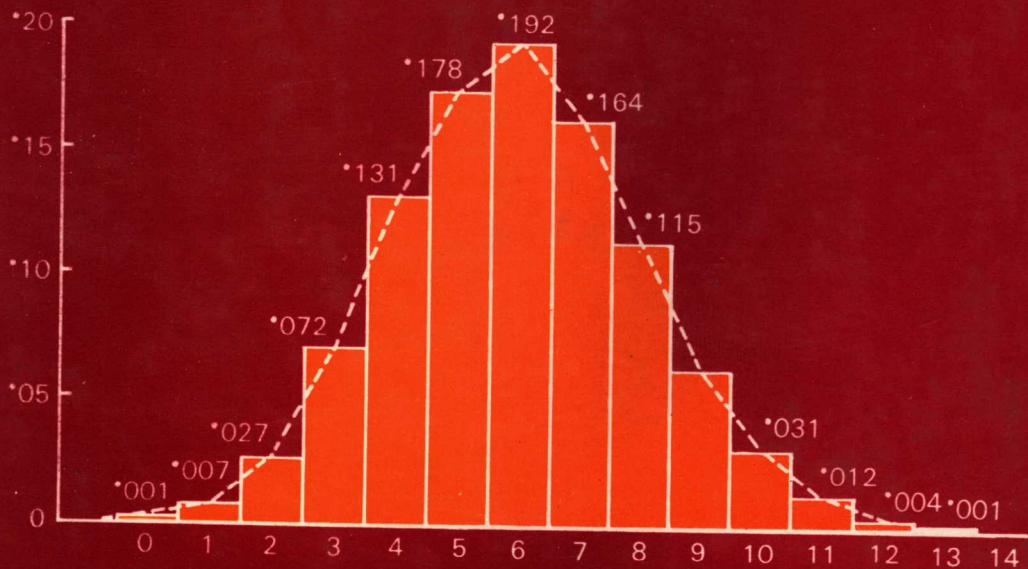


J. MEDHI

STATISTICAL METHODS

AN INTRODUCTORY TEXT



STATISTICAL METHODS

AN INTRODUCTORY TEXT

J. MEDHI

*Emirius Professor of Statistics
Gauhati University*

JOHN WILEY & SONS
NEW YORK • CHICHESTER • BRISBANE • TORONTO • SINGAPORE

First Published in 1992 by
WILEY EASTERN LIMITED
4835/24 Ansari Road, Daryaganj
New Delhi 110 002, India

Distributors:

Australia and New Zealand :
JACARANDA WILEY LIMITED
PO Box 1226, Milton Old 4064, Australia

Canada :
JOHN WILEY & SONS CANADA LIMITED
22 Worcester Road, Rexdale, Ontario, Canada

Europe and Africa :
JOHN WILEY & SONS LIMITED
Baffins Lane, Chichester, West Sussex, England

South East Asia :
JOHN WILEY & SONS (PTE) LIMITED
05-04, Block B, Union Industrial Building
37 Jalan Pemimpin, Singapore 2057

Africa and South Asia :
WILEY EASTERN LIMITED
4835/24 Ansari Road, Daryaganj
New Delhi 110 002, India

North and South America and rest of the world :
JOHN WILEY & SONS INC.
605, Third Avenue, New York, NY 10158, USA

Copyright © 1992 WILEY EASTERN LIMITED
New Delhi, India

Library of Congress Cataloging-in-Publication Data

ISBN 0-470-22085-6 John Wiley & Sons Inc.
ISBN 81-224-0419-7 Wiley Eastern Limited

Printed in India at Taj Press, New Delhi.

To the endearing memory of
my sister NILIMA

na jāyate mriyate vā kadācan

nā 'yam bhūtvā bhavitā na bhūyāh

ajo nityah sasvato' yam purāno

na hanyate hanyamāne sarīre

*The soul is never born nor dies; nor does it exist on coming into being.
For it is unborn, eternal, everlasting and primeval;
even though the body is slain, the soul is not.*

Bhagavadgītā, II, 20

Preface

Statistical methodology has been recognised as an important tool for analysis and interpretation of data in natural, biological, agricultural and engineering sciences as well as in social sciences. The subject has great relevance in several disciplines like economics, commerce, psychology, geography, geology, forestry, agriculture, veterinary, biochemistry, pharmacology etc. in addition to mathematics, statistics, engineering, and business management. Thus there is a very large number of students of several disciplines who need to pursue a course of statistics and statistical methodology at the introductory and intermediate or higher level. The book has been prepared keeping in view the diverse audience with varying orientation.

It is true that there are a number of books at this level. Considering the large, varied and growing audience, it is felt that there is scope for more text books. Texts by different authors have appeal for different sections of this large readership.

The book is divided into 3 parts. Part I deals with Descriptive Statistics, and Part II with Probability and Mathematical Statistics. Part III covers some special topics in Applied Statistics; discussion of some matters of general interest, not usually included in other books at this level, is an added feature of Part III. Mathematics used have been mostly of high school level except in certain places where methods of calculus have been used.

The book, it is expected, would be suitable for use as a text book for a one/two semester course in Introductory Statistics (or Statistical Methods) at US and Canadian Colleges and Universities. Part II alone could be used as a text for a one-semester beginners' level course in Probability and Statistics. This part would also be useful for research workers requiring knowledge of statistical methodology for their research work.

This book has been written with a view to present a qualitative understanding of the subject rather than present a surfeit of formulas and of manipulative examples/exercises only. It is hoped that this approach and objective would be appreciated by a large section of the teachers and students. The theory has been explained in some detail, supplemented with notes, where considered necessary. Several worked examples have been provided to bring home how the theory works and how the methodology could be applied. A large number of exercises are given, answers to most of which (for all Chapters of Part II and Chapter 12 of Part III) are provided. Emphasis has been on the understanding of the basic theoretical content and methodology.

A number of computer programs are given. The computer programs have been presented in some detail keeping the beginners in view. Students who may have access to some computer would find them very interesting. Research workers and analysts who would have to undertake statistical analysis of data would find the programs useful.

The text has grown from the author's long 40 years' experience of teaching and research, in India and abroad in Canada and USA. He has learnt a lot from the students; this has been useful in presenting the material.

The author's motivation was further heightened by the very complimentary and encouraging review (in the *American Mathematical Monthly*, December 1982) that recommended his earlier text *Stochastic Processes as the clear choice* as a text book. This prompted him further to prepare a book in statistics at the introductory level in a similar attractive and readable manner.

I have received encouragement and assistance from a large number of friends and colleagues in this endeavour. The work was undertaken at the request of a friend, Mr. A. Machwe of Wiley Eastern Ltd., My one-time students and colleagues, Profs. S. K. Dutta of Arya Vidyapith College, D. Chakravarty of Handique Girls' College and D. Nath of Gauhati University have rendered enormous help, by going through the manuscript as well as by assisting in the preparation of the solutions of the exercises. S. K. Dutta also read the proofs along with me with meticulous care and offered valuable suggestions which led to improvement of the text at several places. The Computer programs have been prepared with the expertise of and help rendered by S. K. Sinha, of Department of Computer Science, Gauhati University. L. Choudhury, Dept of Statistics, Gauhati University and P. C. Jha, Dept of O.R., Delhi University have also been of help.

The Times of India have kindly accorded permission to reproduce some figures appearing in the esteemed daily. The Biometrika Trust, London have been kind enough to permit reproduction of the statistical tables. I am indeed thankful to them all.

My elder son Deepankar Medhi, of Computer Telecommunications Program, University of Missouri-Kansas and my elder daughter Shakuntala Choudhury of AT&T Technology Systems, Bedminster, NJ have rendered immense help.

Lastly, it is my wife Prity, who bore the tedium of a 'Professor's wife' with great patience and understanding and whose constant support and encouragement made the writing possible.

I hope teachers and students of different disciplines would consider the book and give it a trial. Suggestions and comments received toward improvement of the book in later editions will be highly appreciated.

June 1992

U.N. Bezbarua Road,
Silpukhuri West,
Guwahati-781 003,
Assam, INDIA
Telephone : 91(361)32856

JYOTIPRASAD MEDHI

Contents

<i>Preface</i>	<i>v</i>
Part I. Descriptive Statistics	
Chapter 1. INTRODUCTION	3
1.1. The meaning of Statistics	3
1.2. History of Statistics	4
1.3. The role, scope and limitations of Statistics	5
1.4. Some Basic Concepts	6
1.4.1. Population and Sample	6
1.4.2. Parameter and Statistic	7
EXERCISES-1	7
Chapter 2. COLLECTION OF DATA	8
2.1. Introduction	8
2.2. Methods of collection of primary data	9
2.3. Framing of Questionnaire or Schedule	11
2.4. 1981 Population Census Schedules	11
EXERCISES-2	12
Chapter 3. PRESENTATION AND CLASSIFICATION OF DATA	13
3.1. Introduction	13
3.2. Discrete and Continuous Variables	15
3.3. Frequency Distributions	15
3.4. Selection of the number of classes or groups and the class limits	18
3.4.1. General Rules for construction of frequency distribution	19
3.5. Graphical Representation	20
3.6. Cumulative Frequency Distribution and Ogives	24
3.7. Bivariate Frequency Distributions	26
3.8. Tabulation of Data	30
3.9. Other Forms of Representation	31
3.9.1. Line Graphs	31
3.9.2. Geometric Forms	34
3.9.3. Pictorial Diagram or Pictogram	46
EXERCISES-3	47
Chapter 4. MEASURES OF LOCATION AND DISPERSION	53
4.1. Introduction	53
4.2. The Arithmetic Mean	53
4.2.1. The Arithmetic Mean of Grouped Data	54
4.2.2. Properties of the Arithmetic Mean	55

- 4.2.3. The merits and demerits of the Arithmetic Mean as a measure of location 58
 - 4.3. The Median 58
 - 4.3.1. The median of Grouped Data 58
 - 4.3.2. Graphical determination of median 59
 - 4.3.3. Merits and demerits of the median as a measure of location 60
 - 4.4. The Mode 60
 - 4.4.1. Calculation of Mode 61
 - 4.4.2. Continuous Distribution 62
 - 4.4.3. Empirical formula 62
 - 4.4.4. Merits and demerits of mode as a measure of location 62
 - 4.5. The Geometric and Harmonic means 64
 - 4.5.1. Merits and demerits of Geometric mean and Harmonic mean 64
 - 4.6. Other measures of location : Quartiles, Deciles and Percentiles 65
 - 4.7. Measures of Variation or Dispersion 68
 - 4.7.1. The range 69
 - 4.7.2. The semi-interquartile range 69
 - 4.7.3. The mean Deviation 70
 - 4.8. The Variance and the Standard Deviation 72
 - 4.8.1. Ungrouped Data 72
 - 4.8.2. Grouped Data 74
 - 4.8.3. Computation of variance and standard deviation—short method 76
 - 4.8.4. Interpretation of the Standard Deviation 81
 - 4.8.5. Chebyshev's Lemma or Rule (for sample) 81
 - 4.8.6. Properties of Standard Deviation 82
 - 4.8.7. Uses of Standard Deviation 83
 - 4.8.8. Sheppard's correction for variance 83
 - 4.8.9. Absolute and relative dispersion 84
 - 4.9. Moments of Higher Order 85
 - 4.9.1. Ungrouped data 85
 - 4.9.2. Relation between moments m_r and m_r' 86
 - 4.10. Other Descriptive Measures : Skewness and Kurtosis 88
 - 4.10.1. Skewness and its measurement 88
 - 4.10.2. Kurtosis and its measurement 89
- EXERCISES-4 91

Chapter 5. INDEX NUMBERS

94

- 5.1. Introduction, Meaning and Definition 94
- 5.2. Uses of Index Numbers 94
- 5.3. Price Relatives, Quantity Relatives and Value Relatives 95
 - 5.3.1. Price Relatives 95
 - 5.3.2. Quantity Relatives 96
 - 5.3.3. Value Relatives 96
 - 5.3.4. Properties of Relatives 96
- 5.4. Link and Chain Relatives 97
- 5.5. Problems involved in the Construction of Index Numbers 99
- 5.6. Cost of Living Index Numbers 102
- 5.7. Methods of construction of Index Numbers : Formulas 103
 - 5.7.1. Weighted Aggregates 103

- 5.7.2. Aggregate of Price Relatives 106
- 5.7.3. Comparison of Laspeyre's and Paasche's Index Numbers 106
- 5.8. Quantity Index Numbers 108
- 5.9. Tests for Index Numbers 109
- 5.10. Examples of Index Number Computation 111
- 5.11. Sources or Components of Error in an Index Number 115
- 5.12. Limitations of Index Numbers 116
- 5.13. Some Important Index Numbers 117
- EXERCISES-5 120

Part II. Probability and Mathematical Statistics

- Chapter 6. ELEMENTS OF PROBABILITY THEORY 127**
 - 6.1. Introduction 127
 - 6.2. Definitions of probability : Different approaches 127
 - 6.2.1. Classical Definition of Probability 128
 - 6.2.2. Relative Frequency and Statistical Regularity 129
 - 6.3. Axiomatic Approach to Probability 131
 - 6.4. Probability of a Simple Event 132
 - 6.5. Probability of a Composite Event 134
 - 6.6. Addition Rule 136
 - 6.7. Multiplication rule : Conditional Probability 139
 - 6.7.1. Number of sample points in a combination of events or sets 143
 - 6.7.2. Discrete sample space 143
 - 6.8. Bayes' Formula or Theorem 145
 - EXERCISES-6 148
- Chapter 7. RANDOM VARIABLE AND PROBABILITY DISTRIBUTION 151**
 - 7.1. Discrete Random Variable 151
 - 7.2. Expected Value of a Random Variable 156
 - 7.2.1. Expected Value of a Function of the Random Variable X 157
 - 7.3. Continuous Random Variable 161
 - 7.4. Standard Probability Distributions 163
 - 7.5. Bernoulli Probability Distribution 164
 - 7.6. Binomial Distribution 168
 - 7.6.1. The mean and variance of binomial distribution 172
 - 7.6.2. Skewness and Kurtosis of binomial distribution 174
 - 7.7. The Hypergeometric Distribution 175
 - 7.7.1. The mean and variance of hypergeometric distribution 178
 - 7.7.2. Hypergeometric distribution for large N 178
 - 7.7.3. Extension of hypergeometric distribution to more than 2 categories 179
 - 7.8. Geometric Distribution 180
 - 7.8.1. Another way of defining Geometric Distribution 182
 - 7.9. Poisson Distribution 182
 - 7.9.1. Mean and variance of Poisson distribution 186
 - 7.10. Normal Distribution 188
 - 7.10.1. Calculation of probabilities : Use of tables 190
 - 7.10.2. Properties of normal distribution 195
 - 7.10.3. Importance of normal distribution 197
 - EXERCISES-7 200

Chapter 8.	ELEMENTS OF SAMPLING THEORY	204
8.1.	Introduction 204	
8.2.	Sampling with and without replacement 205	
8.3.	Sampling distribution of the sample mean : Sampling with replacement 205	
8.4.	Sampling distribution of the (sample) mean 217	
8.5.	Sampling distribution of proportion 218	
8.6.	Standard Errors 222	
	EXERCISES-8 223	
Chapter 9.	CORRELATION AND REGRESSION	226
9.1.	Introduction 226	
9.2.	Scatter Diagram 226	
9.3.	The Coefficient of Correlation 227	
9.4.	Linear Regression 235	
9.5.	Fitting of Regression Line 236	
9.5.1.	The Method of Least Squares 237	
9.5.2.	Explained and Unexplained Variation : Coefficient of Variation 241	
9.6.	Two regression lines : Relation with r 244	
9.7.	Statistical Model 246	
9.8.	Correlation and Regression from Grouped Data 248	
9.9.	Further Discussion on Regression 252	
9.9.1.	Non-linear Regression 252	
9.9.2.	Multiple Linear Regression 253	
	EXERCISES-9 255	
Chapter 10.	STATISTICAL INFERENCE	259
10.1.	Introduction 259	
10.2.	Point Estimation 260	
10.3.	Estimation of the Population Mean : Point Estimation 261	
10.4.	Interval Estimation : Estimation of Mean 264	
10.5.	Estimation of the Parameter p 267	
10.6.	Small Sample Result 268	
10.7.	Testing of Statistical Hypothesis 269	
10.8.	Large Sample Theory 272	
10.8.1.	Test for an assumed mean 272	
10.9.	Test for an Assumed Proportion 275	
10.10.	Comparison of Means of Two Samples 275	
10.11.	Comparison of Proportions from Two Samples 276	
	EXERCISES-10 278	
Chapter 11.	FURTHER TESTS OF SIGNIFICANCE	281
11.1.	Small Sample Theory 281	
11.2.	Student's t -distribution 281	
11.2.1.	Test for an assumed mean 282	
11.2.2.	Comparison of means of two samples 283	
11.3.	The F -distribution 285	
11.4.	The Chi-square Distribution 286	
11.4.1.	Testing a hypothetical value of σ 287	
11.5.	Large Sample Test 288	
11.5.1.	Goodness of fit test 288	

- 11.5.2. Some observations on the use of χ^2 test 291
- 11.5.3. Goodness of fit of distributions 291
- 11.6. Contingency Tables 293
 - 11.6.1. Application of χ^2 test 294
- EXERCISES-11 297

Part III. Special Topics : Applied Statistics

Chapter 12.	TIME SERIES ANALYSIS	303
12.1.	Introduction 303	
12.2.	Characteristic Movements in a Time Series 304	
12.3.	Time Series Models 304	
12.4.	Measurement of Trend 305 <ul style="list-style-type: none"> 12.4.1. Inspection Method 306 12.4.2. Method of Moving Averages 306 12.4.3. Fitting of Mathematical Curves 308 	
12.5.	Secular Trend 309 <ul style="list-style-type: none"> 12.5.1. Linear Trend 309 12.5.2. Non-linear Trend 314 	
12.6.	Seasonal Movements 319 <ul style="list-style-type: none"> 12.6.1. Method of Simple Averages 320 12.6.2. Ratio to Trend Method 322 12.6.3. Ratio to Moving Average Method 324 12.6.4. Method of Link Relatives 331 	
12.7.	Cyclical Movement 333 <ul style="list-style-type: none"> 12.7.1. Residual Method 334 12.7.2. Other Methods 334 	
12.8.	Irregular Movements 335	
12.9.	Long Cycles : Kondratiev Waves 335	
	EXERCISES-12 337	
Chapter 13.	DEMOGRAPHY	342
13.1.	Introduction 342	
13.2.	Definitions 342	
13.3.	Birth and Death Rates 344 <ul style="list-style-type: none"> 13.3.1. Crude Rates 344 13.3.2. Specific Death Rates 345 13.3.3. Infant Mortality Rate 347 13.3.4. Adjusted Measures of Mortality : Standardised Death Rate 348 	
13.4.	Life Tables 350 <ul style="list-style-type: none"> 13.4.1. Basic Assumptions in construction of a Life Table 350 13.4.2. Description of Various Columns of a Life Table 350 13.4.3. Construction of Life Tables 351 13.4.4. Complete and Abridged Life Tables 353 13.4.5. Applications of the Life Tables 354 	
13.5.	Fertility and its Measurement 355 <ul style="list-style-type: none"> 13.5.1. Measures of Fertility 355 	
13.6.	Reproduction Rates 358	
13.7.	Sources of Demographic Data in India 360 <ul style="list-style-type: none"> A. Population Census 360 	/

	B. Civil Registration 362	
	C. Sample Surveys 362	
	D. Family Welfare Programmes 363	
13.8.	Population Growth 363	
	13.8.1. Factors of Population Growth 363	
	13.8.2. Theory of Demographic Transition 364	
	13.8.3. Demographic Scenario of India 365	
	13.8.4. Indian situation vis-a-vis Theory of Demographic Transition 365	
13.9.	Concluding Remarks 366	
	13.9.1 Malthusian Theory of Population 366	
	13.9.2 Current World Scenario : Malthus and Mother Nature 366	
	EXERCISES-13 368	
Chapter 14.	SAMPLE SURVEYS	370
14.1.	Introduction 370	
14.2.	Advantage of Sample Surveys over Complete Census 370	
14.3.	Errors in Sample Surveys 372	
14.4.	Various Stages in a Sample Survey 373	
14.5.	Types and Methods of Sampling 375	
	14.5.1. Sampling Methods 376	
14.6.	Simple Random Sampling 376	
	14.6.1. Method of Selection of a simple random sample 377	
	14.6.2. Determination of the sample size 377	
14.7.	Stratified Random Sampling 380	
	14.7.1. Variance and s.e. of weighted sample mean in a stratified random sample 381	
	14.7.2. Allocation of sample size between strata 383	
	14.7.3. Proportional Allocation 383	
	14.7.4. Optimum Allocation 384	
	Neyman Allocation 384	
	Optimum Allocation under Fixed Cost 387	
	Optimum Allocation with given error at minimum cost 388	
	14.7.5. Comparison of stratified random sampling with simple random sampling 389	
14.8.	Systematic Sampling 390	
14.9.	Other Methods 391	
14.10.	Sample Surveys in Practice in India 392	
	EXERCISES-14 394	
	ANSWERS TO EXERCISES	397
	BIBLIOGRAPHY AND REFERENCES	405
	APPENDICES	
	A. Tables	407
	1. The Normal Probability function 408	
	2. Percentage Points of the χ^2 -distribution 412	
	3. Percentage points of the t -distribution 413	
	4. Percentage points of the F -distribution 415	
	5. Individual terms of the Poisson distribution 419	
	B. Computer Programs	423
	INDEX	435

PART I
Descriptive Statistics

“Measure, measure, measure;
measure again and again;
measure the difference,
and the difference of the difference.”

Dictum of Galileo Galilei
(1564-1642)

Introduction

1.1. THE MEANING OF STATISTICS

“The age of chivalry is gone: that of sophisters, economists, and calculators has succeeded”, thus said Burke.¹ Calculations and quantifications have pervaded all spheres. This is the age of facts, figures and statistics. One is faced with statistics everywhere: in newspapers and other media, home and in connection with one’s at work in office. A citizen has to understand the implication and significance of statistics he comes across. “Statistical thinking will be one day as necessary for efficient citizenship as the ability to read and write” observed Wells². That day seems to have arrived. For effective participation in a democracy and for performing one’s task efficiently, one has to understand the underlying meaning of the statistical data that he frequently encountered.

Apart from this, statistics and statistical methodology have increasingly been used in all disciplines of social sciences, pure and applied sciences, as well as in several areas of humanities. It covers all fields of study wherein quantitative data appear and need analysis.

The term *statistics* is used in two senses. As a plural noun it is used for denoting and referring to numerical and quantitative information e.g., labour statistics, vital statistics, and so on. As a singular noun, the term *statistics* is used to denote the science of collecting, analysing and interpreting numerical data relating to an aggregate of individuals. We thus have the following definitions of statistics:

Encyclopaedia Britannica (1969): As is commonly understood now a days, statistics is a mathematical discipline concerned with the study of masses of quantitative data of any kind.

Encyclopaedia Americana (1968): As a name of a field of study, statistics refers to the science and art of obtaining and analyzing quantitative data with a view to make sound inferences in the face of uncertainty.

International Encyclopaedia of Social Sciences (1968): Statistics deals with the inferential process, in particular, with the planning and analysis of experiments and surveys, with the nature of observational errors and sources of variability that obscure underlying patterns with the efficient summarizing of sets of data.

It is to be noted that statistics primarily deals with phenomena in which the occurrences of the event under study cannot be predicted with certainty.

- 1 Edmund Burke (1729-1797) was a great British writer and orator, who is well known for his attack on Pitt and Governor General Warren Hastings for their policies and actions in India.
- 2 Herbert George Wells (1866-1946) was a celebrated British thinker and writer, author of novels and science fiction.

4 Statistical Methods

Numerical data obtained in the face of uncertainty constitute statistical data. Suppose we study the heights of students aged 10 years in a large city. We cannot predict with certainty the height of an individual student: the data of heights of students constitute statistical data. Further we note that there will be variation in heights of students, though they are of the same age. *Uncertainty and variability* are two characteristics of statistical data. Not all quantitative data are statistics. Suppose we prepare a multiplication table by writing the product of 5 and 8 with the integers 1, 2, 3,... in a tabular form. The quantitative data so obtained are not statistical as there is no uncertainty involved in obtaining such data. Further, individual measurements are, not statistical. Statistics is concerned with masses of numerical data. We may summarise as follows:

Statistics is a collection of concepts and methods used for collection, analysis and interpretation of masses of numerical data relating to a certain area of investigation as well as for drawing valid conclusions in situations in the face of uncertainty and variability. Statistics in this sense denotes statistical methodology.

1.2. HISTORY OF STATISTICS

Historically speaking, collection of data began as early as recorded history. It is stated that even the ancient Babylonians collected data on population. Such data were considered important to the state. The word *statistics* was derived from the latin word '*status*' meaning state. Statistics was thought of as mass of data relating to state. The word census is of Roman origin and dates back to the first or second century A.D.: the magistrates in Rome were required to prepare registers of the population by which the state could determine the liability of adult males for military service as well as for imposition of taxes. Apart from census of population and wealth, census of land was also taken. As reported by the Greek historian Herodotus, a census of all lands in Egypt was taken as early as in 1400 B.C. There are reports of collection of data on population, wealth and land also in other countries, such as Greece and China. In Kautilya's *Artha Shastra*, there is reference to collection of such data in ancient India. Gradually the sphere of data collection and census spread in several other directions; and now it encompasses a variety of areas. The word statistics is thus used today in a much broader context.

"Observations on the London bill of mortality" of 1662 marked the start of tabulation and publication of data. Along with progress in collection of data, progress was being achieved in the development of a framework for analysis of data through mathematics. These two aspects developed independently as two parallel streams. The theory of probability was emerging as a potentially rich field of mathematical investigations. The organisations and workers dedicated to collection and tabulation of data created an increasing need for scientific analysis of data. One of the pioneers who recognised the need for analysis of collected data and the capability provided by mathematical framework in such analysis was a Belgian astronomer Quetelet³, by name. He studied mathematics and probability under the renowned French mathematicians Fourier⁴ and

3 Adolphe Quetelet (1796-1874), was a Belgian mathematician, astronomer and statistician.

4 Joseph Fourier (1768-1830) was a French mathematician.

Laplace⁵ and his work comprised of collection and tabulation of data in the Royal Belgium Observatory. He visualised the possibility of fusing the two streams, the collection of data and the analysis of data through mathematical concepts, into a general body of techniques, now known as statistical methodology or statistics, for short. Incidentally Quetelet was the first to formulate the concept of the 'average man' (*l'homme moyen*).

1.3. THE ROLE, SCOPE AND LIMITATIONS OF STATISTICS

The three main components of statistics are the collection of data, analysis of data and interpretation of facts therefrom; statistical methodology is our guide in all these spheres. At the stage of collection of data, statistics indicates what statistical methodology is to be used for the collection of data, so that logical conclusions can be drawn therefrom later. The branch is known as *sampling* or *experimental design*. After this first stage, statistics is concerned with methodologies to be adopted for summarizing the data and to obtain its salient features from the vast mass of original data. This branch of statistics is termed as *descriptive statistics*. The final stage is the one in which inferences are drawn and logical conclusions arrived at. The branch of statistics containing the relevant statistical methodology for this is termed *statistical inference*.

The scope of statistics is indeed very vast. Apart from helping elicit an intelligent assessment from a body of figures and facts, statistics is an indispensable tool for any scientific enquiry — right from the stage of planning the enquiry to the stage of conclusion. It applies to almost all sciences: pure and applied, physical, natural, biological, medical, agricultural and engineering. It also finds applications in social and management sciences, in commerce, business and industry.

Of the social sciences, economics leans most heavily on statistical methods for analysis of data relating to micro as well as to macro economics, from demand analysis to National Income analysis. The impact of mathematics and statistics has led to the development of new disciplines like Econometrics and Economic Statistics. There is wide scope of statistics in handling and analysing data relating to socio-economic, demographic and political processes.

Statistics is now considered an indispensable tool in analysis of activities relating to business, commerce and industry. For example, statistical methods are useful in determining trends of growth in business, in making proper production plans after consumer and market surveys and in the process of production itself, while statistical concepts and methods of quality control and reliability prove invaluable in industrial engineering.

Statistics figures prominently in agricultural and biological sciences, in determining the precise role of various factors in the growth and development of the organism or plant under study. In medical sciences, the effect of drugs on individuals is one area where statistical methods are widely used. Applications of statistical methods to education and psychology have led to development of a new discipline 'Psychometry'. In fact in almost all investigations wherefrom quantitative data is obtained, whatever the subject, statistical analysis is applicable.

5 Pierre-Simon Laplace (1749-1827) was a French mathematician, astronomer and physicist.