CHAPTER-I

INTRODUCTION
CHAPTER - 1

INTRODUCTION

1.1 GENERAL INTRODUCTION

Statistics is a branch of science, that can be used in almost all fields science, where quantitative analysis can be applied. One may observe its application in various branches of Biological and Life Sciences such as Biology, Medicine, Agriculture, Anthropology etc. Some of its interaction with Biology, Medicine, Agriculture, Anthropology gave the births of various special branches as Biostatistics, Medico Statistics, Agricultural Statistics, Anthropometry etc. Statistics plays a vital role in Biological and Medical Sciences.

Biostatistics or Biometrics is a growing filed of statistics. The word ‘Biometrics’ derived from two Greek Words Bios (Life) and Matron (measurement). Thus Biometrics means measurement of life.

In recent years a great deal of research has been directed to the modeling and measurement of biological relationships. Biometrics is the art and science of using statistical methods for the measurement of biological relationships.

Biostatistics many be defined as the quantitative analysis of biological phenomena based on the concurrent development of theory and observation related by appropriate methods of statistical inference.

According to Prof P.V.Sukhatme ‘Biostatistics is an applied scientific discipline rather than a basic or fundamental study. Its roots lie in mathematics, but it’s branches touch all areas of biology, medicine, nutrition and public health’.
Now-a-days, Biostatistics has a wide number of applications. It covers applications and contributions not only from health, medicine, and nutrition but also from fields such as agriculture, genetics, zoology, Botony, biochemistry, demography, epidemiology, anthropology, and many others.

Bioassays and Quantitative Genetics can be considered as two main specialized areas of Biostatistics.

In the present research study an attempt has been made by developing some advanced techniques for Bioassays and Quantitative genetics.

1.2 STATEMENT OF THE PRESENT RESEARCH STUDY

The association between statistical techniques and biological theories was first studied by Francis Galton (1822-1921), known as Father of Biometry, in his work on 'Regression'. Prof Karl Pearson (1857-1936) was greatly motivated by the researches of Galton and he established Department of Applied Statistics (First in the world) with the Greatest Statistical Laboratory in England in year 1911. He started issue of 'Biometrika' Journal. William S. Gosset (1876-1937), A student of Karl Pearson, propounded a number of statistical applications to the biology under the pen name 'student'. His contributions to the biostatistics are also worth processing.

Biometry is a branch of statistics in which various mathematical and statistical tools can be applied to various biological research problems. Much theoretical and empirical research in Biostatistics is developed to the study of the effect of one set of variables on the another. Much of these studies are based on the analysis of experimental data obtained in the laboratory.
During the last few years there has been a rapid growth of interest in the statistical study of biological research problems. Only a few quantitative research works concerning biology and medicine could utilize the application of certain new or sophisticated statistical methods. The field of biometric models is still a new fertile research area for statisticians and others.

One of these chief characteristics of all living beings including human beings is that no two individuals are exactly alike. There can be seen differences even in identical twins due to environmental influences. Variations exist not only between individuals but also within individuals from time to time.

Some important types of variation found in biological and medical sciences, which are given by: biological variability, observer variability, Instrument variability, Environment variation, Incomplete information, Incomplete Medical knowledge, and controlling variation. Generally, one may examine the effects of different factors responsible for these variations. A large part of statistical efforts is devoted to devising and implementing strategies to keep the variations under control. Control of variations is obtained by choosing an appropriate design of study. For instance, in epidemiological studies, generally one may use retrospective, prospective or cross-sectional designs. For clinical trials and animal experiments, the various experimental designs such as Randomized Block Design, Factorial Design and Cross over design may be used. In fact, Randomisation is a prerequisite for constructing any statistical design. By using some scientific methods one may estimate the parameters relating to effects of each source of variation.
Another strategy is to choose a design such that the effect of all known factors of variation like, age, sex, body mass index, education etc., can be directly estimated. Analysis of variance (ANOVA) technique may be used to analyse these variations. Many multiple range tests such as Fisher's test, l.s.d. test, Tukey test, Duncan's Multiple range test, Scheffe's test, studentised range test, Dunnet test etc may be used as multiple comparison procedures in this regard.


1.3 MULTIVARIATE STATISTICAL TOOLS FOR BIOSTATISTICS

Researchers in the Biological and Medical Sciences frequently collect measurements on several variables. Multivariate statistical analysis is concerned with statistical tools for describing and analysing these multivariate data. Multivariate Statistical concepts and tools are analogous (generalization) to univariate statistical concepts and methods. The relationships between many variables make multivariate analysis which is an inherently difficult subject to understand by researchers is the Biological and Medical Sciences. More mathematics is required to derive multivariate statistical techniques for making inferences than in a univariate setting. Modern computer software packages such
as SPSS version-13, SAS, QSB, RATS etc., readily provide the numerical results to rather complex multivariate statistical analysis. It should be noted that many multivariate statistical methods are based upon an underlying probability distribution known as the "Multivariate Normal Distribution".

**Multivariate statistical analysis as an Analogous to Univariate Statistical Analysis**

The following table gives the generalizations of univariate statistical concepts and measures as multivariate statistical tools which are having a wide number of applications in analysing various research problems in the Biological and Medical Sciences.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Arithmetic Mean</td>
<td>Mean Vector</td>
</tr>
<tr>
<td>2</td>
<td>Standard deviation or variance</td>
<td>Dispersion matrix or variance covariance matrix; generalized variance</td>
</tr>
<tr>
<td>3</td>
<td>Simple correlation</td>
<td>Multiple correlation / partial correlation</td>
</tr>
<tr>
<td>4</td>
<td>Simple correlation coefficient</td>
<td>Multiple correlation coefficient / partial correlation; correlation matrix</td>
</tr>
<tr>
<td>5</td>
<td>Square of simple correlation coefficient ($r^2$)</td>
<td>Coefficient of Multiple determination ($R^2$) adjusted $R^2$ or $R^2$, statistic</td>
</tr>
<tr>
<td>6</td>
<td>Simple linear Regression Model</td>
<td>Multiple linear Regression Model</td>
</tr>
<tr>
<td>7</td>
<td>Univariate probability distributions</td>
<td>Multivariate probability distributions</td>
</tr>
<tr>
<td>8</td>
<td>Normal distribution</td>
<td>Multivariate normal distribution</td>
</tr>
<tr>
<td>9</td>
<td>Student’s t-test statistic</td>
<td>Hotelling’s $T^2$ - test statistic</td>
</tr>
<tr>
<td>10</td>
<td>Chi-square Test</td>
<td>Bartlett’s Test</td>
</tr>
<tr>
<td>11</td>
<td>Comparing several sample means</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>--------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(a) ANOVA for one-way classified data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(b) ANOVA for Two-way classified data</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>ANOVA Technique</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Kruskal - wallis Test and Friedman's Test for ranked data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comparing several multivariate mean vectors</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(a) Multivariate ANOVA (MANOVA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>for one-way classified data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(b) Two-way MANOVA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analysis of dispersion</td>
<td></td>
</tr>
</tbody>
</table>

**SPECIAL MULTIVARIATE STATISTICAL TOOLS**

1. Multivariate Multiple linear regression model or sets of linear Regression Models.

2. Seeming Unrelated Regression Equations (SURE) Model.

3. Path coefficients Analysis.

4. Principal components Analysis

5. Factor Analysis

6. Canonical Correlation Analysis.

7. Discriminant Analysis and Mahalanobis D2 – Statistic.


9. Multivariate Time Series Analysis

10. Multiple Logistic and Multiple Poisson Regression Models.
1.4 OBJECTIVES OF THE PRESENT RESEARCH STUDY

In the literature about Biostatistics, many statisticians have applied univariate statistical tools, whereas a few statisticians have applied multivariate statistical tools at elementary level.

In the present study, an attempt has been made by proposing some modified advanced techniques for analysing research problems in biological and medical sciences.

The specific objectives of the present study are:

(i) to develop some statistical procedures by modifying the existing advanced techniques in biostatistics;

(ii) to describe some trend tests for Binomial and Poisson Counts and Rates for Biostatistics;

(iii) to propose some advanced Logistic and Poisson Regression Models for Biostatistics;

(iv) to develop some new tests for comparing sequential contingencies using odds Ratio and coefficients of contingency;

and (v) to describe the various existing statistical tools for the biostatistics.

1.5 ORGANIZATION OF THE PRESENT RESEARCH STUDY

The organization of the present research work itself shows how the objectives of the present research study have been achieved within the described frame work.

Chapter-I is any introductory one. It contains the general introduction about Biostatistics and statement of the present research study besides some
multivariate statistical tools for Biostatistics. It also brings out the organization of the present research study; and chapter scheme besides the objectives of the present research study.

Chapter-II describes the various aspects of Bioassays such as meaning, history and structure of Bioassays; Fundamental theorems; Behren's distribution and it's applications; Quantitative. Dose-Response Relationships; Parallel Line and Slope Ratio Assays; Quantal Response Bioassays and Probit and Logit analyses. It also contains some basic aspects of Quantitative Genetics.

Chapter-III deals with the existing univariate statistical tools in Biostatistics. It presents a brief review about basic statistical tools for Biostatistics.

Chapter-IV gives a brief review about the existing Inferential Multivariate statistical tools for Biostatistics such as Hotelling's $T^2$, Bartlett's test; Analysis of Dispersion, MANOVA, Cluster Analysis, Principal Components Analysis, Factor Analysis; Canonical Correlation Analysis; Path Coefficient Analysis; and Discriminant Analysis.

Chapter-V proposes some advanced and new statistical techniques for Biostatistics such as modified Trend tests for Biostatistics; Tests for the equality between Poisson Counts and tests for comparing sequential contingencies across k groups in a Manifold Contingency table using log odds Ratio and coefficients of contingency and a trend test based on reverse arrangement. Univariate and Multivariate Logistic and Poisson regression models have been specified and estimated for Biostatistics in this chapter.
Chapter-VI epitomizes the conclusions based on the present research study. Besides Conclusions, some directions have been given for future research in the line of the present research study.

Several relevant research articles regarding the present research study have been presented under the title ‘BIBLIOGRAPHY’.

1.6 CHAPTER SCHEME

The contents of the present research study are presented under the following heads:

CHAPTER – I : INTRODUCTION
CHAPTER – II : SOME ASPECTS OF BIOASSAYS
CHAPTER – III : A BRIEF ABOUT EXISTING UNIVARIATE STATISTICAL TOOLS IN BIOSTATISTICS
CHAPTER – IV : A BRIEF ABOUT MULTIVARIATE STATISTICAL TOOLS IN BIOSTATISTICS
CHAPTER – V : SOME NEW ADVANCED TECHNIQUES FOR BIOSTATISTICS
CHAPTER – VI : CONCLUSIONS

BIBLIOGRAPHY