INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeib Road
Ann Arbor, Michigan 48106
SINGH, Umed, 1938-
CONTRIBUTIONS TO STATISTICAL STUDIES OF
COMPARTMENTAL MODELS.

The Ohio State University, Ph.D., 1975
Biostatistics

Xerox University Microfilms, Ann Arbor, Michigan 48106
CONTRIBUTIONS TO STATISTICAL STUDIES OF COMPARTMENTAL MODELS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Umed Singh, B.A(Hon's), M.A., M.S.

* * * * *

The Ohio State University
1975

Reading Committee:
Dr. D. R. Whitney
Dr. J. S. Rustagi
Dr. W. R. Harvey

Approved By

Dr. J. S. Rustagi
Adviser
Department of Biostatistics
ACKNOWLEDGMENTS

I am grateful to Professor Jagdish S. Rustagi for his guidance and encouragement during the preparation of this dissertation, and for taking interest in my professional development. I thank Professors D. Ransom Whitney and Walter R. Harvey for being on my reading committee and for their helpful suggestions.

I am under obligation to the authorities of Haryana Agricultural University, Hissar, India who deputed me for my present studies at The Ohio State University. My special thanks are due to Dr. P. S. Lamba who took keen interest in me.

I take this opportunity to express a deep sense of gratitude to my mother, Mrs. Dharam Kaur, who sustains my spirits in a distant land far away from home. I will fail in my duties if I do not express a deep sense of love to my wife, Mrs. Sudha Singh, who sacrificed so much of her personal comforts to enable me to complete my present studies.
VITA

June 8, 1938 ....... Born-Mandaura, India
1959.................. B.A(Hon's), Panjab University
1962 ................. M.A., Panjab University
1964-1966 ............. Diploma, Institute of Agricultural Research Statistics
1966-1971 ............ Statistician, Haryana Agricultural University
1971-1972 ............ U.S.A.I.D Fellow
1972-1975 ............ Teaching Associate, Department of Statistics The Ohio State University
1973 .................. M.S., The Ohio State University

FIELDS OF STUDY

Major Field:

Studies in Biostatistics. Professor Jagdish S. Rustagi
Studies in Genetics. Professors Harvey and Young
TABLE OF CONTENTS

- ACKNOWLEDGMENTS ........................................ ii
  VITA ................................................................ iii
  LIST OF TABLES ............................................. vi
  LIST OF FIGURES ........................................... vii
  INTRODUCTION ............................................. 1

CHAPTER

I. EXPERIMENTAL STUDIES .................................. 8

II. COMPARTMENTAL MODELS ............................... 19
   Introduction
   Deterministic Theory
   Some Statistical Considerations

III. DETERMINATION OF INITIAL ESTIMATES .............. 42
   Introduction
   General Estimation Procedure
   Simultaneous Estimation By Partial Totals
     For Compartmental Models
   Examples
   Consistency of the Estimators
   Approximate Test For the Number of Compartments
   Comparison With Other Methods

IV. DIFFERENCE EQUATION APPROACH TO COMPARTMENT ANALYSIS... 57
   Introduction
   Derivation of Difference Equation Model For
     Unequally Spaced Time Series
   Estimation
   Testing Hypotheses
   Examples
V. ANTITONIC REGRESSION ESTIMATION IN NONLINEAR MODELS ...66

Introduction
Antitonic Regression Over a Simply Ordered Finite Set
Antitonization of Estimates
Examples

VI. INFERENCE IN STOCHASTIC COMPARTMENT ANALYSIS ..........78

Introduction
Solution of Differential Equation
Population Sizes and Associated Probability Distribution
Joint Probability Distribution of the Numbers in the System at Different Times
Least Squares Estimates

VII. SUMMARY ................................................................. 92

REFERENCES .................................................................97
# LIST OF TABLES

<table>
<thead>
<tr>
<th>No. of Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>13</td>
</tr>
<tr>
<td>1.2</td>
<td>15</td>
</tr>
<tr>
<td>1.3</td>
<td>16</td>
</tr>
<tr>
<td>1.4</td>
<td>17</td>
</tr>
<tr>
<td>1.5</td>
<td>18</td>
</tr>
<tr>
<td>3.1</td>
<td>50</td>
</tr>
<tr>
<td>3.2</td>
<td>51</td>
</tr>
<tr>
<td>5.1</td>
<td>68</td>
</tr>
<tr>
<td>5.2</td>
<td>75</td>
</tr>
<tr>
<td>6.1</td>
<td>84</td>
</tr>
<tr>
<td>No. of Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
</tr>
<tr>
<td>1.1</td>
<td>10</td>
</tr>
<tr>
<td>1.2</td>
<td>10</td>
</tr>
<tr>
<td>1.3</td>
<td>11</td>
</tr>
<tr>
<td>2.1</td>
<td>20</td>
</tr>
<tr>
<td>2.2</td>
<td>20</td>
</tr>
<tr>
<td>2.3</td>
<td>21</td>
</tr>
<tr>
<td>2.4</td>
<td>37</td>
</tr>
<tr>
<td>5.1</td>
<td>69</td>
</tr>
</tbody>
</table>
INTRODUCTION

In recent years, a number of issues have caused increasing attention to be focused on an area of medical therapeutics which has come to be known as "bioavailability." Some of these issues are the rising cost of medical care, the increasing involvement of Federal, state and local governments in financing such care, control of medical treatment, cost factors associated with the researching of new drugs and so on.

The most general definition of bioavailability includes the study of the factors which influence and determine the amount of active drug which is transferred from the administered dose to the site of pharmacologic action as well as the rate at which it is transferred there. However, in common usage today, "bioavailability" is often used as a shortened term for "comparative bioavailability"; that is, the relative bioavailability of active drug from two or more formulations.

At one time it was assumed that chemical equivalents were necessarily therapeutic equivalents. It has been shown than this is not true for a number of drugs, see Lancet (19/2), but the extent of the problem is unknown, and thus the current high interest in bioavailability.

The most direct way to assess therapeutic equivalence of two formulations of the same active ingredient is by a clinical efficacy trial. Efficacy trials are difficult and expensive to carry out,
and bioavailability trials are an attempt to infer therapeutic equivalence without doing efficacy trials. The assumption is that biological equivalents are also therapeutic equivalents.

Various measures have been proposed for assessing bioavailability. Peak concentrations and the time to peak contain information about the rate and extent of absorption and provide a measure of bioavailability. The area under the concentration-time curve is a favorite measure of bioavailability. Bioavailability is also assessed by estimating and comparing the parameters of pharmacokinetic models. No single parameter of the concentration-time curve is the best indicator of bioavailability. In most cases more than one parameter is considered, e.g., concentration at two or more sampling times following dosing, or peaks, time to peak and area. Separate univariate analysis of each parameter may be inappropriate because of probability considerations or because of difficulties of interpretation. Various methods for analyzing serum concentrations observed at more than one sampling time have been provided by Winer (1962), Grizzle and Allan (1969), and Snee (1972). Rahlfs and Bedall (1971) have compared the above methods of analysis and have attempted to define the conditions under which each is the most appropriate.

The fitting of pharmacokinetic models as an adjunct to bioavailability studies is becoming common, see Swarbrick (1970), Wagner (1971). In the enthusiasm for this quantitative tool of pharmacology and biopharmaceutics, too little attention has been given to the mathematical difficulties of estimation of parameters, to the properties
of the estimates obtained, or to the possible consequences of the use of an incorrect model. Unless a proper pharmacokinetic model has been formulated in the development program of a drug, the additional contribution which a pharmacokinetic model can make to a comparative bioavailability study may not justify the effort of formulating the model at that point. In this dissertation we consider the problem of formulating a correct pharmacokinetic model, its solution, the estimation of parameters and testing of hypotheses.

The motivation of pharmacokinetic studies of drugs is to understand such drug-induced behaviour as absorption rate, speed of drug action, rates of synthesis of metabolites, the transport of drugs through bodily system and the rates at which they are eliminated. The study of drug-induced behavior has been greatly facilitated in recent years by the use of radioactive tracer materials. The mathematical description of biological and biochemical processes by means of tracer is called compartment analysis. The concept of compartment analysis assumes that a system may be divided into homogeneous subsystems, or 'compartments'. Various characteristics of the system are determined by observing the movement of tracer material.

The concepts of pharmacokinetics and compartment analysis was first introduced by Teorell (1937). Prior authors had derived pharmacokinetic equations but had not clearly stated the relationship of the equations to a particular model or set of assumptions. For example, Widmark and Tandberg (1924) derived mathematical equations of drug accumulation for the one compartment model with
rapid intravenous injection when doses are administered at uniform intervals of time. More recently Wagner (1971) made an excellent contribution in developing the philosophy and formulating the theory related to compartment models and bioavailability.

Usually the theory is applied to describe the movement of a population of tracer molecules. Since the individual molecules are infinitesimal in size, nearly all the early literature has made the implicit assumption of a deterministic flow pattern. Papers by Zilversmit et al. (1943) and Shappard and Householder (1951) laid the main foundation of the deterministic theory of compartmental system. More recently Sheppard (1962) and Rescigno and Segre (1966) provide comprehensive reviews of the theory. Estimation in the deterministic model utilizes the ordinary least squares technique, and recently 'simultaneous equation' estimation for time-wise uncorrelated error has been introduced to compartment models in a series of papers by Beauchamp and Cornell (1966,1968).

On the other hand, stochastic compartment analysis, which assumes probabilistic behavior of the tracer particles, has been slow to develop. Bartholomay (1958) was one of the early authors to introduce the stochastic assumption, and solved the one-compartment model. Other authors such as Uppuluri, Feder and Shenton (1967) considered a one-compartment system in which radioactive materials were injected at unit time intervals to compensate for material that had decayed. Two models for injection and decay were considered to study the limiting behavior to the random sequences. Rustagi (1964; 1965) introduced
compartment models to describe the stochastic behavior of trace substances with special applications to air pollution problems. Matis and Hartley (1971) considered a general compartment system with probabilistic behavior of the tracer particles. They derived the distribution theory and derived the estimation procedures for the compartmental system.

The formulation of mathematical models in pharmacokinetics has been discussed by Berman and Schoenfeld (1956) and Sheppard (1962). It is assumed that there are fixed transition probabilities or turnover rates from one compartment to another, and the whole system is assumed to be in steady state. The turnover rates are also assumed to be proportional to the amounts of material in the compartments.

A p-compartment system is assumed to be described by a system of first-order, linear ordinary differential equations with constant coefficients called compartment equations. The solution of these differential equations is known to be a linear combination of exponentials. For these non-linear models, obtained as solution of differential equations, application of the least squares method results in equations which are in general soluble only by iteration. For the successful implementation of an iterative process, it is necessary to have good initial estimates of the parameters appearing non-linearly in the model. In recent years, investigations have been made by Agha (1971), Cornell (1962) and Ross (1969, 1970) to find a simple and direct procedure for estimating the non-linear parameters.
Cornell (1962) has proposed a general method which provides a simple and direct procedure for estimating the non-linear parameters and the method is based on independent partial totals of the sample observations but it has the disadvantage that the estimators obtained are not of exponential parameters but of some integral power of the parameters. Agha (1971) has proposed an alternative method which overcomes the disadvantage of Cornell's method but utilizes dependent partial totals of the sample observations. Foss (1969) suggested a reasonable computer oriented technique for obtaining the initial estimates. His method arrives at the initial estimates by a least square "peeling off" technique.

In this dissertation, we examine the various aspects of compartment models. In chapter I, we describe a few recent experimental studies using tracer compounds. The formulation of compartmental models and the problems associated with models are discussed in chapter II. A new difference equation approach to compartmental models in pharmacokinetics and bioavailability is introduced in this chapter. In many cases, the difference equation gives improved estimates. We propose a new method for the estimation of initial estimates which overcomes the drawbacks of Cornell's method in chapter III. A numerical comparison is also made with other methods. In most of the compartmental models, there is a monotonic decreasing function over time. These considerations lead to the study of antitonic regression estimation in non-linear models which is discussed in chapter IV.
A survey of results on antitonic regression is also made here. Choosing of antitonic weights in decay-type data situations is examined with reference to real as well as simulated data. A new test for the number of compartments for a drug in the system based on a difference equation model is proposed in chapter V. A method for the estimation of parameters appearing exponentially in the model is also examined fully in this chapter.

In chapter VI, we discuss the model and least squares estimation from the time series data on stochastic compartmental analysis. The results obtained by Matis and Hartley (1971) are simplified through a different approach with the same set of assumptions.
CHAPTER I
EXPERIMENTAL STUDIES

1.0 Introduction

Numerous studies have been made in the general area of biological radiation, pharmacokinetics, growth and bioavailability where the intake and output of a given substance in the biological system is studied over a period of time. The problems of interest in such investigations are concerned with the procedures to evaluate rates of intake and output, thereby determining the amount retained or absorbed in the biological system. In this chapter, we describe a few recent experimental studies dealing with the above problem. The basic process involves administration of a tracer compound or a drug to a biological system and taking observations at equal or unequal intervals of time on the amount in the system and on the rate of output through various media depending on the biological object. In humans, for example, such observations are made only through measurement of the substance in blood, urine or feces or through biopsy. In animals, the measurements can be more precisely made on various body organs since they can be sacrificed for these studies.

In pharmacokinetic the studies of drugs are conducted for understanding rates of absorption, speed of drug action, rates of synthesis of metabolites and the rates of transport of drug through the biological system. Also studied are the rates at which drugs are eliminated and the factors which influence these parameters.
A biological system is regarded as being composed of many systems which are often known as compartments. For example, in humans the compartments of interest may be blood, tissue, bones, and other organs such as kidney, liver, brain, etc. Some problems may only be concerned with a single compartment; others may involve several compartments and interactions among them. The structure of the biological system and the connections of compartments are known a priori in most experiments. Assumptions of such structures have to be made in other experiments.

Two types of studies are described here. One type concerns observations collected at equally spaced time intervals, following the administration of a compound, and another concerns observations collected at unequal time intervals. Examination of various statistical considerations in some of these studies will be made in later chapters. Such studies are also made in many other areas such as in physical, social and behavioral sciences and we include two studies of that nature.

1.1 Experimental Studies

I Albumin study

This study was concerned with the detection of albumin in human organs. The experiment was conducted with the administration of $^{131}$I labelled albumin. An intravenous injection of $3.14 \times 10^7$ counts per minute was given to a subject. The blood concentration of counts per minute per millilitre of blood was measured -
at various times after injection. The data are reported by Cornfield; Steinfeld and Greenhouse [1960] and are described in figures 1 and 2.

The injected dose was $3.14 \times 10^7$ counts per minute and the amount of concentration 20 minutes after injection was observed to be 7431 counts/minutes/ml blood. Taking into account a 9 percent reduction which is usually allowed for computing the volume to account for the difference between peripheral venous hematocrit and whole body hematocrit, the volume of vascular compartment into which the albumin had distributed itself by 20 minutes must be 3845 ml. Using the graph it becomes clear that the injected albumin has distributed itself to tissues other than blood. The volume of the vascular compartment is the volume of body fluids which hold albumin in solution at the same concentration as the blood. If the blood
concentration at any subsequent time is multiplied by $38^{45}$, we have an estimate of the total amount of labelled albumin in the vascular compartment at that time. The amount of albumin excreted in the urine was also measured and it indicated that not all of the albumin remaining in the body is in the vascular compartment. Figure 3 gives the percentage of the labelled albumin in the vascular compartment.

![Graph: Percent of Label Remaining in Body in the Intravascular Compartment](chart.png)

Note that by the fourth day less than half the body albumin is to be found in the vascular compartment and that the fraction seems to stabilize at a level of about $0.42$. It seems reasonable to conclude, therefore, that when distribution is completed approximately $42$ percent of the administered albumin is in the vascular compartment and $58$ percent outside it, at least in this patient. These results can be applied to normal albumin in the body so as to estimate its distribution in the body.
II Digoxin Experiment.

In many situations, the clinician desires to know the distribution of Digoxin among the various compartments of the human body. It is not completely known as to how many compartments are needed to explain the behavior of Digoxin distribution. This experiment was conducted to compare a two- and a three compartment model and to evaluate the transfer rates among the compartments. Digoxin was administered to five healthy male volunteers. A rapid intravenous injection of one mg of the drug was administered to the individuals. Blood samples were withdrawn repetitively over a period of seventy-two hours and samples were analysed using an \(^{125}\) radioimmunassay. The data are reported by Kramer et al. (1974).

Appropriate equations describing two- and three compartments, open models (with excretion) were fitted to the experimental data. It was found by Kramer et al. that the three compartment model is the simplest pharmacokinetic model consistent with the data observed in this experiment. The results obtained by fitting the two- and three compartment models to the data were compared as shown in table 1.1.

It was found by the authors that in four out of five subjects the reduction in residual sum of squares was significant due to the addition of a third compartment. Therefore, they hypothesized the system composed of three compartments.
Table 1.1
Comparison of Fits of Two- and Three Compartment Models to Digoxin Serum Level-Time Data

<table>
<thead>
<tr>
<th>Subjects</th>
<th>TWO COMPARTMENT</th>
<th>RESIDUAL SUM OF SQUARES</th>
<th>D.F</th>
<th>THREE COMPARTMENT</th>
<th>RESIDUAL SUM OF SQUARES</th>
<th>D.F</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>262.2</td>
<td>17</td>
<td></td>
<td>202.0</td>
<td>15</td>
<td></td>
<td>2.21</td>
</tr>
<tr>
<td>2</td>
<td>51.5</td>
<td>17</td>
<td></td>
<td>32.4</td>
<td>15</td>
<td></td>
<td>4.22**</td>
</tr>
<tr>
<td>3</td>
<td>307.0</td>
<td>16</td>
<td></td>
<td>74.4</td>
<td>14</td>
<td></td>
<td>22.0**</td>
</tr>
<tr>
<td>4</td>
<td>289.0</td>
<td>16</td>
<td></td>
<td>72.7</td>
<td>14</td>
<td></td>
<td>20.8**</td>
</tr>
<tr>
<td>5</td>
<td>321.0</td>
<td>16</td>
<td></td>
<td>135.0</td>
<td>14</td>
<td></td>
<td>9.66**</td>
</tr>
</tbody>
</table>

III Tropicamide Study

There are situations in drug-absorption analysis when direct determination of the time course of variation of drug levels in a body compartment (biophase) are not convenient or practicable. In such cases we examine the performance of drug-absorption analysis entirely from pharmacological data. The pharmacological method of drug-absorption analysis assumes identical behaviour of drug concentration levels in body fluids or tissue compartments and pharmacological time course data. A study was undertaken by Smolen and Schoenwald (1971) to develop a theoretical basis for the performance of drug-absorption analysis from data obtained from the observation of the time course of pharmacological response intensity following the administration of intravenous injection of a drug. Tropicamide was chosen as one of the model drugs because of the
relative ease of following its mydriatic activity. The experiment was conducted on rabbits with various solutions of Tropicamide. Four rabbits were selected on the basis of observed similarities in their pupillary response behaviour. Intravenous dosing of the rabbits was performed into their marginal ear veins. The doses ranged from 13.7 to 416.5 mcg/kg and were contained in fluid volumes of 0.3 to 0.5 ml which were rapidly injected. The mydriatic response intensity, I, was related to measured pupillary diameters at any time following the administration of the drug \( d_t \) and at time zero \( d_0 \) by

\[
I = \frac{d_t - d_0}{d_t}
\]

Measurements of the rabbits pupillary diameters were made at 10 minute intervals for 0-70 minutes. Some function \( f(I) \) which related the observed mydriatic response intensity I to the quantity of drug in the system was also calculated. I and \( f(I) \), given in table 1.2, are based on averages of the four rabbits.

Drug levels in blood were also measured and the experimenters claimed that the above data also give the same information.
Table 1.2
Data on Mydriatic Response Intensity I
and Quantity of Drug $f(I)$ at Various Times

<table>
<thead>
<tr>
<th>Dose mcg/kg</th>
<th>13.7</th>
<th>25.0</th>
<th>50.0</th>
<th>66.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>.23</td>
<td>12.7</td>
<td>.32</td>
<td>23.2</td>
</tr>
<tr>
<td>10</td>
<td>.15</td>
<td>6.7</td>
<td>.23</td>
<td>12.7</td>
</tr>
<tr>
<td>20</td>
<td>.09</td>
<td>3.6</td>
<td>.14</td>
<td>6.2</td>
</tr>
<tr>
<td>30</td>
<td>.05</td>
<td>1.8</td>
<td>.08</td>
<td>3.1</td>
</tr>
<tr>
<td>40</td>
<td>.02</td>
<td>0.7</td>
<td>.04</td>
<td>1.4</td>
</tr>
<tr>
<td>50</td>
<td>.01</td>
<td>0.3</td>
<td>.02</td>
<td>0.7</td>
</tr>
<tr>
<td>60</td>
<td>.01</td>
<td>0.3</td>
<td>.01</td>
<td>0.3</td>
</tr>
<tr>
<td>70</td>
<td>.00</td>
<td>0.0</td>
<td>.01</td>
<td>0.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>167</th>
<th>250</th>
<th>334</th>
<th>416.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$f(I)$</td>
<td>I</td>
<td>$f(I)$</td>
</tr>
<tr>
<td>0</td>
<td>.70</td>
<td>160.7</td>
<td>.84</td>
</tr>
<tr>
<td>10</td>
<td>.54</td>
<td>82.7</td>
<td>.63</td>
</tr>
<tr>
<td>20</td>
<td>.42</td>
<td>42.7</td>
<td>.49</td>
</tr>
<tr>
<td>30</td>
<td>.31</td>
<td>21.7</td>
<td>.37</td>
</tr>
<tr>
<td>40</td>
<td>.20</td>
<td>10.2</td>
<td>.26</td>
</tr>
<tr>
<td>50</td>
<td>.12</td>
<td>5.1</td>
<td>.17</td>
</tr>
<tr>
<td>60</td>
<td>.07</td>
<td>2.7</td>
<td>.10</td>
</tr>
<tr>
<td>70</td>
<td>.04</td>
<td>1.4</td>
<td>.05</td>
</tr>
</tbody>
</table>
IV Sulphate Experiment

The study of metabolic pattern of sulphate in humans is important in several contexts. It is not completely known whether radioactive sulphate is retained in the body or is excreted in the urine. A study was undertaken by Galambos and Cornell (1962) to develop a mathematical model to describe sulphate metabolism in humans. A patient was given intravenously 100 μc of S-35 labeled sodium sulphate of high specific activity. Data on radioactive counts of blood and urine were collected at 0.33, 1, 2, 3, 5, 8, 12, 24, and 72 hours following the injection.

The data on proportions $y_{1j}$ and $y_{2j}$ of radioactive tracer in two different compartments at various times $t_j$ are reported by Galambos and Cornell (1962) and are reproduced in table 1.3.

Table 1.3
Data on Proportions $y_{1j}$ and $y_{2j}$ of Radioactive Tracer in Two Different Compartments at Various Times $t_j$

<table>
<thead>
<tr>
<th>$j$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_j$</td>
<td>0.33</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>12</td>
<td>24</td>
<td>48</td>
<td>72</td>
</tr>
<tr>
<td>$y_{1j}$</td>
<td>0.84</td>
<td>0.79</td>
<td>0.64</td>
<td>0.55</td>
<td>0.44</td>
<td>0.27</td>
<td>0.12</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>$y_{2j}$</td>
<td>0.03</td>
<td>0.10</td>
<td>0.14</td>
<td>0.21</td>
<td>0.30</td>
<td>0.40</td>
<td>0.54</td>
<td>0.66</td>
<td>0.71</td>
</tr>
</tbody>
</table>

The primary interest for their study was to develop a mathematical model to explain the metabolic pattern of sulphate in humans. The authors concluded that a two-compartment model explains the pattern of sulphate in humans. Beauchamp and Cornell (1966) also analysed the same experiment using another statistical technique involving
simultaneous non-linear estimation procedure and concluded that a two-compartment model fits to the observed data.

V Studies in Physical Sciences

Determination of neutron diffusion parameters in Beryllium is of importance in many practical situations. An experiment was conducted by deSaussure and Silver (1959) with the object of determining the neutron diffusion parameters in room-temperature Beryllium. The data on counts describing the decay of the neutron density in a medium-size assembly of beryllium was observed. Observations were made at equally spaced time intervals of time 0.1 milliseconds. Table 1.4 describes the data of the neutron density.

Table 1.4
Decay of the Neutron Density in a Medium-sized Assembly of Beryllium

<table>
<thead>
<tr>
<th>t = 0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>y(t)</td>
<td>100145</td>
<td>78005</td>
<td>60305</td>
<td>46485</td>
<td>36205</td>
<td>28275</td>
<td>21705</td>
<td>16955</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>t = 9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>y(t)</td>
<td>10085</td>
<td>7835</td>
<td>6165</td>
<td>4782</td>
<td>3780</td>
<td>2915</td>
<td>2249</td>
<td>1752</td>
</tr>
</tbody>
</table>

Cornell (1962) examined this experiment and found that a single compartment describes the data completely.

Another experiment was conducted to study the distribution of background pulses generated in a proportional counter by neutron interaction with walls and gas plus pulses due to circuit noise. The logarithm of frequencies and pulse heights t are displayed in table 1.5.
Table 1.5

Logarithms $y(t)$ of Frequencies of Pulse Heights $t$
Generated in a Proportional Counter

<table>
<thead>
<tr>
<th>$t$ = 0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y(t) =$ 10.430</td>
<td>4.703</td>
<td>2.327</td>
<td>1.140</td>
<td>0.615</td>
<td>0.325</td>
<td>0.170</td>
<td>0.117</td>
</tr>
</tbody>
</table>

$y(t) = \begin{array}{cccccccc}
0.05 & 0.040 & 0.046 & 0.022 & 0.036 & 0.021 & 0.018 & 0.016
\end{array}$

A two compartment model was fitted to the above experiment by Cornell (1962) and was found adequate. We examine these experiments in later chapters and propose various alternative procedures. Many of the proposed procedures give improved results.
CHAPTER II

COMPARTMENTAL MODELS

2.1. Introduction

In the study of biological systems, many problems require that the system be considered to be composed of several subsystems called compartments. The studies in pharmacokinetics and bioavailability are specially designed so as to determine the rates of exchange of chemical compound or drug among various compartments. In many biological studies, radioactive tracers are used and the study of the system as composed of several compartments becomes very explicit. For example in animal studies actual measurements in various compartments can be made through sacrificing the animal. In humans also such studies have proved useful since the compounds can be traced in the organs. The behavior of a compartmental system is easily expressed in terms of a system of first order linear differential equation. The study of a complex system with the help of mathematical models has received serious attention in recent years by several workers. Some of the early references in this area are Berman and Schoenfeld (1956) and Sheppard (1962). The concept of dividing a biological system into a number of fixed compartments is merely an aid in analysis, since the various states or sites contain finer structure. Sheppard (1962) gives examples of such compartmentalized
systems as follows:

(i) **Catenary System**: This system involves \( p \) compartments that are arranged in a chain-like manner such that each compartment has non-zero transition rates only with the compartments adjacent to it as is pictured in figure (2.1)

![Figure (2.1) Catenary System](image)

(ii) **Mammilary System**: This system involves \((p - 1)\) peripheral compartments which have transition rates with a central compartment but no transition between the \((p - 1)\) peripheral compartments as illustrated in figure (2.2).

![Figure (2.2) Mammilary System](image)

### 2.2. Deterministic Theory

We consider now a general system of \( p \) interconnected compartments. Figure (2.3) represents a general \( p = 4 \) compartment system with \( p^2 = 16 \) parameters.
Squares representing compartments. A 4 compartment system

We assume that the flux direction has its own channel associated with it, so that the flux from compartment $i$ to compartment $j$ is generally unequal to the flux from $j$ to $i$. Further it is assumed that material may be introduced into each compartment from the external source at a steady rate into the system. The net flux into any compartment equals the net outflux at all times.

Initially a known amount of labelled material is injected into one compartment. Because of the interactions among the compartments, the labelled material is distributed and is, in principle, observable "henceforth" in all compartments. The labelled material that is introduced at time zero creates an observable transient of labelled material that is designed to provide information about the steady-state characteristic of the system (e.g., the material fluxes among the compartments).

We assume that the amount of labelled material introduced is so small as to leave unaltered the steady-state behavior of the system. The ratio of the labelled material to unlabelled material in a compartment is very small compared to one at all times. The total amount of material in a compartment consists of a homogeneous
A mixture of labelled and unlabelled material. Homogeneity implies that the behaviour of the labelled material is representative of the behaviour of the unlabelled material in a compartment. Homogeneity also implies that when labelled material enters a compartment, it is instantaneously "mixed". Further we assume that the turnover rates are independent of time. This assumption is known as the assumption of first order rate constants in pharamokinetic literature.

We define the following

$$\rho(t) = (\rho_1(t), \rho_2(t), \ldots, \rho_p(t))'$$

be the total amount of material in the system, with $\rho_i(t)$ being the total amount in compartment $i$ at time $t$.

$\lambda_{ij}$ is the material transport rate from compartment $j$ to compartment $i$ per unit amount of material in the $j$th compartment; $(i \neq j)$ $\lambda_{ij}$ is also called the fractional turnover rate or kinetic rate constant of compartment $j$ with respect to $i$.

$(-\lambda_{ii})$ is the rate at which the total amount of material in the $i$th compartment is replaced, or the total turnover rate of the $i$th compartment.

$$L = \begin{pmatrix}
\lambda_{11} & \lambda_{12} & \lambda_{13} & \cdots & \lambda_{1p} \\
\lambda_{21} & \lambda_{22} & \lambda_{23} & \cdots & \lambda_{2p} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\lambda_{p1} & \lambda_{p2} & \lambda_{p3} & \cdots & \lambda_{pp}
\end{pmatrix}$$
By the assumption of constant turnover rates, we get the following differential equation to represent the $i$th compartment.

$$\frac{d P_i(t)}{dt} = \lambda_{1i} P_1(t) + \lambda_{2i} P_2(t) + \cdots + \lambda_{pi} P_p(t), \quad i = 1, 2, \ldots, p.$$ 

The above system of differential equations can be written in matrix notation as

$$\frac{d \mathbf{p}(t)}{dt} = \mathbf{L} \mathbf{p}(t) \quad \ldots \quad (2.1)$$

Equations (2.1) are called compartment equations. Implicit in Eq. (2.1) is the assumption that the transit time for material flux between any two compartments is negligible. Note that the elements of the matrix $\mathbf{L}$ must satisfy the conditions:

$$\lambda_{ij} > 0, \quad i, j = 1, 2, \ldots, p; \quad i \neq j \quad \ldots \quad (2.2)$$

$$\sum_{i=1}^{p} \lambda_{ij} \leq 0, \quad j = 1, 2, \ldots, p \quad \ldots \quad (2.3)$$

From Eqs. (2.2) and (2.3) it follows that the diagonal elements $\lambda_{ii}$ are subject to the condition

$$\lambda_{ii} \leq 0, \quad i = 1, 2, \ldots, p, \quad \ldots \quad (2.4)$$

and the equality sign holds only if $\lambda_{ij} = 0$ for all $j$ with $i$ fixed. The excretion rate per unit volume from the $i$-th compartment is defined as
Definition 2.1. A compartment $i$ is called leaky if $\lambda_{0i} > 0$.  

Definition 2.2. A compartment $i$ is called leakproof if $\lambda_{0i} = 0$.  

Definition 2.3. If every compartment is leakproof, the system of $p$ compartments is said to be a closed system, that is, isolated from its environment.  

Definition 2.4. If at least one compartment is leaky, the system of $p$ compartments is said to be an open system.  

The absolute amount of labelled material in a compartment at time $t > 0$ is not generally observable, but the specific activity of the material is. When the material is labelled by a radioactive marker, the specific activity is the amount of radioactivity (in microcuries, counts per minute, etc.) per unit amount of the material in question. Therefore, it is desirable to express compartment equations in terms of $\eta = \eta(t)$ defined by  

$$\eta(t) = \mathcal{D} \eta(t)$$  

(2.6)  

where $\eta = (\eta_1(t), \eta_2(t), \ldots, \eta_p(t))'$ is a column vector; $\eta_i(t)$ is the ratio of labelled to unlabelled material in the $i$th compartment, and $\mathcal{D}$ is a diagonal matrix with diagonal elements $D_i$ representing the total amount of material (labelled and unlabelled) in the $i$th compartment. $D_i$ is referred to as the compartment size or volume of distribution.
Note that if \( \eta_i(t) \) is a fluid concentration, \( D_i \) is indeed a volume. For radioactive markers, \( \eta_i(t) \) is directly proportional to the specific activity of the labelled material. We assume that \( D_i \) are constants.

From the definition of \( D \) and the statement of the problem it follows that all the diagonal elements of \( D \) are positive so that the inverse \( D' \) exists. Combining Eq. (2.6) with Eq. (2.1) we express the Eq. (2.6) as follows:

\[
\frac{d \eta(t)}{dt} = D \frac{d \eta(t)}{dt} = L D \eta(t)
\]

or

\[
\frac{d \eta(t)}{dt} = D^{-1} L \eta(t) = \gamma \eta(t)
\]

Let \( \gamma = (v_{ij}) \), we note that \( v_{ii} = d_{ii} \), non-diagonal elements \( v_{ij} \) give the rate of transport of material from the \( j \)th compartment to the \( i \)th compartment, per unit amount of material in the \( i \)th compartment.

The problems in the study of a compartmental system are the following: Given some knowledge of the compartmental system in the form of an experimental determination of some of the components of the vector \( \eta(t) \), we want to find the matrices \( D \) and \( \gamma \) for which \( \eta(t) \) is the solution to Eq. (2.7).

Observe that if the vector \( \eta(t) \) is completely determined, the
matrix $\mathbf{V}$ is uniquely determined. Usually, not every compartment is accessible to measurement, the data are incomplete and $\mathbf{V}$ may no longer be uniquely determined. The solution to the equations (2.7) is given in terms of the following. It is well known that eigenvalues of $\mathbf{V}$ are negative. Let $-\lambda_1, -\lambda_2, \ldots, -\lambda_p$ be these eigenvalues. Assume that they are all distinct.

Let

$$
\mathbf{M} = \begin{pmatrix}
-\lambda_1 & 0 & \cdots & 0 \\
0 & -\lambda_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & -\lambda_p
\end{pmatrix}
$$

Further let

$$
\mathbf{\gamma}(t) = (e^{-\lambda_1 t}, e^{-\lambda_2 t}, \ldots, e^{-\lambda_p t})'.
$$

The solution of the Eq. (2.7) is given by

$$
\mathbf{\gamma}(t) = \mathbf{A} \mathbf{\gamma}(t)
$$

The matrix $\mathbf{A} = (\alpha_{ij})$ is a constant matrix of order $p$.

To express the matrix $\mathbf{A}$, we proceed as follows.

$$
\frac{d}{dt} \mathbf{\eta}(t) = \mathbf{A} \frac{d}{dt} \mathbf{\gamma}(t) = \mathbf{A} \mathbf{M} \mathbf{\gamma}(t)
$$

Since

$$
\frac{d}{dt} \mathbf{\gamma}(t) = (-\lambda, e^{-\lambda_1 t}, -\lambda_2 e^{-\lambda_2 t}, \ldots, -\lambda_p e^{-\lambda_p t})',
$$

$\mathbf{M} \mathbf{\gamma}(t)$
\[ \frac{d \eta(t)}{dt} = \sim \eta(t) = \sim V \sim \eta(t) \]  

\[ \sim AM = \sim VA \]  

or \[ V = AM^{-1} \]

This equation expresses the fact that the columns of \( A \) are the eigen vectors of \( V \).

Since the initial conditions of the Eq. (2.9) give \( p \) constraints as follows,

\[ \eta(0) = A \sim \eta(0) \]  

So only \((p^2 - p)\) elements of \( A \) can be independently determined. In addition there are \( p \) parameters in \( M \), so, in total there are \( p^2 \) independent parameters. However, when only one compartment is observed we have information on only \((2p - 1)\) parameters. This can be seen through the solution of the one compartment equation, having one initial condition, giving \( p - 1 \) independent parameters as well as \( p \) parameters of \( M \).

Suppose now a compartment is observed at various times giving information on \((2p - 1)\) parameters: the \( p \) eigen values \(-\lambda_i\) and the elements of a row of \( A \) only \((p - 1)\) of which are independent because of \((2.12)\). Observation on an additional compartment yields the values of only \((p - 1)\) additional parameters because the \( \lambda_i \)
are the same, regardless of which compartment is observed. Let $K$ be the number of parameters that are not determined by observation so that if $j$ compartments are observed, then $k = (p - 1)(p - j)$. There is a $k$-parameter family of matrices $V$ that are consistent with the data.

Berman and Schoenfeld (1956) have suggested an ad hoc iterative procedure for the determination of $V$. For details, reference can be made to their paper. An alternative suggestion of Berman (1960) is to make a least squares fit of the data directly to the matrix $V$, although it is recognized that this procedure suffers from the same unsatisfactory features of the ad hoc, iterative procedure of Berman and Schoenfeld.

Mathematically, the only convincing solution to this dilemma is to find the entire class of compatible matrices $V$. If the number of compartments is small, this is feasible. In a situation where the experimenter is interested in determining a unique matrix $V$, he can do this by imposing additional constraints. These constraints may be suggested by physiological, biological, or chemical considerations of the compartment models. For example, he assumes a given compartment is leak proof; or a given transport rate between two compartments is zero. With $p$-compartments the mammillary and catenary systems are the simple ones. They have $(2p - 1)$ non-zero turnover rate constants and the experimenter observes only one compartment. These two simple systems are the commonly hypothesized systems by the experimenter and
have been extensively studied in literature. Recently, two theorems have been proved by Shah (1974) to guide the experimenter in collecting sufficient data to estimate the turnover rate constants in this case and are stated below without proof.

Theorem 1

If the data are available for all time from one compartment of the \( p \)-compartment catenary system having \( 2p - 1 \) unknown kinetic parameters, then for a single dose, the first order rate constants are not uniquely estimable in any catenary system having more than two compartments.

Theorem 2

If the data are available for all time from any one compartment of a \( p \)-compartment non-catenary system having a maximum of \( (2p - 1) \) unknown rate constants and having \((1,p)\)th element of \( \mathbf{y}_j \neq 0 \) for \( j = 1,2,\ldots,p - 1 \), then the rate constants of the system are uniquely estimable.

Example 1

Consider \( p = 2 \) and let us assume that from the data, \( a_{11} \) and \( a_{12} \) are the estimates of the parameters \( \alpha_{11} \), \( \alpha_{12} \), \( \lambda_1 \), \( \lambda_2 \) respectively. The matrix \( \mathbf{Y} \) consists of three unknown rate constants namely \( v_{21} \), \( v_{12} \) and \( v_{02} \). The roots of the matrix \( \mathbf{Y} \) are given by the following equations

\[
 v_{21} + v_{12} + v_{02} = \hat{\lambda}_1 + \hat{\lambda}_2
\]  

\[ \ldots \] (2.13)
where  \( v_{02} = \hat{\lambda}_{1} \lambda_{2} \) ... (2.14)

Using (2.11) we have

\[ v_{12} = -(a_{11} \hat{\lambda}_{1} + a_{12} \hat{\lambda}_{2}) \] ... (2.15)

It can be seen from the equations (2.13) - (2.15) that \( v_{21}, v_{12} \) and \( v_{02} \) are now uniquely estimatable. That is when \( p = 2 \), there is a unique solution of the matrix \( V \).

**Example 2**

Consider \( p = 3 \) and assume that \( a \)'s and \( \lambda \)'s can be estimated from the data. The matrix \( V \) consists of five unknown parameters, namely \( v_{21}, v_{12}, v_{32}, v_{23}, v_{03} \). The characteristic roots of the matrix \( V \) are then given as follows.

\[ v_{21} + v_{32} + v_{03} + v_{12} + v_{23} = \hat{\lambda}_{1} + \hat{\lambda}_{2} + \hat{\lambda}_{3}, \] ... (2.16)

\[ v_{21} v_{32} + v_{32} v_{03} + v_{21} v_{03} + v_{21} v_{23} + v_{12} v_{03} = \hat{\lambda}_{1} \hat{\lambda}_{2} + \hat{\lambda}_{2} \hat{\lambda}_{3} + \hat{\lambda}_{3} \hat{\lambda}_{1}, \] ... (2.17)

\[ v_{21} v_{32} v_{03} = \hat{\lambda}_{1} \hat{\lambda}_{2} \hat{\lambda}_{3}. \] ... (2.18)

Using (2.11) we have further,

\[ v_{21} v_{32} = \sum_{i} a_{i1} \hat{\lambda}_{i} \] ... (2.19)

Thus there are only four equations (2.16) - (2.19) for determining five unknowns making it impossible to estimate all the rate constants of the experiment. In this way theorems 1 and 2 help guide the
researcher in planning a pharmacokinetic experiment.

2.3. Some Statistical Considerations

In the preceding section we have discussed the mathematical theory of deterministic models. The phenomena under study needs further investigation as the mathematical theory is not sufficient to account for the observed data completely. To estimate parameters in these models, it is necessary to introduce probabilistic considerations. The most obvious, and indeed the classical, way of introducing statistical considerations involves assuming that each observed point consists of two components, a "true" component, whose behaviour is described by the mathematical theory, and another which is "random" superimposed on the true one. This second component can also include the errors of measurements.

Model I

Suppose we observe one compartment, then we assume the model as follows.

\[ y(t) = \sum_{k=1}^{p} \alpha_k e^{-\lambda_k t} + \varepsilon(t), \quad \ldots \quad (2.20) \]

where

\[ y(t) \] are observations at times \( t = 0,1,2,\ldots,n \).

We assume that the mean and variance of errors \( \varepsilon(t) \) are given by

\[ E[\varepsilon(t)] = 0; \quad \text{Var}[\varepsilon(t)] = \frac{1}{\kappa_t}, \]
and the errors are uncorrelated, that is

$$E[e(t) \cdot e(t')] = 0, \quad t \neq t', \quad t, t' = 0, 1, 2, \ldots, n. \quad \ldots \quad (2.21)$$

The method of least squares is generally used for estimating values of the rate constants from experimental observations. The least squares computations, have several unusual features when applied to linear combination of exponentials. We shall enumerate some of these. They have also been discussed by Cornfield et al. (1960).

Firstly from the point of view of computation, the usual iterative procedures fail to converge. This behaviour is a consequence of an unusual feature of the sum of squares surface, $S$, where

$$S = \sum_{t=0}^{n} w_t [y(t) - \sum_{k=1}^{p} \alpha_k \lambda_t^k]^2 \quad \ldots \quad (2.22)$$

as it possesses more than one minima. In case $p = 2$ by interchanging the estimated values of $\lambda_1$ and $\lambda_2$ to $\lambda_2$ and $\lambda_1$, and the estimated values of $\alpha_1$ and $\alpha_2$ to $\alpha_2$ and $\alpha_1$, we get the same minima of $S$ and therefore $S$ has two minima. The existence of two minima implies the existence of a stationary point in $S$ at some other intermediate position. Therefore all partial derivatives are zero at this point as well as at the points of the other two minima. For the iterative procedures, the needed initial values of the $\lambda$'s, if selected in the neighborhood of the intermediate stationary point may lead to solutions which do not give absolute minima.

There are various ways of avoiding the difficulties introduced
by the existence of the intermediate stationary points. One involves computing \( S \) in Eq. (2.22) for a sufficiently large number of combinations of \( \lambda_i \), so that the minimum point is physically identified. In a later chapter, we will present a method of finding good initial estimates of \( \lambda_i \)'s.

Model II

Difference Equation Model

In the model given by Eq. (2.20) it is assumed that measurements are made on the same individual for various time intervals. It is quite reasonable to expect that errors in successive measurements are correlated. Now we assume that errors on successive measurements are correlated. For our discussion we assume that the compartmental system is a catenary system.

The compartmental equations given by (2.7) are simultaneous linear differential equations in \( p \) unknowns. \((p - 1)\) unknowns may be eliminated to yield the \( p \)th order linear differential equations with constant coefficients. The \( p \)th order linear differential equation in the \( i \)-th unknown is given by

\[
\frac{d^p \eta_i(t)}{dt^p} + \left( \sum_{k=1}^{p} \lambda_k \right) \frac{d^{p-1} \eta_i(t)}{dt^{(p-1)}} + \left( \sum \lambda_k \lambda_{k'} \right) \frac{d^{p-2} \eta_i(t)}{dt^{(p-2)}} + \ldots + \lambda_1 \lambda_2 \ldots
\]

\[\lambda_p \eta_i(t) = 0 \quad \ldots \quad (2.23)\]

The particular solution of the above \( p \)th order differential
equation in $\eta_i(t)$ is given in (2.9) as follows.

$$\eta_i(t) = \sum_{k=1}^{p} \alpha_{ik} \lambda_k^t \quad i = 1,2,\ldots,p$$  \hspace{1cm} (2.24)

The above solution contains $(2p - 1)$ unknowns and $\alpha_{ik}$ are dependent among themselves for each $i$ according to Theorem 1. We can estimate only $(p + 1)$ parameters uniquely for a catenary system when data are available from one compartment. That is, we can estimate only $p$ eigenvalues $\lambda_k$'s and one $\alpha_{ik}$. Hence if our primary interest is in estimating $\lambda_k$'s most efficiently, and assume observations are available at equispaced time intervals, we can use the following approach.

Approximate the $p$th order differential equation (2.23) by the $p$th order difference equation. For notational convenience let us drop the subscript $i$ and assume that observations are recorded at successive time intervals on the $i$th compartment. Then the $p$th order difference equation for the $i$th compartment is

$$\Delta^p \eta(t) + \sum_{k=1}^{p} \lambda_k \Delta^{p-1} \eta(t) + \ldots + \lambda_1 \lambda_2 \ldots \lambda_p \eta(t) = 0 \quad \ldots \quad (2.29)$$

where $\Delta \eta(t) = \eta(t + 1) - \eta(t)$ and $\frac{d^j \eta(t)}{dt^j}$ is replaced by $\Delta^j \eta(t)$;

$$\Delta^j \eta(t) = \Delta^{j-1} \eta(t + 1) - \Delta^{j-1} \eta(t) \quad i = 1,2,\ldots,p$$

$$\Delta^0 = 1$$
Replacing expectation by its observable random value in (2.25),
we have the following difference model,

\[ \Delta^P y(t) + \sum_{k=1}^{P} \lambda_k \Delta^{P-1} y(t) + (\sum \lambda_k \lambda_2 \ldots \lambda_P) y(t) = u_{t+p} \]

We assume that \( u_{t+p} \) is the error component,

\[ E[y(t)] = \eta(t) \]

Expanding (2.26), we obtain

\[ y(t + p) + c_1 y(t - 1 + p) + c_2 y(t - 2 + p) + \ldots + c_p y(t) = u_{t+p} \ldots (2.27) \]

where \( c_i, i = 1, 2, \ldots, p \), are linear combinations of

\[ \sum_{k=1}^{P} \lambda_k, \sum_{k=1}^{P} \lambda_k \lambda_2 \ldots \lambda_P \], given by the following system of equations

\[ \begin{bmatrix}
1 & 0 & \ldots & 0 & 0 & 0 & 0 \\
-1 & 1 & \ldots & 0 & 0 & 0 & 0 \\
-1 & -1 & \ldots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
(-1)^{p-2} & (-1)^{p-3} & \ldots & 1 & 0 & 0 & 0 \\
(-1)^{p-1} & (-1)^{p-2} & \ldots & 1 & 0 & 0 & 0 \\
(-1)^{p} & (-1)^{p-1} & \ldots & 1 & -1 & 1 & 1
\end{bmatrix} \begin{bmatrix}
1 \\
\lambda_1 \\
\lambda_2 \\
\vdots \\
\lambda_{p-2} \\
\lambda_{p-1} \\
\lambda_p
\end{bmatrix} = \begin{bmatrix}
1 \\
c_1 \\
c_2 \\
\vdots \\
c_{p-2} \\
c_{p-1} \\
c_p
\end{bmatrix} \ldots (2.28)
where, for \( r = 1, 2, \ldots, p \), the elementary symmetric functions \( \Lambda_r \) equals the sum of all possible products

\[
\Lambda_r = \sum \lambda_1 \lambda_2 \cdots \lambda_r
\]

summation is over \( \binom{p}{r} \) different combinations.

In abbreviated form (2.29) can be written as

\[
Q \Lambda = C
\]

where \( C = (1, c_1, c_2, \ldots, c_p)' \)

\( \Lambda = (1, \Lambda_1, \Lambda_2, \ldots, \Lambda_p)' \)

and

\[
Q = \left((q_{ij})_{ij}\right)_{(p+1)(p+1)}
\]

with \( q_{ij} = 0 \) if \( i < j \)

\( = 1 \) if \( i = j \)

\( = (-1)^{i-j(p+1-j)} \) if \( i > j \)

Notice that,

\[
1 + \sum_{i=1}^{p} c_i = \lambda_1 \lambda_2 \cdots \lambda_p
\]

(2.30)

In chapter IV, we shall investigate this model further.

Model III

Stochastic Compartmental Models

So far we considered deterministic flow pattern among
compartments. In this section we consider a discrete population of particles in a steady state compartmental system where the transitions are stochastic. The importance of the study of stochastic compartmental models have been stressed by many authors, for a recent development, see Rustagi (1964, 1965).

Development in stochastic compartmental analysis are more recent. Bartholomay (1958) studied a one-compartmental stochastic model. Cornfield et al (1960) also pointed that the stochastic compartmental model is more realistic and should be investigated. Matis and Hartley (1971) considered probability distribution theory of general p-compartmental models and developed estimation procedure. Uppuluri, Feder and Shenton (1967) considered stochastic one-compartment models for multiple dosing and random decay for continuous population. Purdue (1974) studied the stochastic theory of two-compartmental systems. We consider here the case of a finite tracer population in a compartmental system with probabilistic flow.

Differential Equations Describing p-Compartment Stochastic Model

Consider a p-compartment general system, where every compartment is connected to each other and to the system exterior. In this system, there are \( p^2 \) parameters. Figure (2.4) represents a general p-compartment system with \( p^2 = 16 \) parameters.

![A 4 Compartment System](image)
For \( 0 < \tau \leq t \), let

\[
\nu_{ji} \Delta + o(\Delta) = P_{ji}(\tau, \tau + \Delta) = P_{ij}[\text{an individual in compartment } i \text{ at time } \tau \text{ will be in compartment } j \text{ at time } \tau + \Delta], \ i \neq j
\]

\[
1 + v_{ii} \Delta + o(\Delta) = P_{ii}(\tau, \tau + \Delta), \ i, j = 1, 2, \ldots, p. \quad \ldots \quad (2.31)
\]

\[
\nu_{ii} \Delta + o(\Delta) = P_{ii}(\tau, \tau + \Delta) = \Pr[\text{an individual in compartment } i \text{ at time } \tau \text{ will be in exterior of the system at time } \tau + \Delta]
\]

\[
\nu_{ij} \quad \text{are transition intensities or turnover rates as defined in (2.7) and } o(\Delta) \text{ represents any function such that } \lim_{\Delta \to 0} \frac{o(\Delta)}{\Delta} = 0. \text{ The probability of more than one migration in } \Delta \text{ is } o(\Delta). \text{ We assume that } v_{ij} \text{ are independent of time.}
\]

Let \[ v_{ii} = - \sum_{j \neq i} v_{ji} \quad \ldots \quad (2.33) \]

For convenience we introduce the transition intensity matrix

\[
V = \begin{pmatrix}
  v_{11} & v_{21} & \cdots & v_{p1} \\
  v_{12} & v_{22} & \cdots & v_{p2} \\
  \vdots & \vdots & \ddots & \vdots \\
  v_{1p} & v_{2p} & \cdots & v_{pp}
\end{pmatrix} \quad \ldots \quad (2.33)
\]

and

\[
U = (v_{01}, v_{02}, \ldots, v_{0p})', \quad \ldots \quad (2.34)
\]
For a time interval \((0,t)\) for \(0 \leq t < \infty\), let

\[
P_{ji}(t) = P_{ji}(0,t), \; i,j = 1,2,\ldots,p; P_{oi}(t) = P_{oi}(0,t),
\]

\[i = 1,2,\ldots,p\]

with \(P_{ii}(0) = 1\), \(P_{oi}(0) = 0\), \(P_{ji}(0) = 0\); \(i,j = 1,2,\ldots,p\).

and further let

\[
P(t) = \begin{pmatrix}
P_{11}(t) & P_{21}(t) & \cdots & P_{pl}(t) \\
P_{12}(t) & P_{22}(t) & \cdots & P_{pl}(t) \\
\vdots & \vdots & \ddots & \vdots \\
P_{1p}(t) & P_{2p}(t) & \cdots & P_{pp}(t)
\end{pmatrix} \ldots (2.35)
\]

\[
P(t) = (P_{o1}(t) P_{o2}(t) \cdots P_{0p}(t))' \ldots (2.36)
\]

with \(P(0) = I\) and \(P(0) = Q\)

Assumption 1. The system is closed: Whatever may be \(t \geq 0\)

and for every \(i\), we have

\[
\sum_{j=0}^{p} P_{ji}(t) = 1 \ldots (2.37)
\]

so that the intensities and the transition probabilities have

the relations:

\[
v_{ji} = \frac{d}{dt} P_{ji}(t)\big|_{t=0} \quad i,j = 1,2,\ldots,p .
\]

\[
v_{oi} = \frac{d}{dt} P_{oi}(t)\big|_{t=0} .
\]
Assumption 2. The matrix $V_s$ is of rank $p$ and the matrix $U_s$ is not a zero matrix. Therefore none of the compartments $i$ is an absorbing compartment, and there is excretion to the system exterior.

Consider $\tau < t < t + \Delta$

then,

$$P_{ki}(\tau, t + \Delta) = P_{ki}(\tau, t)P_{kk}(t, t + \Delta) + \sum_{j \neq k} P_{ji}(\tau, t)P_{kj}(t, t + \Delta)$$

... (2.38)

Using (2.31) the Eq. (2.38) can be expressed as

$$\frac{P_{ki}(\tau, t + \Delta) - P_{ki}(\tau, t)}{\Delta} = P_{ki}(\tau, t)v_{kk} + \sum_{j \neq k} P_{ji}(\tau, t)$$

$$\frac{P_{ki}(t, t + \Delta)}{\Delta} + o(\Delta)$$

Taking limit as $\Delta \to 0$, we have

$$\frac{\partial}{\partial t} P_{ki}(\tau, t) = \sum_{j=1}^{p} P_{ji}(\tau, t)v_{kj}$$

... (2.39)

Since $v_{ji}$ do not depend on time, the system of differential equations (2.39) are

$$\frac{d}{dt} P_{kk}(t) = \sum_{j} P_{ji}(t)v_{kj}$$

... (2.40)

These equations are known as Kolmogorov forward differential equations.

The corresponding matrix equation describing the $p$-compartment
stochastic model is

\[ D \tilde{P}(t) = \tilde{P}(t) \tilde{V}_s \]

or

\[ (D - \tilde{V}_s') \tilde{P}(t) = 0 \quad \ldots \quad (2.41) \]

where

\[ D = \left( \begin{array}{cc} \frac{d}{dt} & 0 \\ \frac{d}{dt} & 0 \\ 0 & \frac{d}{dt} \end{array} \right) \]

and \( \tilde{P}'(t) \) and \( \tilde{V}_s' \) are the transposed matrices of \( \tilde{P}(t) \) and \( \tilde{V}_s \) respectively.

In chapter VI, we discuss the solution of (2.41) for determining the population sizes in \( p \)-compartments, along with the resultant probability distributions. The probability generating function of the joint distribution of the population sizes of all the compartments at time \( t \) will be obtained. Least squares estimation of the parameters \( v_{ji} \), compartmental analysis regression function and the covariance kernel of observation will also be obtained.
CHAPTER III
DETERMINATION OF INITIAL ESTIMATES

3.1 Introduction

The methods of estimation, specially method of maximum likelihood for estimating parameters in the models of the compartmental systems proposed earlier, require iterative procedures. For successful implementation of iterative procedure, one needs 'good' initial values, of parameters appearing in a non-linear fashion in the model, for example (2.20). Sometimes the initial estimates may provide the most consistent estimates of the parameters if facilities for computation of iterative least squares estimates are not available.

In this chapter, we show that the method of partial totals may be used to fit linear combinations of any number of exponentials using data taken at equally spaced intervals of time. Also we show that the method of partial totals though not generally efficient, is consistent. For this method, it is not necessary to assume that all observations have the same distribution, therefore it can be utilized sometimes when other methods of estimation of such non-linear models are inappropriate.

The estimation procedure is developed in detail and is illustrated for two simple models. The consistency of the estimators is considered and a numerical comparison with methods proposed by other authors is made. The method for estimation of parameters appearing non-linearly in the models when several compartments are observed simultaneously,
is also described. This method resembles those of Cornell (1962) and Agha (1971), but we construct different partial totals.

3.2 General Estimation Procedure

Consider the model of equation (2.20) such that

$$E[Y(t)] = \sum_{k=1}^{\infty} \alpha_k e^{-\lambda_k t}, \ t=0,1,2,...,n$$  \hspace{1cm} (3.1)

We assume that $\alpha_k \neq 0$, $\lambda_k > 0$ for all $k$ and $\lambda_i \neq \lambda_j$. Also to implement the procedure, we assume $n = 2mp-1$. Let $\rho_k = \exp(-\lambda_k)$, then (3.1) is given by

$$E[Y(t)] = \sum_{k=1}^{p} \alpha_k \rho_k^t$$ \hspace{1cm} (3.2)

We assume that the observations are specified only at equally spaced values of $t$. Number of observations is equal to, $n+1 = 2pm$, that is $m$ times the number of parameters in the model. Notice that no assumptions are being made about error variances and the independence of the errors $e(t)$ in the model (2.20). The estimation procedure is as follows: Partitioning the sample values into $2p$ sums, $s_h$, given by,

$$s_h = \sum_{i=0}^{2p(m-1)+h} y(h+2pi)$$ \hspace{1cm} (3.3)

$$h=1,2,...,2p.$$

These partial totals, $s_h$, have expectations $S_h$ given by

$$E[s_h] = S_h = \sum_{h=1}^{2p} E[Y(t)] = \sum_{t=h+2pi}^{2p(m-1)+h} \left( \sum_{k=1}^{p} \alpha_k \rho_k^t \right)$$

$$= \sum_{k=1}^{p} \alpha_k \rho_k^h \sum_{i=0}^{2p(m-1)+h} \frac{1-\rho_k}{1-\rho_k^{2p}}$$ \hspace{1cm} (3.4)
Since $\alpha_k$ are distinct, it is easy to verify that the polynomial, $S_n$, satisfies the pth order difference equation

$$\sum_{i=1}^{p+1-1} (-1)^{p+1-i} \prod_{e=i}^{i-1} S_{e+1} = 0 \quad e=0,1,2,...,p-1$$

(3.5)

where, for $r=1,2,...,p$ the elementary symmetric functions $A_r$ equals the sum of all possible products,

$$A_r = \Sigma(\alpha_{k_1} \alpha_{k_2} \cdots \alpha_{k_r})$$

(3.5)

summation is over $\binom{p}{r}$ different combinations. Assume $A_0 = 1$.

Replacing $S_n$ by the corresponding observed partial totals $s_n$ in (3.5), we obtain estimators $\hat{A}_j$ of the $A_j$ from the equations

$$\sum_{i=1}^{p+1-1} (-1)^{p+1-i} \prod_{e=i}^{i-1} s_{e+1} = 0 \quad e=0,1,...,p-1.$$

(3.6)

Let

$$\begin{bmatrix}
1 & 2 & j & \cdots & p \\
1 & s_2 & s_j & \cdots & s_p \\
s_2 & s_3 & s_{j+1} & \cdots & s_{p+1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & s_p & s_{p+1} & s_{j+p-1} & s_{2p-1}
\end{bmatrix}$$

and

$$\begin{bmatrix}
1 & 2 & p+1-j & \cdots & p \\
1 & s_2 & s_{p+1} & s_p & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & s_p & s_{p+1} & s_{2p} & \cdots \\
1 & s_{2p-1} & s_{2p} & \cdots & s_{2p-1}
\end{bmatrix}$$

(3.6)
That is, \( \Lambda_j \) is obtained by replacing the \((p+1-j)\)th column by the column vector \((s_{p+1}, s_{p+2}, \ldots, s_{2p})'\). Then the solution of (3.6) is given by

\[
\Lambda_j = (-)^{j+1} \frac{|\Lambda_j|}{|\Lambda|}, \quad j=1,2,\ldots,p
\]  

(3.7)

Since the \( \Lambda_j \) estimate the elementary symmetric functions of the \( \rho_k \), estimators \( r_k \) of the \( \rho_k \) is given by the \( p \) roots of the equation

\[
x^n - \sum_{j=1}^{p} \Lambda_j x^{n-1} + \Lambda_2 x^{n-2} + \cdots + (-1)^p \Lambda_p = 0
\]

(3.8)
or equivalently by

\[
x^n - \frac{|\Lambda_1|}{|\Lambda|} x^{n-1} - \frac{|\Lambda_2|}{|\Lambda|} x^{n-2} - \cdots - \frac{|\Lambda_p|}{|\Lambda|} = 0
\]

(3.9)

For estimators \( \lambda_k \) of the \( \lambda_k \), we take \( \lambda_k = - \log r_k \). The estimators \( a_k \) of \( \sigma_k \) are then obtained by solving any \( p \) equations of the set

\[
\sum_{k=1}^{p} r_k \frac{s_{k}^{2p+m}}{1-r_{k}^{2p+1}} \frac{a_k}{s_h} = s_h, \quad h=1,2,\ldots,2p
\]

(3.10)

The solutions of equations (3.9) and (3.10) give the estimates of the parameters of the model (2.20).

Consider the modified linear combination of exponentials with expectations

\[
W[Y(t)] = a_0 + \sum_{k=1}^{p} a_k \alpha_k^t \quad t=0,1,2,\ldots,n
\]

(3.11)

Let \( n = (2p+1)m-1 \), so that there are \( n+1 = (2p+1)m \) observations.

As for equation (3.2), compute \((2p+1)\) partial totals with expectations

\[
S_h^x = \max_0 + \sum_{k=1}^{s} h \frac{(1-\rho_k^{(2p+1)m})}{(1-\rho_k^{2p+1})} \quad (3.12)
\]

From the \( S_h^x \) we form the differences
Utilizing the same procedure as before for $S_h^*$ and $s_h^*$ the solution for the estimators $\hat{A}_r$ of the $A_r$ is the same in terms of the $s_h^*$ as that given by (3.7) in terms of the $s_h$ and $a_k$ will be obtained in the same manner as before. Estimator $a_o$ of $\sigma_o$ is determined by

$$m a_o = s^*_h - \sum_{k=1}^{2p+1} a_k r^h \frac{(1-r_k(2p+1))}{(1-r_k)}$$

There are many compartmental systems where exponentials are well separated in time, that is when $\lambda_i > \lambda_j$ ($i > j$, $i,j=1,2,...,p$), yield data known as 'decay type' data. In such a case some modification is needed in forming partial sums. Partial sums are formed sequentially with first $h$ or $6p$ or $8p$ observations for model (3.1). The initial estimates so obtained perform better than the consideration of whole set of data. Similarly for the consideration of model given by (3.11) we use sequentially $2(2p+1)$, or $3(2p+1)$ or $4(2p+1)$ observations. In practical situations it may not be feasible to collect data at equi-spaced time intervals after some stage of collection of data. In such situations the modified sequential estimation procedure of partial sums still works.

3.3 Simultaneous Estimation By Partial Totals For Compartmental Models.

We consider a p-compartment system where data are collected from each compartment. The p compartment system is represented by a set of p regression equations

$$Y_i(t) = \alpha_{io} + \sum_{k=1}^{p} \alpha_{ik} e^{-\lambda_k t} + \epsilon_i(t) \text{ for } i=1,2,...,p \text{ and } t=0,1,2,...,n-1.$$

(3.14)
In (3.14) \( e_i(t) \) represent random variables associated with the 
t-th observation on the i-th compartment and \( \alpha_{ik} \)'s and \( \lambda_k \)'s are parameters.

Since the parameters \( \lambda_1, \lambda_2, ..., \lambda_p \) appear in each one of the re-
gression equations of (3.14), we make use simultaneously of the ob-
servations on all the equations being studied to estimate these 
parameters. In this section we present the partial totals estimation 
technique for the above model (3.14) in two cases:

Case I

when \( \alpha_{10} = 0 \) for all \( i = 1, 2, ..., p \)
and let \( N = (p+1)m \).

Case II

when \( \alpha_{10} \neq 0 \) for \( i = 1, 2, ..., p \)
and let \( N = (p+2)m \).

We assume that \( p \) and \( m \) are positive integers, the coefficients are 
real numbers, and the exponents \( \lambda_k \) are distinct positive real numbers
in the above model. Assume that observations are taken at equally 
spaced time intervals. The estimation of the \( \lambda_1 \)'s involves the appli-
cation of a partial totals approach similar to that presented in sec-
tion 3.2. First we consider the case I.

Group the observations from each compartment into \((p+1)\) groups 
each containing \( m \) observations. Then the following partial totals 
are formed:

\[
s_{ih} = \sum_{t=h+(p+1)j}^{h+(p+1)(m-1)} y_i(t) 
\]

where \( h=1, 2, ..., (p+1) \)
\( j=0, 1, ..., (m-1) \)
\( i=1, 2, ..., p \)
Let $S_{ih} = E(s_{ih})$, where

$$
S_{ih} = \sum_{h=1}^{p} \alpha_{ik}^{h} \frac{(1-C_{ik})^{h}}{C_{ik}^{(p+1)}} \quad (3.16)
$$

and $\alpha_{ik}^{h} = e^{-\lambda}, h=1,2,\ldots,p+1$

Since $\alpha_{ik}^{h}$ are distinct, the polynomials, $S_{ih}$, satisfy the $p$th order difference equation.

$$
\sum_{h=1}^{p+1} (-1)^{2p+1-h} A_{p+1-h} S_{ih} = 0, \quad i=1,2,\ldots,p. \quad (3.17)
$$

The elementary symmetric functions $\Lambda_{r}$, $r=1,2,\ldots,p$, are defined in (3.5).

Replacing $S_{ih}$ by $s_{ih}$ in (3.17), then solving the set of equations (3.17) we obtain estimates $\hat{\Lambda}_{j}$ and $\hat{\Lambda'}_{j}$ are given by

$$
\hat{\Lambda}_{j} = (-1)^{j+1} \frac{|A_{x}|}{|A_{y}|}, \quad j=1,2,\ldots,p \quad (3.18)
$$

Here $A_{x}$ is a $p \times p$ matrix whose $j$th column is $(s_{1j}, s_{2j}, \ldots, s_{pj})'$ and $A_{y}$ is a matrix obtained by replacing the $(p+j)$th column of $A_{x}$ by the column vector $(s_{1, p+1}, s_{2, p+1}, \ldots, s_{p, p+1})'$.

Since $\hat{\Lambda}_{j}$ estimate the elementary symmetric functions of the $\alpha_{ik}$, the estimators $\hat{\alpha}_{ik}$ of $\alpha_{ik}$ are obtained by finding the $p$ roots of the equation

$$
x^{p} - \frac{|A_{x}|}{|A_{x}|} x^{p-1} - \frac{|A_{y}|}{|A_{y}|} x^{p-2} - \ldots - \frac{|A_{x}|}{|A_{x}|} = 0 \quad (3.19)
$$

For case II, we group the observations from each compartment into $(p+2)$ groups each containing $m$ observations, and form the partial totals, $s_{ih}^{x} = s_{ih}^{x} - s_{ih}^{x+1}$; for $i=1,2,\ldots,p$

$$
h=1,2,\ldots,p+1$$
where

\[ s_{ih}^x = \sum_{j=1}^{h+(p+2)} y_j(t), \]  

\[ s_{ih} = E[s_{ih}^x] \]  

and

\[ s_{ih}^{x+1} = s_{ih}^x - s_{ih}^{x+1} \]

The following difference equation is satisfied for each \( i \) by the \( s_{ih}^{x+1} \):

\[ \sum_{h=1}^{p+1} \frac{(-1)^{p+1-h}}{p+1-h} s_{ih}^{x+1} = 0 \text{ for } i=1,2,\ldots,p \]  

Equations (3.21) are the same as (3.17) except that \( s_{ih}^{x+1} \)'s have been substituted for \( s_{ih} \)'s. As in case I, the estimators \( r_i \) of the \( \rho_i \)'s are obtained using \( s_{ih}^{x+1} \)'s instead of the \( s_{ih} \)'s, \( i=1,2,\ldots,p \) and \( h=1,2,\ldots,p+1 \).

Now to obtain the estimators for \( \alpha_{ik} \)'s, \( r_i \) are substituted in place of \( \alpha_i \) in (3.14) giving \( p \) regression equations. These are linear in the unknown coefficients \( \alpha_{ik} \) and give their estimate using the weighted least squares procedure.

### 3.4 Numerical Examples

In this section the application of the method discussed above is made to data described in tables 1.4 and 1.5 in chapter 1. Other methods due to Aglu (1971), Cornell (1962) and Foss (1969) are compared numerically.

**Example 3.1**

Consider the case where \( p=1 \). Counts describing the decay of the
neutron density in a medium-size assembly of beryllium have been fitted by the various methods

(i) Cornell's estimators of $p$ and $\alpha$ are given by

$$r_c = 0.77606$$
$$a_c = 100,043$$

(ii) Agha's estimators of $p$ and $\alpha$ are given by

$$r_a = 0.77592$$
$$a_a = 100,089$$

(iii) Estimators based on equations (3.9) and (3.10) are given by

$$r = 0.77588$$
$$a = 100,156$$

The residual sum of squares $\Sigma[y(t) - a \cdot r^2]$ for the three estimation methods are as:

<table>
<thead>
<tr>
<th>Method</th>
<th>Res. S.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agha</td>
<td>282,380</td>
</tr>
<tr>
<td>Cornell</td>
<td>315,265</td>
</tr>
<tr>
<td>Proposed</td>
<td>263,196</td>
</tr>
</tbody>
</table>

It is obvious that the proposed method gives a considerable reduction in the residual sum of squares.

**Example 3.2**

Consider the case when $p=2$. We apply the various estimation methods to the data in table 1.5. The estimators of the parameters and the residual sum of squares $\Sigma[y(t) - a_1 r_1 + a_2 r_2]^2$ are given in the following table 3.1.

<table>
<thead>
<tr>
<th>Table 3.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>Agha</td>
</tr>
<tr>
<td>Cornell</td>
</tr>
<tr>
<td>Foss</td>
</tr>
<tr>
<td>Proposed</td>
</tr>
</tbody>
</table>
Notice that the reduction in residual sum of squares due to the proposed method is drastic and it compares very well with Agha's method.

The iterative maximum likelihood estimators of $\rho_1$, $\rho_2$ and $\alpha_1$ in three different sets of initial estimates were obtained. The number of iterations needed to get the results are given in Table 3.2.

<table>
<thead>
<tr>
<th>Initial Estimates Method</th>
<th>Maximum Likelihood Estimates $\hat{\alpha}_1$</th>
<th>$\hat{\rho}_1$</th>
<th>$\hat{\rho}_2$</th>
<th># of Iterations</th>
<th>Res. M. S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agha</td>
<td>8.3093</td>
<td>0.5188</td>
<td>0.1860</td>
<td>7</td>
<td>0.000,35161</td>
</tr>
<tr>
<td>Cornell</td>
<td>4.1305</td>
<td>0.3000</td>
<td>0.5519</td>
<td>21</td>
<td>0.000,45862</td>
</tr>
<tr>
<td>Proposed</td>
<td>8.3093</td>
<td>0.5188</td>
<td>0.1860</td>
<td>9</td>
<td>0.000,35161</td>
</tr>
</tbody>
</table>

Notice that the iterative maximum likelihood estimators of the parameters are same for the initial values provided by Agha's method and the method proposed. However, with the proposed method initial values it needed 2 more iterations for the convergence of estimators. Unfortunately Cornell's method does not seem to provide the right answer inspite of 21 iterations and the residual sum of squares is also considerably high.

5 Consistency of the Estimators

The partial-totals estimators are not in general unbiased since they are solution of polynomials, they are consistent estimators. The proof of consistency follows along the lines of Cornell (1962) and we outline it below for our case. Consider the models (3.2). Suppose the errors are independently distributed for all $t$ and are identically distributed for all $t$ in the same group. Define group means $\bar{y}_h = \frac{\bar{s}_h}{m}$. Replacing the sum $s_h$ by the corresponding $\bar{y}_h$ in equation (3.3) we have by law of large numbers.
\[ p \lim_{m \to \infty} \bar{y}_h = \lim_{m \to \infty} \frac{1}{m} \sum_{t=1}^{2p(m-1)+h} \rho^t, \quad i=0,1,2,\ldots,m-1 \quad (3.22) \]

Letting \( m \to \infty \), keeping the domain for \( t \) constant in length, say \( T \), the \( m \) observations included in the \( h \)-th group are made for

\[ t = \frac{hT}{2p}, \frac{hT + T}{m}, \frac{hT + 2T}{2p}, \ldots, \frac{hT + \frac{2p(m-1)}{2}}{m} \]

Now (3.22) may be written as

\[ p \lim_{m \to \infty} \bar{y}_h = \lim_{m \to \infty} \frac{1}{2p} \sum_{i=1}^{hT} \alpha_k \rho^i_T \quad (3.23) \]

The expression on right is definite integral equal to

\[ \varphi = \frac{1}{T} \sum_{k=1}^{P} \alpha_k (1-\rho^T), \quad (3.24) \]

At the point \( \tau \), \( r_k = \rho_k^T \) and \( \lim_{m \to \infty} a_k = \alpha_k \) for all \( k \).

Then with the \( \alpha_k \) distinct as specified by our model, derivative of all orders of the estimators \( r_k \) and \( a_k \) are continuous in the neighborhood of \( \tau \) and Slutsky's Theorem ref Cramér (1946, p. 255) is applied to show that \( r_k \) and \( a_k \) converge in probability to \( \rho_k \) and \( \alpha_k \), respectively, for all \( k \). Thus \( r_k \) and \( a_k \) are consistent estimators of \( \rho_k \) and \( \alpha_k \), respectively, for all \( k \) when \( m \to \infty \). This is also true for the model with constant term \( a_0 \) added.

Consistency of \( \rho_i \)'s in the above case can be proved by the application of the Chebycheff inequality, refer Cramer (1946, p. 253). Since \( \alpha_{ik} \) are continuous functions of the \( \rho_k \)'s, \( \alpha_{ik} \) are also consistent estimators of \( \alpha_{ik} \).

In case of errors are correlated, we use the weighted least squares with known covariance matrix \( \Sigma \) of errors or the consistent estimator.
of $\Sigma$ if the covariance matrix $\Sigma$ is unknown. In that case also the estimators $\hat{a}_{ik}$ are consistent.

Regarding estimators of $\alpha_i$ as functions of $\bar{y}_{ih} = \frac{S_h}{m}$ or $\bar{y}'_{ih}$, we can demonstrate the consistency of these estimators in the case of simultaneous estimation as $m \to \infty$. We need the following assumptions:

Assumptions

(i) For each value of $i=1,2,...,p$, the random variables $\varepsilon_i(t)$, for all values of $t$, are uncorrelated with $E[\varepsilon_i(t)] = 0$.

(ii) For each value of $i$ and $h$ the random variables $\varepsilon_i(t)$ associated with the corresponding observations $y_i(t)$ in $s_{ih}$ or $s'_{ih}$ have common variance.

(iii) For each value of $i$ and $h$ the domain of the independent variable is of constant length, $T$, for $s_{ih}$ or $s'_{ih}$ where $i=1,2,...,p$ and $h=1,2,...,p+1$.

(iv) For $g = (g_{ik})$, a $p \times p$ matrix, $|g| \neq 0$

3.6 Approximate Test For The Number Of Compartments

A question of considerable interest relates to the number of compartments required to specify a system. This is equivalent to asking how many exponential terms of the form of equations (2.20) are required to describe a given set of experimental results, no exact way of answering this question is known. By analogy with methods appropriate to the fitting of polynomials, one is tempted to test the hypothesis of an additional term. This is generally accomplished by testing for significance the reduction in the sum of squares obtained
by fitting the two additional constants implied by an additional term. There is an important difference here, however. A polynomial specified by 2p constants can be passed exactly through 2p arbitrary points, but a linear combination of p exponentials specified by 2p real numbers cannot be. Each constant in such a polynomial can effect a larger reduction in the sum of squares than can each constant in a linear combination of exponentials. This is simply an intuitive way of saying that the usual theory, which covers the case in which the unknown constants appear linearly, does not provide an exact treatment for cases such as the present, in which half of the constants appear non-linearly.

It seems clear that level of significance obtained by applying the usual theory to the reduction in sum of squares effected by an additional exponential term provides an upper limit to the correct value, so that a component found significant by the usual theory can be considered significant. The converse is false, however, and the magnitude of the error involved in acting as if it were true is unknown.

3.7 Comparison With Other Methods

For the non-linear model given by equations (4.1) and (4.9), application of the Least Squares method results in equations which are in general soluble only by iteration. Most methods for curve fitting to a linear combination of exponentials require initial estimates of the parameters. The degree of accuracy required for these initial estimates increases with p, the number of exponentials to be fitted. Iterative methods using initial estimates include (i) those developed by Worsley (1964) involving the least squares process, whereby the non-linear is linearized by expanding in a Taylor series; (ii) the
method utilizing maximum likelihood developed by Grard (1962).

Earlier methods for obtaining the initial values were the following: (i) 'by eye' determinations, graphical methods and results from similar experiments by Perl (1960); (ii) the method of partial totals by Cornell (1962); (iii) the method of dependent partial totals by Agha (1971); and (iv) the 'peeling-off' method by Foss (1969).

Cornell's method provides a simple and direct procedure for estimating the non-linear parameters and is based on independent partial totals of the sample observations, it has the disadvantage that the estimators obtained are not of $p_k$, but of some integral power of $p_k$. Agha's method overcomes the disadvantage of Cornell's method of providing estimator of some integral power of $p_k$, but utilizes dependent partial totals of the sample observations. Also it assumes that all $\alpha_k > 0$ in model given by equation (3.1). Foss' method is computer oriented and arrives at the initial estimates by a least square 'peeling-off' technique. The proposed method in this dissertation provides estimates of $p_k$ rather than some integral power of $p_k$ as in the case of Cornell's method. Secondly, it utilizes independent partial totals as against the use of over-lapping partial totals in the case of Agha's method. If computing facilities are not available for obtaining maximum likelihood estimates, the method provides a systematic and relatively simple estimation procedure. Also in most of the cases this method yields a smaller residual sum of squares than other methods, as has been seen by examples (3.1) and (3.2). In case of log transformation of the observations, it is generally noticed that for large $t$, the curve is approximately a straight line. In such situations the modified form of the proposed method based on sequential estimation technique performs
better over the other methods. The proposed method of course, require equally spaced data as required by Cornell and Agha methods.
CHAPTER IV

DIFFERENCE EQUATION APPROACH TO COMPARTMENT ANALYSIS

4.1 Introduction

In chapter II, we introduced the difference equation approach for the compartmental system so as to utilize equi-spaced time series data. In this chapter, a similar model is introduced for unequally spaced time series observations. The approach is similar to the one proposed earlier except that here derivatives of the function are expressed in terms of divided differences, for reference, see Whittaker and Robinson (1944). We also consider the problem of estimation of the eigenvalues of the constant matrix specifying the set of differential equations describing the compartmental system. Tests for the number of compartments governing the system are also discussed. Data from three experiments discussed in the literature are analyzed through the difference equation approach.

4.2 Derivation of Difference Equation Model for Unequally Spaced Time Series.

Consider the differential equation for p-compartmental system as given by (2.23). Suppose the observations are made at times \( t = t_0, t_1, \ldots, t_n \). We assume that the intervals \( t_1 - t_0, t_2 - t_1, \ldots, t_n - t_{n-1} \) are not necessarily equal. In place of ordinary differences approximating derivatives divided differences are introduced. The
quantities

\[
\eta(t_i, t_{i-1}) = \frac{\eta(t_i) - \eta(t_{i-1})}{t_i - t_{i-1}}
\]  

(4.1)

\(i = 1, 2, \ldots, n\)

are called divided differences of the first order. Similarly

\[
\eta(t_i, t_{i-1}, t_{i-2}) = \frac{\eta(t_i, t_{i-1}) - \eta(t_{i-1}, t_{i-2})}{t_i - t_{i-2}}
\]  

(4.2)

\(i = 2, 3, \ldots, n\)

are called divided differences of the second order. The divided differences of higher orders are formed in the same way.

To the first order of approximation, the derivatives of \(\eta(t)\) in terms of its divided differences are given as

\[
\frac{d \eta(t_i)}{dt} = \eta(t_{i+1}, t_i)
\]  

(4.3)

\[
\frac{d^2 \eta(t_i)}{dt^2} = 2 \cdot \eta(t_{i+2}, t_{i+1}, t_i)
\]  

(4.4)

\[
\frac{d^3 \eta(t_i)}{dt^3} = 3 \cdot \eta(t_{i+3}, t_{i+2}, t_{i+1}, t_i)
\]  

(4.5)

and so on.

Approximating the derivatives of \(\eta(t)\) in (2.23) by the corresponding divided differences, we obtain the corresponding difference equation as follows.
where \( i = 0, 1, 2, \ldots, n-p \)

We consider the following difference equation model for (4.6) when \( y(t) \) is observed.

\[
p! \eta(t_{i+p}, t_{i+p-1}, \ldots, t_i) + (p-1)! \eta(t_{i+p-1}, \ldots, t_i) + \ldots + \eta(t_i) = 0
\]

\[(4.6)\]

It is assumed that \( u_{t_{i+p}} \) is the error,

\[
E[u_{t_{i+p}}] = 0, \text{ Variance of } u_{t_{i+p}} = \sigma^2 \text{ and } E[Y(t_i)] = \eta(t_i) \text{ and }
\]

\[
\frac{E[Y(t_i) - Y(t_{i-1})]}{t_i - t_{i-1}} = E[Y(t_i, t_{i-1})] = \eta(t_i, t_{i-1})
\]

\[
E[Y(t_i, t_{i-1}, t_{i-2}) - Y(t_{i-1}, t_{i-2})] = E[Y(t_i, t_{i-1}, t_{i-2})] = \eta(t_i, t_{i-1}, t_{i-2})
\]

and so on.

**4.3 Estimation**

**Case I Equi-spaced time series observations**

The estimates of the constants \( c_i \) in the model (2.29) can be made by the method of least squares or weighted least squares with appropriate assumptions on errors \( u_t \).
Minimizing the expression

\[
\sum_{t=0}^{n-p} \left[ y(t+p) + c_1 y(t+p-1) + \ldots + c_p y(t) \right]^2 ,
\]

(4.8)

where the number of observations \( n+1 \), leads to the usual estimate of the parameters

\[
\mathbf{c}_0 = (c_1, c_2, \ldots, c_p)',
\]
given by

\[
\hat{\mathbf{c}}_0 = -A^{-1} \mathbf{A}_0
\]

Here,

\[
A = \sum_{t=0}^{n-p} \chi(t+p-1)\chi'(t+p-1)
\]

(4.10)

\[
A_0 = \sum_{t=0}^{n-p} y(t+p)\chi(t+p-1)
\]

(4.11)

\[
\chi(t+p-1) = [y(t+p-1), y(t+p-2), \ldots, y(t)]'
\]

(4.12)

These estimates are also maximum likelihood estimates when it is assumed that \( u_i, i = p, p+1, \ldots, n \) are normally distributed with means 0 and variance \( \sigma^2 \). It is generally assumed for simplicity in these models that the first \( p \) observations \( y(0), y(1), \ldots, y(p-1) \) are known without error. The maximum likelihood estimate of \( \sigma^2 \) is given by

\[
\hat{\sigma}^2 = \frac{1}{n+1} \sum_{t=0}^{n-p} \left[ y(t+p) + \mathbf{c}_0' \chi(t+p-1) \right]^2
\]

(4.13)

and an unbiased estimate of \( \sigma^2 \) is given by

\[
\hat{s}^2 = (n+1)\hat{\sigma}^2 / (n-2p+1)
\]

(4.14)

The properties of these estimates have been extensively studied in the literature. It is also well known that the estimates \( \mathbf{c}_0 \) are consistent and asymptotically efficient, Anderson (1971).
Case II Unequally spaced time-series data

The estimates of the parameters $\Lambda_1$ in the model (4.7) are also obtained by the method of least squares or weighted least squares when we assume $u_{1+p}$ are independently distributed with means zero and a known variance structure. However, $y(t_{i+p}, t_{i+p-1}, \ldots, t_1)$, $y(t_{i+p-1}, t_{i+p-2}, \ldots, t_1), \ldots y(t_1)$ are constructed first from the observations and these divided differences are treated as observations for fitting the model (4.7). Again, we assume that first $p$ observations are known without error. The properties of these estimates are like the properties of $\hat{\Lambda}_0$ mentioned earlier.

4.4 Testing Hypotheses

An important question in pharmacokinetics is concerned with finding the number of compartments which explain a given set of data. For example, we may be required to test the hypothesis whether the number of compartments in the system is $(p-1)$ or $p$.

Case I Equi-spaced Observations

The hypothesis that there are $(p-1)$ compartments can be formulated as in (2.30) by testing

$$H_0: \sum_{i=1}^{p} c_i = -1 \quad (4.15)$$

or

$$H_0: L^c \hat{\zeta}_0 = -1 \quad (4.16)$$

with $L = (1, 1, \ldots, 1)^{\top}$. $l_{1\times p}$

The test is based on the statistic

$$T = \frac{\sqrt{n+1}(\hat{\zeta}_0^c + 1)}{\sqrt{\hat{\Lambda}^{-1} \hat{\Sigma} \hat{\Lambda}}}, \quad (4.17)$$
where $\Lambda$ is defined in (4.10). The asymptotic distribution of $\hat{\Lambda}_{\sigma^2}^{-1}$ is normal with mean $-1$ and variance $\frac{\sigma^2}{\hat{\Lambda}^{-1}}$. Using an unbiased estimate of $\sigma^2$ given by (4.14), we can use the t-distribution to determine the critical region.

**Case II Unequally Spaced Observations**

Here we test the hypothesis

$$H_0 : \Lambda_p = 0 \text{ vs } H_1 : \Lambda_p \neq 0$$

The test is based on the sequential F-test, which uses fitting the difference equation model (4.7) with $(p-1)$ and $p$ parameters sequentially, Draper and Smith (1966) p. 72.

Suppose the assumptions underlying the difference equation model do not hold, for example the errors $u_t$ are not independent. The model still provides good initial estimates for the usual compartment model where as other methods of initial estimates may fail.

**4.5 Examples**

In this section, the application of the difference equation models is made to experiments II and V of chapter I. For the application of divided difference equation approach to unequally spaced data, we examine subject 1 of Digoxin experiment III described in chapter I.

**Example 4.1**

The model (2.27) is used to describe the logarithm of frequencies and pulse heights in experiment V. The case considered is that of assuming 2, 3, or 4 compartments. The test for them shows that there are three compartments. The test statistic $T$ for the case of three compartments is

$$T_9 = 3.862$$
which is significant at 1% level of significance.

The test statistic $T$ for the case of four compartments is

$$T_8 = 2.072$$

which is non-significant at $\%$ level of significance.

However, in the literature, this experiment has been examined by several authors assuming only two compartments. Using three compartment models with the difference equation approach the following estimates are obtained.

$$\begin{align*}
\hat{c}_1 &= -0.529,148, \\
\hat{c}_2 &= -0.248,875, \\
\hat{c}_3 &= 0.121,009 \\
s^2 &= 0.000,184,6
\end{align*}$$

Example 4.2

Using a two compartment difference-equation model to the pharmacological data for experiment II of chapter I, the estimates of $c_1$ and $c_2$ are as follows.

$$\begin{align*}
\hat{c}_1 &= -0.711,77 \\
\hat{c}_2 &= 0.104,55 \\
s^2 &= 0.000,011,66
\end{align*}$$

The test for the number of compartments shows that there is only one compartment. The test statistic $T$ for the case of two compartments is

$$T_4 = 1.765$$

which is non-significant at 10% level of significance.

Examination of the above experiment by using the usual model (2.20), with two compartments, the iterative maximum likelihood estimates are given by
\[ \exp(-\lambda_1) = \hat{\lambda}_1 = 0.5095 \]
\[ \exp(-\lambda_2) = \hat{\lambda}_2 = 0.5100 \]
\[ s^2 = 0.000,018,14 \]

Since the two eigenvalues \( \lambda_1 \) are practically the same, they imply also that there is one compartmental system confirming our results obtained above through the difference equation approach.

**Example 4.3**

The data on subject 1 of experiment II described in chapter I is at unequal time intervals. We use the divided differences to examine the model. The test shows that there are two compartments for the distribution of Digoxin in this individual. The test statistic \( F \) for the case of two compartments is

\[ F_{1,17} = 9.78 \]

which is significant at 1% level of significance.

The test statistic \( F \) for the case of three compartments is

\[ F_{1,15} = 2.51 \]

which is non-significant at 5% level of significance.

The estimates of the constants \( \lambda_1 \) and \( \sigma^2 \) are as follows.

\[ \hat{\lambda}_1 = 29.6259 \]
\[ \hat{\lambda}_2 = 2.1359 \]
\[ s^2 = 209.67 \]

Using usual compartment models, it is also seen that there are only two compartments. However, the estimates of \( \lambda_1 \) are not comparable. The estimates given by Kramer (1974) are as follows.
\[ \hat{\lambda}_1 = 83.7245 \]
\[ \hat{\lambda}_2 = 0.1953 \]

Note that the above models are not comparable, since the estimates are based on different assumptions.
CHAPTER V
ANTITONIC REGRESSION ESTIMATION IN NON-LINEAR MODELS

In the field of pharmacokinetics we have some situations where a specified dose of a drug is administered by rapid intravenous injection to a subject. It is assumed that drug mixes uniformly and instantaneously into blood. Blood samples are taken at different time intervals and analyzed chemically for drug concentration present at different timings. As we have seen earlier the blood concentration data fit well to a linear combination of exponentials. In such experiments a knowledge of the time $t$ at which blood samples are taken, determines an ordering, partial or total, of the corresponding mean concentrations, $\eta_t$. In most classical approaches to fitting such models, no notice is taken of the ordered observations. In this chapter we give an account of the statistical theory to deal with such problems. We assume that the regression functions are subject to order restriction.

Consider the model in (4.1) given as follows:

Writing $\rho_k = \exp(-\lambda_k), \lambda_k > 0$

$$Y(t) = \sum_{k=1}^{p} \alpha_k \rho_k^t + \varepsilon(t),$$

(5.1)

$0 < \rho_k < 1,$

$\alpha_k > 0,$

$t=0,1,2,\ldots,n,$

Assume again the $\varepsilon(t)$'s, the random errors, are independent with zero means. Let the $(n+1)$ observations be denoted by $\cdots$.
y(0), y(1), ..., y(n) for t = t_o, t_1, ..., t_n. In the above case of blood levels of drug for t_o < t_1 < t_2 < ... < t_n, \(E[Y(t_i)] > E[Y(t_j)]\), t_i < t_j, i < j, i, j = 0, 1, 2, ..., n. If there are no order restriction on \(E[Y(t_j)]\), j = 0, 1, 2, ..., n, the weighted least squares estimates of the parameters in (5.1) are obtained by minimizing

\[ \sum_{j=0}^{n} w_j [y(t_j) - E(Y(t_j))]^2 \]  

(5.2)

Here \(w_j\) are known or unknown weights. For the sake of completeness, we review some definitions and results on regression with order restrictions. For a comprehensive survey, see Barlow and others (1972).

5.2 Antitonic Regression over a Simply Ordered Finite Set

Definition 5.1

The estimates obtained by minimizing (5.2) will be called the basic estimates of the parameters.

Definition 5.2

Let T be the finite set \{t_1, t_2, ..., t_n\} with the simple order \(t_1 < t_2 < ... < t_n\). A real valued function \(f\) on T is antitonic if \(t_i, t_j \in T\) and \(t_i < t_j\) imply \(f(t_i) \geq f(t_j)\). Let \(g\) be a given function on T and \(w\) a given positive function on T. An antitonic function \(g^*\) on T is an antitonic regression of \(g\) with weights \(w\) with respect to the simple ordering \(t_1 < t_2 < ... < t_n\) if it minimizes in the class of antitonic function \(f\) on T the sum

\[ \sum_{t_i \in T} w(t_i) [g(t_i) - f(t_i)]^2 \]  

(5.3)

and \(g^*\) is simply called an antitonic regression of \(g\).
A graphical interpretation of the antitonic regression is illuminating. Assuming the simple ordering $t_1 < t_2 < \ldots < t_n$ we plot the cumulative sums

$$G_k = \sum_{j=1}^{k} g(t_j)w(t_j)$$

against the cumulative sums

$$W_k = \sum_{j=1}^{k} w(t_j), \quad k=1,2,3,\ldots,n.$$ 

That is, we plot the points $P_k = (W_k,G_k)$ for $k=0,1,2,\ldots,n$, $(P_0 = (0,0))$. These points constitute the cumulative sum diagram (CSD) of the given function $g$ with weights $w^i$. The slope of the segment joining $P_{k-1}$ to $P_k$ is $g(t_k)$, $k=1,2,\ldots,n$. It will be seen that the antitonic regression of $g$ is given by the slope of the Least Concave Majorant (LCM) which is the graph of the infimum of all concave functions whose graphs lie above the CSD. The value of the antitonic regression $g^*$ at a point $t_k$ is just the slope of the LCM at the point $P^*_k$ with abscissa $\sum_{j=1}^{k} w(t_j)$. An illustrative example is shown in table 5.1 and Fig 5.1.

Table 5.1 Example of CSD and LCM

<table>
<thead>
<tr>
<th>$j$</th>
<th>$w(t_j)$</th>
<th>$W_j$</th>
<th>$g(t_j)$</th>
<th>$G_j$</th>
<th>$G^*_j$</th>
<th>$g^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2/1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>-4/3</td>
<td>-2</td>
<td>3</td>
<td>1/3</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>5</td>
<td>4/1</td>
<td>2</td>
<td>10/3</td>
<td>1/3</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>7</td>
<td>2/2</td>
<td>4</td>
<td>4</td>
<td>1/3</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>8</td>
<td>-1/1</td>
<td>3</td>
<td>3</td>
<td>-1</td>
</tr>
</tbody>
</table>
Fig. 5.1 Example of CSD and LCM

It is useful to express analytically some of the properties of the LCM. First, the CSD and LCM coincide at $P_k^*$, i.e.

$$G_k^* = G_k$$

Secondly, if for some index $i$ the LCM at $P_{i-1}^*$ lies strictly above the CSD at $P_{i-1}$, then the slopes of the LCM entering $P_{i-1}^*$ from the left and leaving to the right are the same. That is,

$$G_{i-1}^* > G_{i-1} = g_i^* - g_{i-1}^* = 0 \quad i=1,2,\ldots,k.$$  

Finally, if $g^*(t)$ has the constant value $a$ for $t_r < t \leq t_s$ and the constant value $b < a$ for $t_s < t < t_u$, then

$$\sum_{i=j+1}^{s} g(t_i)w(t_i) \geq a, \quad \text{for } r \leq j < s.$$
\[
\frac{\sum_{i=s+1}^{j} g(t_i)w(t_i)}{\sum_{i=s+1}^{j} w(t_i)} = \frac{G(t_j) - G(t_s)}{W(t_j) - W(t_s)} \leq b \text{ for } s < j \leq u.
\]

In particular,

\[g(t_s) \geq g^*(t_s) = a > b = g^*(t_{s+1}) > g(t_{s+1}).\]

These properties yield a simple proof that \(g^*\) is the antitonic regression of \(g\) with weights \(w\).

**Theorem 5.1**

If \(T\) is simply ordered, the slope of the LCM furnishes the antitonic regression of \(g\). Indeed, if \(f\) is antitonic on \(T\) then

\[
\sum_{t} [g(t) - f(t)]^2 w(t) \geq \sum_{t} [g(t) - g^*(t)]^2 w(t) + \sum_{t} [g^*(t) - f(t)]^2 w(t).
\]

(5.4)

The antitonic regression is unique. For the proof of theorem 5.1 one may refer to Barlow et al. (1972).

**The Min-Max Formulas**

Closely related to the graphical representation of the antitonic regression as the slope of the LCM is the formula

\[g^*(t_j) = \min_{k \geq j} \max_{s \leq j} \text{Av}(s,k)\]  

(5.5)

where

\[
\text{Av}(s,k) = \sum_{r=s}^{k} g(t_r)w(t_r) / \sum_{r=s}^{k} w(t_r).
\]

(5.6)
Equivalent formulas are:

\[ g^*(t_j) = \max_{s \leq j} \min_{k > j} Av(s, k) \quad (5.7) \]
\[ g^*(t_j) = \min_{k \leq s} \max_{s \leq j} Av(s, k) \quad (5.8) \]
\[ g^*(t_j) = \max_{s < k} \min_{k > j} Av(s, k) \quad (5.9) \]

One may think of finding the slope, \( g^*(t_j) \), of the LCM graphically as follows. Select \( s \leq j \) and draw that ray proceeding to the right from \( P_s \) which intersects again the graph of partial sums but lies above all others which do so. The slope of this ray is

\[
\max_{s \leq j} \left\{ \frac{\sum_{r=s}^{k} g(t_r)w(t_r)}{\sum_{t=s}^{k} w(t_r)} \right\}
\quad (5.10)
\]

Repeat this for all other choices of \( s \leq j \), selecting finally that for which the resulting maximum slope is as small as possible. A segment of this final ray is part of the graph of the LCM. On the other hand, this is a graphical way of calculating \( g^*(t_j) \) by formula (5.8). Similarly each of the other three formulas may also be regarded as a representation of the slope of the LCM at \( P_s \).

A real valued function \( f \) on \( T \) is isotonic with respect to a simple ordering "<" on \( T \) if \( t_i, t_j \in T, t_i < t_j \) imply \( f(t_i) < f(t_j) \). Suppose it is required to minimize (5.3) in the class of functions \( f \) satisfying \( f(t_1) \geq f(t_2) \geq \ldots \geq f(t_n) \). These functions are antitonic with respect to the order \( t_1 < t_2 < \ldots < t_n \); and isotonic with respect to the reverse order on \( T \):

\[ t_1 > t_2 > \ldots > t_n \quad \text{or} \quad t_n < t_{n-1} < \ldots < t_1 \]
With appropriate relabelling, of course, the problem may be recast so that methods and descriptions above apply directly. In the graphical interpretation, the slopes $g^*$ must now be nondecreasing, and a greatest convex minorant $(GCM)$ replaces the LCM. In formulas (5.5), (5.7), (5.8) and (5.9), max and min are interchanged.

5.3 Antitonicization of Estimates

If the basic estimates of $E[Y(t_j)]$ satisfy the order restrictions, they are antitonic estimates. If not, we replace $y(t_j)$ by their antitonic regression estimates $y^*(t_j)$ with weights $w_j$. These estimates will be better than $y(t_j)$, in the sense of least squares. In the case of exponential family, the antitonic regression of the basic estimate turns out to coincide with the maximum likelihood estimates under the order restrictions.

**Theorem 5.2**

Let $\eta$ be an unknown function on a finite set $T$, known to be antitonic with respect to a simple order on $T$. Let $w(t), t \in T$, be a set of positive weights. Let $g$ be an estimate of $\eta$. Let $g^*$ be the antitonic regression of $g$ with weights $w$. Then

$$\sum_{t} [\eta(t) - g^*(t)]^2 w(t) \leq \sum_{t} [\eta(t) - g(t)]^2 w(t) \quad (5.11)$$

**Proof:**

This is an immediate consequence of inequality (5.4), valid for all antitonic $f$:

$$\sum_{t} [g(t) - f(t)]^2 w(t) \geq \sum_{t} [g(t) - g^*(t)]^2 w(t) + \sum_{t} [g^*(t) - f(t)]^2 w(t), \quad (5.12)$$
since \( \eta \) is antitonic and
\[
\sum_{t} [g(t) - g^*(t)]^2 w(t) \geq 0;
\]
f is simply replaced in (5.4) by \( \eta \).

The procedure for finding antitonic regression is as follows:

For a given set of observation \( u(t^i), i=0,1,2,...,n \), we first determine antitonic regression \( y^*(t^i) \) of \( y(t^i) \). Having determined the antitonic regression, the estimates of the parameters \( \rho_k \) and \( \alpha_k \) in the model (5.1) are determined. Note that \( \rho_k \) and \( \alpha_k \) are common to each \( \eta(t^i) \). The antitonic estimates of \( \rho_k \) and \( \alpha_k \), are given by minimizing the following.

\[
\sum_{i=0}^{n} [y^*(t^i) - \sum_{k=1}^{p} \alpha_k \rho_k^t] w(t^i) \quad (5.13)
\]

Choice of Weights \( w(t^i) \)

Various schemes of choosing \( w(t^i) \) for the model (5.1) are examined under the assumptions of

i) equal error variances and

ii) unequal error variances.

Case I Equal Error Variances

Let \( E[t^2(t^i)] = \sigma^2 \) for all \( t^i \). In usual least squares analysis we take \( w(t^i) \) to be the same for all \( t^i \)'s but for the model under discussion it may not be the best choice. In this model the coefficient of variation which is an index of the reliability of data rapidly increases with the increase in the value of \( t^i \) and then \( t^i \) is sufficiently large the error component dominates the fixed component in (5.1). For large values of \( t^i \) (depending upon the magnitude of \( \sigma^2 \) ) the observations contain almost no information.
about the parameters $\rho_k$ and $\alpha_k$. In such cases it becomes very essential to have antitonic weights with respect to the independent variable $t$. Obviously the choice of antitonic weights puts maximum reliance on the initial observation and least reliance on the last observation. If we disregard that there are $p$ fixed components in the model given by equation (5.1) then the simplest antitonic weighing scheme in case $t_i$ are equi-space, is given by

$$w_i = a^{i+1}$$

(5.15)

where $a$ is given by,

$$\sum_{i=0}^{n} w_i = \frac{a(1-a^n)}{1-a} = 1$$

(5.16)

Hence given the number of observations we can uniquely determine the weights to be assigned to each level of $t$. The above weighing scheme was found to work better in many cases and gave smaller residual sum of squares over the scheme of equal weights. Let us assume that $\rho_1 > \rho_2 > ... > \rho_p$ then the contribution of the $p$th component becomes negligible first as the level of $t$ is increased followed by the next lowest $\rho_{p-1}$ component. Finally beyond certain level of $t$ the contribution of only the slowest moving component is appreciable. Obviously the increase in the level of $t$ beyond certain level depending on the magnitude of slowest moving component does not contain much information about the parameters and it is not advisable to derive a weighing scheme as a function of $p$. If the parameters are well separated in time then the above proposed weighing scheme is enough. However, if the parameters are not well separated then we may choose weights directly proportional to the observed magnitude of observations, i.e.
Choice of weights according to equation (5.17) amounts to the selection of weights inversely proportional to the coefficient of variation as error variance is assumed to be constant for all \( t \).

**Example 5.1**

In order to examine the performance of weighted least squares estimation over the usual least squares estimation procedure, we consider the set of simulated data given in table 5.2 obtained by simulated regression model given in (5.1) for \( p = 2 \).

<table>
<thead>
<tr>
<th>Simulated Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t )</td>
</tr>
<tr>
<td>( y(t) )</td>
</tr>
<tr>
<td>( t )</td>
</tr>
<tr>
<td>( y(t) )</td>
</tr>
<tr>
<td>( t )</td>
</tr>
<tr>
<td>( y(t) )</td>
</tr>
<tr>
<td>( t )</td>
</tr>
<tr>
<td>( y(t) )</td>
</tr>
</tbody>
</table>

The error mean squares were found by fitting the model (5.1) with \( p = 2 \) and are given below for the two weighing schemes.

<table>
<thead>
<tr>
<th>Weighing Scheme</th>
<th>Error Mean Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal Wts</td>
<td>0.000,735,94</td>
</tr>
<tr>
<td>Unequal Wts</td>
<td>0.000,624,26</td>
</tr>
</tbody>
</table>
That is the reduction in Error Mean Squares is 0.000,111,68 which is about 18% by using the scheme (5.15).

**Example 5.2**

We examine the data given in table 2.5. Three following weighing schemes, were adopted and the error mean squares are as below:

<table>
<thead>
<tr>
<th>Weighing Scheme</th>
<th>Error Mean Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal Weights</td>
<td>0.000,449,15</td>
</tr>
<tr>
<td>$w_i = a^i$</td>
<td>0.000,351,61</td>
</tr>
<tr>
<td>$w_i = ay(i)$</td>
<td>0.000,121,35</td>
</tr>
</tbody>
</table>

A reduction in error mean squares varying from 28% to 73% is accomplished by choosing appropriate antitonic weights.

**Example 5.3**

In the experiment (2.4) we have the following data on proportions $y_i$ of radioactive tracer from a two compartment model.

$$
\begin{array}{cccccccc}
 t_i & 2 & 3 & 5 & 8 & 12 & 24 & 48 & 72 \\
y_i & 0.84 & 0.79 & 0.64 & 0.55 & 0.44 & 0.27 & 0.12 & 0.06 \\
\end{array}
$$

Three weighing schemes were adopted and the error mean squares are as below:

<table>
<thead>
<tr>
<th>Weighing Scheme</th>
<th>Error Mean Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal Weights</td>
<td>0.000,819,33</td>
</tr>
<tr>
<td>$w_i = a^i$</td>
<td>0.000,115,73</td>
</tr>
<tr>
<td>$w_i = ay_i$</td>
<td>0.000,053,29</td>
</tr>
</tbody>
</table>

As is clear from the above figures, a considerable reduction in error mean squares is accomplished by choosing appropriate antitonic weights.
Case II Unequal Error Variances

Errors $e(t)$ are assumed to be independent with variances $\sigma^2_t$. In the usual least squares analysis, we minimize

$$\sum \left[ \frac{y(t) - E(y(t))}{\sigma_t} \right]^2$$

that is, the weights $w_t$ are $\frac{1}{\sigma_t^2}$, $t=0,1,2,...,n$. In the case of equal variance the choice of antitonic weights is motivated by the index of coefficient of variation, in this case the choice of $w_t = \frac{1}{\sigma_t^2}$ takes care of this fact and no new scheme of weights is proposed.
6.1 Introduction

In chapter II we introduced the problem of stochastic compartmental analysis. The differential equations describing p-compartmental stochastic model were derived. In this chapter, we consider the problems of obtaining the solution of the differential equations for giving transition probabilities, population sizes and the associated probability distributions. The covariance kernel of observations is derived and are also provided the least squares estimates of the transition intensities. The problems have been studied by Matis and Hartley (1971) though their approach is entirely different and very lengthy. The approach adopted here resembles the approach of Chiang (1968) for 'illness-death' process.

6.2 Solution of Differential Equation

The differential equation for the p-compartmental system as given in (2.41) is considered. In case p = 1, the differential equation

\[ \frac{dP(t)}{dt} = P(t)Y_s \]  

(6.0)

is an ordinary first order differential equation with a constant coefficient. Formally, the solution of (6.0) is given by

\[ P(t) = e^{Bt}P(0) \]  

(6.1)
Defining the matrix exponential

\[ e^{Vt} = \sum_{n=0}^{\infty} \frac{(Vt)^n}{n!}, \quad (6.2) \]

(6.1) provides a solution of the matrix differential equation (2.41) as

\[ P(t) = e^{Vt} P(0) \quad (6.3) \]

where \( P(0) = 1 \).

It can be shown that the matrix series in (6.2) converges uniformly in \( t \), therefore, we can take the derivative of the infinite sum in (6.1) with respect to \( t \) term by term. We see that (6.1) satisfies the matrix differential equation (6.0) as follows:

\[ \frac{d}{dt} e^{Vt} P(0) = e^{Vt} \frac{d}{dt} \left[ \sum_{n=0}^{\infty} \frac{V^n t^n}{n!} P(0) \right] \]

\[ = \sum_{n=1}^{\infty} \frac{V^n t^{n-1}}{(n-1)!} P(0) = e^{Vt} P(0) V = e^{Vt} P(t) \]

The formal solution (6.1), however, is not very useful from a practical point of view. For the purpose of application we need explicit functions for the individual transition probabilities \( P_{ji}(t) \) that will satisfy the differential equation. Such functions depend on the eigen values of \( V_s \). Let \( \delta_1, \delta_2, \ldots, \delta_p \) be the eigen values of \( V_s \) and let \( A_{ij}(\ell) \) be the cofactor of the matrix \( A(\ell) = (\delta_\ell I - V_s) \).

Let \( T_j(k) = (A_{1j}(\ell), A_{2j}(\ell), \ldots, A_{pj}(\ell))^t \) be an eigen vector of \( V_s \) for \( \delta_j \).
The matrix of eigen vectors,

\[ T(k) = [T_1(k), T_2(k), \ldots, T_p(k)] \]  \hspace{1cm} (6.4)

diagonalizes \( V \)

\[ T^{-1}(k) V T(k) = \delta \]  \hspace{1cm} (6.5)

where

\[ \delta = \begin{bmatrix} \delta_1 & 0 & \cdots & 0 \\ 0 & \delta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta_p \end{bmatrix} \]

(6.5) may be written in the form

\[ V_s = T(k) \delta T^{-1}(k) \]

from which we compute

\[ V_s^2 = [T(k) \delta T^{-1}(k)][T(k) \delta T^{-1}(k)] = T(k) \delta^2 T^{-1}(k) \] \hspace{1cm} (6.6)

Induction shows that in general

\[ V_s^n = T(k) \delta^n T^{-1}(k) \], \quad n = 1, 2, \ldots \] \hspace{1cm} (6.7)

Using (6.7), we have

\[ e_s = \sum_{n=0}^{\infty} \frac{V_s^n}{n!} = \sum_{n=0}^{\infty} \frac{T(k) \delta^n T^{-1}(k)^n}{n!} = T(k) \left( \sum_{n=0}^{\infty} \frac{\delta^n}{n!} \right) T^{-1}(k) \] \hspace{1cm} (6.8)

Since \( \delta \) is a diagonal matrix,

\[ \delta^n = \begin{bmatrix} \delta_1^n & 0 & \cdots & 0 \\ 0 & \delta_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta_p^n \end{bmatrix} \]
and hence

\[
\begin{vmatrix}
\delta_{n_1}^n \\
\sum_{n=0}^{\infty} \frac{n^n}{n!}
\end{vmatrix}
\]

Using (6.8) and (6.9), solution (6.1) can be written as below.

\[
e^{\frac{\alpha t}{a}} F(x) = T(k)H(t)T^{-1}(k)P(x)
\]

Since \( P(x) = I \), the particular solution for the differential equation is

\[
P(k) = T(k)H(t)T^{-1}(k)
\]

Expanding (6.10) we obtain the explicit solution for the transition probabilities below.

\[
P_{ij}(t) = \sum_{k=1}^{p} A_{ki}(t) \frac{T_i^j(k)}{T(k)} e^{\frac{\alpha t}{a}}, \ i,j = 1,2,\ldots,p
\]

Equation (6.11) holds true whatever may be \( k=1,2,\ldots,p \). The second form of the solution is in terms of eigen values as below.
where \( A'(\ell) \) is the transpose of \( A(\ell) \) (6.11) and (6.12) are also given in Chiang (1968).

### Transition Probabilities Leading to Excretion \( P_{oi}(t) \)

The probability \( P_{oi}(t) \) that an individual in compartment \( i \) at time zero will be out of the system at time \( t \) and the corresponding transition probability matrix \( P_{o}(t) \) can be derived from probabilities \( P_{ji}(t) \). An individual in compartment \( i \) may reach the exterior of the system directly from compartment \( i \) or by way of some other compartment \( j \), \( j \neq i \). Since an individual in exit at time \( t \) may have reached that state at any time prior to \( t \), let us consider an infinitesimal time interval \((\tau, \tau + \Delta \tau)\) for a fixed \( \tau \), \( 0 < \tau \leq t \). The probability of an event \( A_i \), where \( A_i \) is such that an individual in compartment \( i \) at time zero will reach out of the system in the interval \((\tau, \tau + \Delta \tau)\) is given by

\[
P(A_i) = P_{ii}(\tau) \text{\#}_{oi} \Delta \tau + \sum_{j=1}^{p} \sum_{j \neq i} P_{ji}(\tau) \text{\#}_{oj} \Delta \tau
\]

(6.13)

As \( \tau \) varies over the interval \((0, t)\) the events \( A_i \) are mutually exclusive. Hence

\[
P_{oi}(t) = \int_{0}^{t} \left( \sum_{j=1}^{p} P_{ji}(\tau) \text{\#}_{oj} \right) d\tau
\]

(6.14)

and the corresponding matrix equation

\[
P_{O}(t) = \int_{0}^{t} P(t) \text{\#} d\tau
\]

(6.15)
where the intensity matrix \( \mathcal{I} \) is defined below

\[
\mathcal{I} = \begin{pmatrix} \nu_{01} & \nu_{02} & \cdots & \nu_{0p} \\
\end{pmatrix}.
\]

The integrand is a matrix and by the integral is meant the matrix of integrals.

Substituting (6.10) in (6.15) and integrating we obtain

\[
P_0(t) = \int_0^t T(k) \mathcal{H}(\tau) T^{-1}(k) \mathcal{I} \, d\tau
\]

\[
= \mathcal{I}^T T(k) \mathcal{H}(t) - \mathcal{I}^T T^{-1}(k) \mathcal{I}
\]

(6.16)

The individual transition probability \( p_{ij}(t) \) can be obtained either from (6.14) or by expanding (6.16). It is given by

\[
p_{0i} = \sum_{\ell=1}^{p} \sum_{j=1}^{p} \lambda_{k\ell}(\mathcal{I}) \frac{T_{\ell\ell}(k)}{T(k)} \delta_{\ell j} \delta \left( e^{\delta t} - 1 \right) \nu_{ij}
\]

(6.17)

6.3 Population Sizes and Associated Probability Distribution

An individual in compartment \( i \) at time \( o \) must be either in one of the \( p \)-compartments or in the exterior of the system at time \( t \). Consequently the corresponding transition probabilities add to one, so that

\[
\sum_{j=1}^{p} p_{ij}(t) + p_{0i}(t) = 1
\]

(6.18)

(6.18) may be used to derive the probability distribution of population sizes in \( p \)-compartments at any time \( t \). Let \( N_i(o) \) individuals be in compartment \( i \) at time \( o \). Let \( N_i(t) \) be a random variable specifying the number of individuals in compartment \( i \) at time \( t \). The initial size of the population is \( N(o) = \sum_{i=1}^{p} N_i(o) \).
The joint distribution of the random variables \( N_i(t) \) \( i=1,2,...,p \), is obtained. \( N_i(t) \) can be characterized according to their state of time \( o \). This is expressed by the formula

\[
N_i(t) = N_{i'1}(t) + N_{i'2}(t) + ... + N_{i'p}(t)
\] (6.19)

where \( N_{ij}(t) \) is the number of individuals in compartment \( i \) who were in compartment \( j \) at time \( o \). On the other hand, each of the \( N_i(o) \) individuals in compartment \( i \) at time \( o \) must be in one of the compartments or in the exterior of the system at time \( o \). That is, at any instant \( t \),

\[
N_i(o) = \sum_{j=1}^{p} N_{ij}(t) + N_{oi}(t)
\] (6.20)

From (6.18), we have

\[
1 = \sum_{j=1}^{p} P_{ij}(t) + P_{oi}(t),
\]

therefore, given \( N_i(o) \), the random variables on the right side of (6.20) have joint multinomial distribution.

We describe the random variables \( N_{ij}, N_{oi}(t), i,j=0,1,2,...,p \) in table 6.1.

### Table 6.1

Location of individuals in \( p \)-compartments at time \( t \) according to initial compartment \( i \) at time \( o \).

<table>
<thead>
<tr>
<th>Compartment at time ( o )</th>
<th>Compartment at time ( t )</th>
<th>Exterior ( o )</th>
<th>Initial population sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>( N_{11}(t) ) ( N_{21}(t) ) ( N_{31}(t) ) ... ( N_{p1}(t) )</td>
<td>( N_{o1}(t) )</td>
<td>( N_{1}(o) )</td>
</tr>
<tr>
<td>2</td>
<td>( N_{12}(t) ) ( N_{22}(t) ) ( N_{32}(t) ) ... ( N_{p2}(t) )</td>
<td>( N_{o2}(t) )</td>
<td>( N_{2}(o) )</td>
</tr>
<tr>
<td>3</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>( p )</td>
<td>( N_{1p}(t) ) ( N_{2p}(t) ) ( N_{3p}(t) ) ... ( N_{pp}(t) )</td>
<td>( N_{op}(t) )</td>
<td>( N_{p}(o) )</td>
</tr>
</tbody>
</table>
The behavior of the $N_i(0)$ units in distributing themselves among the various compartments is independent of $N_j(0)$ units originating in compartment $j \neq i$. Hence in table (7.1) for $N_{ij}(t)$, the rows are independent. However, the total (over all $p$ origins) number of units in compartment $j$, $N_j(t)$ is not independent of the other marginal totals, hence the columns are dependent.

For each $i$ the random variables $N_{ij}(t)$, $j=0,1,2,\ldots,p$ have a multinomial distribution. Their probability generating function (pgf) is given by

$$E[s_{1i}N_{1j}(t)s_{2j} \cdots s_{pj}N_{pj}(t)s_{0j} | N_i(0)]$$

$$= [p_{11}(t)s_1 + p_{21}(t)s_2 + \cdots + p_{pi}(t)s_i + p_{01}(t)s_0]^{N_i(0)}$$

Therefore the pgf of the joint probability distribution for the population sizes of all the compartments at time $t$ is

$$P[s_{11},s_{21},\ldots,s_{p1},s_{01}] = E[s_{11}N_{11}(t)s_{21}N_{12}(t) \cdots s_{p1}N_{p1}(t)s_{01}N_{01}(t) | N_1(0),N_2(0),\ldots,N_p(0)]$$

$$= \prod_{i=1}^{p} [p_{11}(t)s_1 + p_{21}(t)s_2 + \cdots + p_{pi}(t)s_i + p_{01}(t)s_0]^{N_i(0)}$$

The joint probability distribution is then given by

$$P[N_1(t)=n_1,N_2(t)=n_2,\ldots,N_p(t)=n_p,N_o(t)=n_o | N_1(0),N_2(0),\ldots,N_p(0)]$$

$$= \sum \prod_{i=1}^{p} \frac{N_i(0)!}{n_1! \cdots n_i!} p_{1i}^{n_1}(t)p_{2i}^{n_2}(t) \cdots p_{oi}^{n_o}(t)$$

(6.23)

where the summation is taken over all possible values of $n_{ji}$.
i, j = 1, 2, ..., p, so that

\[ n_{j1} + n_{j2} + \ldots + n_{pj} = n_j \quad j = 0, 1, 2, \ldots, p. \]

The cumulant generating function, using \( s_i = e^{\theta_i} \), is given by

\[
k(\theta_1, \theta_2, \ldots, \theta_p, t) = \log P(e^{\theta_1}, e^{\theta_2}, \ldots, e^{\theta_p}, e^{\theta_0})
\]

\[
= \sum_{i=1}^{p} N_i(0) \log \left( \sum_{j=1}^{p} P_{ji}(t) e^{\theta_j} \right)
\]

(6.24)

Noting that \( s_0 = 1 \) and consequently \( \theta_0 = 0 \), (6.24) is equivalent to

\[
= \sum_{i=1}^{p} N_i(0) \log \left[ \sum_{j=1}^{p} e^{\theta_j} P_{ji}(t) + P_{0i}(t) \right]
\]

or

\[
= \sum_{i=1}^{p} N_i(0) \log \left[ \sum_{j=1}^{p} e^{\theta_j} P_{ji}(t) + (1 - \sum_{j=1}^{p} P_{ji}(t)) \right]
\]

or

\[
k(\theta_1, \ldots, \theta_p, t) = \sum_{i=1}^{p} N_i(0) \log \left[ \sum_{j=1}^{p} (e^{\theta_j-1}) P_{ji}(t) + 1 \right]
\]

(6.25)

This is the cumulant generating function also obtained by Matis and Hartley (1971) by different procedure. Their approach is lengthy and much complicated.

The expected number of individuals in each compartment at time \( t \) can be computed directly from (6.21).

\[
E[N_j(t) | N_1(0), \ldots, N_p(0)] = \sum_{i=1}^{p} N_i(0) P_{ji}(t)
\]

(6.26)

The corresponding variances and covariances are
\[
\text{Cov}\{N_k(t), N_j(t)\} = -\sum_{i=1}^{p} N_i(o) P_{ki}(t) P_{ji}(t) \\
\]

\[
\sigma^2_{N_j(t)} = \sum_{i=1}^{p} N_i(o) P_{ji}(t) [1 - P_{ji}(t)] \\
\]

\[
\sigma^2_{N_k(t)} = \sum_{i=1}^{p} N_i(o) P_{oi}(t) [1 - P_{oi}(t)] \\
\]

The population becomes extinct at time \( t \) if \( N(t) = N_2(t) = \ldots = N_p(t) = 0 \). The probability of extinction at time \( t \) may be obtained from (6.21) by setting \( s_1 = s_2 = \ldots = s_p = 0 \) and \( s_o = 1 \).

The expected values of \( N_j(t) \) given by \( \lambda_j(t) \) is below.

Let \( \bar{M}(t) = [\bar{\mu}_1(t), \bar{\mu}_2(t), \ldots, \bar{\mu}_p(t)] \)

\[
\bar{M}(t) = \begin{bmatrix}
P_{11}(t) & P_{12}(t) & \cdots & P_{1p}(t) \\
P_{21}(t) & P_{22}(t) & \cdots & P_{2p}(t) \\
\vdots & \vdots & \ddots & \vdots \\
P_{p1}(t) & P_{p2}(t) & \cdots & P_{pp}(t) \\
\end{bmatrix}
\begin{bmatrix}
N_1(o) \\
N_2(o) \\
\vdots \\
N_p(o) \\
\end{bmatrix}
\]

or

\[
\bar{M}(t) = P'(t) \bar{M}(o) \\
\]

Differentiating with respect to \( t \) both sides of (6.28)

\[
\frac{d}{dt} M(t) = P'(t) \bar{M}(o) \\
\]

\[
= \lambda_{\bar{s}} P'(t) \bar{M}(o) \\
\lambda_{\bar{s}} P'(t) \bar{M}(o) = 0 \\
\]

\[
= \lambda_{\bar{s}} \bar{M}(t) \\
\]

Equation (6.29) is identical to the deterministic equations of a general \( p \)-compartment system. That is, when \( N_i(o), i = 1,2,\ldots,p \), are given constants and the probabilities are intensities, the
A stochastic model is the deterministic one.

In practical situations, one observes the number of individuals excreted to the exterior of the system as observations on individual compartments are either impossible or difficult. Let $N_T(t)$ denote the number of individuals remaining in the system at time $t$. Clearly $N_T(t) = \sum_{i=1}^{p} N_i(t)$ and

$$E[N_T(t)] = \mu_T(t) = \sum_{j=1}^{p} \sum_{i=1}^{p} N_i(o) P_{ji}(t)$$

$$= \sum_{i=1}^{p} N_i(o) \sum_{j=1}^{p} P_{ji}(t)$$

$$= \sum_{i=1}^{p} N_i(o)[1-P_{oi}(t)]$$

$$= N(o) - \sum_{i=1}^{p} N_i(o) P_{oi}(t)$$

or

$$\mu_T(t) = N(o) - \mu_o(t) \quad (6.30)$$

6.4 Joint Probability Distribution of the Numbers in the System at Different Times.

Let, for a given $u$, $N_T = [N_T(t_1), N_T(t_2), \ldots, N_T(t_u)]$ be the number of individuals in the system at $t_1, t_2, \ldots, t_u$ times respectively. The pgf of the joint probability distribution of $N_T$ is given by

$$G_{N_T}[N(o)(s_1, s_2, \ldots, s_u)] = E[s_1^{N_T(t_1)} s_2^{N_T(t_2)} \ldots s_u^{N_T(t_u)} | N(o)] \quad (6.31)$$

where $|s_i| < 1$ for $i = 1, 2, \ldots, u$

To derive an explicit formula for the pgf (6.31), we use the identity

$$E[s_1^{N_T(t_1)} s_2^{N_T(t_2)} \ldots s_u^{N_T(t_u)} | N(o)] = E[s_1^{N_T(t_1)} \ldots s_{u-1}^{N_T(t_{u-1})} E[s_u^{N_T(t_u)} | N(t_{u-1})], \ldots, N(t_u)| N(o)] \quad (6.32)$$
Using the Markov property, (6.31) becomes

\[ N_T(t_1) \ldots N_T(t_u-1) N_T(t_u) = E[s_1 \ldots s_{u-1} E[s_u | N_T(t_u-1)] N(o)] \] (6.33)

Repeated application of the same process gives

\[ G_{N_T|N(o)}(s_1, s_2, \ldots, s_u) = (1 - (1 - P_{o1})(1 - s_1) + (1 - P_{o2})s_1(1 - s_2) + \]

\[ (1 - P_{o3})s_1s_2(1 - s_3) + \ldots \]

\[ + (1 - P_{ou})s_1s_2 \ldots s_{u-1}(1 - s_u)] N(o) \] (6.34)

Here \( P_{oi} \) is the probability that the individual in the system moves out of the system in the interval \((o, t_i), i = 1, 2, \ldots, u.\)

(6.34) is then used for deriving the joint probability function of \( N_T \) and their moments are given below.

\[ \sigma_i^2 = \text{Var}[N_T(t_i)] = \text{Var}[\sum_{k=1}^{p} N_k(t_i)] \]

\[ = \text{Var}[N(o) - N_0(t_i)] \]

\[ = \sum_{k=1}^{p} N_k(o)P_{ok}(t_i)[1 - P_{ok}(t_i)] \] (6.35)

The covariance \( \sigma_{ij} \) of the process at two different times, \( t_i \) and \( t_j \) is obtained as usual and is given by

\[ \sigma_{ij} = \sum_{k=1}^{p} N_k(o)P_{ok}(t_j)[1 - P_{ok}(t_i)] \] (6.36)

These results show that for a given \( u, N_T(t_1), \ldots, N_T(t_u) \) for the process form a chain of binomial distributions. For a given \( t, N_K(t) \) can be regarded as a mixture of multinomial distributions. Indeed \( N_T(t_i) - N_T(t_j) \) for various intervals of \( t \) may also be regarded
as a mixture of multinomial distributions, where the kth component results from the \( N_k(o) \) units placed in the kth compartment. Equations (6.35) and (6.36) may be combined into the following result:

**Proposition I.** Let \( \sigma_{ij} = \text{Cov}[N_T(t_i), N_T(t_j)] \) be covariance kernel of the process describing the total number of individuals in the system at times \( t_j \) and \( t_i \) such that \( t_j \geq t_i \). Then

\[
\sigma_{ij} = \sum_{k=1}^{p} N_k(o) P_{ok}(t_j)[1-P_{ok}(t_i)]
\]

### 6.5 Least Squares Estimates

It has been shown that the distribution of \( [N_{11}(t), N_{21}(t), \ldots, N_{p1}(t)] \) for \( i = 1, 2, \ldots, p \) is multinomial with parameters \( N_i(o), P_{11}(t), P_{21}(t), \ldots, P_{p1}(t) \). Least squares estimates of the transition probability rates or \( v_{ji} \) from output data can be obtained.

When data are available only on the total number of individuals in the system at various times, a two stage least squares procedure is given by Matis and Hartley (1971). The procedure is described briefly below.

Assuming the model,

\[
\chi = g(x, \Omega) + \xi
\]

where

(i) \( \chi = [N_T(t_1), N_T(t_2), \ldots, N_T(t_u)]' \)

(ii) \( \Omega = p^2 \)-parameter vector of the rates

\[
[v_{o1}, v_{21}, v_{31}, \ldots, v_{pp}]'
\]
(iii) \[ g(t, \Omega) = [\mu_T(t_1), \mu_T(t_2), \ldots, \mu_T(t_u)]' \]

(iv) \[ \xi = [\epsilon_1, \epsilon_2, \ldots, \epsilon_u]' \]

Further assume \( E(\xi) = \Omega, \) \( \text{Cov}(\xi) = \Sigma = (\sigma_{ij}) \). Expanding, 
\( g(t, \Omega) \) by a first order Taylor series about some initial estimates of the parameter vector, \( \Omega_0 \) and using the modified Gauss-Newton technique, Hartley (1961) we get,

\[ \chi - g(\xi, \Omega_0) = \mathbf{G} \Delta \Omega_0 + \epsilon \]  \[ (6.38) \]

Here

(i) \[ \mathbf{G}_0 = \left[ \frac{\partial g(t_i, \Omega)}{\partial \nu_k} \right]_{\Omega = \Omega_0} \] is the \( \exp^2 \) matrix of first partials of the elements \( g(t_i, \Omega), i = 1, 2, \ldots, u \), with respect to the parameters \( \nu_k, k = 1, 2, \ldots, p^2 \),

(ii) \[ \Delta = [\Delta_1, \Delta_2, \ldots, \Delta_{p^2}]' \] is the \( p^2 \)-vector of the differences \( \nu_k - \nu_{0k} \),

The best linear unbiased estimates (BLUE) of the \( d_{ij} \) are found by minimizing the generalized sum of squared deviations, \( \xi \Sigma^{-1} \xi \).

By Aitken theorem, (Goldberger [1964] p. 233) the BLUE \( \Delta_0 \) of \( \Delta \) is given by

\[ \Delta_0 = \left[ G_o \Sigma^{-1} G_o \right]^{-1} G_o \Sigma^{-1} [\chi - g(\xi, \Omega_0)] \]  \[ (6.39) \]

For the iterative procedure, initially one assumes \( \Sigma = \mathbb{I} \), for estimating the parameters \( \nu_{j1} \). These estimates are used to estimate \( \xi \) which in turn maybe used for another iteration to estimate the parameters. The numerical procedure is continued until the parameter estimates converge. Such two-stage least squares procedures are commonly used in practice for other kinds
of nonlinear models. In view of estimation of probabilities in this model, minimization with constants of the \( \nu_{ji} \geq 0 \) is more appropriate. Procedures using mathematical programming methods can be applied in many cases of the above type.
COMPARTMENTAL MODELS ARE WIDELY USED IN PHARMACEUTICAL RESEARCH FOR QUANTITATIVE STUDY OF THE KINETICS OF ABSORPTION, DISTRIBUTION, METABOLISM AND EXCRETION OF A DRUG. THE DETERMINISTIC COMPARTMENTAL MODELS FOR THE STEADY STATE SYSTEM ARE DESCRIBED BY A SYSTEM OF LINEAR FIRST ORDER DIFFERENTIAL EQUATIONS. THE SOLUTION OF THESE DIFFERENTIAL EQUATIONS IS SHOWN TO BE LINEAR COMBINATIONS OF THE SAME EXPONENTIAL PARAMETERS. ALSO THE NUMBER OF INDEPENDENT REGRESSION EQUATIONS AS SOLUTION OF DIFFERENTIAL EQUATIONS IN COMPARTMENT ANALYSIS IS SHOWN TO BE THE SAME AS THE NUMBER OF EXPONENTIAL PARAMETERS.

THE LEAST SQUARES COMPUTATION HAS SEVERAL UNUSUAL FEATURES WHEN APPLIED TO LINEAR COMBINATIONS OF EXPONENTIALS. THE MOST UNUSUAL ASPECT IS THE FREQUENT FAILURE OF THE ITERATIVE COMPUTATION SCHEMES TO CONVERGE. SECONDLY, THE ITERATIVE PROCESS CONVERGES BUT THE RESULTING ESTIMATORS MAY NOT BE THE LEAST SQUARES ESTIMATES. THIS BEHAVIOR IS A CONSEQUENCE OF AN UNUSUAL FEATURE OF THE SUM OF SQUARES SURFACE, NAMELY, THAT IT POSSESSES MORE THAN ONE minimum WHENEVER THERE ARE MORE THAN ONE COMPARTMENT. THE PRESENCE OF MORE THAN ONE minimum IMPLIES THE EXISTENCE OF OTHER STATIONARY POINTS ON THE SUM OF SQUARES SURFACE. THE CHOICE OF INITIAL ESTIMATES IN THE NEIGHBORHOOD OF A STATIONARY POINT LEADS TO ESTIMATES WHICH ARE NOT LEAST SQUARES ESTIMATES.

IN THIS DISSERTATION ESTIMATION PROCEDURE FOR INITIAL ESTIMATES IS DEVELOPED UNDER THE ASSUMPTION OF EQUALLY SPACED VALUES OF THE
independent variable. This method utilizes independent partial totals of the observations and provides a simple and direct procedure for estimating the parameters. It is shown that the method provides consistent estimates of the parameters. Also the method has been extended for simultaneous estimation of parameters. The procedure is illustrated with two numerical examples from the literature. It is found that there is considerable reduction in residual sum of squares as compared to other methods.

The compartmental equations are simultaneous linear first order differential equations in $p$ unknowns where $p$ is the number of compartments assumed in the system. $(p-1)$ unknowns may be eliminated to yield the $p$th order linear differential equation with constant coefficients. A difference equation model is proposed for the estimation of exponential parameters. This model is derived by approximating the $i$th order differential coefficient, in the $p$th order linear differential equation, by the $i$th order differences, $i=1,2,...,p$. This model is similar to time series models which have been extensively studied in literature. The estimates obtained are maximum likelihood estimates when it is assumed that the errors are normally distributed. When the assumptions underlying the difference equation model do not hold, the model still provides good initial estimates for the usual compartmental model whereas other methods fail to provide initial estimates.

A question of considerable interest relates to the number of compartments required to specify a system. So far no exact tests were available for hypothesizing the number of compartments. Tests
based on difference equation model are derived for determining the number of compartments for equispaced observations as well as unequally spaced observations.

The application of the difference equation model is made to three experiments reported in the literature. This model provides a good fit to the observed experimental data.

Experimental observations from an individual compartment in the case of single intravenous dosing, form a decay-type data. The problem of optimum choice of empirical weights for obtaining least squares estimates under order restrictions is examined. It is observed that the choice of antitonic weights in the case of models with constant error variance perform better, resulting in reduced residual sum of squares. The optimum choice of weights happens to be the weights proportional to the absolute magnitudes of the observations. These findings have been confirmed by examining both real and simulated data.

A discrete population of particles in a steady state system having several compartments is considered. Stochastic models are developed in terms of transition probabilities giving the probabilities of transfer among compartments. Efficient estimation of these transition rate parameters requires the associated distribution theory which is developed by providing a compact analytic solution to the compartmental problem.

Often in practice, individual compartments are inaccessible for observation and instead time series data are available only on the passage of material to the system exterior. The covariance kernel of such observations is derived for utilization in parameter estimation.
The problems of estimation of parameters relating to bioavailability, and of formulation of the statistical model, have been examined in this dissertation. The study has been confined to only steady state systems, that is, the systems for which the transfer rates are independent of time. However, several problems in the study of bioavailability and pharmacokinetics need further investigation. Some of these problems are pointed out below.

The classical compartmental analysis assumes that if the dose is progressively raised the amount of drug transferred to different compartments is directly proportional to dose. In many situations a nonlinear compartmental model is assumed if the amount of drug transferred to compartments is a curvilinear function of the dose. A statistical study of the nonlinear compartmental models is needed. Another problem of interest is the theoretical investigation of multicompartmental models where multiple doses are considered. Extensive study is needed for the development of experimental designs, determination of sample size and sampling times in the study of these problems in bioavailability and pharmacokinetics. The stochastic compartmental models are needed for continuous populations.
REFERENCES


Grard, F., "A general program for statistical analysis using the maximum-likelihood method," "Malik' Program. Lawrence Radiation Laboratory, Univ. of California, Berkeley, California. UCRL-10133 (1962).


Shah, B.K., "Data analysis problems in the area of clinical and pharmacokinetics research with application to comparison of oral and intravenous doses," Presented at the spring meeting of the Biometric Society at Florida State University, (1974).


