STATISTICAL METHODS AND THEIR APPLICATIONS–I

STATISTICAL METHODS AND THEIR APPLICATIONS–I

By

M. HAJ MEERAL
M.Sc., M.Phil., P.G.D.C.A.
Department of Mathematics,
Mohamed Sathak College of Arts and Science,
Sholinganallur, Chennai, Tamil Nadu.

M. RAVITHAMMAL
M.Sc., M.Phil.
Department of Mathematics,
Mohamed Sathak College of Arts and Science,
Sholinganallur, Chennai, Tamil Nadu.

UNIVERSITY SCIENCE PRESS
(An imprint of Laxmi Publications Pvt. Ltd.)
BANGALORE ● CHENNAI ● COCHIN ● GUWAHATI ● HYDERABAD
JALANDHAR ● KOLKATA ● LUCKNOW ● MUMBAI ● RANCHI
NEW DELHI ● BOSTON, USA

Published by

UNIVERSITY SCIENCE PRESS

(*An imprint of Laxmi Publications Pvt. Ltd.)*

113, Golden House, Daryaganj,

New Delhi-110002

Phone : 011-43 53 25 00

Fax : 011-43 53 25 28

www.laxmipublications.com

info@laxmipublications.com

Price : ₹ **175.00** *Only* *First Edition* : 2013

OFFICES

✆ **Bangalore**	080-26 75 69 30	✆ **Chennai**	044-24 34 47 26
✆ **Cochin**	0484-405 13 03, 23 77 00 44	✆ **Guwahati**	0361-251 38 81, 251 36 69
✆ **Hyderabad**	040-24 65 23 33	✆ **Jalandhar**	0181-222 12 72
✆ **Kolkata**	033-22 27 43 84	✆ **Lucknow**	0522-220 99 16
✆ **Mumbai**	022-24 91 54 15, 24 92 78 69	✆ **Ranchi**	0651-220 44 64

USM-9672-175-STAT METHOD APPL-I-MEE **C—**

Typeset at: Balaji Graphics, New Delhi. *Printed at:*

CONTENTS

Unit I : Statistics **1—45**

Introduction 1
Limitations of Statistics 1
Characteristics of Statistics 1
Scope of Statistics 2
Collection of Data 2
Secondary Data 3
Introduction for Classification 3
Presentation of Statistical Data 5
Textual Presentation 5
Tabular Presentation or Tabulation 5
Types of Tabulation 7
Values of a Variable and Frequency 8
Frequency Distribution 9
Types of Charts, Diagrams 17
Diagrams 19
Frequency Polygon 28
Frequency Curve 29
Cumulative Frequency Curve (Ogive) 32
Lorenz Curve 36
Exercise 43

Unit II : Measures of Central Tendency or Average **46—77**

Meaning 46
Definition 46
Types of Average 46
Arithmetic Mean (or) Mean 46
Geometric Mean 53
Merits and Demerits of G.M. 53
Merits and Demerits of H.M. 56
Median 58
Exercise 75

Unit III : Measures of Dispersion **78—119**

Measures of Dispersion 78
Mean Deviation (or) Average Deviation 86
Skewness & Kurtosis 104
Exercise 115

Unit IV : Probability **120—146**

Experiment 120
Theorem 1 124
Theorem 2 125
Theorem 3 125
Theorem 4 126
Theorem 5 127
Theorem 6 128
Permutations and Combinations 136
Exercise 146

Unit V : Random Variable **147—209**

Introduction for Random Variable 147
Mathematical Expectation 161
Moment Generating Function 190
Tchebchev's Inequality 201
Exercise 208

Appendix **210—214**

Question Papers **215—221**

Index **222—223**

PREFACE

This book **'Statistical Methods and their Applications-I'** provides a simple, clear and exhaustive treatment of the subject in a simple non-mathematical language.

This book is addressed primarily to the undergraduate students of Computer Science and Biochemistry. It is written strictly in accordance with the University of Madras Syllabus for B.Sc. Computer Science and B.Sc. Biochemistry.

Each chapter is discussed elaborately. Chapterisation is drawn in a logical sequence. Each chapter is followed by a list of questions and a set of exercise. Worked-out problems are provided wherever necessary. University question papers are appended to the book. Thus, the book is adapted to suit the requirements of the modern students.

Suggestions for the improvement of the book are welcome.

— Authors

SYLLABUS

STATISTICAL METHODS AND THEIR APPLICATIONS-I (III-SEMESTER)

UNIT - I

Nature and Scope of Statistical Methods and Their Limitations — Classifications, Tabulation and Diagrammatic Representation of various types of statistical data — Frequency Curves and Ogives — Graphical determination of percentiles quartiles and their properties — Merits and Demerits.

UNIT - II

Measures of Location — Arithmetic Mean, Median, Mode, Geometric Mean, Harmonic Mean and their properties — Merits and Demerits.

UNIT - III

Measures of Dispersion — Range, Mean Deviation, Quartile Deviation, Standard Deviation, Coefficient of Variation, Skewness and Kurtosis and their properties.

UNIT - IV

Probability of an event — Finitely additive probability sopace addition and multiplication theorems — Independence of events — Conditional Probability — Bayes Theorem.

UNIT - V

Concepts of Random Variable — Mathematical expectation — Moments of random variable (raw and central moments) — Moment generating function — Chebychev's inequality — Simple Problems.

1 STATISTICS

INTRODUCTION

The term statistical data means numerical statements of facts whereas, the term statistical method suggests the principles and techniques employed in collecting analysing and interpreting such data.

Statistics may be defined as

(*i*) the collection of data

(*ii*) presentation of data

(*iii*) analysis of data and

(*iv*) interpretation of numerical data.

By the efficient application of such methods, useful deductions are made and statistical relationships that governing the data are suitably formulated.

LIMITATIONS OF STATISTICS

1. **Statistics does not study individuals:** Since statistics deals with aggregates of facts, the study of individual measurement lies outside the scope of statistics.
2. **Statistics does not study qualitative phenomena:** It deals with a set of numerical data, it is applicable only to the quantitative aspect of problem. Qualitative phenomena like honesty, poverty, wisdom etc., which cannot be expressed numerically are capable of direct statistical analysis.
3. **Statistical laws can never be regarded as universal truths:** Statistical analysis talk only interms of probability and not in terms of certainty.
4. Statistics only simplifies and helps the analysis of the certain quantitative facts. But the real background of the data may not be reflected through these facts.
5. The results and conclusions obtained by statistical treatment need not be accurate.
6. **Statistics is most useful but only of great value to those who understand its proper use.** Wrong and Fallacious conclusions will be obtained if the statistical methods are used by people without training and experience. The misuse of statistics has given rise to the distruct of the subject.

CHARACTERISTICS OF STATISTICS

1. In statistics all decisions are to be expressed in quantitative units.
2. The main concern of statistics lies with an aggregate of objects and facts. An individual item has no bearing in statistical science.

3. Statistical data are collected always with a pre-determined purpose from a definite field of enquiry.
4. In everyfield of enquiry there is a large number of factors. Each one of these factors contributes to the final data collected.
5. Statistics because of its nature cannot be called an exact science.
6. Statistics should be presented in such a manner that cause and effect—link can be established.
7. The statistical enquiry passes through four stages;
   (*a*) Collection of data
   (*b*) Classification and tabulation of data
   (*c*) Analysis of data
   (*d*) Interpretation of data.

SCOPE OF STATISTICS

The scope of statistics is ever expanding that it has permeated almost every fact of our lives. It is applicable in almost every field like trade, industry, commerce, biology, state, economics, research, natural science etc., Statistics must be tailored to fit the percuralities of each concrete problem to which it is applied. Therefore it is vital part of the economic and business life of the commodity.

COLLECTION OF DATA

Collection of data is the first step in any statistical investigation. Statistical data are of two types. viz., Primary and secondary. Primary data are those which are collected for the first time and are original in character. Secondary data are those which have already been collected, tabulated and presented in some form by some one else for some purpose.

On the basis of the classification of data made above we have two methods of data collection.

1. Primary method
2. Secondary method

Primary Data

Merits and Demerits of Primary Data

Primary data have two advantages compared with secondary data. Primary data are (*a*) truthful and (*b*) purposive, However, primary data have, two disadvantages compared with secondary data. They are as follows (*i*) it takes time to collect data by the primary method (*ii*) its collection involves much expenditure of money. The investigator has to define before collecting statistical data.

1. Nature, object and scope of the enquiry
2. Availability of finance
3. Availability of time
4. Degree of accuracy desired.

Methods of Collecting Primary Data

Primary data may be obtained by adopting any one of the following methods

1. Direct personal investigation
2. Indirect oral investigation

3. Information through local correspondents
4. Mailed questionaire method and
5. Schedules sent through enumerators.

SECONDARY DATA

Secondary data are those which have already been collected, tabulated and presented in some form by some one else for some purpose.

If we take primary data to resemble 'raw materials' then we have to take secondary data to resemble 'finished products'.

Sources

The various sources of secondary data may be divided into two broad categories are noted below:

1. Published sources
2. Unpublished sources

Published data

There are three important sources of published data as noted below:

1. Official publications
2. Semi official publications
3. Private publications.

Official publications include reports and official publications of the Government, official publications of international bodies such as the U.N.O, I.F.C etc. Semi official publications include reports and publications of Municipalities, Corporations, District Boards, Life Insurance Corporation of India.

Private publications include of trade and professional bodies as the federations of Indian, Institute of Chartered Accountants, Publications of research bureaus, research scholars etc.

Unpublished data

All statistical material is not always published. There are various sources of unpublished data. These include records maintained by various Government and private offices, studies made by research bureaus, labour bureaus, trade association, chambers of commerce, research workers etc.

Before using the secondary data, one should be careful in noticing that the data are

1. Reliable
2. Suitable and
3. Adequate.

INTRODUCTION FOR CLASSIFICATION

After having collected and edited the data, the next important step is to organise it in a systematic manner. The first step in the analysis and interpretation of data is classification and tabulation. The process of arranging data in groups or classes according to resemblances and similarities is technically called classification.

Classification

Classification can be defined as the process of arranging or bringing together all the enumerated individuals or items under separate head or classes according to some common characteristics possessed by them.

When individuals or items are classified in accordance with some common non-measurable characteristics possessed by them, they are said to form a statistical class. But when the same individuals or items are classified on the basis of some common measurable characteristics possessed by them, they are said to form a statistical group.

Types of Classification

There are four types of classification

(*i*) Classification on a qualitative basis:

Classification according to some non-measurable characteristics such as nationality, religion, occupation etc., belongs to this class. Such classifications may also be termed as class.

(*ii*) Classification on a quantitative basis:

If the data are classified on the basis of a phenomenon which is capable of quantitative measurement like age, height, weight etc., is termed as quantitative classification. The quantitative phenomenon under study is known as variable and hence this classification is also sometimes called classification by variables..

(*iii*) Classification on a time basis (or chronological basis):

When statistical data is classified according to the time of its occurance, the type of classification is known as chronological classification. Data regarding sales of firm, population, imports and exports etc., are always subjects to chronological classification.

(*iv*) Classification on a geographical basis:

It is a classification based on geographical region. If the existing political boundaries are taken as the basis, the classification may be done by states, districts or talukas.

Objectives of Classification

(*i*) To condence the mass of data in such a manner that similarities and dissimilarities can be readily apprehended.

(*ii*) To facilitate comparison.

(*iii*) To pinpoint the most significant feature of the data at a glance.

(*iv*) To prepare the data for tabulation.

(*v*) To study the relationships.

PRESENTATION OF STATISTICAL DATA

Statistical data can be presented in three different ways.

(*i*) Textual presentation

(*ii*) Tabular presentation

(*iii*) Graphical presentation

TEXTUAL PRESENTATION

Numerical data presented in a descriptive form are called textual presentation.

Disadvantages of Textual Presentation

(*i*) The text is lengthy.

(*ii*) There are repetition of same words in the text.

(*iii*) The presentation has been made in such a way that comparison between the corresponding figures in two time periods is difficult and is not feasible at glance.

(*iv*) It is difficult to grasp the salient points, if the text is lengthy and consequently it is difficult job for an investigator to arrive at appropriate conclusions.

TABULAR PRESENTATION OR TABULATION

The orderly or systematics presentation of numerical data in rows and columns designed to clarify the problem under consideration and to facilitate the comparison between the figures. Tabulation is thus a form of presentation of quantitative data in condensed and concise form.

Tabulation can also be defined as the process of condensing classified data in the form of a table so that it may be more easily understand and so that any comparisons involved may be more readily made.

Objectives of Tabulation

1. To clarify the object of investigation
2. To simplify complex data
3. To depict trend
4. To economise space
5. To clarify the characteristics of data
6. To facilitate comparison
7. To detect errors and omissions in the data
8. To facilitate statistical processing
9. To help reference

Significance of Tabulation

The significance of tabulation as follows:

(*i*) Tabulation simplifies complex data when data are tabulated all unecessary details are repetitions are avoided.

(*ii*) Tabulation facilitates comparison. Since table is divided into various parts and for each part there are totals and subtotals, the relationship between different parts of data can be studied much more easily with the help of a table than without it.

(*iii*) Tabulation gives an identity to the data when the data are arranged in a table with a little and a number, they can be distinctly identified and can be used as a source reference in the interpretation of a problem.

(*iv*) Tabulation reveals pattern within the figures which cannot be seen in the narrative form.

Precautions

The following points may be kept in view while tabulating data:

(*i*) The table should suit the size of the paper usually with more rows than columns.

(*ii*) In all tables, the captions and stubs should be arranged in some systematic order.

(*iii*) The unit of measurement should be clearly defined and given in table.

(*iv*) Figures should be rounded to avoid unnecessary details in the table and footnote to this effect should be given.

(*v*) A column entitled "miscellaneous column" should be added for data which do not fit in the classification made.

Components (or) Parts of a Table

1. Title

A titles of a table is self-explanatory. It briefly describes the contents of the table and is therefore placed at the head of the table concerned. A table number is given just above the title for ready reference.

2. Stub

In a table, the extreme left-hand column containing the heading (or descriptors of the rows with its own heading is known as stub.

3. Captions

These are the heading of the columns and sub-columns other than the portion of the stub in the upper part of the table.

4. Body

The body is the main part of the table except the title, stub and captions. It contains numerical information which are arranged in the table according to the descriptions of the rows and columns given in the stub and captions.

Requisities of a Good Statistical Table

(*i*) **Suit the purpose:** A table should be in keeping with the object of statistical enquiry.

(*ii*) **Scientifically prepared:** The table should be prepared in a systematic and logically organised manner. Simple and compact so that it is readily comprehensible. It should be free from all sorts of overlappings and ambiguities.

(*iii*) **Clarity:** A table should be easily understandable, complete and self explanatory.

(*iv*) Manageable size.

(*v*) Columns and rows should be numbered.

(*vi*) Suitably approximated

(*vii*) Units: The unit designation should be given at the top of the table below the title such as price in rupees and weight in tones.

(*viii*) Averages and totals

(*ix*) Logical arrangements of items. Items may be arranged

(*a*) Alphabetically

(*b*) Geographically

(*c*) Chronologically

(*d*) Conventionally in order to magnitude in ascending or descending order.

TYPES OF TABULATION

1. Simple Tabulation

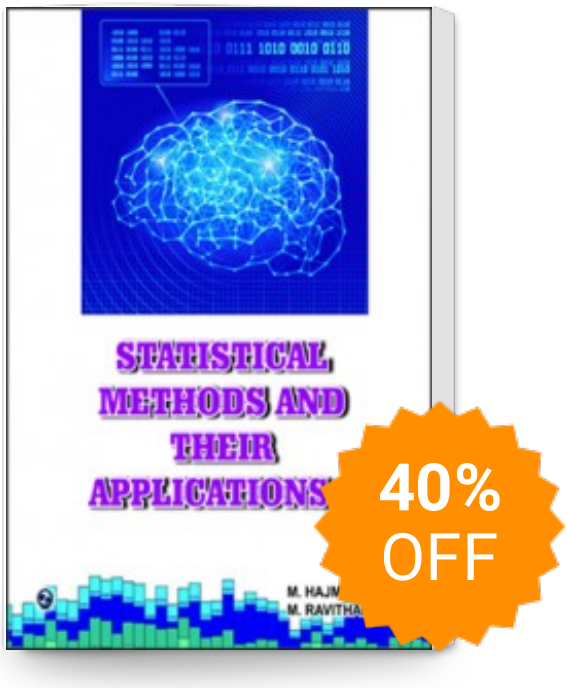
In this type the number of measurements of the items are placed just below the heading indicative of characteristics. This type is also called one-way table.

Example:

Progress of co-operative societies in India

Years	*No. of societies*	*No. of members*
1970–71	212.35	17452.85
1971–72	245.76	17834.50
1972–73	284.60	18263.70
1973–74	298.85	18621.20

Statistical Methods And Their Applications-I By M. Haj Meeral,M. Haj Meeral



Publisher : Laxmi Publications ISBN : 9789381159590

Author : M. Haj Meeral,M. Haj Meeral

Type the URL : http://www.kopykitab.com/product/3483