample. It is, so to say, a lottery method in which individual units are picked up from the whole group not deliberately but by some mechanical process. Here it is blind chance alone that determines whether one item or the other is selected. The results obtained from probability or random sampling can be assured in terms of probability i.e., we can measure the errors of estimation or the significance of results obtained from a random sample, and this fact brings out the superiority of random sampling design over the deliberate sampling design. Random sampling ensures the law of Statistical Regularity which states that if on an average the sample chosen is a random one, the sample will have the same composition and characteristics as the universe. This is the reason why random sampling is considered as the-best technique of selecting a representative sample.

Random sampling from a finite population refers to that method of sample selection which gives each possible sample combination an equal probability of being picked up and each item in the entire population to have an equal chance of being included in the sample. This applies to sampling without replacement i.e., once an item is selected for the sample, it cannot appear in the sample again (Sampling with replacement is used less frequently in which procedure the element selected for the sample is returned to the population before the next element is selected. In such a situation the same element could appear twice in the same sample before the second element is chosen). In brief, the implications of random sampling (or simple random sampling) are:

c. It gives each element in the population an equal probability of getting into the sample; a all choices are independent of one another.
d. It gives each possible sample combination an equal probability of being chosen.

Keeping this in view we can define a simple random sample (or simply a random sample) from a finite population as a sample which is chosen in such a way that each of the $^{N}C_{n}$ possible sample has the same probability, $1/^{N}C_{n}$, of being selected. To make it more clear we take a certain finite population consisting of six elements (say *a, b,* c, *d, e, f)* i.e., $N = 6$. Suppose that we want to take a sample of size $n = 3$ from it. Then there are $^{6}C_{3} = 20$ possible distinct samples of the required size, and they consist of the elements *abc, abd, abe, abf, acd, ace, acf, ade, adf, aef, bcd, bce, bcf, bde, bdf, bef, cde, cdf, cef,* and *def* If we choose one of these samples in such a way that each has the probability 1/20 of being chosen, we will then call this a random sample.

HOW TO SELECT A RANDOM SAMPLE?

With regard to the question of how to take a random sample in actual practice, we could, in simple cases like the. one above, write each of the possible samples on a slip of paper, mix these slips thoroughly in a container and then draw as a lottery either blindfolded or by rotating a drum or by any other similar device. Such a procedure is obviously impractical, if not altogether impossible in complete problems of sampling. In fact, the practical utility of such a method is very much limited.

Fortunately, we can take a random sample in a relatively easier way without taking the trouble of enlisting all possible samples on paper-slips as explained above. Instead of this, we can write the name of each element of a finite population on a slip of paper, put the slips of paper so prepared into a box or a bag and mix them thoroughly and then draw (without looking) the required number of slips for the sample one after the other without replacement. In doing so we must make sure that in successive drawings each of the remaining elements of the population has the same chance of being selected. This procedure will also result in the same probability for each possible sample. We can verify this by taking the above example. Since we have a finite population of 6 elements and we want to select a sample of size 3, the probability of drawing anyone element for our sample in the first draw is *3/6,* the probability of drawing one more element in the second draw is *2/5,* (the first element drawn is not replaced) and similarly the probability of drawing one more element in the third draw is *1/4.* Since these draws are independent, the joint probability of the three elements which constitute our sample is the product of their individual probabilities and this works out to 3/6 x 2/5 x 114 = 1120. This verifies our earlier calculation.

Even this relatively easy method of obtaining a random sample can be simplified in actual practice by the use of random number tables. Various statisticians like Tippett. Yates, Fisher have prepared tables of ran om numbers which can be used for selecting a random sample. Generally, Tippett's random number tables are used for the purpose. Tippett gavel 0400 four figure numbers. He selected 41600 digits from the census reports and combined them ilto fours to give his random numbers which may be used to obtain a random sample.

We can illustrate the procedure by an example. First of all we reproduce the first thirty sets of Tippett's numbers

2952	6641	3992	9792	7979	5911
3170	5624	4167	9525	1545	1396
7203	5356	1300	2693	2370	7483
3408	2769	3563	6107	6913	7691
0560	5246	1112	9025	6008	8126

Suppose we are interested in taking a sample of 10 units from a population of 5000 units, bearing numbers from 3001 to 8000. We shall select 10 such figures from the above random numbers which are not less than 3001 and not greater than 8000. If we randomly decide to read the table numbers left to right, starting from the first row itself, we obtain the following numbers: 6641, 3992, 7979, 5911, 3170, 5624, 4167, 7203, 5356, and 7483.

The units bearing the above serial numbers would then constitute our required random sample.

One may note that it is easy to draw random samples from finite populations with the aid of m number tables only when lists are available and items are readily numbered. But in some ions it is often impossible to proceed in the way we have narrated above. For example, if we to estimate the mean height of trees in a forest, it would not be possible to number the trees, and random numbers to select a random sample. In such situations what we should do is to select trees for the sample haphazardly without aim or purpose, and should treat the sample as a sample for study purposes.

RANDOM FROM AN INFINITE UNIVERSE

So for we have talked about random sampling, keeping in view only the finite populations. But what about random sampling in context of infinite populations? It is relatively difficult to explain the concept m sample from an infinite population. However, a few examples will show the basic characteristic of such a sample. Suppose we consider the 20 throws of a fair dice as a sample from the hypothetically infinite population which consists of the results of all possible throws of the dice. If the probability of getting a particular number, say 1, is the same for each throw and the 20 throws all independent, then we say that the sample is random. Similarly, it would be said to be sampling from an infinite population if we sample with replacement from a finite population and our sample would be considered as a random sample if in each draw all elements of the population have the same probability of being selected and successive draws happen to be independent. In brief, one can say that the selection of each item in a random sample from an infinite population is controlled by the same probabilities and that successive selections are independent of one another.

COMPLEX RANDOM SAMPLING DESIGNS

Probability sampling under restricted sampling techniques, as stated above, may result in comp random sampling designs. Such designs may as well be called 'mixed sampling designs' for many such designs may represent a combination of probability and non-probability sampling procedure selecting a sample. Some of the popular complex random sampling designs are as follows:

(i) Systematic sampling: In some instances, the most practical way of sampling is to select every *i*th item on a list. Sampling of this type is known as systematic sampling. An element of randomness is introduced into this kind of sampling by using random numbers to pick up the unit with which to start. For instance, if a 4 per cent sample is desired, the first item would be selected randomly from the first twenty-five and thereafter every 25th item would automatically be included in the sample. Thus, in systematic sampling only the first unit is selected randomly and the remaining units of the sample are selected at fixed intervals. Although a systematic sample is not a random sample in the strict sense of the term, but it is often considered reasonable to treat systematic sample as if it were a random sample.

Systematic sampling has certain plus points. It can be taken as an improvement over a simple random sample in as much as the systematic sample is spread more evenly over the entire population. It is an easier and less costlier method of sampling and can be conveniently used even in case of large populations. But there are certain dangers too in using this type of sampling. If there is a hidden periodicity in the population, systematic sampling will prove to be an inefficient method of sampling For instance, every 25th item produced by a certain production process is defective. If we are to select a 4% sample of the items of this process in a systematic manner, we would either get all defective items or all good items in our sample depending upon the random starting position. If all elements of the universe are ordered in a manner representative of the total population, i.e., the population list is in random order, systematic sampling is considered equivalent to random sampling. But if this is not so, then the results of such sampling may, at times, not be very reliable. In practice systematic sampling is used when lists of population are available and they are of considerable length.

(ii) Stratified sampling: If a population from which a sample is to be drawn does not constitute homogeneous group, stratified sampling technique is generally applied in order to obtain a representative sample. Under stratified sampling the population is divided into several sub-populations that are individually more homogeneous than the total population (the different sub-populations are c 'strata') and then we select items from each stratum to constitute a sample. Since each stratum is more homogeneous than the total population, we are able to get more precise estimates for each stratum and by estimating more accurately each of the component parts, we get a better estimate the whole. In brief, stratified sampling results in more reliable and detailed information.

The following three questions are highly relevant in the context of stratified sampling:

d) How to form strata?
e) How should items be selected from each stratum?
f) How many items be selected from each stratum or how to allocate the sample size of each stratum?

Regarding the first question, we can say that the strata be formed on the basis of common characteristic(s) of the items to be put in each stratum. This means that various strata be formed in such a way as to ensure elements being most homogeneous within each stratum and most heterogeneous between the different strata. Thus, strata are purposively formed and are usually based on past experience and personal judgement of the researcher. One should always remember that careful consideration of the relationship between the characteristics of the population and the characteristics to be estimated are normally used to define the strata. At times, pilot study may be conducted for determining a more appropriate and efficient stratification plan. We can do so by taking small samples of equal size from each of the proposed strata and then examining the variances within and among the possible stratifications, we can decide an appropriate stratification plan for our inquiry.

In respect of the second question, we can say that the usual method, for selection of items for the sample from each stratum, resorted to is that of simple random sampling. Systematic sampling can be used if it is considered more appropriate in certain situations.

Regarding the third question, we usually follow the method of proportional allocation under which the sizes of the samples from the different strata are kept proportional to the sizes of the strata. That if P_i represents the proportion of population included in stratum i, and n represents the total sample size, the number of elements selected from stratum i is $n.\ P_i$. To illustrate it, let us suppose that we want a sample of size $n = 30$ to be drawn from a population of size $N = 8000$ which is divided into three strata of size $N_1 = 4000$, $N_2 = 2400$ and $N_3 = 1600$. Adopting proportional allocation, we shall the sample sizes as under for the different strata:

For strata with $N_i = 4000$, we have $P_i = 4000/8000$ hence $n_1 = n.\ PI = 30\ (4000/8000) = 15$

Similarly, for strata with $N2 = 2400$, we have

$n_2 = n.\ P_2 = 30\ (2400/8000) = 9$, and

for strata with $N3 = 1600$, we have

$n_3 = n.\ P_3 = 30\ (1600/8000) = 6$.

Thus, using proportional allocation, the sample sizes for different strata are 15,9 and 6 respectively which is in proportion to the sizes of the strata viz., 4000 : 2400 : 1600. Proportional allocation is considered most efficient and an optimal design when the cost of selecting an item is equal for each stratum, there is no difference in within-stratum variances, and the purpose of sampling happens to be to estimate the population value of some characteristic. But in case the purpose happens to be to compare the differences among the strata, then equal sample selection from each stratum would be more efficient even if the strata differ in sizes. In cases where strata differ not only in size but also in variability and it is considered reasonable to take larger samples from the more variable strata and smaller samples from the less variable strata, we can then account for both (differences in stratum size and differences in stratum variability) by using disproportionate sampling design by requiring:

$$n_1/N_1\ \sigma_1 = n_2/N_2\ \sigma_2 = \ldots\ldots\ldots\ldots = n_k/N_k\ \sigma_k$$

Where $\sigma_1, \sigma_2, \ldots$ and σ_k denote standard deviations of the k strata, $N_1, N_2, \ldots, N_k$ denote the sizes of the k strata and $n_1, n_2 \ldots, n_k$ denote the sample sizes of k strata. This is called *'optimum allocation'* in the context of disproportionate sampling. The allocation in such a situation results in the following formula for determining the sample sizes different strata:

$$n.\ N_i\ \sigma_i$$

n_i=.--------------------------------------- for i = 1, 2, ... and k.

$N_1\sigma_1 + N_2\sigma_2 + ... + N_k\sigma_k$

We may illustrate the use of this by an example.

Illustration 1

A population is divided into three strata so that N_1 = 5000, N_2 = 2000 and N_3 = 3000. Respective

standard deviations are:

a_1 = 15, a_2 = 18 and o , = 5.

How should a sample of size n = 84 be allocated to the three strata, if we want optimum allocation using disproportionate sampling design?

Solution: Using the disproportionate sampling design for optimum allocation, the sample sizes for different strata will be determined as under:

Sample size for strata with N_1 = 5000

84(5000) (15)

n_1 =---

(5000) (15) + (2000) (18) + (3000) (5)

= 6300000/126000 = 50

Sample size for strata with N_2 = 2000

84(2000) (15)

n_2 =---

(5000) (15) + (2000) (18) + (3000) (5)

=3024000/126000=24

Sample size for strata with N_3 = 3000

84(3000) (15)

n_3=---

(5000) (15) + (2000) (18) + (3000) (5)

=1260000/126000=10

In addition to differences in stratum size and differences in stratum variability, we ma differences in stratum sampling cost, then we can have cost optimal disproportionate sampling by requiring

$$\frac{n_1}{N_1\,\sigma_1\sqrt{C_1}} = \frac{n_2}{N_2\,\sigma_2\sqrt{C_2}} = \cdots = \frac{n_k}{N_k\,\sigma_k\sqrt{C_k}}$$

where

C_1 = Cost of sampling in stratum 1

C_2 = Cost of sampling in stratum 2

C_k = Cost of sampling in stratum *k*

and all other terms remain the same as explained earlier. The allocation in such a situation results in the following formula for determining the sample sizes for different strata:

$$n_i = \frac{n \cdot N_i\sigma_i/\sqrt{C_i}}{N_1\sigma_1/\sqrt{C_1} + N_2\sigma_2/\sqrt{C_2} + \ldots + N_k\sigma_k/\sqrt{C_k}} \quad \text{for } i = 1, 2, \ldots, k$$

It is not necessary that stratification be done keeping in view a single characteristic. Populations are often stratified according to several characteristics. For example, a system-wide survey designed to determine the attitude of students toward a new teaching plan, a state college system with 20 colleges might stratify the students with respect to class, see and college. Stratification of this type is known as *cross-stratification*, and up to a point such stratification increases the reliability of estimates and is much used in opinion surveys.

From what has been stated above in respect of stratified sampling, we can say that the sample so constituted is the result of successive application of purposive (involved in stratification of items) and random sampling methods. As such it is an example of mixed sampling. The procedure wherein we first have stratification and then simple random sampling is known as stratified random sampling.

(iii) Cluster sampling: If the total area of interest happens to be a big one, a convenient way in which a sample can be taken is to divide the area into a number of smaller non-overlapping areas and then to randomly select a number of these smaller areas (usually called clusters), with the ultimate sample consisting of all (or samples of) units in these small areas or clusters.

Thus in cluster sampling the total population is divided into a number of relatively small subdivisions which are themselves clusters of still smaller units and then some of these clusters are randomly selected for inclusion in the overall sample. Suppose we want to estimate the proportion of machine- parts in an inventory which are defective. Also assume that there are 20000 machine parts in the inventory at a given point of time, stored in 400 cases of 50 each. Now using a cluster sampling, we would consider the 400 cases as clusters and randomly select' *n'* cases and examine all the machine- parts in each randomly selected case.

Cluster sampling, no doubt, reduces cost by concentrating surveys in selected clusters. But certainly it is less precise than random sampling. There is also not as much information in *'n'* observations within a cluster as there happens to be in *'n'* randomly drawn observations. Cluster sampling is used only because of the economic advantage it possesses; estimates based on cluster Samples are usually more reliable per unit cost.

(iv)Area sampling: If clusters happen to be some geographic' subdivisions, in that case cluster sampling is better known as area sampling. In other words, cluster designs, where the primary sampling unit represents a cluster of units based on geographic area, are distinguished as area sampling. The plus and minus points of cluster sampling are also applicable to area sampling.

(v)Multi-stage sampling: Multi-stage sampling is a further development of the principle of cluster sampling. Suppose we want to investigate the working efficiency of nationalised banks in India and we want to take a sample of few banks for this purpose. The first stage is to select large primary sampling unit such as states in a country. Then we may select certain districts and interview all banks in the chosen districts. This would represent a two-stage sampling design with the ultimate sampling units being clusters of districts.

If instead of taking a census of all banks within the selected districts, we select certain towns and interview all banks in the chosen towns. This would represent a three-stage sampling design. If instead of taking a-census of all banks within the selected towns, we randomly sample banks from each selected town, then it is a case of using a four-stage sampling plan. If we select randomly at all stages, we will have what is known as 'multi-stage random sampling design'.

Ordinarily multi-stage sampling is applied in big inquires extending to a considerable large geographical area, say, the entire country. There are two advantages of this sampling design viz., (a) It is easier to administer than most single stage designs mainly because of the fact that sampling frame under multi-stage sampling is developed in partial units. (b) A large number of units can be sampled for a given cost under multistage sampling because of sequential clustering, whereas this is not possible in most of the simple designs.

(vi)Sampling with probability proportional to size: In case the cluster sampling units do not have the same number or approximately the same number of elements, it is considered appropriate to use a random selection process where the probability of each cluster being included in the sample is proportional to the size of the cluster. For this purpose, we have to list the number of elements in each cluster irrespective of the method of ordering the cluster. Then we must sample systematically the appropriate number of elements from the cumulative totals. The actual numbers selected in this way do not refer to individual elements, but indicate which clusters and how many from the cluster are to be selected by simple random sampling or by systematic sampling. The results of this type of sampling are equivalent to those of a simple random sample and the method is less cumbersome and is alsorelatively less expensive. We can illustrate this with the help of an example.

Illustration 2

The following are the number of departmental stores in 15 cities: 35, 17, 10, 32,70,28,26, 19, 26, 66, 37, 44, 33, 29 and 28. If we want to select a sample of 10 stores, using cities as clusters selecting within clusters proportional to size, how many stores from each city should be chose (Use a starting point of 10).

Solution: Let us put the information as under (Table 4.1):

Since in the given problem, we have 500 departmental stores from which we have to select a sample of 10 stores, the appropriate sampling interval is 50. As we have to use the starting point of 10, so we add successively increments of 50 till 10 numbers have been selected. The numbers, thus obtained are: 10,60,110,160,210,260,310,360,410 and 460 which have been shown in the last column of the table (Table 4.1) against the concerning cumulative totals. From this we can say that two stores should be selected randomly from city number five and one each from city number 1,3, 7, 9, 10, 11, 12, and 14. This sample of 10 stores is the sample with probability

(vii) Sequential sampling: This sampling design is some what complex sample design. The ultimate size of the sample under this technique is not fixed in advance, but is determined according to mathematical decision rules on the basis of information yielded as survey progresses. This is usually adopted in case of acceptance sampling plan in context of statistical quality control. When a particular lot is to be accepted or rejected on the basis of a single sample, it is known as single sampling; when the decision is to be taken on the basis of two samples, it is known as double sampling and in case the decision rests on the basis of more than two samples but the number of samples is certain and decided in advance, the sampling is known as multiple sampling. But when the number of samples is more than two but it is neither certain nor decided in advance, this type of system is often referred to as sequential sampling. Thus, in brief, we can say that in sequential sampling, one can go on taking samples one after another as long as one desires to do so.

CONCLUSION

From a brief description of the various sample designs presented above, we can say that normally one should resort to simple random sampling because under it bias is generally eliminated and the sampling error can be estimated. But purposive sampling is considered more appropriate when the universe happens to be small and a known characteristic of it is to be studied intensively. There are situations in real life under which sample designs other than simple random samples may be considered better (say easier to obtain, cheaper or more informative) and as such the same may be used. In a situation when random sampling is not possible, then we have to use necessarily a sampling design other than random sampling. At times, several methods of sampling may well be used in the same study.

Measurement and Scaling Techniques

MEASUREMENT SCALES

In our daily life we are said to measure when we use some yardstick to determine weight, height, or some other feature of a physical object. We also measure when we judge how well we like a song, a painting or the personalities of our friends. We, thus, measure physical objects as well as abstract concepts. Measurement is a relatively complex and demanding task, specially so when it concerns qualitative or abstract phenomena. By measurement we mean the process of assigning numbers to objects or observations, the level of measurement being a function of the rules under which the numbers are assigned.

It is easy to assign numbers in respect of properties of some objects, but it is relatively difficult in respect of others. For instance, measuring such things as social conformity, intelligence, or marital adjustment is much less obvious and requires much closer attention than measuring physical weight, biological age or a person's financial assets. In other words, properties like weight, height, etc., can be measured directly with some standard unit of measurement, but it is not that easy to measure properties like motivation to succeed, ability to stand stress and the like. We can expect high accuracy in measuring the length of pipe with a yard stick, but if the concept is abstract and the measurement tools are not standardized, we are less confident about the accuracy of the results of measurement.

Technically speaking, measurement is a process of mapping aspects of a domain onto other aspects a range according to some rule correspondence n measuring, we devise some form of scale in the range (in terms of set theory, range may refer to some set) and then transform or map the of objects from the domain (in terms of set theory, domain may refer to some other set) scale. For example, in case we are to find the male to female attendance ratio while conducting a study of persons who attend some show, then we may tabulate those who come to the show according to sex. In terms of set theory, this process is one of mapping the observed physical properties of those corning to the show (the domain) on to a sex classification (the range). The rule of correspondence is: If the object in the domain appears to be male, assign to "0" and if female assign to "I". Similarly, we can record a person's marital status as 1,2,3 or 4, depending on whether the person is single, married, widowed or divorced. We can as well record "Yes or No" answers to a question as "0" and "I" (or as 1 and 2 or perhaps as 59 and 60). In this artificial or nominal way, categorical data (qualitative or descriptive) can be made into numerical data and if we thus code the various categories, we refer to the numbers we record as nominal data. *Nominal data* are numerical in name only, because they do not share any of the properties of the numbers we deal in ordinary arithmetic. For instance if we record marital status as 1,2, 3, or 4 as stated above, we cannot write $4 > 2$ or $3 < 4$ and we cannot write $3 - 1 = 4 - 2$, $1 + 3 = 4$ or $4 + 2 = 2$.

In those situations when we cannot do anything except set up inequalities, we refer to the data as *ordinal data*. For instance, if one mineral can scratch another, it receives a higher hardness number and on Mohs' scale the numbers from 1 to 10 are assigned respectively to talc, gypsum, calcite, fluorite, apatite, feldspar, quartz, topaz, sapphire and diamond. With these numbers we can write 5 > 2 or 6 < 9 as apatite is harder than gypsum and feldspar is softer than sapphire, but we cannot write for example 1O - 9 = 5 - 4, because the difference in hardness between diamond and sapphire is actually much greater than that between apatite and fluorite. It would also be meaningless to say that topaz is twice as hard as fluorite simply because their respective hardness numbers on Mobhs' scale are 8 and 4. The greater than symbol (i.e., » in connection with ordinal data may be used to designate "happier than" "preferred to" and so on.

When in addition to setting up inequalities we can also form differences, we refer to the data as *interval data*. Suppose we are given the following temperature readings (in degrees Fahrenheit): 58°,63°,70°,95°, 110°, 126° and 135°. In this case, we can write 100° > 70° or95° < 135° which simply means that 110° is warmer than 70° and that 95° is cooler than 135°. We can also write for example 95° - 70° = 135° - 110°, since equal temperature differences are equal in the sense that the same amount of heat is required to raise the temperature of an object from 70° to 95° or from 110 to 135°. On the other hand, it would not mean much if we said that 126° is twice as hot as 63°, even though 126° + 63° = 2. To show the reason, we have only to change to the centigrade scale, where the first temperature becomes 5/9 (126 - 32) = 52°, the second temperature becomes 5/9 (63-32) = 17° and the first figure is now more than three times the second. This difficulty arises from the fact that Fahrenheit and Centigrade scales both have artificial origins (zeros) i.e., the number 0 of neither scale is indicative of the absence of whatever quantity we are trying to measure.

When in addition to setting up inequalities and forming differences we can also form quotients (i.e., when we can perform all the customary operations of mathematics), we refer to such data as *ratio data*. In this sense, ratio data includes all the usual measurement (or determinations) of length height, money amounts, weight, volume, area, pressures etc.

The above stated distinction between nominal, ordinal, interval and ratio data is important for the nature of a set of data may suggest the use of particular statistical techniques'. A researcher has to be quite alert about this aspect while measuring properties of objects or of abstract concepts.

MEASUREMENT SCALES

From what has been stated above, we can write that scales of measurement can be considered in terms of their mathematical properties. The most widely used classification of measurement scales are: (a) nominal scale; (b) ordinal scale; (c) interval scale; and (d) ratio scale.

(a) Nominal scale: Nominal scale is simply a system of assigning number symbols to events in order to label them. The usual example of this is the assignment of numbers of basketball players in order to identify them. Such numbers cannot be considered to be associated with an ordered scale for their order is of no consequence; the numbers are just convenient labels for the particular class of events and as such have no quantitative value. Nominal scales provide convenient ways of keeping track of people, objects and events. One cannot do much with the numbers involved. For example, one cannot usefully average the numbers on the back of a group of football players and come up with a meaningful value. Neither can one usefully compare the numbers assigned to one group with the numbers assigned to another. The counting of members in each group is the only possible arithmetic operation when a nominal scale is employed. Accordingly, we are restricted to use mode as the measure of central tendency. There is no generally used measure of dispersion for nominal scales. Chi-square test is the most common test of statistical significance that can be utilized, and for the measures of correlation, the contingency coefficient can be worked out.

Nominal scale is the least powerful level of measurement. It indicates no order or distance relationship and has no arithmetic origin. A nominal scale simply describes difference between things by assigning them to categories. Nominal data are, thus, counted data. The scale wastes any information that we may have about varying degrees of attitude, skills, understandings, etc. In spite of all this, nominal scales are still very useful and are widely used in surveys and other *ex-post-facto* research when data are being classified by major sub-groups of the population.

(b) Ordinal scale: The lowest level of the ordered scale that is commonly used is the ordinal scale. The ordinal scale places events in order, but there is no attempt to make the intervals of the scale equal in terms of some rule. Rank orders represent ordinal scales and are frequently used in research relating to qualitative phenomena. A student's rank in his graduation class involves the use of an ordinal scale. One-has to be very careful in making statement about scores based on ordinal scales. For instance, if Ram's position in his class is 10 and Mohan's position is 40, it cannot be said that Ram's position is four times as good as that of Mohan. The statement would make no sense at all. Ordinal scales only permit the ranking of items from highest to lowest. Ordinal measures have no absolute values, and the real differences between adjacent ranks may not be equal. All that can said that one person is higher or lower on the scale than another, but more precise comparisons cannot be made.

Thus the use of an ordinal scale implies a statement of 'greater than' or 'less than' (an equality statement is also acceptable) without our being able to state how much greater or less. The real difference between ranks 1 and 2 may be more or less than difference between ranks 5 and 6. Since the numbers of this scale have only a rank meaning, the appropriate measure of central tendency is the median. A percentile or quartile measure is used for measuring dispersion. Correlations are restricted to various rank order methods. Measures of statistical significance are restricted to the non parametric methods.

(c) Interval scale: In the case of interval scale, the intervals are adjusted in terms of some rule that has been established as a basis for making the units equal. The units are equal only in so far as one accepts the assumptions on which the rule is based. Interval scales can have an arbitrary Zero, but it is not possible to determine for them what may be called an absolute zero or the unique origin. The primary limitation of the interval scale is the lack of a true zero; it does not have the capacity to measure the complete absence of a trait or characteristic. The Fahrenheit scale is an example of an interval scale and shows similarities in what one can and ca not do with it. One can say that an increase in temperature from 30° to 40° involves the same increase in temperature as an increase from 60° to 70°, but one cannot say that the temperature of 60° is twice as warm as the temperature of 30° because both numbers are dependent on the fact that the zero on the scale is set arbitrarily at the temperature of the freezing point of water. The ratio of the two temperatures, 30° and 60, means nothing because zero is an arbitrary point.

Interval scales provide more powerful measurement than ordinal scales for interval scale also incorporates the concept of equality of interval. As such more powerful statistical measures can be used with interval scales. Mean is the appropriate measure of central tendency, while standard deviation is the most widely used measure of dispersion. Product moment correlation techniques are appropriate and the generally used tests for statistical significance are the't' test and 'F' test.

(d) Ratio scale: Ratio scales have an absolute or true zero of measurement. The term 'absolute zero' is not as precise as it was once believed to e. We can conceive of an absolute zero of length and similarly we can conceive of an absolute zero of time. For example, the zero point on a centimeter scale indicates the complete absence of length or height. But an absolute zero of temperature theoretically unobtainable and it remains a concept existing only in the scientist's mind. The number of minor traffic-rule violations and the number of incorrect letters in a page of type script represent scores on ratio scales. Both these scales have absolute zeros and as such all minor traffic violations and all typing errors can be assumed to be equal in significance. With ratio scales involved one can make statements like "Jyoti's" typing performance was twice as good as that of "Reetu." The ratio involved does have significance and facilitates a kinds of comparison which is not possible in case of an interval scale.

Ratio scale represents the actual amounts of variables. Measures of physical dimensions such as weight, height, distance, etc. are examples. Generally, all statistical techniques are usable with ratio scales and all manipulations that one can carry out with real numbers can also be carried out with ratio scale values. Multiplication and division can be used with this scale but not with other scales mentioned above. Geometric and harmonic means can be used as measures of central tendency and coefficients of variation may also be calculated.

Thus, proceeding from the nominal scale (the least precise type of scale) to ratio scale (the most precise), relevant information is obtained increasingly. If the nature of the variables permits, the researcher should use the scale that provides the most precise description. Researchers in physical sciences have the advantage to describe variables in ratio scale form but the behavioural sciences are generally limited to describe variables in interval scale form, a less precise type of measurement.

Sources of Error in Measurement

Measurement should be precise and unambiguous in an ideal research study. This objective, however, is often not met with in entirety. As such the researcher must be aware about the sources of error in measurement. The following are the possible sources of error in measurement.

(a) Respondent: At times the respondent may be reluctant to express strong negative feelings or it is just possible that he may have very little knowledge but may not admit his ignorance. All this reluctance is likely to result in an interview of 'guesses.' Transient factors like fatigue, boredom, anxiety, etc. may limit the ability of the respondent to respond accurately and fully.

(b) Situation: Situational factors may also come in the way of correct measurement. Any condition which places a strain on interview can have serious effects on the interviewer-respondent rapport. For instance, if someone else is present, he can distort responses by joining in or merely by being present. If the respondent feels that anonymity is not assured, he may be reluctant to express certain feelings.

(c) Measurer: The interviewer can distort responses by rewording or reordering questions. His behaviour, style and looks may encourage or discourage certain replies from respondents. Careless mechanical processing may distort the findings. Errors may also creep in because of incorrect coding, faulty tabulation and/or statistical calculations, particularly in the data-analysis stage.

(d) Instrument: Error may arise because of the defective measuring instrument. The use of complex words, beyond the comprehension of the respondent, ambiguous meanings, poor printing, inadequate space for replies, response choice omissions, etc. are a few things that make the measuring instrument defective and may result in measurement errors. Another type of instrument deficiency is the poor sampling of the universe of items of concern.

Researcher must know that correct measurement depends on successfully meeting all of the problems listed above. He must, to the extent possible, try to eliminate, neutralize or otherwise deal with all the possible sources of error so that the final results may not be contaminated.

Tests of Sound Measurement

Sound measurement must meet the tests of validity, reliability and practicality. In fact, these are the three major considerations one should use in evaluating a measurement tool. "Validity refers to the extent to which a test measures what we actually wish to measure. Reliability has to do with the accuracy and precision of a measurement procedure ... Practicality is concerned with a wide range of factors of economy, convenience, and interpretability ..." We briefly take up the relevant details concerning these tests of sound measurement.

1. Test the Validity

Validity is the most critical criterion and indicates the degree to which an instrument measures what it is supposed to measure. Validity can also be thought of as utility. In other words, validity is the extent to which differences found with a measuring instrument reflect true differences among those being tested. But the question arises: how can one determine validity without direct confirming knowledge? The answer may be that we seek other relevant evidence that confirming the answers we have found with our measuring tool. What is relevant, evidence often depends upon the nature of the research problem and the judgement of the researcher. But one can certainly consider three types of validity in this connection: (i) Content validity; (ii) Criterion-related validity and (iii) Construct validity,

(i) ***Content validity*** is the extent to which a measuring instrument provides adequate coverage of the topic under study. If the instrument contains a representative sample of the universe, the content validity is good. Its determination is primarily judgemental and intuitive. It can also be determined by using a panel of persons who shall judge how well the measuring instrument meets the standards, but there is no numerical way to express it.

(ii) ***Criterion-related validity*** relates to our ability to predict some outcome or estimate the existence of some current condition. This form of validity reflects the success of measures used for some empirical estimating purpose. The concerned criterion must possess the following qualities:

Relevance: (A criterion is relevant if it is defined in terms we judge to be the proper measure.)

Freedom from bias: (Freedom from bias is attained when the criterion gives each subject an equal opportunity to score well.),

Reliability: (A reliable criterion is stable or reproducible.)

Availability: (The information specified by the criterion must be available.)

In fact, a Criterion-related validity is a broad term that actually refers to (i) *Predictive validity* and *(ii) Concurrent validity*. The former refers to the usefulness of a test in predicting some future performance whereas the latter refers to the usefulness of a test in closely relating to other measures of known validity. Criterion-related validity is expressed as the coefficient of correlation between test scores and some measure of future performance or between test scores and scores on another measure of known validity.

(iii) ***Construct validity*** is the most complex and abstract. A measure is said to possess construct validity to the degree that it confirms to predicted correlations with other theoretical propositions. Construct validity is the degree to which scores on a test can be accounted for by the explanatory constructs of a sound theory. For determining construct validity, we associate a set of other propositions with the results received from using our measurement instrument. If measurements on our devised scale correlate in a predicted way with these other propositions, we can conclude that there is some construct validity.

If the above stated criteria and tests are met with, we may state that our measuring instrument is valid and will result in correct measurement; otherwise we shall have to look for more information and/or resort to exercise of judgement.

2. Test of *Reliability*

The test of reliability is another important test of sound measurement. A measuring instrument is reliable if it provides consistent results. Reliable measuring instrument does contribute to validity, but a reliable instrument need not be a valid instrument. For instance, a scale that consistently overweighs objects by five kgs., is a reliable scale, but it does not give a valid measure of weight. But the other way is not true i.e., a valid instrument is always reliable. Accordingly reliability is not as valuable as validity, but it is easier to assess reliability in comparison to validity. If the quality of reliability is satisfied by an instrument, then while using it we can be confident that the transient and situational factors are not interfering.

Two aspects of reliability viz., stability and equivalence deserve special mention. The *stability aspect* is concerned with securing consistent results with repeated measurements of the same person and with the same instrument. We usually determine the degree of stability by comparing the results of repeated measurements. The *equivalence aspect* considers how much error may get introduced by different investigators or different samples of the items being studied. A good way to test for the equivalence of measurements by two investigators is to compare their observations of the same events. Reliability can be improved in the following two ways:

i. By standardising the conditions under which the measurement takes place i.e., we must ensure that external sources of variation such as boredom, fatigue, etc., are minimised to the extent possible. That will improve stability aspect.

ii. By carefully designed directions for measurement with no variation from group to group, by using trained and motivated persons to conduct the research and also by broadening the sample of items used. This will improve equivalence aspect.

3. Test of *Practicality*

The practicality characteristic of a measuring instrument can be judged in terms of economy, convenience and interpretability. From the operational point of view, the measuring instrument ought to be practical i.e., it should be economical, convenient and interpretable. *Economy* consideration suggests that some trade-off is needed between the ideal research project and that which the budget can afford. The length of measuring instrument is an important area where economic pressures are quickly felt. Although more items give greater reliability as stated earlier, but in the interest of limiting the interview or observation time, we have to take only few items for our study purpose. Similarly, data-collection methods to be used are also dependent at times upon economic factors. *Convenience* test suggests that the measuring instrument should be easy to administer. For this purpose one should give due attention to the proper layout of the measuring instrument. For instance, a questionnaire, with clear instructions (illustrated by examples), is certainly more effective and easier to complete then one which lacks these features. *Interpretability* consideration is specially important when persons other than the designers of the test are to interpret the results. The measuring instrument, in order to be interpretable, must be supplemented by (a) detailed instructions for administering the test; (b)scoring keys; (c) evidence about the reliability and (d) guides for using the test and for interpreting results.

TECHNIQUE OF DEVELOPING MEASUREMENT TOOLS

The technique of developing measurement tools involves a four-stage process, consisting of the Following:

a) Concept development;
b) Specification of concept dimensions;
c) Selection of indicators; and
d) Formation of index.

The first and foremost step is that of *concept development* which means that the researcher should arrive at an understanding of the major concepts pertaining to his study. This step of concept development is more apparent in theoretical studies than in the more pragmatic research, where the fundamental concepts are often already established.

The second step requires the researcher to specify the *dimensions of the concepts* that the developed in the first stage. This task may either be accomplished by deduction i.e., by adopting a more or less-intuitive approach or by empirical correlation of the individual dimensions with the total concept and/or the other concepts. For instance, one may think of several dimensions such as product reputation, customer treatment, corporate leadership, concern for individuals, sense of social responsibility and so forth when one is thinking .about the image of a certain company.

Once the dimensions of a concept have been specified, the researcher must *develop indicators* for measuring each concept element. Indicators are specific questions, scales, or other devices by which respondent's knowledge, opinion, expectation, etc., are measured. As there is seldom a perfect measure of a concept, the researcher should consider several alternatives for the purpose. The use of more than one indicator gives stability to the scores and it also improves their validity.

The last step is that of combining the various indicators into an index, i.e., *formation of an* *index.* When we have several dimensions of a concept or different measurements of a dimension we may need to combine them into a single index. One simple way for getting an overall index is to provide scale values to the responses and then sum up the corresponding scores. Such an overall index would provide a better measurement tool than a single indicator because of the fact that an "individual indicator has only a probability relation to what we really want to know." This way we must obtain an overall index for the various concepts concerning the research study.

Scaling

In research we quite often face measurement problem (since we want a valid measurement but may not obtain it), specially when the concepts to be measured are complex and abstract and we do not possess the standardised measurement tools. Alternatively, we can say that while measuring attitudes and opinions, we face the problem of their valid measurement. Similar problem may be faced by a researcher, of course in a lesser degree, while measuring physical or institutional concepts. As such we should study some procedures which may enable us to measure abstract concepts more accurately. This brings us to the study of scaling techniques.

Meaning of Scaling

Scaling describes the procedures of assigning numbers to various degrees of opinion, attitude and other concepts. This can be done in two ways viz., (i) making a judgement about some characteristic of an individual and then placing him directly on a scale that has been defined in terms of that characteristic and (ii) constructing questionnaires in such a way that the score of individual's responses assigns him a place on a scale. It may be stated here that a scale is a continuum, consisting of the highest point (in terms of some characteristic e.g., preference, favourableness, etc.) and the lowest point along with several intermediate points between these two extreme points. These scale-point positions are so related to each other that when the first point happens to be the highest point the second point indicates a higher degree in terms of a given characteristic as compared to the third point and the third point indicates a higher degree as compared to the fourth and so on. Numbers for measuring the distinctions of degree in the attitudes/opinions are, thus, assigned to individuals corresponding to their scale-positions. All this is better understood when we talk about scaling technique(s). Hence the term 'scaling' is applied to the procedures for attempting to determine quantitative measures of subjective abstract concepts. Scaling has been defined as a "procedure for the assignment of numbers (or other symbols) to a property of objects in order to impart some of the characteristics of numbers to the properties in question."

Scale Classification Bases

The number assigning procedures or the scaling procedures may be broadly classified on one or more of the following bases: (a) subject orientation; (b) response form; (c) degree of subjectivity; (d) scale properties; (e) number of dimensions and (f) scale construction techniques. We take up each of these separately.

(a) Subject orientation: Under it a scale may be designed to measure characteristics of the respondent who completes it or to judge the stimulus object which is presented to the respondent. In respect of the former, we presume that the stimuli presented are sufficiently homogeneous so that the between- stimuli variation is small as compared to the variation among respondents. In the latter approach, we ask the respondent to judge some specific object in terms of one or more dimensions and we presume that the between-respondent variation will be small as compared to the variation among the different stimuli presented to respondents for judging.

(b) Response form: Under this we may classify the scales as categorical and comparative. Categorical scales are also known as rating scales. These. scales are used when a respondent scores Some object without direct reference to other objects. Under comparative scales, which are also known as ranking scales, the respondent is asked to compare two or more objects. In this sense the respondent may state that one object is superior to the other or that three models of pen rank in order 1,2 and 3. The essence of ranking is, in fact, a relative comparison of a certain property of two or more objects.

(c) Degree of subjectivity: With this basis the scale data may be based on whether we measure subjective personal preferences or simply make non-preference judgements. In the former case, the respondent is asked to choose which person he favours or which solution he would like to see employed, whereas in the latter case he is simply asked to judge which person is more effective in some aspect or which solution will take fewer resources without reflecting any personal preference.

(d) Scale properties: Considering scale properties, one may classify the scales as nominal, ordinal, interval and ratio scales. Nominal scales merely classify without indicating order, distance or unique origin. Ordinal scales indicate magnitude relationships of 'more than' or 'less than', but indicate no distance or unique origin. Interval scales have both order and distance values, but no unique origin ratio scales possess all these features.

(e) Number of dimensions: In respect of this basis, scales can be classified as 'unidimensional' and 'multidimensional' scales. Under the former we measure only one attribute of the respondent or object, whereas multidimensional scaling recognizes that an object might be described better by using the concept of an attribute space of 'n' dimensions, rather than a single-dimension continuum.

(f) Scale construction techniques: Following are the five main techniques by which scale be developed.

i. *Arbitrary approach:* It is an approach where scale is developed on *ad hoc* basis. This is the most widely used approach. It is presumed that such scales measure the concepts for which they have been designed, although there is little evidence to support such an assumption

ii. *Consensus approach:* Here a panel of judges evaluate the items chosen for inclusion in the instrument in terms of whether they are relevant to the topic area and unambiguous implication.

iii. *Item analysis approach:* Under it a number of individual items are developed into testswhich is given to a group of respondents. After administering the test, the total scores calculated for every one. Individual items are then analysed to determine which items discriminate between persons or objects with high total scores and those with low scores

iv. *Cumulative scales* are chosen on the basis of their conforming to some ranking of items with ascending and descending discriminating power. For instance, in such a scale the endorsement of an item representing an extreme position should also result in the endorsement of all items indicating a less extreme position.

v. *Factor scales* may be constructed on the basis of intercorrelations of items which indicate that a common factor accounts for the relationship between items. This relationship is typically measured through factor analysis method.

Important Scaling Techniques

We now take up some of the important scaling techniques often used in the context of research specially in context of social or business research.

Rating scales: The rating scale involves qualitative description of a limited number of aspects of a thing or of traits of a person. When we use rating scales (or categorical scales), we judge an object in absolute terms against some specified criteria i.e., we judge properties of objects without reference to other similar objects. These ratings may be in such forms as "like-dislike", "above average, average below average", or other classifications with more categories such as "like very much-like somewhat-neutral-dislike somewhat-dislike very much"; "excellent-good-average- below average-poor", "always-often-occasionally-rarely-never", and so on. There is no specific rule whether to use a two-points scale, three-points scale or scale with still more points. In practice three to seven points scales are generally used for the simple reason that more points on a scale provide an opportunity for greater sensitivity of measurement.

Rating scale may be either a graphic rating scale or an itemized rating scale.

i. *The graphic rating scale* is quite simple and is commonly used in practice. Under it the various points are usually put along the line to form a continuum and the rater indicates his rating by simply making a mark (such as "r) at the appropriate point on a line that runs from one extreme to the other. Scale-points with brief descriptions may be indicated along the line, their function being to assist the rater in performing his job. The following is an example of five-points graphic rating scale when we wish to ascertain people's liking or disliking any product:

This type of scale has several limitations. The respondents may check at almost any position along the line which fact may increase the difficulty of analysis. The meanings of the terms like "very much" and "some what" may depend upon respondent's frame of reference so much so that the statement might be challenged in terms of its equivalency. Several other rating scale variants (e.g., boxes replacing line) may also be used.

ii. The *itemized rating scale* (also known as numerical scale) presents a series of statements from which a respondent selects one as best reflecting his evaluation. These statements are ordered progressively in terms of more or less of some property. An example of itemized scale can be given to illustrate it.

Suppose we wish to inquire as to how well does a worker get along with his fellow workers? In such a situation we may ask the respondent to select one, to express his opinion, from the following:

- He is almost always involved in some friction with a fellow worker.
- He is often at odds with one or more of his fellow workers.
- He sometimes gets involved in friction.
- He infrequently becomes involved in friction with others.
- He almost never gets involved in friction with fellow workers.

The chief merit of this type of scale is that it provides more information and meaning to the rater, and thereby increases reliability. This form is relatively difficult to develop and the statements may not exactly what the respondent would like to express.

Rating scales have certain good points. The results obtained from their use compare favourably with alternative methods. They require less time, are interesting to use and have a wide range of applications. Besides, they may also be used with a large number of properties or variables. But their Value for measurement purposes depends upon the assumption that the respondents can and do make good judgements. If the respondents are not very careful while rating, errors may occur. Three types of errors are common viz., the error of leniency, the error of central tendency and the error of hallo effect. The error of leniency occurs when certain respondents are either easy raters or hard raters. When reluctant to give extreme judgements, the result is the error of central tendency. The error of hallo effect or the systematic bias occurs when the rater carries over a generalized impression of the subject from one rating to another. This sort of error takes place when we conclude for example, that a particular report is good because we like its form or that someone is intelligent because he agrees with us or has a pleasing personality. In other words, hallo effect is likely to appear when the rater is asked to rate many factors, on a number of which he has no judgement.

Ranking scales: Under ranking scales (or comparative scales) we make relative judgements against other similar objects. The respondents under this method directly compare two or more objects and make choices among them. There are two generally used approaches of ranking scales viz.

(a) Method of paired comparisons: Under it the respondent can express his attitude by making a choice between two objects, say between a new flavour of soft drink and an established brand of drink. But when there are more than two stimuli to judge, the number of judgements required in a paired comparison is given by the formula:

$$N = \frac{n(n-1)}{2}$$

Where N = number of judgements

n = number of stimuli or objects to be judged.

For instance, if there are ten suggestions for bargaining proposals available to a workers union, there are 45 paired comparisons that can be made with them. When N happens to be a big figure, there is the risk of respondents giving ill considered answers or they may even refuse to answer. We can reduce the number of comparisons per respondent either by presenting to each one of them only a sample of stimuli or by choosing a few objects which cover the range of attractiveness at about equal intervals and then comparing all other stimuli to these few standard objects. Thus, paired-comparison data may be treated in several ways. If there is substantial consistency, we will find that if X is preferred to Y, and Y to Z, then X will consistently be preferred to z. If this is true, we may take the total number of preferences among the comparisons as the score for that stimulus,

It should be remembered that paired comparison provides ordinal data, but the same may be converted into an interval scale by the method of the *Law of Comparative Judgement* developed by L.L. Thurstone. This technique involves the conversion of frequencies of preferences into a table of proportions which are then transformed into Z matrix by referring to the table of area under the normal curve. J.P. Guilford in his book "Psychometric Methods" has given a procedure which relatively easier. The method is known as the *Composite Standard Method* and can be illustrated under:

Suppose there are four proposals which some union bargaining committee is considering. The committee wants to know how the union membership ranks these proposals. For this purpose a sample of 100 members might express the views as shown in the following table:

Table 5.1: Response Patterns of 100 Members' Paired Comparisons of 4 Suggestions for Union Bargaining Proposal Priorities

	Suggestion			
	A	B	C	D
A	—	65	32	20
B	40	—	38	42
C	45	50	—	70
D	50	20	98	—
TOTAL	165	135	168	132

Rank order	2	3	1	4
M_p	0.5375	0.4625	0.5450	0.455
Z_j	0.09	(-).09	0.11	(-).11
R_j	0.20	0.02	0.22	0.00

Comparing the total number of preferences for each of the four proposals, we find that C is the most popular, followed by *A*,*B* and *D* respectively in popularity. The rank order shown in the above table explains all this.

By following the composite standard method, we can develop an interval scale from the paired- comparison ordinal data given in the above table for which purpose we have to adopt the following steps in order:

i). Using the data in the above table, we work out the column mean with the help of the formula given below:

$$M_p = \frac{C + .5(N)}{nN} = \frac{165 + .5(100)}{4(100)} = .5375$$

where

M_p = the mean proportion of the columns
C = the total number of choices for a given suggestion
n = number of stimuli (proposals in the given problem)
N = number of items in the sample.
ii). The column means have been shown in the M_p row in the above table.
The Z values for the M_p are secured from the table giving the area under the normal curve. When the M_p value is less than .5, the Z value is negative and for all M_p values higher than .5, the Z values are positive.' These Z values are shown in Z_j row in the above table.

iii) As the Z_j values represent an interval scale, zero is an arbitrary value. Hence we can eliminate negative scale values by giving the value of zero to the lowest scale value (this being (-).11 in our example which we shall take equal to zero) and then adding the absolute value of this lowest scale value to all other scale items. This scale has been shown in R_j row in the above table.

Graphically we can show this interval scale that we have derived from the paired-comparison data using the composite standard method as follows:

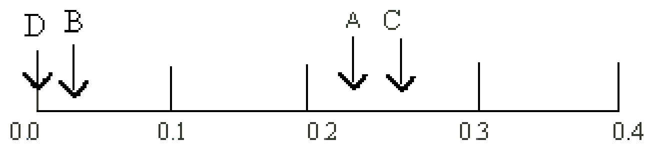


Fig. 5.2

(b) **Method of rank order:** Under this method of comparative scaling, the respondents are asked to rank their choices. This method is easier and faster than the method of paired comparisons stated above. For example, with 10 items it takes 45 pair comparisons to complete the task, whereas the method of rank order simply requires ranking of 10 items only. The problem of transitivity (such as A prefers to *B*, *B* to C, but C prefers to *A)* is also not there in case we adopt method of rank order Moreover, a complete ranking at times is not needed in which case the respondents may be asked to rank only their first, say, four choices while the number of overall items involved may be more than four, say, it may be 15 or 20 or more. To secure a simple ranking of all items involved we simply total rank values received by each item. There are methods through which we can as well develop an interval scale of these data. But then there are limitations of this method. The first one is that data obtained through this method are ordinal data and hence rank ordering is an ordinal scale with all its limitations. Then there may be the problem of respondents becoming careless in assigning ranks particularly when there are many (usually more than 10) items.

Scale Construction Techniques

In social science studies, while measuring attitudes of the people we generally follow the technique of preparing the opinionnaire (or attitude scale) in such a way that the score of the individual responses assigns him a place on a scale. Under this approach, the respondent expresses his agreement or disagreement with a number of statements relevant to the issue. While developing such statements, the researcher must note the following two points:

(i) That the statements must elicit responses which are psychologically related to the attitude being measured;

(ii) That the statements need be such that they discriminate not merely between extremes attitude but also among individuals who differ slightly.

Researchers must as well be aware that inferring attitude from what has been recorded in opinionnaires has several limitations. People may conceal their attitudes and express socially acceptable opinions: They may not really know how they feel about a social issue. People may be unaware of their attitude about an abstract situation; until confronted with a real situation, they may be unable to predict their reaction. Even behaviour itself is at times not a true indication of attitude. For instance when politicians kiss babies, their behaviour may not be a true expression of affection toward infants. Thus, there is no sure method of measuring attitude; we only try to measure the expressed opinion and then draw inferences from it about people's real feelings or attitudes.

With all these limitations in mind, psychologists and sociologists have developed several scale construction techniques for the purpose. The researcher should know these techniques so as to develop an appropriate scale for his own study. Some of the important approaches, along with the corresponding scales developed under each approach to measure attitude are as follows:

Table 5.2: Different Scales for Measuring Attitudes of People

Name of the scale construction approach	*Name of the scale developed*
1. Arbitrary approach	Arbitrary scales
2 Consensus scale approach	Differential scales (such as Thurstone Differential scale)
3. Item analysis approach	Summated scales (such as Likert Scale)
4. Cumulative scale approach	Cumulative scales (such as Guttman's Scalogram)
5. Factor analysis approach	Factor scales (such as Osgood's Semantic Differential, Multi-dimensional Scaling, etc.)

A brief description of each of the above listed scales will be helpful.

Arbitrary Scales

Arbitrary scales are developed on *ad hoc* basis and are designed largely through the researcher's own Subjective selection of items. The researcher first collects few statements or items which he believes are unambiguous and appropriate to a given topic. Some of these are selected for inclusion in the measuring instrument and then people are asked to check in a list the statements with which they agree.

The chief merit of such scales is that they can be developed very easily, quickly and with relatively less expense. They can also be designed to be highly specific and adequate. Because of these benefits, such scales are widely used in practice.

At the same time there are some limitations of these scales. The most important one is that we do not have objective evidence that such scales measure the concepts for which they have been developed. We have simply to rely on researcher's insight and competence.

Differential Scales (or Thurstone-type Scales)

The name of L.L. Thurstone is associated with differential scales which have been developed using consensus scale approach. Under such an approach the selection of items is made by a panel of judges who evaluate the items in terms of whether they are relevant to the topic area and unambiguous in implication. The detailed procedure is as under:

- The researcher gathers a large number of statements, usually twenty or more, that express various points of view toward a group, institution, idea, or practice (i.e., statements belonging to the topic area).

- These statements are then submitted to a panel of judges, each of whom arranges them in eleven groups or piles ranging from one extreme to another in position. Each of the judges is requested to place generally in the first pile the statements which he thinks are most unfavourable to the issue, in the second pile to place those statements which he thinks are next most unfavourable and he goes on doing so in this manner till in the eleventh pile he puts the statements which he considers to be the most favourable.

- This sorting by each judge yields a composite position for each of the items. In case of marked disagreement between the judges in assigning a position to an item, that item is discarded.

- For items that are retained, each is given its median scale value between one and eleven as established by the panel. In other words, the scale value of anyone statement is computed as the 'median' position to which it is assigned by the group of judges.

- A final selection of statements is then made. For this purpose a sample of statements, whose median scores are spread evenly from one extreme to the other is taken. The statements so selected, constitute the final scale to be administered to respondents. The position of each statement on the scale is the same as determined by the judges.

After developing the scale as stated above, the respondents are asked during the administration of the scale to check the statements with which they agree. The median value of the statements that they check is worked out and this establishes their score or quantifies their opinion. It may be noted that in the actual instrument the statements are arranged in random order of scale value. If the values are valid and if the opinionnaire deals with only one attitude dimension, the typical respondent will choose one or several contiguous items (in terms of scale values) to reflect his views. However, at times divergence may occur when a statement appears to tap a different attitude dimension.

The Thurstone method has been widely used for developing differential scales which are utilised to measure attitudes towards varied issues like war, religion, etc. Such scales are considered most appropriate and reliable when used for measuring a single attitude. But an important deterrent to their use is the cost and effort required to develop them. Another weakness of such scales is that the values assigned to various statements by the judges may reflect their own attitudes. The method is not completely objective; it involves ultimately subjective decision process. Critics of this method also opine that some other scale designs give more information about the respondent's attitude in comparison to differential scales.

Summated Scales (or** Likert-type **Scales)

Summated scales (or Likert-type scales) are developed by utilizing the item analysis approach wherein a particular item is evaluated on the basis of how well it discriminates between those persons whose total score is high and those whose score is low. Those items or statements that best meet this sort of discrimination test are included in the final instrument.

Thus, summated scales consist of a number of statements which express either a favourable or unfavourable attitude towards the given object to which the respondent is asked to react. The respondent indicates his agreement or disagreement with each statement in the instrument. Each response is given a numerical score, indicating its favourableness or unfavourableness, and the scores are totaled to measure the respondent's attitude. In other words, the overall score represents the respondent position on the continuum of favourable-unfavourableness towards an issue.

Most frequently used summated scales in the study of social attitudes follow the pattern devised by Likert. For this reason they are often referred to as Likert-type scales. In a Likert scale, the respondent is asked to respond to each of the statements in terms of several degrees, usually five degrees (but at times 3 or 7 may also be used) of agreement or disagreement. For example, when asked to express opinion whether one considers his job quite pleasant, the respondent may respond anyone of the following ways: (i) strongly agree, (ii) agree, (iii) undecided, (iv) disagree, (v) strongly disagree.

We find that these five points constitute the scale. At one extreme of the scale there is strong agreement with the given statement and at the other, strong disagreement, and between them lie intermediate points. We may illustrate this as under:

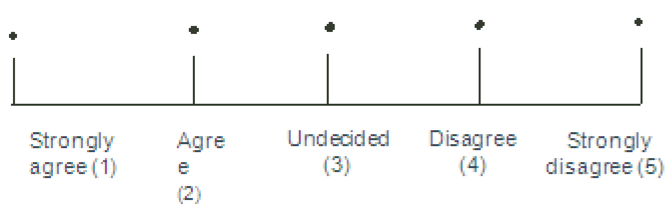


Fig. 5.3

Each point on the scale carries a score. Response indicating the least favourable degree of job satisfaction is given the least score (say 1) and the most favourable is given the highest score say 5). These score-values are normally not printed on the instrument but are shown here just to indicate the scoring pattern. The Likert scaling technique, thus, assigns a scale value to each of the five responses. The same thing is done in respect of each and every statement in the instrument. This way the instrument yields a total score for each respondent, which would then measure the respondent's favourableness toward the given point of view. If the instrument consists of, say 30 statements, the following score values would be revealing.

30 x 5 = 150 Most favourable response possible
30 x 3 = 90 A neutral attitude
30 x 1 = 30 Most unfavourable attitude.

The scores for any individual would fall between 30 and 150. If the score happens to be above 90, it shows favourable opinion to the given point of view, a score of below 90 would mean unfavourable opinion and a score of exactly 90 would be suggestive of a neutral attitude.

Procedure: The procedure for developing a Likert-type scale is as follows:

- As a first step, the researcher collects a large number of statements which are relevant to he attitude being studied and each of the statements expresses definite favourableness or unfavourableness to a particular point of view or the attitude and that the number of favourable and unfavourable statements is approximately equal.
- After the statements have been gathered, a trial test should be administered to a number of subjects. In other words, a small group of people, from those who are going to be studied finally, are asked to indicate their response to each statement by checking one of the categories of agreement or disagreement using a five point scale as stated above.
- The response to various statements are scored in such a way that a response indicative of the most favourable attitude is given the highest score of 5 and that with the most unfavourable attitude is given the lowest score, say, of 1.
- Then the total score of each respondent is obtained by adding his scores that he received for separate statements.
- The next step is to array these total scores and find out those statements which have a high discriminatory power. For this purpose, the researcher may select some part of the highest and the lowest total scores, say the top 25 per cent and the bottom 25 per cent. These two extreme groups are interpreted to represent the most favourable and the least favourable attitudes and are used as criterion groups by which to evaluate individual statements. This way we determine which statements consistently correlate with low favourability and which with high favourability.
- Only those statements that correlate with the total test should be retained in the final instrument and all others must be discarded from it.

Advantages: The Likert-type scale has several advantages. Mention may be made of the import ones.

a) It is relatively easy to construct the Likert-type scale in comparison to Thurstone-type
scale because Likert-type scale can be performed without a panel of judges.
b) Likert-type scale is considered more reliable because under it respondents answer ea
statement included in the instrument. As such it also provides more information and d,
than does the Thurstone-type scale.
c) Each statement, included in the Likert-type scale, is given an empirical test for discriminating
ability and as such, unlike Thurstone-type scale, the Likert-type scale permits the use of
statements that are not manifestly related (to have a direct relationship) to the attitude
being studied.
d) Likert-type scale can easily be used in respondent-centred and stimulus-centred studies
i.e., through it we can study how responses differ between people and how responses
differ between stimuli.
e) Likert-type scale takes much less time to construct, it is frequently used by the students of
opinion research. Moreover, it has been reported in various research studies' that there is
high degree of correlation between Likert-type scale and Thurstone-type scale.

Limitations: There are several limitations of the Likert -type scale as well. One important limitations is that, with this scale, we can simply examine whether respondents are more or less favourable to atopic, but we cannot tell how much more or less they are. There is no basis for belief that the five positions indicated on the scale are equally spaced. The interval between 'strongly agree' and 'agree', may not be equal to the interval between "agree" and "undecided". This means that Likert scale does not rise to a stature more than that of an ordinal scale, whereas the designers of Thurstone scale claim the Thurstone scale to be an interval scale. One further disadvantage is that often the total score of an individual respondent has little clear meaning since a given total score can be secured by a variety of answer patterns. It is unlikely that the respondent can validly react to a short statement on a printed form in the absence of real-life qualifying situations. Moreover, there "remains a possibility that people may answer according to what they think they should feel rather than how they do feel.:" This particular weakness of the Likert-type scale is met by using a cumulative scale which we shall take up later in this chapter.

In spite of all the limitations, the Likert-type summated scales are regarded as the most useful in a situation wherein it is possible to compare the respondent's score with a distribution of scores from some well defined group. They are equally useful when we are concerned with a program of Change or improvement in which case we can use the scales to measure attitudes before and after the programme of change or improvement in order to assess whether our efforts have had the desired effects. We can as well correlate scores on the scale to other measures without any concern for the absolute value of what is favourable and what is unfavourable. All this accounts for the popularity of Likert -type scales in social studies relating to measuring of

Cumulative scales: Cumulative scales or Louis Guttman's Scalogram analysis, like other scales, consist of series of statements to which a respondent expresses his agreement or disagreement. The special feature of this type of scale is that statements in it form a cumulative series. This, in other words, means that the statements are related to one another in such a way that an individual, who replies favourably to say item No.3, also replies favourably to items No.2 and 1, and one who replies favourably to item No.4 also replies favourably to items No.3, 2 and 1, and so on. This being so an individual whose attitude is at a certain point in a cumulative scale will answer favourably all the items on one side of this point, and answer unfavourably all the items on the other side of this point. The individual's score is worked out by counting the number of points concerning the number of statements he answers favourably. If one knows this total score, one can estimate as to how a respondent has answered individual statements constituting cumulative scales. The major scale of this type of cumulative scales is the Guttman's scalogram. We attempt a brief description of the same below

The technique developed by Louis Guttman is known as Scalogram analysis, or at times simply 'scale analysis'. Scalogram analysis refers to the procedure for determining whether a set of items forms a unidimensional scale. A scale is said to be unidimensional if the responses fall into a pattern in which endorsement of the item reflecting the extreme position results also in endorsing all items which are less extreme. Under this technique, the respondents are asked to indicate in respect of each item whether they agree or disagree with it, and if these items form a unidimensional scale, the response pattern will be as under:

Table 5.3: Response pattern in Scalogram analysis

Item no.				**Respondent score**
4	**3**	**2**	**1**	
X	**X**	**X**	**X**	**4**
-	**X**	**X**	**X**	**3**
-	**-**	**X**	**X**	**2**
-	**-**	**-**	**X**	**1**
-	**-**	**-**	**-**	**0**

X = Agree
- = Disagree

A score of 4 means that the respondent is in agreement with all the statements which is indicative of the most favourable attitude. But a score of 3 would mean that the respondent is not agreeable to item 4 but he agrees with all others. In the same way one can interpret other values of the respondents' scores. This pattern reveals that the universe of content is scalable.

Procedure: The procedure for developing a scalogram can be outlined as under:

a) The universe of content must be defined first of all. In other words, we must lay down in clear terms the issue we want to deal within our study.

b) The next step is to develop a number of items relating the issue and to eliminate by inspection the items that are ambiguous, irrelevant or those that happen to be too extreme items.

c) The third step consists in pre-testing the items to determine whether the issue at hand is scalable (The pretest, as suggested by Guttman, should include 12 or more items, while the final scale may have only 4 to 6 items. Similarly, the number of respondents in a pretest may be small, say 20 or 25 but final scale should involve relatively more respondents, say 1OO or more).

In a pretest the respondents are asked to record their opinions on all selected items using
a Likert-type 5-point scale, ranging from 'strongly agree' to 'strongly disagree'. The strongest
favourable response is scored as 5, whereas the strongest unfavourable response as 1. The total score can thus range, if there are 15 items in all, from 75 for most favourable to 15 for
the least favourable.

Respondent opinionnaires are then arrayed according to total score for analysis and
evaluation. If the responses of an item form a cumulative scale, its response category scores should decrease in an orderly fashion as indicated in the above table. Failure to show the said decreasing pattern means that there is overlapping which shows that the item concerned is not a good cumulative scale item i.e., the item has more than one meaning.
Sometimes the overlapping in category responses can be reduced by combining categories.
After analysing the pretest results, a few items, say 5 items, may be chosen.

d) The next step is again to total the scores for the various opinionnaires, arid to rearray them to reflect any shift in order, resulting from reducing the items, say, from 15 in pretest to, say 5 for the final scale. The final pretest results may be tabulated in the form of a table given in Table 5.4.

Table 5.4: The Final Pretest Results in a Scalogram Analysis

Scale type	Item 5	Item 12	Item 3	Item 10	Item 7	Errors per case	Number of cases	Number of errors
5 (perfect)	X	X	X	X	X	0	7	0
4 (perfect)	–	X	X	X	X	0	3	0
(nonseale)	–	X	–	X	X	1	1	1
(nonseale)	–	X	X	–	X	1	2	2
3 (perfect)	–	–	X	X	X	0	5	0
2 (perfect)	–	–	–	X	X	0	2	0
1 (perfect)	–	–	–	–	X	0	1	0
(nonseale)	–	–	X	–	–	2	1	2
(nonseale)	–	–	X	–	–	2	1	2
0 (perfect)	–	–	–	–	–	0	2	0
			n=5				*N*=25	*e*=7

The table shows that five items (numbering 5, 12,3, 10 and 7) have been selected for the final scale. The number of respondents is 25 whose responses to various items have been tabulated along with the number of errors. Perfect scale types are those in which the respondent's answers fit the pattern that would be reproduced by using the person's total score as a guide. *Non-scale types* are those in which the category pattern differs from that expected from the respondent's total score i.e., non-scale cases have deviations from unidimensionality or errors. Whether the items (or series of statements) selected for final scale may be regarded a perfect cumulative (or a unidimensional scale), we have to examine on the basis of the coefficient of reproducibility. Guttman has set 0.9 asthe level of minimum reproducibility in order to say that the scale meets the test of unidimensionality. He has given the following formula for measuring the level of reproducibility:

Guttman's Coefficient of Reproducibility = 1 - $e/n(N)$

where e = number of errors

n = number of items

N = number of cases

For the above table figures,

Coefficient of Reproducibility = 1 - 7/5(25) = .94

This shows that items number 5, 12,3, 10 and 7 iI1 this order constitute the cumulative or unidimensional scale, and with this we can reproduce the responses to each item, knowing only the total score of the respondent concerned. Scalogram, analysis, like any other scaling technique, has several advantages as well as limitations. One advantage is that it assures that only a single dimension of attitude is being measured. Researcher's subjective judgement is not allowed to creep in the development of scale since the scale is determined by the replies of respondents. Then, we require only a small number of items that make such a scale easy to administer. Scalogram analysis can appropriately be used for personal, telephone or mail surveys. The main difficulty in using this scaling technique is that in practice perfect cumulative or unidimensional scales are very rarely found and we have only to use its approximation testing it through coefficient of reproducibility or examining it on the basis of some other criteria. This method is not a frequently used method for the simple reason that its development procedure is tedious and complex. Such scales hardly constitute a reliable basis for assessing attitudes of persons towards complex objects for predicting the behavioural responses of individuals towards such objects. Conceptually, this analysis is a bit more difficult in comparison to other scaling methods.

Factor scales

Factor scales are developed through factor analysis or on the basis of intercorrelations of items which indicate that a common factor accounts for the relationships between items. Factor scales are particularly "useful in uncovering latent attitude dimensions and approach scaling through the concept of multiple-dimension attribute space." More specifically the two problems viz., how to deal appropriately with the universe of content which is multi-dimensional and how to uncover underlying (latent) dimensions which have not been identified, are dealt with through factor scales. An important factor scale based on factor analysis is Semantic Differential (S.D.) and the other one is Multidimensional Scaling. We give below a brief account of these factor scales.

Semantic differential scale: Semantic differential scale or the S.D. scale developed by Charles E. Osgood, G.J. Suci and P.H. Tannenbaum (1957), is an attempt to measure the psychological meanings of an object to an individual. This scale is based on the presumption that an object can have different dimensions of connotative meanings which can be located in multidimensional property space, or what can be called the semantic space in the context of S.D. scale. This scaling consists of a set of bipolar rating scales, usually of 7 points, by which one or more respondents rate one or more concepts on each scale item. For instance, the S.D. scale items for analysing candidates for leadership position may be shown as under:

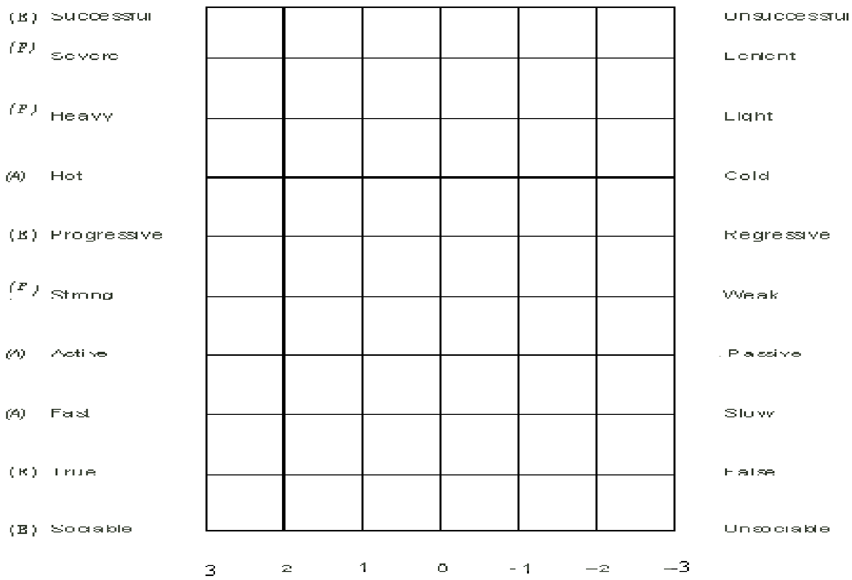


Fig. 5.1

Candidates for leadership position (along with the concept-the 'ideal' candidate) may be compared and we may score them from +3 to -3 on the basis of the above stated scales. (The letters, *E*, *P*, *A* showing the relevant factor viz., evaluation, potency and activity respectively, written along the left side are not written in actual scale. Similarly the numeric values shown are also not written in actual scale.)

Osgood and others did produce a list of some adjective pairs for attitude research purpose and concluded that semantic space is multidimensional rather than unidimensional. They made sincere efforts and ultimately found that three factors, viz., evaluation, potency and activity, contributed most to meaningful judgements by respondents. The evaluation dimension generally accounts for 1/2 and 3/4 of the extractable variance and the other two factors account for the balance.

Procedure: Various steps involved in developing S.D. scale are as follows:

a) First of all the concepts to be studied are selected. The concepts are usually chosen
personal judgement, keeping in view the nature of the problem.

b) The next step is to select the scales bearing in mind the criterion of factor composition and the criterion of scale's relevance to the concepts being judged (it is common practice to use at least three scales for each factor with the help of which an average factor score has to be worked out). One more criterion to be kept in view is that scales should be stable across subjects and concepts.

c) Then a panel of judges are used to rate the various stimuli (or objects) on the various selected scales and the responses of all judges would then be combined to determine the composite scaling.

To conclude, "the S.D. has a number of specific advantages. It is an efficient and easy way to secure attitudes from a large sample. These attitudes may be measured in both direction and intensity. The total set of responses provides a comprehensive picture of the meaning of an object, as well as a measure of the subject doing the rating. It is a standardized technique that is easily repeated, but escapes many of the problems of response distortion found with more direct methods."

Multidimensional scaling: Multidimensional scaling (MDS) is relatively more complicated scaling device, but with this sort of scaling one can scale objects, individuals or both with a minimum of information. Multidimensional scaling (or MDS) can be characterized as a set of procedures for portraying perceptual or affective dimensions of substantive interest. It "provides useful methodology for portraying subjective judgements of diverse kinds." MDS is used when all the variables (whether metric or non-metric) in a study are to be analyzed simultaneously and all such variables happen to be independent. The underlying assumption in MDS is that people (respondents) "perceive a set of objects as being more or less similar to one another on a number of dimensions (usually uncorrelated with one another) instead of only one." Through MDS techniques one can represent geometrically the locations and interrelationships among a set of points. In fact, these techniques attempt to locate the points, given the information about a set of interpoint distances, in space of one or more dimensions such as to best summarise the information contained in the interpoint distances. The distances in the solution space then optimally reflect the distances contained in the input data. For instance, if objects, say X and *Y*, are thought of by the respondent as being most similar as compared to all other possible pairs of objects, MDS techniques will position objects X and *Y* in such a way that the distance between them in multidimensional space is shorter than that between any two other objects.

Two approaches, viz., the metric approach and the non-metric approach, are usually talked about in the context of MDS, while attempting to construct a space containing *m* points such that)12 m(m-1)/2 interpoint distances reflect the input data. The *metric approach to MDS* treats the input data as scale data and solves applying statistical me~hods for the additive constant which

minimises the dimensionality of the solution space. This approach utilises all the information in the data in obtaining a solution. The data (i.e., the metric similarities of the objects) are often obtained data a bipolar similarity scale on which pairs of objects are rated one at a time. If the data reflect exact distances between real objects in an r-dimensional space, their solution will reproduce the Set of interpoint distances. But as the true and real data are rarely available, we require random and systematic procedures for obtaining a solution. Generally, the judged similarities among a Set of objects are statistically transformed into distances by placing those objects in a multidimensional space of some dimensionality.

The *non-metric approach* first gathers the non-metric similarities by asking respondents to rank order all possible pairs that can be obtained from a set of objects. Such non-metric data is then transformed into some arbitrary metric space and then the solution is obtained by reducing the dimensionality. In other words, this non-metric approach seeks "a representation of points in a space of minimum dimensionality such that the rank order of the interpoint distances in the solution space maximally corresponds to that of the data. This is achieved by requiring only that the distances in the solution be monotone with the input data."? The non-metric approach has come into prominence during the sixties with the corning into existence of high speed computers to generate metric solution for ordinal input data.

The significance of MDS lies in the fact that it enables the researcher to study "the perceptual structure of a set of stimuli and the cognitive processes underlying the development of this structure Psychologists, for example, employ multidimensional scaling techniques in an effort to scale psychophysical stimuli and to determine appropriate labels for the dimensions along which these stimuli vary."!" The MDS techniques, infact, do away with the need in the data collection process specify the attribute(s) along which the several brands, say of a particular product, may be compared as ultimately the MDS analysis itself reveals such attribute(s) that presumably underlie the expressed relative similarities among objects. Thus, MDS is an important tool in attitude measurement and the techniques falling under MDS promise "a great advance from a series of unidimensional measurements (e.g., a distribution of intensities of feeling towards single attribute such as colour, taste or a preference ranking with indeterminate intervals), to a perceptual mapping in multidimensional space of objects company images, advertisement brands, etc."

In spite of all the merits stated above, the MDS is not widely used because of the computation complications involved under it. Many of its methods are quite laborious in terms of both the collection of data and the subsequent analyses. However, some progress has been achieved (due to the pioneering efforts of Paul Green and his associates) during the last few years in the use of non-metric MDS in the context of market research problems. The techniques have been specifically applied in “finding out the perceptual dimensions, and the spacing of stimuli along these dimensions, that people, making judgements about the relative similarity of pairs of Stimuli." But, "in the long run, the of MDS will be determined by the extent to which it advances the behavioral sciences."

POSSIBLE QUESTIONS
UNIT V

PART A (20 x 1 = 20 Marks)
Question number 1 – 20 Online examination

PART B (5 X 2 = 10Marks)

1. How are the scaling procedures classified?
2. What are the important criteria of selecting a sampling procedure?
3. What are complex random sampling designs?
4. What are the types of scaling?
5. How will you proceed Thurstone type scales?
6. What are the major considerations in evaluating a measurement tool?

7. What are the different types of sample designs?

8. What is meant by scaling? Elaborate the types of scales

9. How would you differentiate between simple random sampling and complex random sampling designs?

10. What do you mean by 'Sample Design'?

PART C (5 X 6 = 30Marks)

11. How are the scaling procedures classified? Discuss.

12. What are the important criteria of selecting a sampling procedure? Explain.

13. What are complex random sampling designs? Discuss.

14. Explain in detail the scaling measurement techniques and types.

15. How will you proceed Thurstone type scales?

16. What are the major considerations in evaluating a measurement tool?

17. What are the different types of sample designs? Explain in detail.

18. What is meant by scaling? Elaborate the types of scales

19. Explain the various steps involved in sample design.

20. Explain in detail the types of scales.

21. How would you differentiate between simple random sampling and complex random sampling designs? Explain clearly giving examples.

22. What do you mean by 'Sample Design'? What points should be taken into consideration by a researcher in developing a sample design for research project.

23. Explain and illustrate the procedure of selecting a random sample.

24. Describe the different methods of scale construction, pointing out the merits and demerits of each.

25. What is the meaning of measurement in research? What difference does it make whether we measure in terms of a nominal, ordinal, interval or ration scale? Explain giving examples.

26. Point out the possible sources of error in measurement. Describe the tests of sound measurement.

27. Discuss the relative merits and demerits of:
(a) Rating vs. Ranking scales.
(b) Summated vs. Cumulative scales
(c) Scalogram analysis vs. Factor analysis.

28. Under what circumstances stratified random sampling design is considered appropriate? How would you select such sample? Explain by means of an example.

29. Distinguish between:
(a) Restricted and unrestricted sampling;
(b) Convenience and purposive sampling;
(c) Systematic and stratified sampling;
(d) Cluster and area sampling.

Karpagam Academy of Higher Education
Department of Biochemistry
Biostatistics and Research Methodology
UNIT V

UNIT V

1	5	All items in any field of inquiry constitute a --------------------	**Population**	sample	total items	infinite samples	**Population**
2	5	Population census conducted once in a -----------------	century	**decade**	fortnight	two decade	**decade**
3	5	A definite plan for obtaining a sample from a given population is known as	Research design	**sample design**	clear plan	accurate plan	**sample design**
4	5	---------------- is determined before data are collected	Research design	**sample design**	clear plan	accurate plan	**sample design**
5	5	The following is an example of finite universe	Number of stars in the sky	listeners of a specific radio programme	Throwing of a dice	**the population of a city**	**the population of a city**
6	5	The following are infinite universe except	Number of stars in the sky	listeners of a specific radio programme	Throwing of a dice	**the population of a city**	**the population of a city**
7	5	One of the following is the example of infinite universe	the population of a city	the number of workers in a factory	the number of students in a college	**listeners of a specific radio programme**	**listeners of a specific radio programme**
8	5	A --------------- results from errors in the sampling procedures	**systematic bias**	sampling error	designing error	research bias	**systematic bias**
9	5	When the size of the sample increases sampling error ----------	**decreases**	increases	no change	zero	**decreases**
10	5	To study the economic status of a town or village the sampling design used is	probability sampling	**non-probability sampling**	based on each item	based on items at random	**non-probability sampling**
11	5	Probability sampling is also known as	quota sampling	judgement sampling	**random sampling**	purposive sampling	**random sampling**
12	5	One of the following is not a non-probability sampling	deliberate sampling	purposive sampling	judgement sampling	**chance sampling**	**chance sampling**
13	5	The probability of selecting sample of size 3 from a finite population of 6 elements is	**1/20**	3/6	2/6	1/6	**1/20**
14	5	Selecting every i^{th} item on a list is known as	**systematic sampling**	area sampling	multi-stage sampling	cluster sampling	**systematic sampling**
15	5	If a 4% sample is desired, the sample selection will be	one in every 10^{th} item	**one in every 25^{th} item**	one in every 50^{th} item	one in every 4^{th} item	**one in every 25^{th} item**
16	5	If a population from which a sample is to be drawn doesnot constitute a homogenous group the technique is known as	**Stratified sampling**	area sampling	multi-stage sampling	cluster sampling	**Stratified sampling**
17	5	In stratified sampling the different sub populations are called as	sub datas	**strata**	substitutes	random samples	**strata**
18	5	The most efficient and an optimal design of complex random sampling is	**Stratified sampling**	area sampling	multi-stage sampling	cluster sampling	**Stratified sampling**
19	5	Using proportional allocation, the sample sizes for different strata are 15,9 and 6 respectively which is in proportion to the sizes of the stra	3000:4000:1000	**4000:2400:1600**	2400:1600:4000	1800:2200:4000	**4000:2400:1600**
20	5	If clusters happens to be some geographic subdivisions it is known as	Stratified sampling	**area sampling**	multi-stage sampling	cluster sampling	**area sampling**
21	5	___________ is a process of mapping aspects of a domain onto other aspects of a range according to some rule of correspondence	strata	**measurement**	sub strata	clusters	**measurement**
22	5	The data which are numerical in name only and donot share any properties of the numbers we deal in ordinary arithmetic is	**nominal data**	ordinal data	interval data	ratio data	**nominal data**
23	5	In those situations,when we can't do anything except set up inequalities, we refer to the data as	nominal data	**ordinal data**	interval data	ratio data	**ordinal data**
24	5	On Mohs' scale number 5&2 refers to apatite and gypsum.In this 5 > 2 refers to	**apatite is harder than gypsum**	apatite is softer than gypsum	gypsum is softer than apatite	both are equally hard	**apatite is harder than gypsum**
25	5	On Mohs' scale the numbers 6&9 refers to feldspar and sapphire respectively in this 6 < 9 refers to	feldspar is the most hardest	sapphire and feldspar are equally hard	**feldspar is softer than sapphire**	sapphire is softer than feldspar	**feldspar is softer than sapphire**
26	5	is simply a system of assigning number symbols to events in order to label them	**nominal scale**	ordinal scale	interval scale	ratio scale	**nominal scale**
27	5	The assignment of number of basketball players in order to identify them is the usual example of	**nominal scale**	ordinal scale	interval scale	ratio scale	**nominal scale**
28	5	In nominal scale the measure of central tendency used is	mean	median	**mode**	quartiles	**mode**
29	5	Generally used measure of dispersion for nominal scales is	standard deviation	mean deviation	quartile deviation	**no measure of dispersion**	**no measure of dispersion**
30	5	The most common test of statistical significance that can be utilized in nominal scales is	F-test	t-test	Z-test	**chi-square test**	**chi-square test**
31	5	For the measure of correlation ----------- can be worked out for nominal scales	**contingency coefficient**	scatter diagram	Karl-pearson's coefficient of correlation	graphs	**contingency coefficient**
32	5	is the least powerful level of measurement	**nominal scale**	ordinal scale	interval scale	ratio scale	**nominal scale**
33	5	The lowest level of the ordered scale that is commonly used is the -----------------	nominal scale	**ordinal scale**	interval scale	ratio scale	**ordinal scale**
34	5	A students rank in his graduation class involves the use of an -------------	nominal scale	**ordinal scale**	interval scale	ratio scale	**ordinal scale**
35	5	Multiplication and division can only be used with this scale but not with other scales.	nominal scale	ordinal scale	interval scale	**ratio scale**	**ratio scale**
36	5	Measures of central tendency used in ratio scale is	median	mode	arithmetic mean	**geometric and harmonic mean**	**geometric and harmonic mean**
37	5	Researchers in physical sciences have the advantage to describe variables in	nominal scale	ordinal scale	interval scale	**ratio scale**	**ratio scale**
38	5	Researchers in behavioural sciences have the advantage to describe variables in	nominal scale	ordinal scale	**interval scale**	ratio scale	**interval scale**
39	5	The most precise type of scale is	nominal scale	ordinal scale	interval scale	**ratio scale**	**ratio scale**
40	5	refers to the extent to which a test measures what we actually wish to measure	**Validity**	Reliability	Practicality	speed	**Validity**
41	5	has to do with the accuracy and precision of a measurement procedure	validity	**reliability**	practicality	speed	**reliability**
42	5	is concerned with a wide range of factors of economy,convenience and interpretability	validity	reliability	**practicality**	speed	**practicality**
43	5	If the instrument contains a representative sample of the universe,the ---------- is good	Criterion related validity	**content validity**	construct validity	obstruct validity	**content validity**
44	5	enables the researcher to study the perceptual structure of a set of stimuli and the cognitive processes underlying the developm	rational scaling	nominal scaling	**multi-dimentional scaling**	interval scale	**multi-dimentional scaling**
45	5	__________ describes the procedures of assigning numbers to various degrees of opinion,attitude and other concepts	**scaling**	sampling	validity	reliability	**scaling**
46	5	Categorical scales are also known as	nominal scale	ordinal scale	interval scale	**rating scale**	**rating scale**
47	5	Comparative scales are also known as	nominal scale	**ranking scale**	interval scale	ratio scale	**ranking scale**
48	5	Classification without indicating order, distance or unique origin is	**nominal scale**	ordinal scale	interval scale	ratio scale	**nominal scale**
49	5	_________ indicates magnitude relationship of 'more than' or 'less than' but indicate no distance or unique origin	nominal scale	**ordinal scale**	interval scale	ratio scale	**ordinal scale**
50	5	Which scales have both order and distance values but no unique origin	nominal scale	ordinal scale	**interval scale**	ratio scale	**interval scale**
51	5	The most widely used scale construction technique developed on ad hoc basis is	**arbitrary approach**	item analysis approach	consensus approach	factor scales	**arbitrary approach**
52	5	Qualitative description of a limited number of aspects of a thing or of traits of a person is	nominal scale	ordinal scale	interval scale	**rating scale**	**rating scale**
53	5	The scaling technique in the form of 'always-often-occasionally-rarely-never' is	nominal scale	ordinal scale	interval scale	**rating scale**	**rating scale**
54	5	Greater sensitivity of measurement is achieved in rating scale if there is	**more points on a scale**	less points on a scale	only two points	zero points	**more points on a scale**
55	5	If the respondents are either easy raters or hard raters, the type of error occurs is	error of central tendency	**error of leniency**	error of hallo effect	error of reliability	**error of leniency**
56	5	When raters are reluctant to give extreme judgments, the result is the	**error of central tendency**	error of leniency	error of hallo effect	error of reliability	**error of central tendency**
57	5	Composite standard method of paired comparision is given by	**J.P.Guilford**	E.L.Thorndike	M.D.Thurston	R.A.Fisher	**J.P.Guilford**
58	5	It is easy to construct Likert-type scale in comparison to Thurstone type because	only written examination is done	**can be performed without a panel of judges**	only oral questions are asked	It is performed with a panel of judges	**can be performed without a panel of judges**
59	5	An attempt to measure the psychological meanings of an object to an individual is	**semantic differential scale**	ordinal scale	interval scale	ratio scale	**semantic differential scale**
60	5	In which section of dissertation ,we can give our suggestions?	Introduction	Result	**Discussion**	Su mmary	**Discussion**