1976

A Semi-markov Model For The Prognosis Of The Cancer Of The Cervix

Rajendra Kumar Jain

A SEMI-MARKOV MODEL FOR THE PROGNOSIS OF THE CANCER OF THE CERVIX

by

Rajendra Kumar Jain

Department of Epidemiology and Preventive Medicine

Submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Faculty of Graduate Studies

The University of Western Ontario

London, Ontario

July 1976

ABSTRACT

A model based on the Semi-Markov process for predicting the survival of cancer of the cervix patients is proposed. The model has three transient states: under-treatment, recovery and relapse; and two absorbing states: death due to cancer of the cervix and death due to other causes or lost to follow-up. The Semi-Markov model is a generalization of a Markov model in which the transition probability from a given state to any other state depends not only on the given state but also on the duration in that state depending on the next state entered.

The model was tested on 101 patients available from Victoria Hospital Cancer Clinic, London, Ontario. The transition probability matrix and holding-time matrix (the two sets of parameters of this model) were estimated by the maximum likelihood method. The first five years data after diagnosis was used to estimate these parameters. The predictions were then made for eight years from initial observation point. Comparisons of survival rates predicted by the Semi-Markov and Markov models were made with the observed rates. The predicted rates by the Semi-Markov model were in close agreement with the observed rates; whereas the Markov model underestimated the survival rates at all points. Thus, the Semi-Markov model can be used to describe the prognosis of cancer

of the cervix. In addition, it can be used for predicting future survival rates from a limited follow-up. The life table lacks this flexibility.

The effect of changes in the parameters of the model on survival are discussed. This type of analysis provides a tool to clinicians to compare the alternative approach in treatment policy without actually performing the experiment. Various other uses of the model are also presented.

ACKNOWLEDGEMENTS

The author is deeply indebted to his supervisor Dr. S.P. Sangal, for his valuable advice, assistance, guidance and constructive criticism in the preparation of this thesis.

The author wishes to express his sincere thanks and appreciation to Drs. J.M. Wanklin and K.M. Stavraky, for their suggestions, advice, and continuous encouragement throughout the course of this study. Their constructive comments were immensely helpful in improving the quality and content of the final product. The author is grateful to them for finding time in their busy schedule to read the earlier drafts.

The author is also grateful to Dr. J.C. Baskerville for many stimulating discussions and advice from time to time.

The author would like to extend a special word of thanks to Dr. T.A. Watson, Director, Victoria Hospital Cancer Clinic, London, Ontario for providing patient release data.

Finally, the author is grateful to Dr. C. Buck, Chairman, Department of Epidemiology and Preventive Medicine, for providing the necessary facilities and financial support.

This research was supported by a grant from the Department of National Health and Welfare, Canada.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Objectives

Traditionally, the prognosis of chronic diseases, and cancer in particular, has been studied by computing the 'T-year survival rate'*. The method of choice often is the clinical life table as developed by Cutler and Ederer (1958). Clinical trials have also employed T-year survival rates to compare the efficacy of various treatments. This approach requires T-year follow-up of at least some surviving patients. However, if the investigator is interested in the survival beyond T-year, this method does not provide an adequate answer to it. Further, the efficacy of the treatments can also be compared by studying not only the survival but the duration of a relatively 'disease free' period. Clinicians and health planners may be interested in estimating the length of stay distribution of the patients among the various states (viz., disease free and requiring treatment etc.) at various time points in future for planning medical facilities.

---

* T-year survival is used for convenience. However, any time interval can be used without altering any conclusion and meaning.

The clinical life table approach for analysing survival curves is not suitable to answer these questions when only a limited follow-up data is available.

Attempts towards developing methodology to answer these questions were made in the early 40's and 50's. The notable works have been due to Fix & Neyman (1951), Boag (1949) and Berkson & Gage (1952). One of the major problems that was ignored by Fix and Neyman (1951) was the heterogeneity of the patients in terms of their behaviour of relapse and recovery depending upon the duration of the exposure to these conditions. Boag (1949) and Berkson & Gage (1952) used a conceptually different approach to tackle this problem. They divided a group of cancer patients into cured and non-cured groups of patients. The uncured patients were assumed to follow a lognormal distribution by Boag (1949) and an exponential distribution by Berkson & Gage (1952) while the cured patients followed the survival pattern of a general population. However, Berg (1965) noted that this approach did not explain the observed survival pattern. Therefore, the present aim of this thesis is to

1. suggest a Semi-Markov model to overcome the difficulties mentioned above, in particular-
   (a) predict survival rate from the short term data,
   (b) estimate length of duration in various states (viz., disease free and requiring treatment etc.);
2. study cancer of the cervix by the model developed above; and
3. compare this model with other methods.

1.2 Organization

The material of this thesis is organised as follows:

Chapter 2 begins with an introduction of Markov and Semi-Markov process and proceeds to establish the necessary definitions, notations and results which are most relevant to this project. The use of Semi-Markov concepts from this chapter is made in chapter 4 and 5 to study the survival distribution of cancer of cervix patients.

In chapter 3, a critical review of the uses of the Markov model and the other methods in the prognosis of cancer patients are presented.

Chapter 4 is devoted solely to the development of the Semi-Markov model in the analysis of the prognosis of cancer of cervix patients. Estimates based on Semi-Markov model are compared with those obtained by using Markov model developed by Fix and Neyman (1951) and the observed values.

Chapter 5 is concerned with some of the applications of the Semi-Markov model and it is noted that the model is flexible enough to be used in studying the effect of changes in the parameters.

In chapter 6, conclusions, limitations and applications of the model are discussed. Suggestions for future research are also presented.

CHAPTER 2

MARKOV AND SEMI-MARKOV PROCESSES

2.1 Introduction

This chapter contains a summary of those concepts in Markov and Semi-Markov processes which are necessary for the development and application of these processes in the analysis of the prognosis of cancer patients. The basic notations and terminology of Markov and Semi-Markov processes of this chapter will be explained in terms of malignant disease. The contents of this chapter will be used freely throughout this thesis. Detailed treatments of the Markov and Semi-Markov processes are given in Kolmogorov (1950), Doob (1953), Smith (1955), Feller (1957), Pyke (1958, 1961), Parzen (1962), Karlin (1969) and Howard (1971).

2.2 Notation

The following notations will be used extensively:

n = Number of states in the model;

For example, using cancer as the model disease, the states of the patients can be characterized as follows:

S_1 = under-treatment, S_2 = apparent recovery, S_3 = relapse, S_4 = death due to cancer and S_5 = lost from the study or death due to other causes. Therefore, in this example $n = 5$.

p_{ij} = Transition probability;

Conditional probability that the next transition is to state j, given that the last transition was to state i. Thus, for example, in cancer patients the transition probability p_{12} is the probability that a patient presently in state S_1 (under-treatment) will occupy state S_2 (apparent recovery) after the next transition.

τ_{ij} = The holding-time for the transition $i \to j$, the time the patient will spend in state i before making a transition to j. For example, suppose a patient spends 3 months under-treatment (S_1) before he recovers (S_2), then in that case $\tau_{12} = 3$.

$\phi_{ij}(t)$ = Interval transition probability from state i to state j defined as:

Probability that the patient will be in state j at time t given that he was in state i at time zero.

$H_{ij}(t)$ = Holding-time distribution function in state i for a patient who will next be in state j;

Probability that the patient will be in state j at time t before the next transition to state j given an i to j move.

$$H_{ij}(t) = \text{Prob}(\tau_{ij} \leq t)$$

$$= \sum_{t_1 = 0}^{t} \text{Prob}(\tau_{ij} = t_1)$$

$$= \sum_{t_1 = 0}^{t} h_{ij}(t_1);$$

where, $h_{ij}(t_1) = \text{Prob}(\tau_{ij} = t_1)$.

2.3 Definitions

In this section the basic definitions are introduced which will be used throughout this thesis.

2.3.1 Stochastic Process. Consider a random variable such as the occurrence of deaths. We may describe these deaths by a counting function $x(t)$, defined for all $t \geq 0$, which represents the number of deaths that have occured during the time period from 0 to t. For each time t, the value of $x(t)$ is an observed value of a random variable. The family of random variable $\{x(t), t \geq 0\}$ is called a stochastic process.

2.3.2 Markov Process. A discrete parameter stochastic process $\{x(t);\ t = 0,1,2,.....\}$ or a continuous parameter process $\{x(t);\ t \geq 0\}$ constitutes a Markov process if, for any set of n time points $t_1 < t_2 < ..... < t_n$ in the index set of the process and any real numbers $S_1, S_2, ..., S_n$,

$$\text{Prob}\left[x(t_n) = S_n \mid x(t_{n-1}) = S_{n-1}\right]$$

$$= \text{Prob}\left[x(t_n) = S_n \mid x(t_1) = S_1, ..., x(t_{n-1}) = S_{n-1}\right] \quad (2.3.1)$$

This Markovian property has been shown to be equivalent to stating that the conditional probability at time t_n depends on the outcome at time t_{n-1} and on nothing that occurred at any previous time point.

Thus a Markov process can be characterized as follows:

(a) A set of states

$$S = (S_1, S_2, ..., S_n) \quad (2.3.2)$$

which can be considered finite in our discussions;

(b) A probability distribution over the states at each time t

$$p(t) = (p_1(t), p_2(t), ..., p_n(t)) \quad (2.3.3)$$

where $(p_i(t),\ i = 1, 2, ..., n)$ is real, non-negative and subject to the condition

$$\sum_{i=1}^{n} p_i(t) = 1 \quad (2.3.4)$$

for all t;

(c) and a matrix

$$P(t) = [p_{ij}(t)] = \begin{bmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1n}(t) \\ p_{21}(t) & p_{22}(t) & \cdots & p_{2n}(t) \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1}(t) & p_{n2}(t) & \cdots & p_{nn}(t) \end{bmatrix} \qquad (2.3.5)$$

The typical element of $P(t)$ is also real and non-negative and subject to the condition

$$\sum_{j=1}^{n} p_{ij}(t) = 1 \qquad (2.3.6)$$

Such a matrix $P(t)$ is called a transition probability matrix in which the typical element $p_{ij}(t)$ is defined as the conditional probability that an element is in state j at time t given that it was in state i at time $(t-1)$.

2.3.3 Markov Chain. A Markov chain is a Markov process whose state space is countable or finite and for which time is discrete.

Remark. The analysis of a Markov chain consists mainly in calculating the t-step transition probabilities $P^{(t)} = [p_{ij}^{(t)}]$. If the transition probability matrix P is given by equation (2.3.5), where

$$p_{ij}^{(s)} = \text{Prob}\left[\, x(k+s) = j \mid x(k) = i \,\right],$$

$$s \geq 1,$$

and $p_{ij}^{(1)} = p_{ij}$.

Then the Markovian assumption allows us to express t-step transition probabilities in the following form:

$$P^{(t)} = P^t$$

which implies that the numbers $p_{ij}^{(t)}$ may be regarded as the entries in the matrix P^t, the t^{th} power of P.

The Chapman-Kalmogorov equation provides a method for calculating these t-step transition probabilities which are given by

$$p_{ij}(t) = \sum_{k=1}^{n} p_{ik}(r)p_{kj}(t-r) \qquad (2.3.7)$$

for any $r \leq t$. The summation in the equation is over all states of the Markov chain.

2.3.4 Stationary Transition Probabilities. When transition probabilities are independent of t, i.e.,

$$P(t) = P \qquad (2.3.8)$$

Then such transition probabilities are said to be stationary.

2.3.5 Transient State. A state i is said to be transient if and only if starting from state i there is a positive probability that the process may not eventually return to this state. For example in cancer of the cervix, state of recovery will be a transient state.

2.3.6 Absorbing State. A state i is said to be an absorbing state if and only if $p_{ii} = 1$. Death due to cancer of the cervix is an absorbing state.

2.3.7 Semi-Markov Process. Let $\{ x(t); t \geq 0 \}$ be a stochastic process, where $x(t) = i$ denotes that the process is in state i at time t. Suppose that the process has just entered a state, say i, the transition to the next state is made according to the transition probability matrix

$$P = [p_{ij}] \qquad (2.3.9)$$

and the distribution function for the 'wait' of the process in state i given that the next transition will be to state j is denoted by the holding-time matrix

$$H(t) = [h_{ij}(t)]. \qquad (2.3.10)$$

Such a process $x(t)$ based on two sets of parameters $(P, H(t))$ is called a Semi-Markov process.

Remark. If we ignore the random nature of transition times in Semi-Markov process and focus on the transition instants, the process is an ordinary Markov process. However, when the holding-time behaviour is included, the Semi-Markov process will not satisfy the Chapman-Kalmogorov equations unless the holding-times are all exponentially distributed. Kurtz (1971) has shown under reasonable assumptions that a Semi-Markov process

is 'close' to a corresponding Markov process provided the expected waiting time in each state is small.

Example. Consider an illness and death process having three states 0, 1, and 2. We can interpret such a process by thinking of state 1 as health, 0 as illness and 2 as death. We can imagine that a person starting from state 1 will be at state 0, 1, or 2 after t units time. A transition from 1 to 0 corresponds to an 'illness', while a transition from 1 to 2 corresponds to 'death'. The transition probability matrix is:

$$P = \begin{array}{c} \\ \text{Illness (0)} \\ \text{Health (1)} \\ \text{Death (2)} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Illness} & \text{Health} & \text{Death} \\ (0) & (1) & (2) \end{array} \\ \begin{bmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{bmatrix} \end{array}$$

0 and 1 are both transient states and 2 is an absorbing state. This is an example of a simple Markov chain. If we include the holding-time behaviour between transitions in this example, viz., the length of time a person stay in any state depends on both where he is and where he goes, then this is an example of a Semi-Markov process.

2.4 Holding-Time and Waiting-Time Distribution

The concept of holding-time distribution has been introduced in the analysis of Semi-Markov process. The figure 2.4.1 represents a portion of possible realization for a discrete time Semi-Markov process.

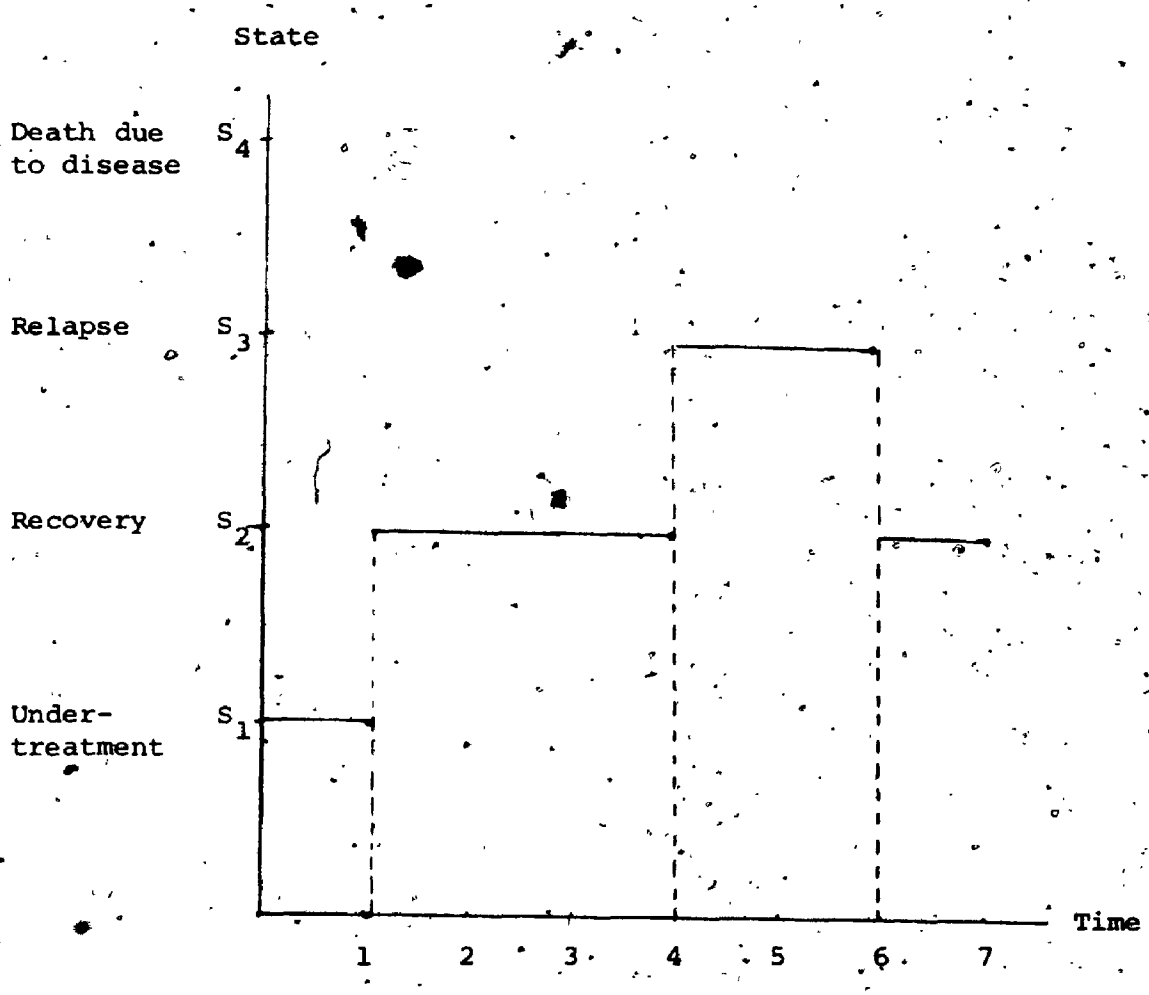


Figure 2.4.1. A possible realization of a Semi-Markov process.

The figure shows that a patient entered in the state of under-treatment (S_1) at time 0, stayed in that state for one time unit, then made a transition to state of recovery (S_2). He stayed in the state of recovery (S_2) for three units and then moved to S_3 'relapse'. The patient stayed in state of relapse (S_3) for two time units and then moved to state of recovery (S_2) when last observed.

Let τ_i be the waiting time in state i. Let $h_i(t_1)$ denote the discrete density function of the unconditional wait in state i. Then we have:

$$h_i(t_1) = \sum_{j=1}^{n} p_{ij} h_{ij}(t_1) = \text{Prob}\,(\tau_i = t_1) \qquad (2.4.1)$$

which is the probability that the patient will spend t_1 time units in state i if we do not know its successor state. The cumulative and complementary cumulative probability distribution for the waiting times are given by:

$$H_i(t) = \text{Prob}\,(\tau_i \leq t) = \sum_{t_1=0}^{t} h_i(t_1)$$

$$= \sum_{t_1=0}^{t} \sum_{j=1}^{n} p_{ij} h_{ij}(t_1) \qquad (2.4.2)$$

$$H_i^c(t) = \text{Prob}\,(\tau_i > t) = \sum_{t_1=t+1}^{\infty} h_i(t_1)$$

$$= \sum_{t_i=t+1}^{\infty} \sum_{j=1}^{n} p_{ij} h_{ij}(t_1) \qquad (2.4.3)$$

2.5 Interval Transition Probability Matrix

In this analysis of Semi-Markov process $(P, H(t))$, a basic statistic to be computed is the interval transition probability, $\phi_{ij}(t)$ which corresponds to the t-step transition probabilities for the Markov process. This quantity $\phi_{ij}(t)$ is the probabilty that the person is in state j at time t, given that he entered state i at time zero. The interval transition probability can be calculated from the parameters of the Semi-Markov processes $(P, H(t)')$ as follows:

A person starting in state i can be in state j at time t if either

1. the person never left state i throughout the interval $(0, t)$, i.e., $i = j$, or
2. the person made at least one move during the interval $(0, t)$, i.e., the person could have made the first move from state i to some state k at time t_1, $0 < t_1 < t$, and finally reached j at time t.

The probabilities of the above two mutually exclusive cases lead us to the following equation:

$$\phi_{ij}(t) = \delta_{ij} H_i^c(t) + \sum_{k=1}^{n} p_{ik} \sum_{t_1=0}^{t} h_{ik}(t_1)\phi_{kj}(t-t_1) \qquad (2.5.1)$$

where, $t \geq 0$, $1 \leq i, j \leq n$,

$$\delta_{ij} = \begin{cases} 1 & \text{when } i = j, \\ 0 & i \neq j. \end{cases}$$

The notation $H_i^c(t)$ is the probability that the person will leave the starting state i at time greater than t.

Matrix Formulation. A matrix formulation of the equation (2.5.1) is necessary in order to simplify the computation. Let us start with defining the following matrices:

$$\Phi(t) = [\phi_{ij}(t)] \qquad (2.5.2)$$

$$P = [p_{ij}] \qquad (2.5.3)$$

We also define the following diagonal matrix:

$$H^c(t) = [\delta_{ij} H_i^c(t)] \qquad (2.5.4)$$

Let us denote a special kind of matrix multiplication by $\square$. If A, B, and C are n by n square matrices, then

$$C = A \square B, \qquad (2.5.5)$$

The above equation implies:

$$\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdot & \cdot & \cdots & \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2n}b_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1}b_{n1} & a_{n2}b_{n2} & \cdots & a_{nn}b_{nn} \end{bmatrix}$$

This means that 'box' operation is an element by element matrix multiplication.

The equation (2.5.1) for estimating interval transition probabilities can be written in matrix form as follows:

$$\Phi(t) = H^{c}(t) + \sum_{t_1=0}^{t} (P \square H(t_1)) \Phi(t-t_1) \qquad (2.5.6)$$

where $t = 0, 1, 2, 3, \ldots$.

$\Phi(t)$ is thus the interval transition probability matrix for the Semi-Markov processes (P, H(t)) in the interval (0,t). At $t = 0$, the above equation reduces to

$$\Phi(0) = I, \qquad (2.5.7)$$

where I is the identity matrix.

Thus the equation (2.5.6) provides a convenient way to calculate the interval transition probabilities from the basic parameters P and H(t) of the Semi-Markov processes.

2.6 Duration Distribution.

The duration measure is defined as how long a person stays in a given state. This quantity has practical significance for managements as well as for physician. For, example, health planners might be interested to know how long a patient stays in the hospital. In order to compute this quantity, we define an indicator variable $u_{ij}(t)$ as follows:

$$u_{ij}(t) = \begin{cases} 1 & \text{when the process is in state } j \text{ at time } t, \\ & \text{given that the process entered state } i \text{ at} \\ & \text{time zero;} \\ 0 & \text{otherwise.} \end{cases}$$

Let v_{ij} = the total time the process will spend in state j if it is started in state i.

we have:

$$v_{ij} = \sum_{t=0}^{\infty} u_{ij}(t) \qquad (2.6.1)$$

and the expected time the person will spend in state j

$$\bar{v}_{ij} = E(v_{ij}) = E\left(\sum_{t=0}^{\infty} u_{ij}(t)\right) = \sum_{t=0}^{\infty} E(u_{ij}(t)) \qquad (2.6.2)$$

The probability that $u_{ij}(t) = 1$ is $\phi_{ij}(t)$. Therefore, we have

$$E(u_{ij}(t)) = \phi_{ij}(t) \qquad (2.6.3)$$

Hence, we can write the equation (2.6.2) using equation (2.6.3) as

$$\bar{v}_{ij} = E(v_{ij}) = \sum_{t=0}^{\infty} \phi_{ij}(t) \qquad (2.6.4)$$

The second moment is:

$$\bar{v}_{ij}^{(2)} = E(v_{ij}^2) = \bar{v}_{ij}(2 \bar{v}_{jj} - 1) \qquad (2.6.5)$$

The variance in the time a patient will spend in state j, given that he started in state i, is given by :

$$\begin{aligned} \text{Var}(v_{ij}) &= E(v_{ij}^2) - (E(v_{ij}))^2 \\ &= \bar{v}_{ij}^{(2)} - (\bar{v}_{ij})^2 \\ &= \bar{v}_{ij}(2 \bar{v}_{jj} - 1 - \bar{v}_{ij}) \qquad (2.6.6) \end{aligned}$$

When i = j, we have

$$\text{Var}(v_{ii}) = \bar{v}_{ii}(\bar{v}_{ii} - 1) \qquad (2.6.7)$$

It may be noted that if $\bar{v}_{ij}$ is fairly large, the standard deviation of length of stay will be almost as large as its expectation.

In order to estimate the mean and variance for the duration of stay in transient states, the matrix approach is used. Let us define a matrix N_μ with element $\bar{v}_{ij}$, then using the

technique of flow graph analysis given by Howard (1971), the total expected time spent in the transient process is just the sum of the expected time spent in each of the transient states. We, therefore, have:

$$N_{\mu} = \left[I - P_0\right]^{-1} M \qquad (2.6.8)$$

where, P_0 = Transition probability matrix corresponding to the portion of transient states,

M = Diagonal matrix of mean waiting times for the transient states,

$\left[I - P_0\right]^{-1}$ = Inverse of the matrix $\left[I - P_0\right]$.

CHAPTER 3

USE OF MARKOV MODEL IN HEALTH FIELD

3.1 Introduction

Various methods have been developed for studying mortality experience of a population. T-year survival rate is the most common measure used for such studies. This is defined as the proportion of a population who survive for at least T-years from some defined time point. Cutler and Ederer (1958) describe the life table approach for calculating the 'T-year survival rate' in clinical situations but such a method does not provide an adequate tool for the early prediction of later results.

Boag (1949) considered a lognormal model to describe the survival distribution for cancer patients which could be used for predictive purposes. Fix and Neyman (1951) proposed a more general stochastic model, using the Markov process in the study of the prognosis of chronic disease and illustrated its use with an example from cancer.

The following sections describe the various methods mentioned above and discuss their advantages and limitations. It is proposed that the Semi-Markov approach will provide a model that will be closer to reality than the ones previously described.

3.2 Life Table Method (or Actuarial Method)

The main objective of the life table is to calculate the probability of surviving various time period reckoned from some initial time. The basic principle underlying this approach is to calculate the probability of dying during each unit time interval. Then the survival to a given time point can be computed by multiplying the compliments of the individual probabilities of dying. The basic data needed for this method are: number of patients at the begining of the interval, patients died due to disease, died due to other causes, and withdrawn alive during the interval. Then the risk of death due to disease during the unit interval is given by :

$$q_i = \frac{d_i}{N_i - \frac{1}{2} w_i} \qquad (3.2.1)$$

where, q_i = Probability of death due to disease,

d_i = Number of deaths due to disease during the interval,

w_i = Number of deaths due to other causes or withdrawn alive during the interval, and

N_i = Number of patients at the beginning of the interval.

It is assumed that times to deaths due to other causes or withdrawal are uniformly distributed. Thus, a patient who has died from other causes during the interval will be at risk of death from disease under study for only half the length of the interval. This accounts for the adjustment in the denominator.

It is obvious from equation (3.2.1) that at least some patients would have to be observed and their status determined at the end of the period for which the survival probability is to be calculated. Thus, the survival in future could not be predicted from observing only a part of the experience.

3.3 Lognormal Model

Boag (1949) tried to overcome this difficulty to some extent by using a lognormal model to describe the survivorship for cancer patients and placed most emphasis on the early prediction of results. He divided the population of patients treated for cancer into cured and non-cured groups. Let c denote the proportion of cured patients (1-c) the proportion of non-cured patients. He assumed that the cured patients had survival similar to that of a general population, while the non-cured patients had survival times given by the lognormal distribution

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{1}{2}\left(\frac{\log t - \mu}{\sigma}\right)^2} \cdot \frac{1}{t} \qquad (3.3.1)$$

where, t is the time to survival; μ and σ are the mean and standard deviation of the natural logarithm of survival times.

The method of maximum likelihood is used to estimate the three parameters μ , σ and c. The explicit expression for the likelihood function is

$$L = \sum_{A} \log \left\{ \frac{(1-c)f(t)}{\sigma} \right\} + \sum_{B,C} \log \{ c + (1-c)q \} + \sum_{D} \log (1-c)q \qquad (3.3.2)$$

where, $q = \int_t^{\infty} f(t)dt$;

A = Group A : Those who have died with cancer present at time t, reckoned from the commencement of treatment;

B = Group B : Those who have died of other disease at time t;

C = Group C : Those who remain alive and symptom free after time t; and

D = Group D : Those who remain alive but with cancer progressing at time t.

Each summation in equation (3.3.2) being extended over all patients in the group whose notation is placed below the summation sign.

In accordance with the principle of maximum likelihood, the values of μ, σ, and c are to be estimated which makes L a maximum. This requires that

$$\frac{\partial L}{\partial \mu} = \frac{\partial L}{\partial \sigma} = \frac{\partial L}{\partial c} = 0 \qquad (3.3.3)$$

where, $\frac{\partial}{\partial \mu}$, $\frac{\partial}{\partial \sigma}$ and $\frac{\partial}{\partial c}$ are the partial derivatives with respect to μ , σ and c respectively.

These equations are solved by an iterative process. Such a process is very cumbersome and in practice, it is found that the sequence of iterations may not converge or may lead to a value of σ which is unrealistically large. An alternative approach suggested by Boag of fixing the value of σ from past experience of other series, is open to criticism on the ground that the final estimate of 'c' is dependent on the value of σ chosen. Also the past experience does not necessarily enable σ to be chosen with sufficient accuracy. It was shown by Berg (1965), Gehan (1969) and Campos (1971) that the fitting of the lognormal model to cancer data was often poor in many instances.

3.4 Markov Model For Recovery, Relapse And Death

A conceptually different approach is used by Fix and Neyman (1951) in formulating their Markov model for recovery, relapse and death. They assumed that a patient suffering from cancer may go through the following exclusive states during the course of follow-up study:

S_0 - State of being under-treatment for cancer;

S_1 - State of being dead immediately following treatment for cancer;

S_2 - State of recovery in which the patient is not under-treatment but under observation;

S_3 - State of being lost after recovery or death due to other causes not related with cancer.

A patient who is under-treatment (state S_0) can leave the state S_0 by dying (transfer to state S_1) or by recoverying (transfer to state S_2). Similarly, the patient can leave state of recovery (state S_2) in only one of two ways, by having a relapse (transfer to state S_0) or by being lost or by dying from other causes not related with cancer (state S_3). The allowable transitions in this model are indicated in Figure 3.4.1.

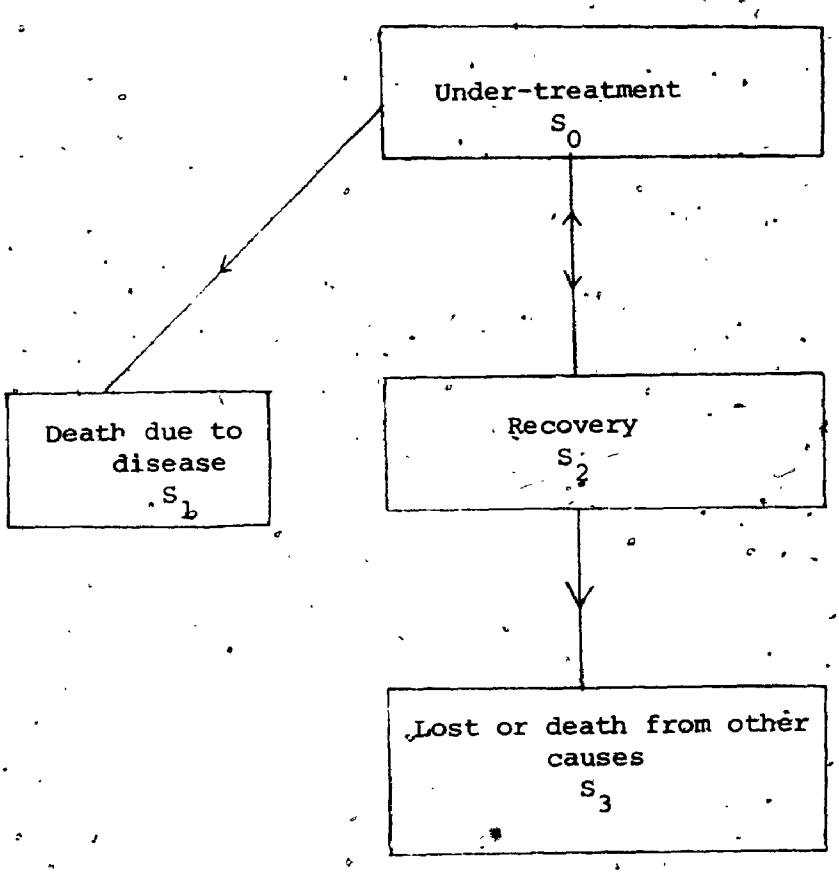


Figure 3.4.1. Transition Diagram for Markov Model

Let $P_i(t)$ be the probability that the patient is found in state i at time t, when i may be 0, 1, 2 and 3. Assuming that the time spent in state S_0 and S_2 are negative exponential random variables, then the forward Kolmogorov equations (equation, 2.3.7) for the model can be written as follows:

$$P_0'(t) = -(q_{01} + q_{02})P_0(t) + q_{20}P_2(t) \qquad (3.4.1)$$

$$P_1'(t) = q_{01}P_0(t) \qquad (3.4.2)$$

$$P_2'(t) = -(q_{20} + q_{23})P_2(t) + q_{02}P_0(t) \qquad (3.4.3)$$

$$P_3'(t) = q_{23}P_2(t) \qquad (3.4.4)$$

where,

$$P_0'(t) = \frac{d(P_0(t))}{dt}; \quad P_1'(t) = \frac{d(P_1(t))}{dt};$$

$$P_2'(t) = \frac{d(P_2(t))}{dt} \text{ and } P_3'(t) = \frac{d(P_3(t))}{dt}.$$

$$q_{ij} = \lim_{\Delta t \to 0} \frac{p_{ij}(t + \Delta t)}{\Delta t}$$

for every pair of states i and j, $i \neq j$, and these q_{ij} are called transition rates.

Using Laplace techniques, the following solutions of the above differential equations are obtained:

$$P_0(t) = \frac{1}{p_1 - p_2} \{ (p_1 + q_{20} + q_{23}) e^{p_1 t} - (p_2 + q_{20} + q_{23}) e^{p_1 t} \} \quad (3.4.5)$$

$$P_1(t) = q_{01} \{ \frac{q_{20} + q_{23}}{p_1 p_2} + \frac{(p_1 + q_{20} + q_{23})}{p_1 (p_1 - p_2)} \cdot e^{p_1 t} - \frac{(p_0 + q_{20} + q_{23})}{p_2 (p_1 - p_2)} \cdot e^{p_2 t} \} \quad (3.4.6)$$

$$P_2(t) = \frac{q_{02}}{p_1 - p_2} (e^{p_1 t} - e^{p_2 t}) \quad (3.4.7)$$

$$P_3(t) = q_{02} q_{23} \{ \frac{1}{p_1 p_2} + \frac{p_2 e^{p_1 t} - p_1 e^{p_2 t}}{p_1 p_2 (p_1 - p_2)} \} \quad (3.4.8)$$

where,

$$p_1 = \frac{- B - \sqrt{B^2 - 4 C}}{2},$$

$$p_2 = \frac{- B + \sqrt{B^2 - 4 C}}{2},$$

$$B = q_{01} + q_{02} + q_{20} + q_{23},$$

and $C = q_{01} q_{20} + q_{01} q_{23} + q_{02} q_{23}$.

The net risk of death from cancer can be obtained by substituting $q_{23} = 0$, to eliminate the risk of being lost to follow-up or dying from other causes. This approach provides estimates which appear to be more meaningful and reasonable than the estimates obtained by the frequently used simple actuarial adjustment for deaths due to other causes or lost to follow-up. To demonstrate the preceding point, transition rates in table 3.4.1 as given by Fix and Neyman (1951) were used to estimate probabilities of dying at various time intervals.

Table 3.4.1

Transition Rates For Two Different Treatments (Fix and Neyman, 1951).

Transition rates	Treatment 1	Treatment 2
q_{01}	1.0	0.2
q_{02}	2.7	0.2
q_{20}	0.2	0.4
q_{23}	0.2	0.1

Using the above rates in equations (3.4.6) and (3.4.8), the probabilities $P_1^{(1)}(t)$ and $P_3^{(1)}(t)$ of dying at time 't' due to cancer and other causes respectively for treatment 1 are given by

$$P_1^{(1)}(t) = 0.42553 - 0.17743\, e^{-0.24376t} - 0.24804\, e^{-3.85624t} \quad (3.4.9)$$

$$P_3^{(1)}(t) = 0.57447 - 0.61323\, e^{-0.24376t} + 0.03876\, e^{-3.85624t} \quad (3.4.10)$$

Similarly for treatment 2, the probabilities are

$$P_1^{(2)}(t) = 0.75 - 0.66667\, e^{-0.2t} - 0.08333\, e^{-0.8t} \quad (3.4.11)$$

$$P_3^{(2)}(t) = 0.25 - 0.33333\, e^{-0.2t} + 0.08333\, e^{-0.8t} \quad (3.4.12)$$

Putting $q_{23} = 0$, the net risk of death due to cancer denoted by P_1^{10} for treatment 1 is given by

$$P_1^{10}(t) = 1.0 - 0.75073\, e^{-0.05195t} - 0.24927\, e^{-3.84805t} \quad (3.4.13)$$

And, for treatment 2, the net risk of death due to cancer denoted by P_1^{20} is given by

$$P_1^{20}(t) = 1.0 - 0.85355\, e^{-0.11716t} - 0.14645\, e^{-0.68284t} \qquad (3.4.14)$$

The net risk of death due to cancer at time 't' as estimated by the actuarial method is given by

$$P_1^{A}(t) = \frac{P_1(t)}{1 - \frac{1}{2} P_3(t)} \qquad (3.4.15)$$

The net risks of death due to cancer calculated by the Markov model and actuarial method at different periods of time are presented in table 3.4.2. The mortality curves for treatment 1 and treatment 2, from the above tables are presented in figures 3.4.2 and 3.4.3 respectively. The crude death rates for both the treatments are lower than the net rates calculated by the two methods. Initially the net rates obtained by the actuarial method are higher than those given by Markov model. However, as the time interval increases, the net rates from Markov model start climbing higher than the actuarial method. In the limit as time approaches infinity, these rates tends to one while for actuarial method, the limit is less than one. Logically, the net risk should approach one as time approaches infinity.

Table 3.4.2

Death Rate of Cancer for Fix & Neyman Data Using

i) Markov Model & ii) Actuarial Method

	Treatment 1			Treatment 2		
t	$P_1^{(1)}(t)$	$P_1^{1O}(t)$	$P_1^{1A}(t)$	$P_1^{(2)}(t)$	$P_1^{2O}(t)$	$P_1^{2A}(t)$
0.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.5	0.23238	0.23215	0.23679	0.09091	0.09092	0.09111
1.0	0.28124	0.28199	0.29522	0.16673	0.16683	0.16795
1.5	0.30167	0.30480	0.32599	0.23102	0.23143	0.23432
2.0	0.31645	0.32327	0.35120	0.28629	0.28737	0.29264
2.5	0.32905	0.34071	0.37415	0.33437	0.33660	0.34455
3.0	0.34013	0.35763	0.39535	0.37656	0.38052	0.39116
3.5	0.34993	0.37411	0.41491	0.41387	0.42016	0.43327
4.0	0.35861	0.39015	0.43288	0.44705	0.45626	0.47148
4.5	0.36629	0.40579	0.44935	0.47668	0.48942	0.50623
5.0	0.37308	0.42104	0.46438	0.50322	0.52004	0.53789
5.5	0.37910	0.43587	0.47806	0.52706	0.54847	0.56675
6.0	0.38443	0.45034	0.49048	0.54852	0.57496	0.59308
6.5	0.38915	0.46443	0.50171	0.56785	0.59971	0.61710
7.0	0.39332	0.47816	0.51185	0.58529	0.62288	0.63900
7.5	0.39702	0.49154	0.52099	0.60104	0.64463	0.65897

Continue .....

8.0	0.40029	0.50458	0.52921	0.61526	0.66505	0.67717
8.5	0.40318	0.51728	0.53659	0.62812	0.68425	0.69375
9.0	0.40575	0.52966	0.54321	0.63974	0.70232	0.70884
9.5	0.40802	0.54172	0.54913	0.65025	0.71933	0.72257
10.0	0.41003	0.55343	0.55443	0.65975	0.73535	0.73506
10.5	0.41181	0.56492	0.55916	0.66834	0.75045	0.74641
11.0	0.41338	0.57608	0.56337	0.67612	0.76467	0.75674
11.5	0.41478	0.58694	0.56714	0.68315	0.77808	0.76612
12.0	0.41601	0.59754	0.57049	0.68952	0.79072	0.77463
12.5	0.41710	0.60786	0.57347	0.69527	0.80264	0.78237
13.0	0.41807	0.61791	0.57612	0.70048	0.81387	0.78938
13.5	0.41899	0.62771	0.57848	0.70519	0.82447	0.79575
14.0	0.41968	0.63725	0.58058	0.70946	0.83446	0.80153
14.5	0.42035	0.64655	0.58244	0.71332	0.84388	0.80676
15.0	0.42095	0.65562	0.58410	0.71681	0.85276	0.81151

Figure 3.4.2

Comparison of Crude Rate, Net Rate And Actuarial Rate for Treatment 1

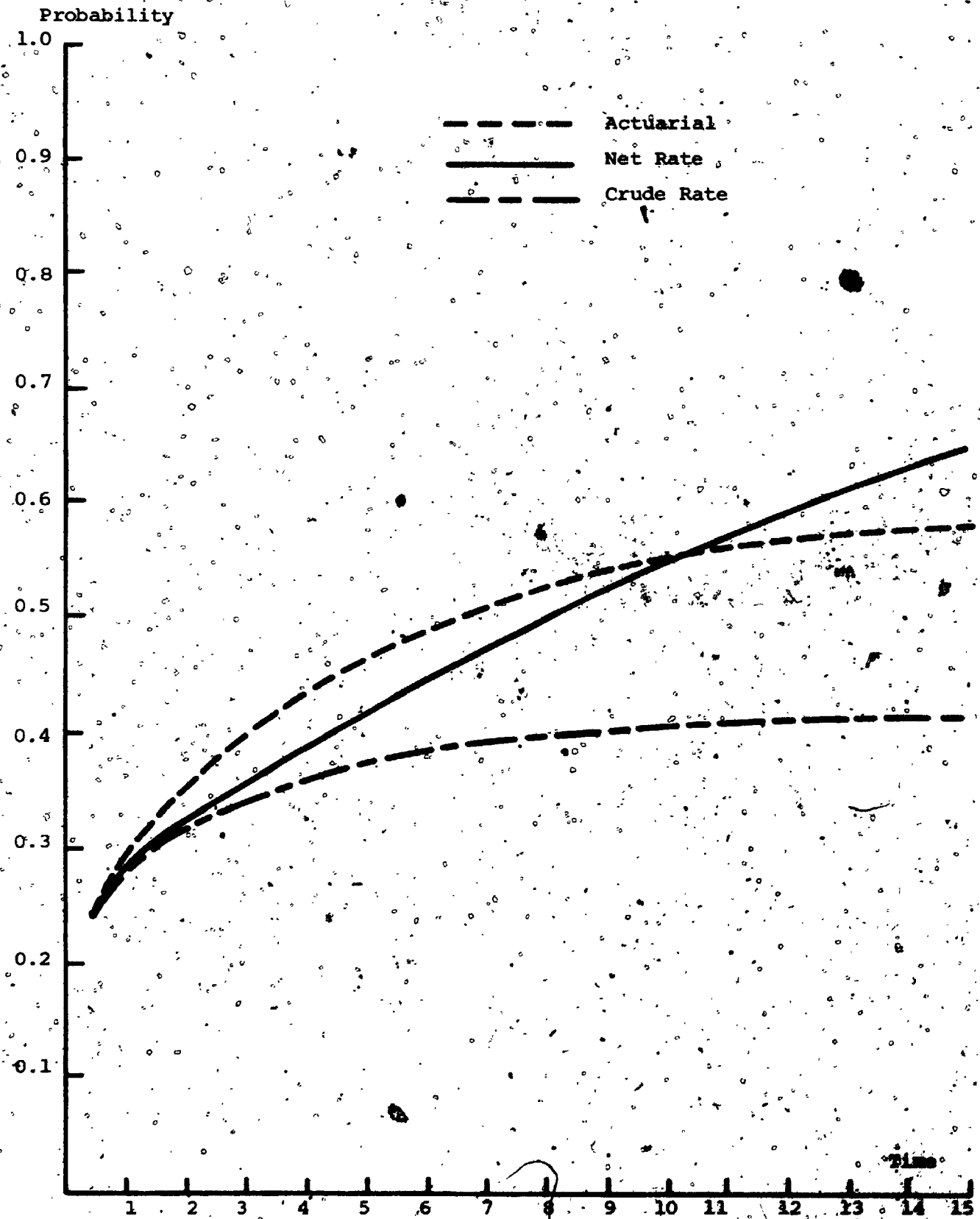
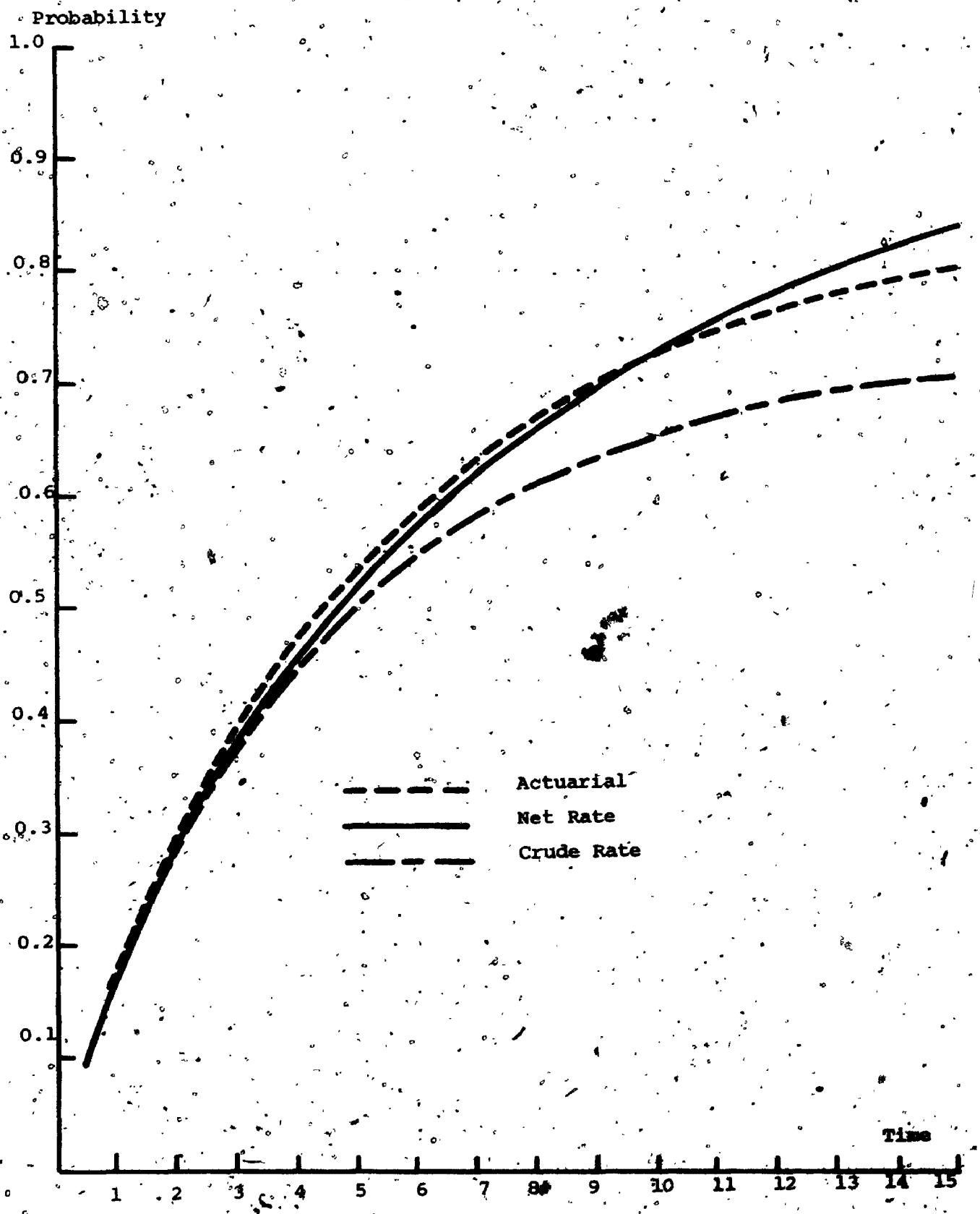


Figure 3.4.3

Comparison of crude Rate, Net Rate And Actuarial Rate for Treatment 2



3.5 Discussion

In the preceding sections, the actuarial method, lognormal model and Markov model have been examined. It is observed that actuarial method could not be used for predictive purposes. The lognormal model proposed by Boag (1949) could be used for the early prediction of later results. But such a model was very cumbersome to use and did not describe the data adequately in several instances observed by Berg (1965), Gehan (1969) and Campos (1971).

With minor variations, Fix's & Neyman's approach of Markov model has been used by several other investigators. Zahl (1955) further discussed this Markov model in more general terms and introduced a time-stationary continuous Markov process with a finite number of states as a model for the follow-up study in cancer. Alling (1958) used such a model to study mortality from tuberculosis. This model of tuberculosis provides predictions of the expected proportion of patients with active tuberculosis, the proportion with arrested tuberculosis at yearly intervals after a certain starting point in time. Sverdrup (1965) also applied the Markov model to the flow of people between different states of health such as 'able', 'disabled' and 'dead'. Marshall and Goldhamer (1965) used the Markov process to characterize the age distributions of the mentally-ill population.

Their approach for estimating the parameters, however, differed somewhat depending upon the availability of the data.

Since the Markov model of Fix and Neyman (1951) assumes that a patient's future change of state depends on his present state of health without knowing its past history, i.e., that the transition probability of going from state 'S_i' (i = 0 or 2) to state 'S_j' (j_c = 1 or 3) depends only on state 'S_i'. This simply implies that holding time density function (section 2.2) for each transient state (S_0 and S_2) is exponentially distributed. However, if patients change their state of health (S_0 and S_2) at irregular intervals of time, then it seems reasonable not to assign equal probabilities of transition to two patients, one of whom had to stay for many days in one state and the other just a few days. In such cases, along with probabilities of transition, distribution of lengths of time between transitions should be considered. A convenient model for such a situation can be given by a Semi-Markov process, in which the time spent in any state follows an arbitrary distribution that depends upon the present state as well as on the next state to be visited. It is this inclusion of the time dimension which makes the Semi-Markov model more realistic than the Markov model of Fix and Neyman (1951). A general discription and an application of Semi-Markov model for cancer of cervix patients are discussed in the next chapter.

CHAPTER 4

MODEL FOR CANCER OF THE CERVIX

4.1 Introduction

Three principal approaches to the study of cancer mortality were discussed in the previous chapter. These methods did not provide an adequate description of the prognosis of the disease. Life table methods are not suitable for prediction of future mortality beyond the observation point. Boag's method did have this capacity. However, it is cumbersome to use and heterogeniety among patients was poorly handled. The use of Markov model by Fix and Neyman (1951) does not take into account the time spent by the patients in various states which is an important contributing factor in the survival pattern. In addition, none of these methods provides any estimate of the expectation of duration of remaining relatively free of disease or being under-treatment. This information, along with survival rates is important for comparing efficacy of two treatments. In order to plan health facilities and allocate resources efficiently, health planners need to know expected numbers of patients in various states of recovery-relapse death system.

It is proposed that a model based on Semi-Markov process will provide answers to some of the problems associated with the current methods in use. Such a model will be developed for cancer of the cervix. It will be demonstrated that this model provides a better representation of the actual prognosis of the cancer of the cervix as compare to Markov model of Fix and Neyman (1951). Boag's method will not be considered any further due to the difficulties for estimating the parameters and it was found by Berg (1965), Gehan (1969) and Campos (1971) that the model often did not describe the data adequately.

4.2 Description of Semi-Markov Model

Let us consider a woman in recovery-relapse-death system of the cancer of the cervix. At a given time, a woman can be classified to belong to one of the following exclusive states:

S_1 - Under-treatment

S_2 - State of recovery in which the patient is not under-treatment but under observation;

S_3 - State of relapse;

S_4 - Death due to cancer of the cervix;

S_5 - State of being lost after recovery or death due to other causes not related to cancer of the cervix.

In the model all patients enter the study in state S_1 (i.e. under-treatment). After a course of treatment a patient may recover, in which case she is said to have transitted to state S_2 or may die due to cancer of the cervix, in that case she transits to S_4. Theoretically, a patient under-treatment can die from other causes or be lost to follow-up. However, in our model such a transition was not considered, since cases of this nature were not available for validating the model. Thus, the transition from S_1 to S_5 is not possible in this model. A patient who is under recovery (S_2) may have a relapse (S_3) or may die due to other causes or lost to follow-up (S_5). Further, a patient who is in the state of relapse may go back to recovery (S_2) or may die due to cancer of the cervix (S_4). There are no cases from relapse (S_3) to death due to other causes or lost to follow-up (S_5). Thus, transition from S_3 to S_5 was not included in the model. It may be noted that this model consists of three transient states, namely, under-treatment (S_1), recovery (S_2) and relapse (S_3) and two absorbing states S_4 and S_5 corresponding to death due to cancer of the cervix and death due to other causes or lost to follow-up respectively. The model is graphically represented in figure 4.2.1. The arrows indicate the direction of the possible transitions between the states.

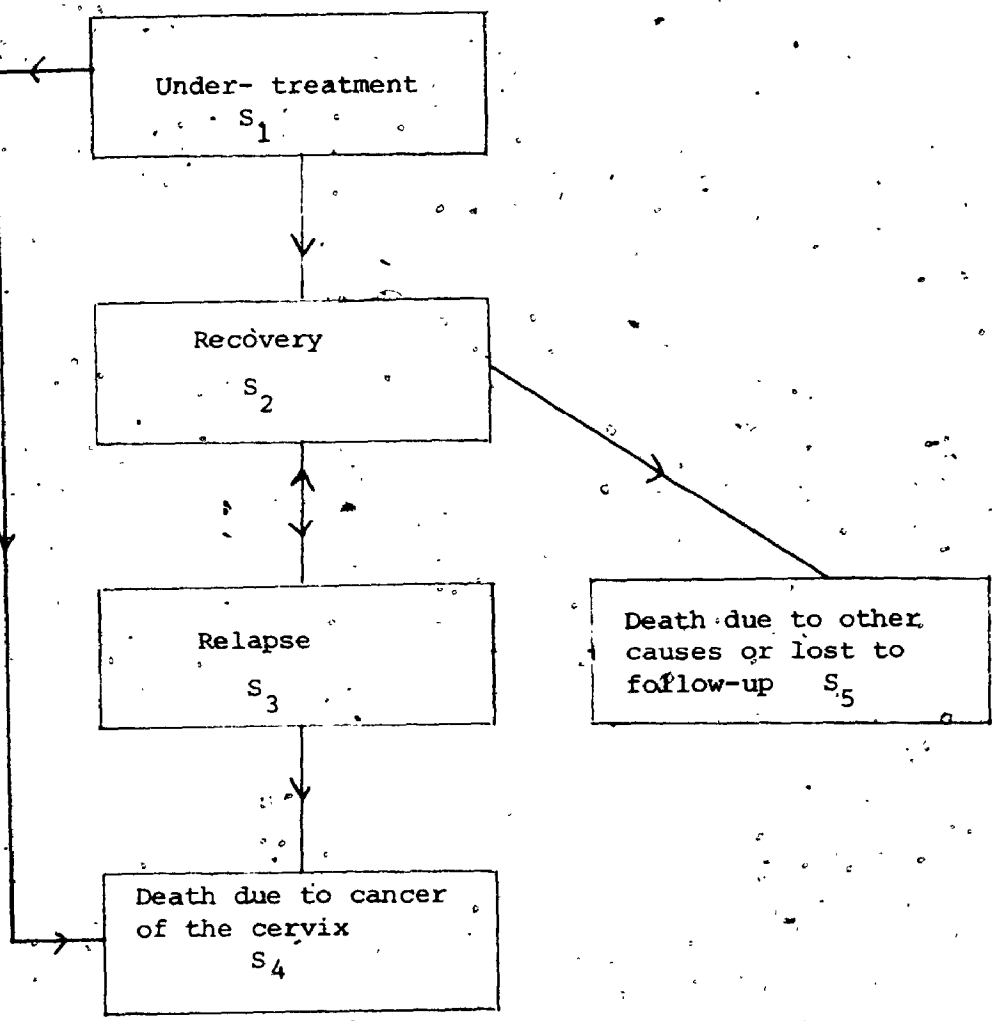


Figure 4.2.1 Graphical Representation of the Transitions in cancer of the Cervix.

So far the description of this model corresponds to ordinary Markov process. However, the mortality experience of a woman suffering from cancer of the cervix depends not only on the transitions from one state to the other but also on the duration of stay in a given state depending on the next state visited. The model should therefore, account for these durations of stay. The knowledge of the sequence in which these states are visited and the length of the time spent in each state at each visit will then characterize the survival distribution in the course of follow-up. It is logical to consider that the length of time spent under treatment state (S_1) is a random variable which depends upon a number of factors such as treatment conditions, biological variations and individual characteristics. In the same manner, the length of stay of a woman in each of the other two transient states of the model are also random variables. With the introduction of this element of duration of stay, the model for prognosis of cancer of the cervix as specified above corresponds to Semi-Markov process. Though a graphic representation of a general Semi-Markov model was given in figure 2.4.1, an example to explain the description of the follow-up of a cancer of the cervix patient by the Semi-Markov model is presented here in figure 4.2.2 to keep the discussion self-contained.

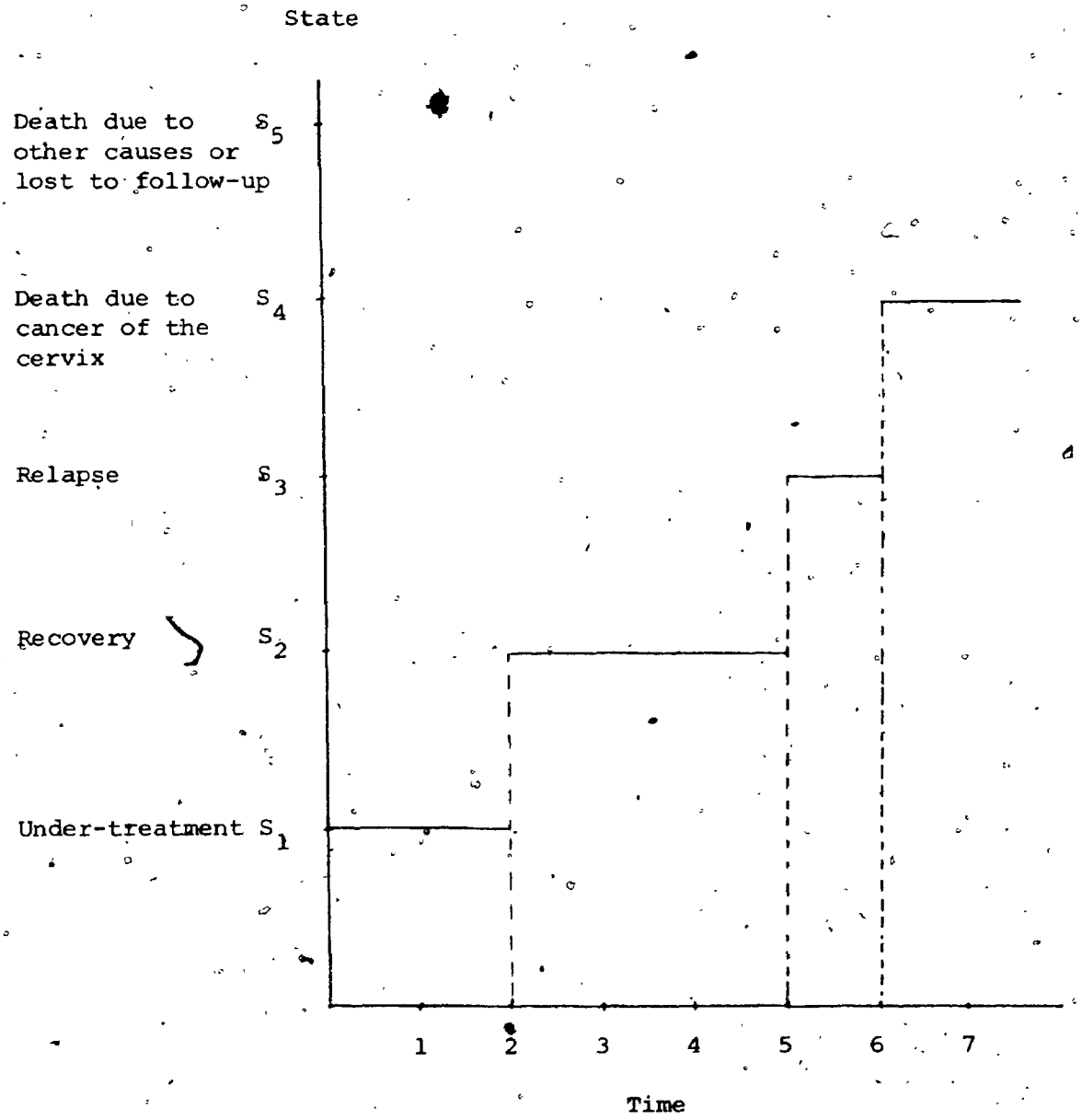


Figure 4.2.2 Semi- Markov Representation of the Follow-up of a Hypothetical Patient of Cancer of the Cervix.

Thus, a patient who is under-treatment (S_1) may stay for one time unit in that state before going to the state of recovery (S_2) where the patient might stay 3 time units before getting a relapse (S_3). The patient now remains in state of relapse for one time unit before finally dying due to disease (S_4) and remains there permamently. In terms of mathematical notations, the patient is absorbed to state S_4. In a follow-up study, a patient can visit a transient state several times depending upon the frequency of recovery and relapse.

In order to predict the future outcome of the patients in various states, this model requires to estimate the transition probability matrix, holding-time matrix and interval transition matrix. The background material for estimating these parameters is presented in chapter 2.

4.3 Description of the Data

In order to illustrate the application of the foregoing model for cancer of the cervix, the data for all previously untreated patients admitted to and treated at Victoria Hospital Cancer Clinic, London, Ontario between 1965 and 1966 have been collected. Each patient is followed upto 8 years. The following information from the individual medical files of the patients was extracted on the form attached in appendix 1:

1. Age of the patient at the time of admission;
2. Stage of the disease;
3. Type of treatment used (which in our cases is the usual radium therapy combined with cobalt 60);
4. Time spent in each state of recovery system (viz., under-treatment, apparent recovery, and relapse); and
5. Time of death due to cancer of the cervix and due to other causes not related to cancer of the cervix.

The appraisal of the stage of the cancer of the cervix was done according to the International classification which is based on objective and often on the surgeon's subjective estimate. The four classifications are:

Stage I: The carcinoma is strictly confined to the cervix.

Stage II: The carcinoma extends beyond the cervix but has not reached the pelvic wall.

Stage III: The carcinoma involves either the lower third of the vagina or has reached the pelvic wall, i.e., there is no cancer free space between the tumour and the pelvic wall.

Stage IV: The carcinoma involves the bladder or the rectum, or both.

It is important to note the distinction between stage and state. Stage refers to severity of the disease whereas state refers to the recovery-relapse-death system.

The term 'recovery' is difficult to define for cancer patients. Sorensen (1958) defined recovery as follows: Patients designated as recovered means that current diagnostic methods have not shown signs of cancer. We shall use this definition in this study. Relapse is defined as the reappearance of cancer as diagnosed by the physician.

The primary treatment of carcinoma of the cervix in our cases is the usual intracavitary radium application combined with cobalt 60 radiation and the length of treatment usually varies from 28 days to 35 days.

The choice of a proper starting point in follow-up studies is obviously important. Often comparisons between studies are complicated by different choices of starting points and ambiguity in their definition. Since initiation of the treatment for cancer is clear cut and unambigiously recorded, it was taken to be the start of follow-up from which all mortality and other variables were reckoned.

The total number of cases of canter of the cervix in this study is 101. There were only four cases of stage IV which have been excluded from the data. The distribution of cases according to the stage of the disease is given in table 4.3.1.

Table 4.3.1

Distribution of Patients Treated for Cancer of the Cervix According to Stage of Disease

STAGE	NUMBER OF CASES
I	36
II	46
III	19
TOTAL	101

In order to prevent the various age-groups from being too small and to maintain homogeneity with respect to the mortality within age-group, the cases have been divided in the present study into three age-groups; less than 40 years; 40 to 60 years and greater than 60 years. The distribution of the cases and deaths from cancer of the cervix and due to other causes according to various age-groups are presented in table 4.3.2.

Table 4.3.2

Distribution of the Number of Patients Treated and Number of Deaths due to Cancer of Cervix and Other Causes According to Age-groups

Age-Group Years	Number of Patients	Number of deaths due to cancer of the cervix	Number of deaths due to other causes
< 40	23	8	0
40 to 60	55	22	2
≥ 60	23	14	4
TOTAL	101	44	6

The deaths from other causes are not equally divided between the various ages. Instead, as might be expected, they increase with age. The effect of deaths due to other causes becomes increasingly important in the higher age-groups, particularly since these groups are often small.

The table 4.3.3 represents the observed number of deaths due to cancer of the cervix and due to other causes or lost to follow-up during each year of observation.

Table 4.3.3

Distribution of Deaths Due to Cancer of Cervix and Due to Other Causes or Lost to Follow-up for All Patients

Year after Treatment	Number of patients alive at the beginning of the year	Number of deaths during the year due to	
		cancer of the cervix	other causes
0 -	101	15	0
1 -	86	10	2
2 -	74	8	0
3 -	66	2	0
4 -	64	4	1
5 -	59	4	1
6 -	54	1	0
7 -	53	0	2

4.4 Estimation of Transition Probability Matrix

A patient who is suffering from cancer of the cervix, on any day can be considered to belong in one of the five states listed in section 4.2 of the model. Let p_{ij} be the probability that a patient who entered state S_i (i = 1,2,3,4,and 5) on the last transition will be in state S_j (j = 1,2,3,4 and 5) on the next transition. The transition probability p_{ij} must satisfy the following relation:

$$p_{ij} \geq 0, \quad i, j = 1,2,3,4 \text{ and } 5 \qquad (4.4.1)$$

$$\sum_{j=1}^{5} p_{ij} = 1; \quad i = 1,2,3,4 \text{ and } 5 \qquad (4.4.2)$$

The maximum-liklihood technique for estimating transition probability p_{ij} given by Anderson and Goodman (1957) was used. Maximum-likelihood estimates of the p_{ij} are calculated by dividing the number of times patients move from a state S_i (i = 1,2,3,4 and 5) to state S_j (j = 1,2,3,4 and 5). by the total number of transitions from S_i to all the states of the system. The estimates of the p_{ij} are then given by

$$\hat{p}_{ij} = \frac{l_{ij}}{\sum_{j=1}^{5} l_{ij}} \qquad (4.4.3)$$

where l_{ij} is the number of transitions from state S_i to S_j. The number of transitions l_{ij} are obtained from counting the transition from S_i to S_j from observed data.

Initial five years follow-up data were used to estimate these transitions and transition probability matrices for all the patients combined. These matrices were also estimated for the three stages of the disease and the three age-groups separately. These matrices are presented in the following tables 4.4.1 to 4.4.7.

Table 4.4.1

Transition Probability Matrix for Stage I Patients

Estimated from 5 years Follow-up

	S_1	S_2	S_3	S_4
S_1	0	0.9167	0	0.0833
S_2	0	0.6970	0.3030	0
S_3	0	0.2000	0	0.8000
S_4	0	0	0	1.0000

Note: There are no deaths due to other causes or lost to follow-up for stage I cases; thus there is no state S_5 for stage I patients.

Table 4.4.2

Transition Probability Matrix for Stage II patients

Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.9565	0	0.0435	0
S_2	0	0.4771	0.4091	0	0.1138
S_3	0	0.1111	0	0.8889	0
S_4	0	0	0	1.0000	0
S_5	0	0	0	0	1.0000

Table 4.4.3

Transition Probability Matrix for Stage III Patients

Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.6842	0	0.3158	0
S_2	0	0.1538	0.7692	0	0.0770
S_3	0	0	0	1.0000	0
S_4	0	0	0	1.0000	0
S_5	0	0	0	0	1.0000

Table 4.4.4

Transition Probability Matrix for All Patients Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.8911	0	0.1089	0
S_2	0	0.5111	0.4222	0	0.0667
S_3	0	0.1503	0	0.8947	0
S_4	0	0	0	1.0000	0
S_5	0	0	0	0	1.0000

Table 4.4.5

Transition Probability Matrix for Age-group Less Than 40 Years Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4
S_1	0	0.9131	0	0.0869
S_2	0	0.6667	0.3333	0
S_3	0	0.1429	0	0.8571
S_4	0	0	0	1.0000

Note: Since there are no deaths due to other causes or lost to follow-up for age-group less than 40 years; thus there is no state S_5 in this case.

Table 4.4.6

Transition Probability Matrix for Age-group Between 40 to 60 Years Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.9273	0	0.0727	0
S_2	0	0.5490	0.4118	0	0.0392
S_3	0	0.1429	0	0.8571	0
S_4	0	0	0	1.0000	0
S_5	0	0	0	0	1.0000

Table 4.4.7

Transition Probability Matrix for Age-group Greater Than 60 Years Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.7826	0	0.2174	0
S_2	0	0.2778	0.5000	0	0.2222
S_3	0	0	0	1.0000	0
S_4	0	0	0	1.0000	0
S_5	0	0	0	0	1.0000

4.5 Estimation of Holding-Time Matrix

The length of stay in state S_i (i =1, 2, 3) given that the next transition to state S_j (j =1, 2, 3, 4, and 5) is a random variable τ_{ij}. Let its probability distribution function be

$$h_{ij}(t_l) = \text{Prob}(\tau_{ij} = t_l) \qquad (4.5.1)$$

Suppose that the patient moves from state S_i (i = 1, 2, 3) to state S_j (j = 1, 2, 3, 4 & 5) each month with probability θ_{ij} and stays in S_i with probability $(1 - \theta_{ij})$. Then, the total length of stay in S_i given that the next transition is to state S_j can be considered to have the geometric distribution

$$h_{ij}(t_l) = (1 - \theta_{ij})^{t_l - 1} \theta_{ij} , \qquad (4.5.2)$$

where, $t_l = 1, 2, 3, \ldots$

This is also called the 'Holding-time' distribution.

It is important to note at this point that if θ_{ij} does not depend on j, then the Semi-Markov process is an ordinary Markov process. That is, the holding-time density function for all transitions out of a state must be identical and therefore must equal the waiting time density function for that state. Thus the discrete-time Markov process must have geometrically distributed waiting times.

The maximum likelihood estimate of θ_{ij} is given by

$$\hat{\theta}_{ij} = \frac{1}{E(\tau_{ij})} \qquad (4.5.3)$$

The parameters $\hat{\theta}_{ij}$ are calculated from the observed data and is estimated by the reciprocal of the average holding-time, τ_{ij}. The holding-time matrix corresponding to transition probability matrix in tables 4.4.1 to 4.4.7 are given in tables 4.5.1 to 4.5.7 respectively.

Table 4.5.1

Holding-Time Matrix for Stage I Patients

Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4
S_1	0	$(0.74)(0.26)^{t-1}$	0	$(0.18)(0.82)^{t-1}$
S_2	0	0	$(0.03)(0.97)^{t-1}$	0
S_3	0	$(0.53)(0.47)^{t-1}$	0	$(0.11)(0.89)^{t-1}$
S_4	0	0	0	0

Table 4.5.2

Holding-Time Matrix for Stage II Patients

Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4	S_5
S_1	0	$(0.84)(0.16)^{t-1}$	0	$(0.08)(0.92)^{t-1}$	0
S_2	0	0	$(0.04)(0.96)^{t-1}$	0	$(0.29)(0.71)^{t-1}$
S_3	0	$(0.18)(0.82)^{t-1}$	0	$(0.14)(0.86)^{t-1}$	0
S_4	0	0	0	0	0
S_5	0	0	0	0	0

Table 4.5.3

Holding-Time Matrix for Stage III Patients

Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4	S_5
S_1	0	$(0.79)(0.21)^{t-1}$	0	$(0.12)(0.88)^{t-1}$	0
S_2	0	0	$(0.05)(0.95)^{t-1}$	0	$(0.02)(0.98)^{t-1}$
S_3	0	0	0	$(0.09)(0.91)^{t-1}$	0
S_4	0	0	0	0	0
S_5	0	0	0	0	0

Table 4.5.4

Holding-Time Matrix for All Stages Combined

Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4	S_5
S_1	0	$(0.79)(0.21)^{t-1}$	0	$(0.12)(0.88)^{t-1}$	0
S_2	0	0	$(0.05)(0.95)^{t-1}$	0	$(0.02)(0.98)^{t-1}$
S_3	0	$(0.26)(0.74)^{t-1}$	0	$(0.11)(0.89)^{t-1}$	0
S_4	0	0	0	0	0
S_5	0	0	0	0	0

Table 4.5.5

Holding-Time Matrix for Age-group Less Than 40 Years

Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4
S_1	0	$(0.69)(0.31)^{t-1}$	0	$(0.13)(0.87)^{t-1}$
S_2	0	0	$(0.04)(0.96)^{t-1}$	0
S_3	0	$(0.10)(0.90)^{t-1}$	0	$(0.14)(0.86)^{t-1}$
S_4	0	0	0	0

Table 4.5.6

Holding-Time Matrix for Age-group Between 40 to 60 Years

Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4	S_5
S_1	0	$(0.86)(0.14)^{t-1}$	0	$(0.08)(0.92)^{t-1}$	0
S_2	0	0	$(0.04)(0.96)^{t-1}$	0	$(0.02)(0.98)^{t-1}$
S_3	$(0.59)(0.41)^{t-1}$	0	0	$(0.23)(0.77)^{t-1}$	0
S_4	0	0	0	0	0
S_5	0	0	0	0	0

Table 4.5.7

Holding-Time Matrix for Age-group Greater Than 60 Years

Estimated From 5 Years Follow-up

	S_1	S_2	S_3	S_4	S_5
S_1	0	$(0.78)(0.22)^{t-1}$	0	$(0.17)(0.83)^{t-1}$	0
S_2	0	0	$(0.04)(0.96)^{t-1}$	0	$(0.03)(0.97)^{t-1}$
S_3	0	0	0	$(0.06)(0.94)^{t-1}$	0
S_4	0	0	0	0	0
S_5	0	0	0	0	0

4.6 Estimation of Proportion of Patients in Various States

The transition probability matrix and holding-time matrix corresponding to the different stages of the disease and various age-groups for cancer of cervix patients have been completely specified in section 4.4 and 4.5. Also, we have seen in chapter 2, the Semi-Markov model can be characterized by transition probability matrix P and holding-time matrix H(t); therefore, we have a complete description of the Semi-Markov model for cancer of the cervix patients. The essential statistic in the Semi-Markov model is the interval transition probability, $\phi_{ij}(t)$ which can be computed by solving equation (2.5.1). If we assume that the holding-time density function for all transitions out of a state is an identical geometric distribution, then $\phi_{ij}(t)$ corresponds to the t-step transition probabilities in Markov model.

The estimation of $\phi_{ij}(t)$ is required to solve the equation (2.5.1). Using the estimated parameters P and H(t) from sections 4.4 and 4.5, the interval transition probabilities corresponding to various stages of the disease and different age-groups are presented in appendix 2 and 3 respectively at various time points.

We define 'State probability' as the probability that a patient is in state S_i (i = 1, 2, 3, 4 and 5) at time t and will be denoted by $\pi_i(t)$; i = 1, 2, 3, 4, and 5. The state probabilities $\pi_i(t)$ in our model must satisfy the relation

$$\sum_{i=1}^{5} \pi_i(t) = 1, \qquad (4.6.1)$$

where t = 0,1,2,3,.....

Let $\Pi(t)$ denote the row vector whose elements are the state probabilities. Then

$$\Pi(t) = [\pi_1(t), \pi_2(t), \pi_3(t), \pi_4(t), \pi_5(t)] \qquad (4.6.2)$$

for t = 0, 1, 2. .....

The elements of this row vector correspond to the proportion of the patients in various states of the system at time 't'. This vector can be estimated by using the following relation:

$$\Pi(t) = \Pi(0)\Phi(t) \qquad (4.6.3)$$

for t = 0, 1, 2, ...

Where $\Pi(0)$ is the vector whose elements are the proportion of patients in different states at the start of the follow-up. In our case, since all the patients enter the study in the state S_1 'under-treatment', $\Pi(0)$ is given by

$$\Pi(0) = [1, 0, 0, 0, 0] \qquad (4.6.4)$$

From equation (4.6.3), the probability that a patient will be in state S_j (j = 1, 2, 3, 4, and 5) at time t can be written as

$$\pi_j(t) = \sum_{i=1}^{5} \pi_i(0)\phi_{ij}(t) \qquad (4.6.5)$$

where, j = 1, 2, 3, 4, and 5;

t = 0, 1, 2, 3, ... ; and

$\phi_{ij}(t)$ is the (i,j)th element of the interval transition matrix.

Inital vector $\Pi(0)$ is known. Thus, using equation (4.6.5), the proportion of patients in various states of the system at different time points can be easily estimated.

4.7 Evaluation of the Model

The main objective of developing the model in previous sections is to provide a tool for estimating the expected proportion of patients in various states of the system such as recovery, relapse, or death etc. Before it can be recommended for practical use, it needs to be evaluated. Generally, the model if based on reasonable assumptions will provide an adequate fit for the observed data which were used to estimate parameters. It is, therefore, important to evaluate its performance not only on these data but also in predicting beyond this set. The parameters of the Semi-Markov model were estimated from 5 years of observed follow-up data. Using this model, the prediction for survival

rates are made upto 8 years since observed data were available upto that time.

The survival rates at different time points are equal to

$$1 - \frac{\pi_4(t)}{1 - \frac{1}{2}\pi_5(t)}$$

Where $\pi_4(t)$ and $\pi_5(t)$ refer to the probability of dying at time t due to the cancer of cervix and other causes respectively.

The survival rates for Markov model were obtained by substituting appropriate holding-time matrix as described in section 4.5. Appendix 5 provides the estimated t-step transition probability matrix used in computation of survival rates for Markov model.

The validity of the model is first tested on over-all data for cancer of the cervix patients. It is further evaluated for prediction in various stages of the disease and in the three age-groups.

Over-all Data

A comparison of observed survival rates and predicted survival rates based on the Semi-Markov model and the Markov model are presented in table 4.7.1.

Table 4.7.1

Observed and Predicted Survival Rates for Cancer of the Cervix for All Patients

Years after treatment	Survival percentage by different methods		
	Semi-Markov	Markov	Observed
1	87.0	85.0	85.1
2	76.1	76.1	75.0
3	68.7	67.2	66.9
4	63.4	59.3	64.8
5	59.9	52.4	60.7
6	57.4	46.4	56.6
7	55.7	41.2	55.5
8	54.5	36.7	55.5

Since the five years observed data is used in the model, the resulting predictions for 5 and 8 years generated by the Semi-Markov model and Markov model are compared with the observed rates by means of chi-square statistic. The chi-square statistic is summarized in table 4.7.2.

Table 4.7.2

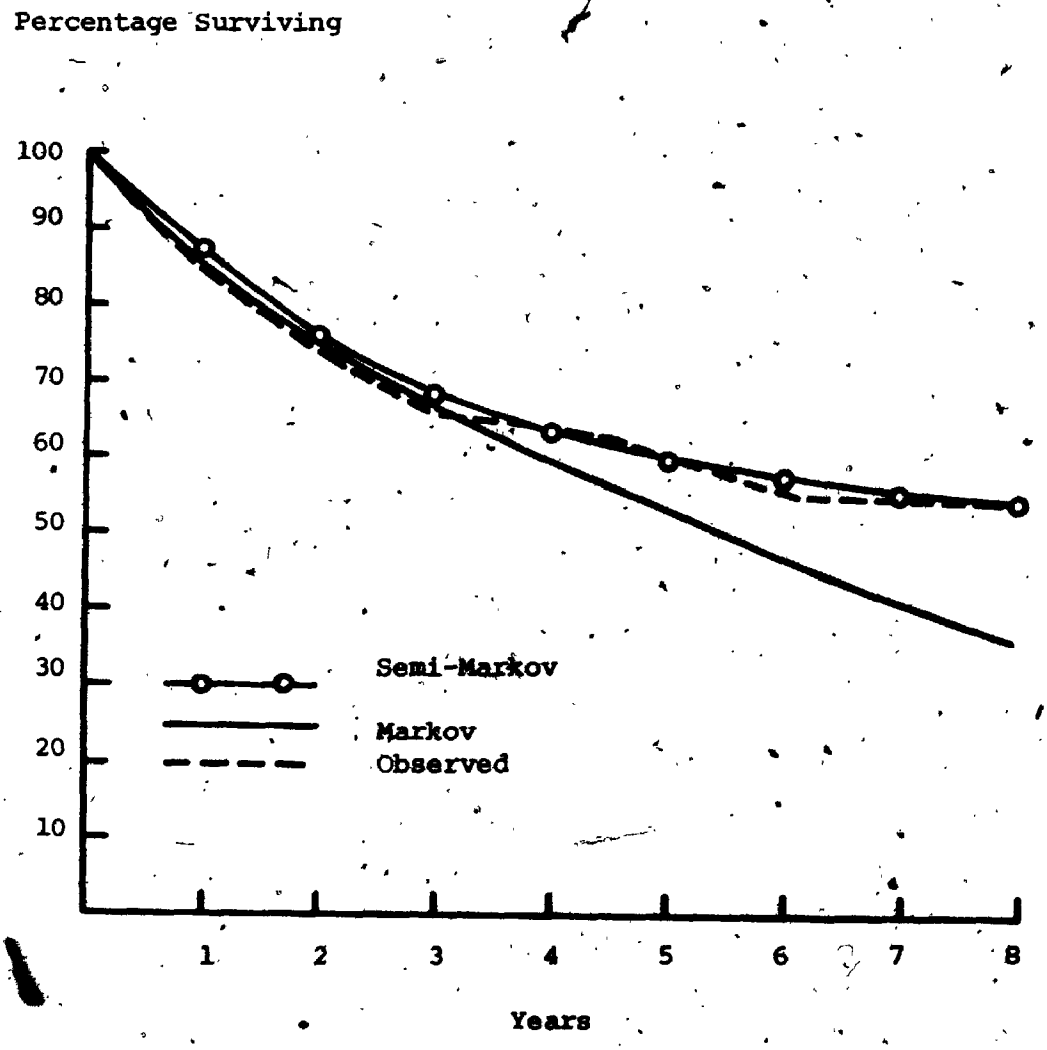
Summary of Chi-square Test Comparing Predictions Between Semi-Markov Model and Markov Model with the Observed Survival Rates for All Patients

Years of Follow-up	Chi-square with one degree of freedom between	
	Semi-Markov & observed	Markov & observed
5 years	0.01717	1.42151
8 years	0.01999	7.19365

The values of chi-square in table 4.7.2 at 8 years suggest that the predictions based on Semi-Markov model are not statistically significant; while for the Markov model, the predictions are highly significant. Further, Markov model does not show statistical significance at 5 years. It should not be surprising because the model is based on 5 years observed data. The figure 4.7.1 shows that Semi-Markov model predicts the survival rates at each year very closely to the observed rates. On the other

Figure 4.7.1

Comparision of Observed and Predicted Survival Rates for Cancer of the Cervix for all Patients



hand, Markov model underestimates the survival rate at each year and underestimation increases with the years of follow-up. The difference increases from 0.1 percent at 1 year to 18.8 percent at year eight.

Stages of the Disease

The adequacy of the model is further tested by grouping the data into different stages of the disease. The tables 4.7.3, 4.7.4 and 4.7.5 present the observed survival rates and the predicted survival rates based on the Semi-Markov and Markov model for stage I, stage II and stage III patients respectively.

Table 4.7.3

Observed and Predicted Survival Rates

For Cancer of the Cervix for Stage I

Years after treatment	Survival percentage by different methods		
	Semi-Markov	Markov	Observed
1	90.1	89.1	88.9
2	84 8	83.3	83.3
3	81 6	77.2	83.3
4	78.1	71.4	80.5
5	74.9	66.1	75.0
6	73.1	61.1	72.2
7	71.7	56.6	69.4
8	70.7	52.3	69.4

Table 4.7.4

Observed and Predicted Survival Rates for Cancer of the Cervix for Stage II

Years after treatment	Survival percentage by different methods		
	Semi-Markov	Markov	Observed
1	90.8	92.7	89.1
2	79.5	85.1	77.9
3	71.5	76.2	68.7
4	66.3	67.5	68.7
5	62.9	59.5	66.4
6	60.7	52.6	61.7
7	59.3	46.5	59.4
8	58.4	41.4	59.4

Table 4.7.5

Observed and Predicted Survival Rates for

Cancer of the Cervix for Stage III

Years after treatment	Survival percentage by different methods		
	Semi-Markov.	Markov	Observed
1	67.5	62.7	68.4
2	47.1	45.9	52.5
3	33.9	31.4	31.5
4	25.7	21.0	26.3
5	20.8	14.3	20.4
6	17.9	10.0	13.6
7	15.1	7.4	13.6
8	14.5	5.7	13.6

The summary of chi-square statistic according to the stage of disease comparing observed survival rates with Semi-Markov and Markov model are given in table 4.7.6.

Table 4.7.6

Summary of Chi-square Test Comparing Predictions Between Semi-Markov model and Markov model With the Observed Survival Rates According to Stage of Disease

Stage of disease	Years of follow-up	Chi-square with one degree of freedom between	
		Semi-Markov & observed	Markov & observed
I	5	0.00064	0.58381
	8	0.01410	2.10389
II	5	0.00179	2.78261
	8	0.12163	0.44779
III	5	0.00071	0.29619
	8	0.00863	0.78553

Prediction of survival rates based on Semi-Markov model appear to be in close agreement with observed rates in each stage of the disease as can be seen from figures 4.7.2, 4.7.3, and 4.7.4. The chi-square test (table 4.7.6) support this observation. The survival curve obtained by Markov model does not approximate the observed experience that well. The tables 4.7.3, 4.7.4 and 4.7.5 show that the survival rates obtained by using Markov model are consistently lower than the Semi-Markov model and observed survival rates except for stage II cases where they are slightly higher during first three years. The discrepancy between observed and Markov model becomes very noticable when predictions are made beyond the data used for estimation. The differences between observed rates and predicted survival rates by Markov model at 8-years are 17.1 percent, 18.0 percent, and 7.9 percent for stage I, stage II and stage III respectively compared with the maximum difference of only 1.3 percent in any stage for Semi-Markov model. Though, the chi-square test did not show statistical significance for stage I and III survival rates, the differences are large enough to suspect the adequacy of Markov model in predicting the survival. It may be due to smaller number of cases in stage I and stage III.

Figure 4.7.2

Comparison of Observed And Predicted Survival Rates for Cancer of the Cervix for Stage I

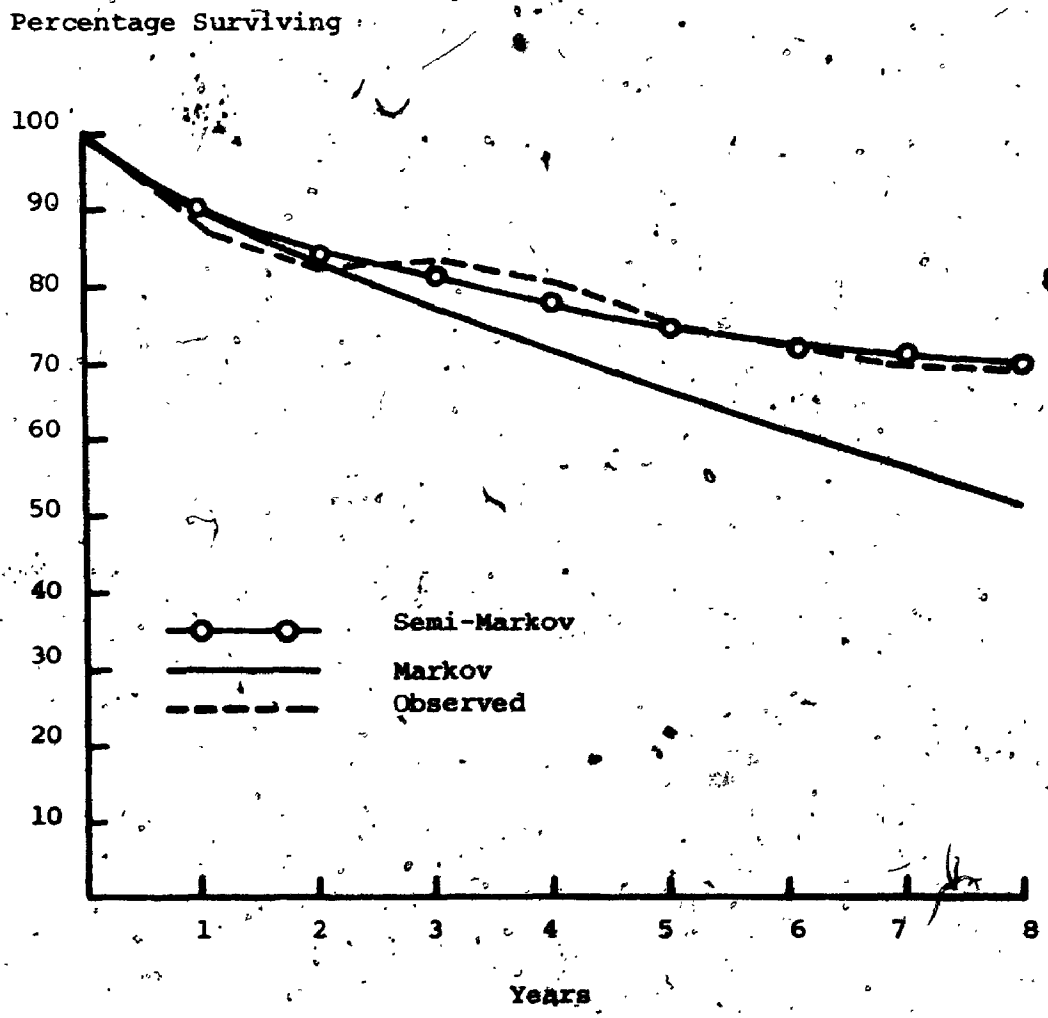


Figure 4.7.3

Comparison of Observed and Predicted Survival Rates for Cancer of the Ceryix for Stage II

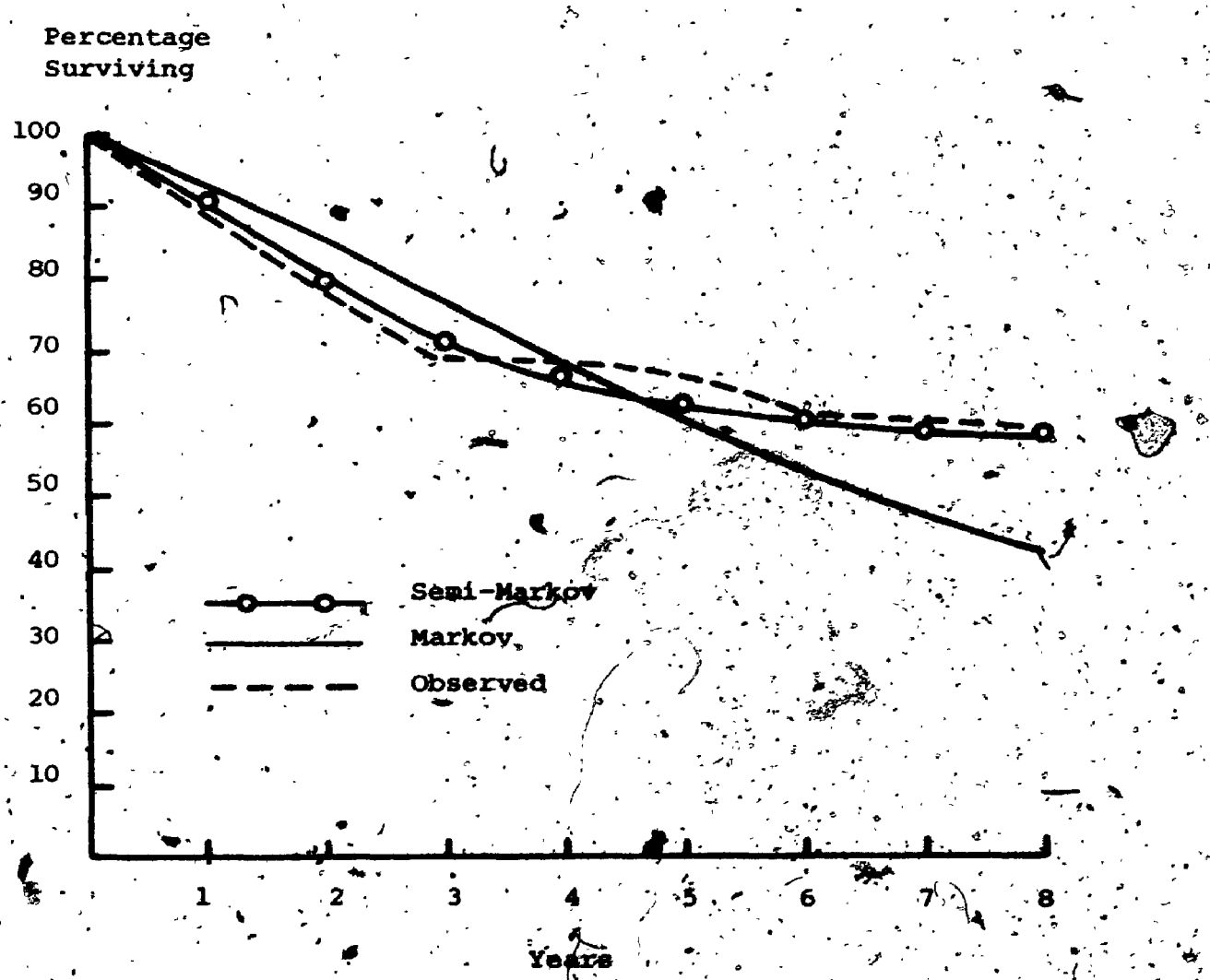
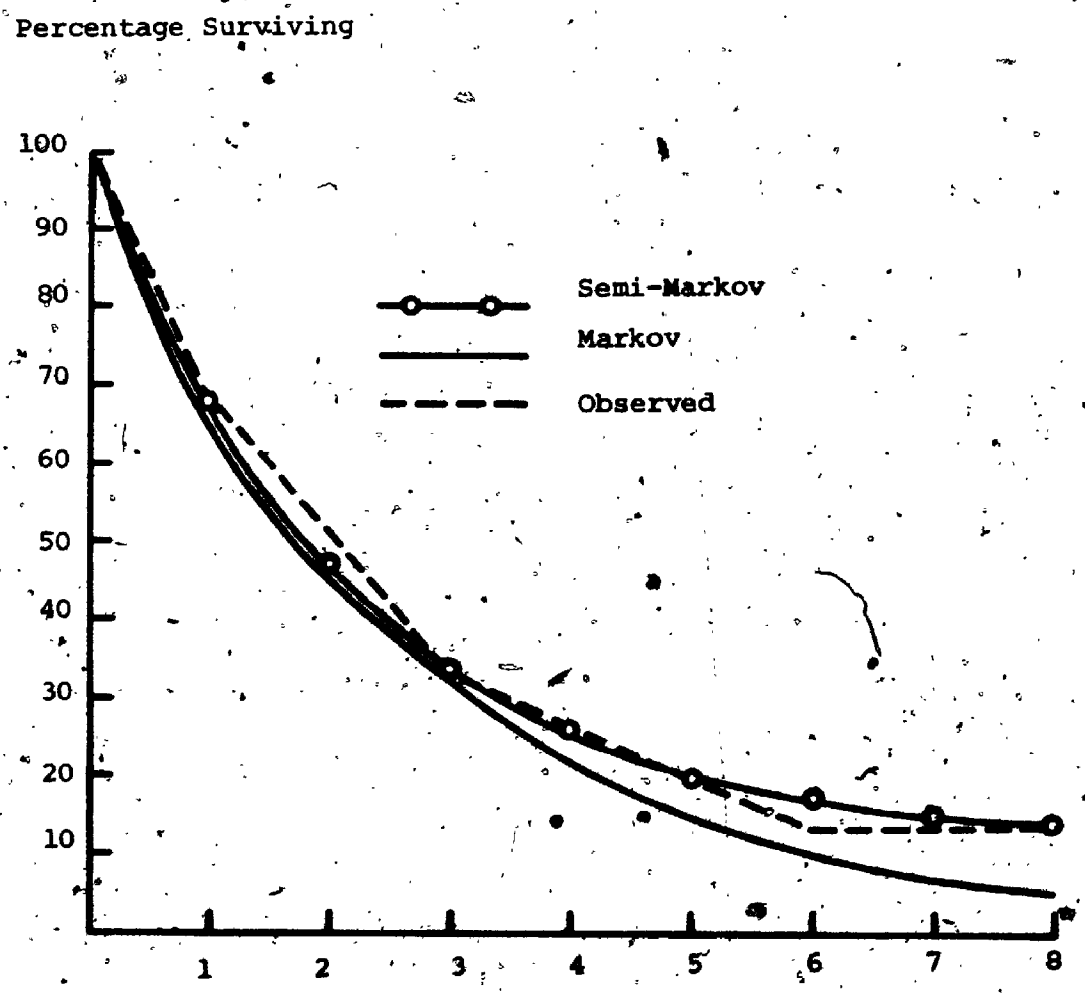


Figure 4.7.4

Comparison of Observed and Predicted Survival Rates for Cancer of the Cervix for Stage III



Age-groups

Since mortality due to cancer depends upon the age, the model is also tested for each of the several relatively homogeneous age-groups. Three age-groups less than 40 years, between 40 to 60 years and greater than 60 years were chosen to test the model. The survival rates for these age-groups, using Semi-Markov model, Markov model and observed rates are given in tables 4.7.7, 4.7.8 and 4.8.9.

Table 4.7.7

Observed and Predicted Survival Rates for Cancer of the Cervix Patients for Age-group Less Than 40 Years

Years after treatment	Survival percentage by various methods		
	Semi-Markov	Markov	Observed
1	88.1	88.6	87.0
2	80.5	82.3	73.9
3	74.8	75.5	73.9
4	71.1	68.9	73.9
5	68.9	63.0	73.9
6	67.0	57.5	69.6
7	66.0	52.5	65.3
8	65.2	47.9	65.3

Table 4.7.8

Observed and Predicted Survival Rates for Cancer of the Cervix Patients for Age-group Between 40-60 Years

Years after treatment	Survival percentage by various methods		
	Semi-Markov	Markov	Observed
1	87.7	90.0	87.3
2	76.4	82.6	78.1
3	69.4	73.9	66.9
4	64.9	65.6	63.2
5	62.2	57.9	63.2
6	60.3	51.1	61.4
7	59.2	45.1	59.5
8	58.5	39.9	59.5

Table 4.7.9

Observed and Predicted Survival Rates for Cancer of the Cervix Patients for Age-group greater than 60 Years

Years after treatment	Survival percentage by various methods		
	Semi-Markov	Markov	Observed
1	76.5	75.2	78.3
2	65.6	67.1	67.1
3	56.5	57.6	53.2
4	48.9	48.5	53.2
5	43.2	40.8	43.0
6	38.9	34.4	36.7
7	35.8	29.3	36.7
8	33.4	25.4	29.6

Chi-square values with one degree of freedom for comparing the observed and predicted survivorship for different age-groups are presented in table 4.7.10.

Table 4.7.10

Summary of Chi-square Test Comparing Predictions Between Semi-Markov Model and Markov Model with the Observed Survival Rates for Different Age-groups

Age group years	Years of follow-up	Chi-square with one degree of freedom between	
		Semi-Markov & observed	Markov & observed
< 40	5	0.10723	0.62945
	8	0.00047	1.41538
40-60	5	0.00699	0.34251
	8	0.00939	4.24167
> 60	5	0.00044	0.02232
	8	0.08157	0.10931

The tables 4.7.7, 4.7.8 and 4.7.9 show that survival rates using Semi-Markov model for age-group less than 40 years, between 40 to 60 years and greater than 60 years are in close agreement with the observed survival rates. The values of chi-square in table 4.7.10 between Semi-Markov model's predictions and observed rates are not significant at 5 percent level. Thus, this model adequately describe the survival within age-groups. The predicted survival rates at 8 years for age-groups between 40 to 60 years, using Markov model is significant at 5 percent level. Though the chi-square values for other two age-groups do not show statistical significance between Markov model's prediction and observed rates. The differences are fairly large as compared with Semi-Markov model. The differences between predicted survival rates obtained by Semi-Markov model and observed rates at 8 years for various age-groups varies by a maximum of 3.8 percent. However, in case of Markov model, the difference ranges between 4.2 to 19.6 percent. Further, it may be noted that in all cases Markov model underestimates the survival probabilities as was the case in over-all and stage-wise estimation. Thus, Semi-Markov model provides a better representation of the prognosis of cancer of the cervix.

Application of the Model on Other Data

It is possible that the Semi-Markov model is adequately representing the set of data used for developing this model. But it may not be suitable for a different set of data. For this reason, a different set of data was used to evaluate the suitability of the

Semi-Markov model in describing the survival experience of cancer of the cervix patients.

The model is applied in a series of 36 case histories for cancer of the cervix patients given by Sorensen (1958). The patients were treated for cervical carcinoma during the period 1922 to 1929 at Radium centre, Copenhegan. In these cases, there were many cases of advanced disease. The results of the model based on five years observation are given in table 4.7.11.

Table 4.7.11

Observed and Predicted Survival Rates for Cancer of the Cervix

Years follow-up	Survival percentage by different methods		
	Semi-Markov	Markov	Observed
1	75.0	65.8	76.1
2	59.8	55.9	69.9
3	51.5	50.5	57.9
4	46.9	46.4	57.9
5	44.4	43.1	48.3
6	42.9	40.1	42.3
7	42.1	37.9	42.3
8	41.6	36.1	42.3
9	41.3	34.5	42.3
10	41.1	33.3	42.3

The table 4.7.11 shows that the survival rates estimated by Semi-Markov model are closer to the observed survival rates than those predicted by Markov model. Markov model, again underestimates the survival rates for this set of data as well. The underestimation increases as the time from the start of follow-up becomes more distant. All of these results in different situations indicate that the Semi-Markov model is a suitable model for representing the survival distribution for cancer of the cervix patients.

4.8 Discussion

On theoretical grounds, it was hypothesized that the Semi-Markov model will be more suitable to describe the course that a patient of cancer of the cervix will take from the time of treatment to the end of follow-up. The main characteristic in this model which makes it more versatile than other models is the fact that it accounts for the duration that a patient stays in a particular state of disease. For example, if a patient stays recovered for a longer time, chances of her relapse are less. However, the model will be of no practical value unless it could be demonstrated that the model adequately describes the observed mortality experience of a group of patients suffering from cancer of the cervix.

The data from Victoria Hospital Cancer Clinic, London were available for an 8-year follow-up. The first five years data were used to estimate the parameters of the model. These were then used to predict future outcome of cancer of the cervix patients,

and comparison was made on the 8-year rates.

There were 101 cancer of the cervix patients. The close agreement between the rates predicted by the Semi-Markov model and the observed rate is evident from figure 4.7.1. Chi-square test suggests that the difference between these two rates is not significant. The rates predicted by the Markov model were significantly different from the observed rate. This implies that the Markov model does not fit the observed data and hence is not an adequate representation of the real prognosis of the disease. Further, the Markov model underestimates the survival rates at all time points. This underestimation increases from 5.5 percent at 4-year to 18.8 percent at 8-year from follow-up. In case of the Semi-Markov model there is a maximum difference of only one percent with the observed rate at any time. The underestimation of survival rates can be explained by the fact that the Markov model assumes geometrical survival distribution which implies a constant force of mortality over time; for example, the probability of surviving another year is the same for a patient who has survived 7 years as one who has survived 3 years. This constant force of mortality is to some extent taken cared in the Semi-Markov model by considering the survival distribution depending both on the state occupied as well as on the next state visited.

After being satisfied that the over-all prediction by Semi-Markov model is adequate, the question was asked whether this model will describe the mortality experience of patients with

various stages of disease.

The survival rates were estimated separately for different stages. The differences in survival rates predicted from the Semi-Markov model with the observed rates vary from 0 to 5 percent. (tables 4.7.3, 4.7.4, and 4.7.5) approximately; and with the Markov model, this difference increases to almost 18 percent. It can be seen from figures 4.7.2, 4.7.3, and 4.7.4 that the Semi-Markov model is adequate for predicting the mortality experience in various stages. It is well-known that the mortality of cancer patients is dependent on age. Therefore, the patients were divided into several homogenous age-groups. The model in these cases also predicted the survival experience to a great degree of accuracy, and in all situations better than Markov model which again underestimates the survival rates in various age-groups (tables 4.7.7, 4.7.8, and 4.7.9).

It may be argued that this model was sufficient the case of this particular set of data, but may not hold true for another set of data. To validate this model for universal use, another set of data was used (Sorenson 1958). The parameters of the model were estimated for these data and predictions were made. It is demonstrated as is evident from the table 4.7.11 in section 4.7 that the Semi-Markov model is in close agreement with the observed rates, whereas the Markov model is again underestimating the survival probabilities of these patients. The procedure used in

Semi-Markov model for estimating the parameters is less cumbersome than the iterative procedure used by Boag (1949) in case of lognormal model where in certain cases the iterative process does not converge. Also, in many instances, the lognormal model was not suitable to describe the survivorship as discussed by Berg (1965), Gehan (1969) and Campos (1971). In summary, it can then be concluded that the Semi-Markov model describes the mortality experience more adequately than any other models that exist so far.

CHAPTER 5

FURTHER USES OF THE MODEL

5.1 Introduction

Most of the cancer studies emphasize the survivorship as the end point and ignore the duration of remaining relatively free of disease, viz., 'recovery' which is an important dimension of the broader picture of the prognosis of the cancer. This component could differ widely for different treatments while the survivorship may remain the same. Therefore, the treatment for which the duration of recovery state is longer should logically be considered superior to others. However, this aspect of prognosis has not received much attention from investigators. Lack of proper methodology might have been responsible to a great extent. The Semi-Markov model developed in the previous chapters attempts to bridge this gap to some extent. The model can also be used to estimate the survival rates from different points of time, viz., from the time of relapse and recovery etc. In addition, its use in estimating the impact of changes in the parameters of the model will be demonstrated for the patients with cancer of the cervix.

5.2 Length-of-stay Distribution in Transient States

The proposed Semi-Markov model for cancer of the cervix has three transient states: under-treatment (S_1), recovery (S_2) and relapse (S_3) and two absorbing states: death due to cancer of the cervix (S_4) and death due to other causes or lost to follow-up (S_5). Let the random variable v_{ij} represent the number of months that a patient will spend in the transient state S_j (j = 1, 2, and 3) if he moved from the transient state S_i (i = 1, 2, and 3). For example, a patient might have to spend on average six months in the state of relapse (S_3) if he has moved from the state of recovery (S_2). The probability mass functions of random variables v_{ij} (i,j = 1, 2, & 3) are completely determined by the transition probabilities and holding-time distributions of the Semi-Markov process. The expression for mean and variance for v_{ij} (i,j = 1, 2, & 3) in terms of the parameters of the Semi-Markov model are given in section 2.6.

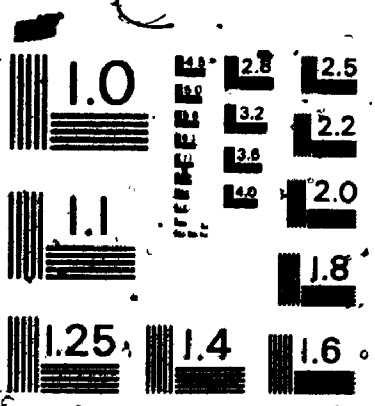
Let N_μ be a 3 by 3 matrix with elements $\bar{v}_{ij}$ (i,j = 1,2,3):

$$N_\mu = [\bar{v}_{ij}] \qquad (5.2.1)$$

where, $\bar{v}_{ij} = E(v_{ij})$.

The above matrix N_μ is estimated by the following matrix equation:

$$N_\mu = (I - P_o)^{-1} M \qquad (5.2.2)$$



where, P_o is a 3x3 transition probability matrix corresponding to the transient states; M, a 3x3 diagonal matrix of mean waiting time for the transient states, and I is the identity matrix. Details have been previously given in section 2.6 of chapter 2.

The variance of v_{ij} (i,j = 1, 2, & 3) is given by

$$\text{Var}(v_{ij}) = \bar{v}_{ij}(2\bar{v}_{jj} - 1 - \bar{v}_{ij}) \qquad (5.2.3)$$

Let the standard deviation matrix for v_{ij} (i,j = 1, 2, and 3) be denoted by N_σ with elements σ_{ij}:

$$N_\sigma = [\sigma_{ij}], \quad i,j = 1,2, \&3. \qquad (5.2.4)$$

Where, $\sigma_{ij} = \sqrt{\text{Var}(v_{ij})}$

In order to obtain N_μ and N_σ matrices for the expected duration of stay, the matrix P_o, the inverse of the matrix $(I - P_o)$ and matrix M were estimated from the observed data which are given in appendix 4. These matrices were then used to estimate the matrices N_μ and N_σ. The table 5.2.1 presents the mean and standard deviation matrices for all the patients together. Similar matrices, for the three stages of the disease are given in tables 5.2.2, 5.2.3 and 5.2.4.

Table 5.2: 1

Mean and Standard Deviation Matrices for v_{ij} for All Patients

$$N_{\mu} = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 2.074 & 91.351 & 7.229 \\ 0 & 102.519 & 8.113 \\ 0 & 10.794 & 8.540 \end{bmatrix} \end{array}$$

$$N_{\sigma} = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 1.493 & 101.459 & 7.959 \\ 0 & 102.018 & 8.041 \\ 0 & 45.671 & 8.025 \end{bmatrix} \end{array}$$

Table 5.2.2

Mean and Standard Deviation Matrices v_{ij} for Stage I

$$N_\mu = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 1.702 & 200.483 & 9.050 \\ 0 & 218.699 & 9.871 \\ 0 & 43.740 & 9.971 \end{bmatrix} \end{array}$$

$$N_\sigma = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 1.093 & 199.983 & 9.366 \\ 0 & 218.198 & 9.358 \\ 0 & 131.054 & 9.358 \end{bmatrix} \end{array}$$

Table 5.2.3

Mean and Standard Deviation Matrices for v_{ij}

for Stage II

$$N_\mu = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 1.658 & 79.773 & 5.792 \\ 0 & 83.400 & 6.055 \\ 0 & 9.265 & 7.740 \end{bmatrix} \end{array}$$

$$N_\sigma = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 1.045 & 82.842 & 7.094 \\ 0 & 82.899 & 7.143 \\ 0 & 38.084 & 7.223 \end{bmatrix} \end{array}$$

Table 5.2.4

Mean and Standard Deviation Matrices for v_{ij}

for Stage III

$$N_{\mu} = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 3.562 & 24.362 & 6.623 \\ 0 & 34.606 & 9.681 \\ 0 & 0 & 10.650 \end{bmatrix} \end{array}$$

$$N_{\sigma} = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 1.045 & 33.422 & 9.518 \\ 0 & 35.103 & 10.139 \\ 0 & 0 & 10.137 \end{bmatrix} \end{array}$$

The tables 5.2.1 to 5.2.4 give the mean and standard deviation for the expected duration of stay in the transient states; viz., under-treatment, recovery and relapse. Knowing these, it is possible to write the confidence statement concerning the mean of the duration of stay in transient states. The standard error of the mean $\bar{v}_{ij}$ is $\sigma_{ij} / \sqrt{n}$, where n is the number of patients observed. Using normal tables, $(1 - \alpha)$ level confidence interval for $\bar{v}_{ij}$ then can easily be calculated.

It is seen from table 5.2.2 that a patient who entered in the study suffering with stage I carcinoma of the cervix is expected to stay in the recovery state for about 200.5 months with standard deviation 199.9 months after treatment; while for stage II patients the mean and standard deviation for the length of stay in recovery state are 79.7 and 82.8 months respectively (table 5.2.3) and for stage III cases, the corresponding mean and standard deviation are 24.4 and 33.4 months respectively. The average length of stay in the recovery state decreases with the increasing severity of the disease.

The coefficient of variation for the length of stay in the recovery state after treatment corresponding to stage I, stage II, and stage III are 99.7 percent, 103.9 percent and 136.9 percent respectively. These coefficient of variations suggest that the duration of stay distribution in the recovery state is highly variable. Similarly, the duration of stay in other states is also highly variable. Thus, these must be highly skewed.

5.3 Influence of Relapse

Kaplan (1962) suggested the use of computing survival from the time of relapse in the prognosis of Hodgkin's disease. It therefore, seems desirable to examine the survival experience from the time of relapse for cancer of the cervix patients. The Semi-Markov model provides an easy method since these probabilities are already available as a by-product in the mortality experience.

The survival rates from the time of relapse can be directly obtained from the interval transition probability matrices as described in section 4.6. The survival rates at 3 monthly interval from the time of relapse are given in table 5.3.1.

Table 5.3.1

Survival Rates Based on the Semi-Markov Model From the Time of Relapse According to the Stage of Cancer of the Cervix

	Survival rates with respect to stages			
Months	I	II	III	All Patients
3	77.0	67.9	74.4	73.6
6	60.7	47.4	55.3	54.9
9	48.8	33.9	41.2	41.2
12	40.3	25.3	30.6	32.3
15	34.1	19.6	22.8	25.5
18	29.6	15.9	16.9	20.6
21	26.3	13.2	12.6	17.1
24	23.9	11.7	9.4	14.6

The survival rates of the patients in the three stages of the disease are plotted against time in figure 5.3.1 to examine the time trends. It appears mortality experience of stage III patients is uniformly worse than stage I or II patients. This predicted out-come corresponds to the expectation based on clinical judgement. Further, at 2 years after relapse, 24 percent patients in stage I of disease survive as compared to only 9 percent in stage III with about 12 percent in stage II. This demonstrates that the survival experience of stage I patients even after relapse is better than those of stage III.

Another aspect of comparing survival rates could be from the time of recovery instead of measuring survival from the time of treatment. This type of analysis seems to be important since the observed survival curves from the time of treatment in figure 5.3.2 shows that the survival during first year is better for stage II cases than stage I cases. But if the survival is plotted from the time of recovery as in figure 5.3.3, it is seen that survival for stage I is uniformly better than stage II. The reason might be that more patients die during treatment for stage I cases. Therefore, it also seems worth-while to measure survival from the time of recovery.

Figure 5.8.1

Survival Curves from the Time of Relapse According to Stage of Disease Based on Semi-Markov Model

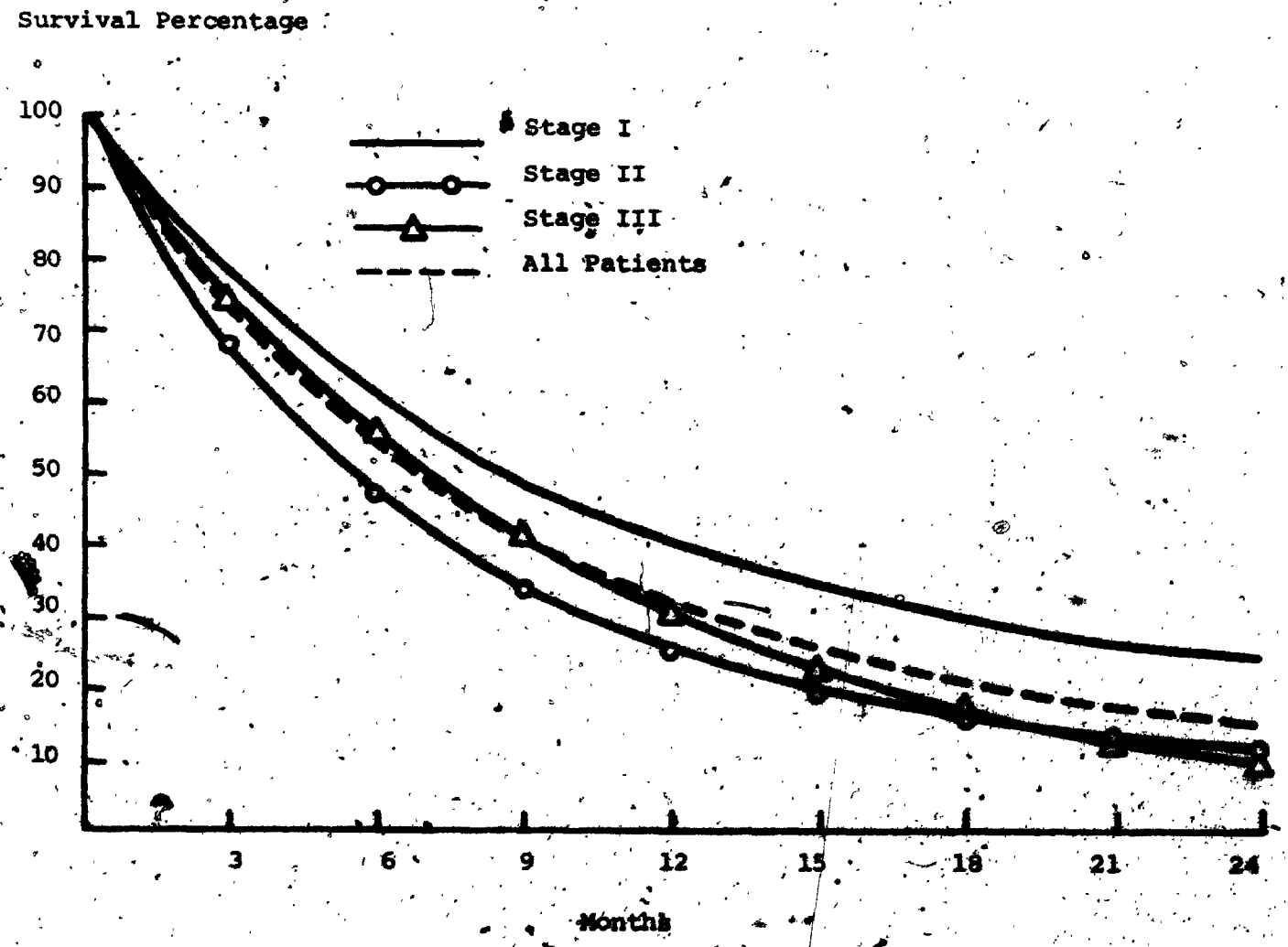


Figure 5.3.2

Observed Survival Curves for Cancer of Cervix Patients According to the Stage of Disease from the Time of Treatment

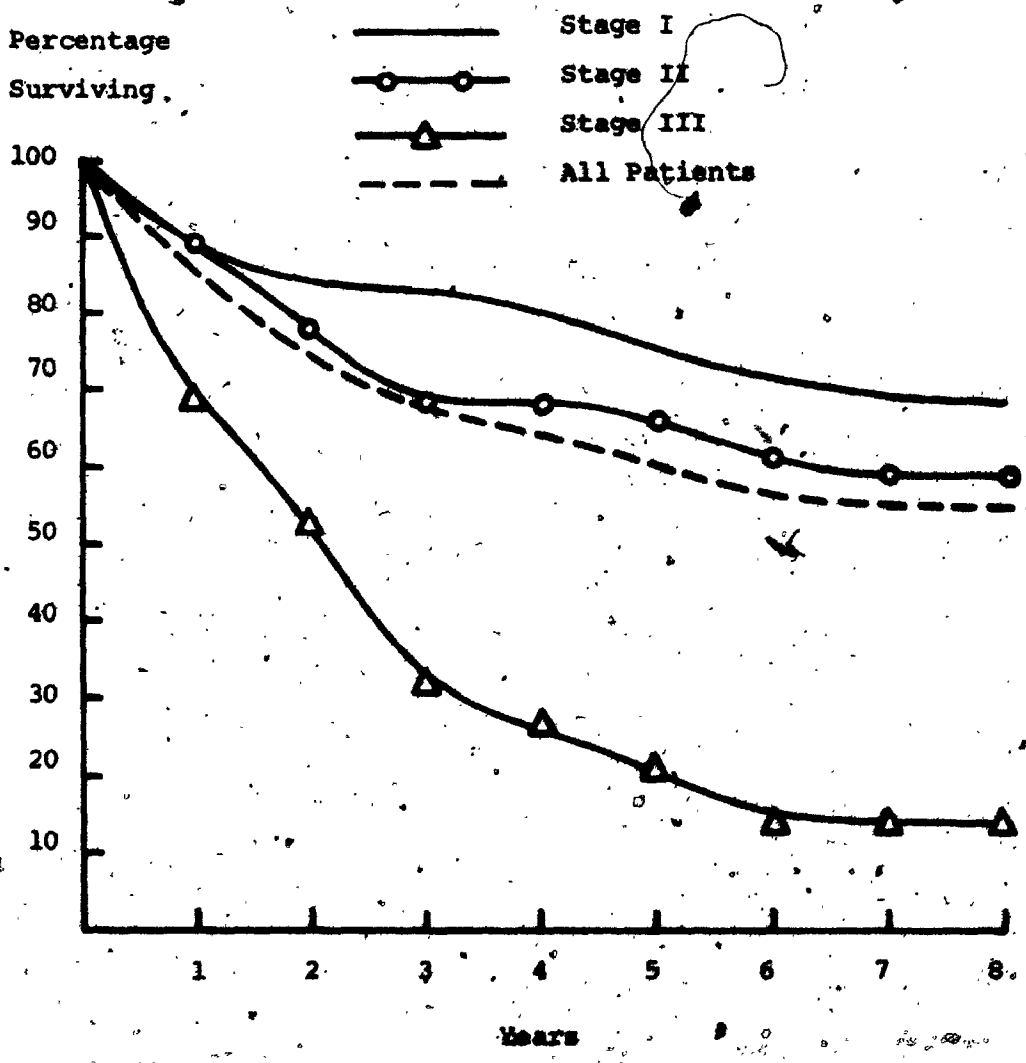
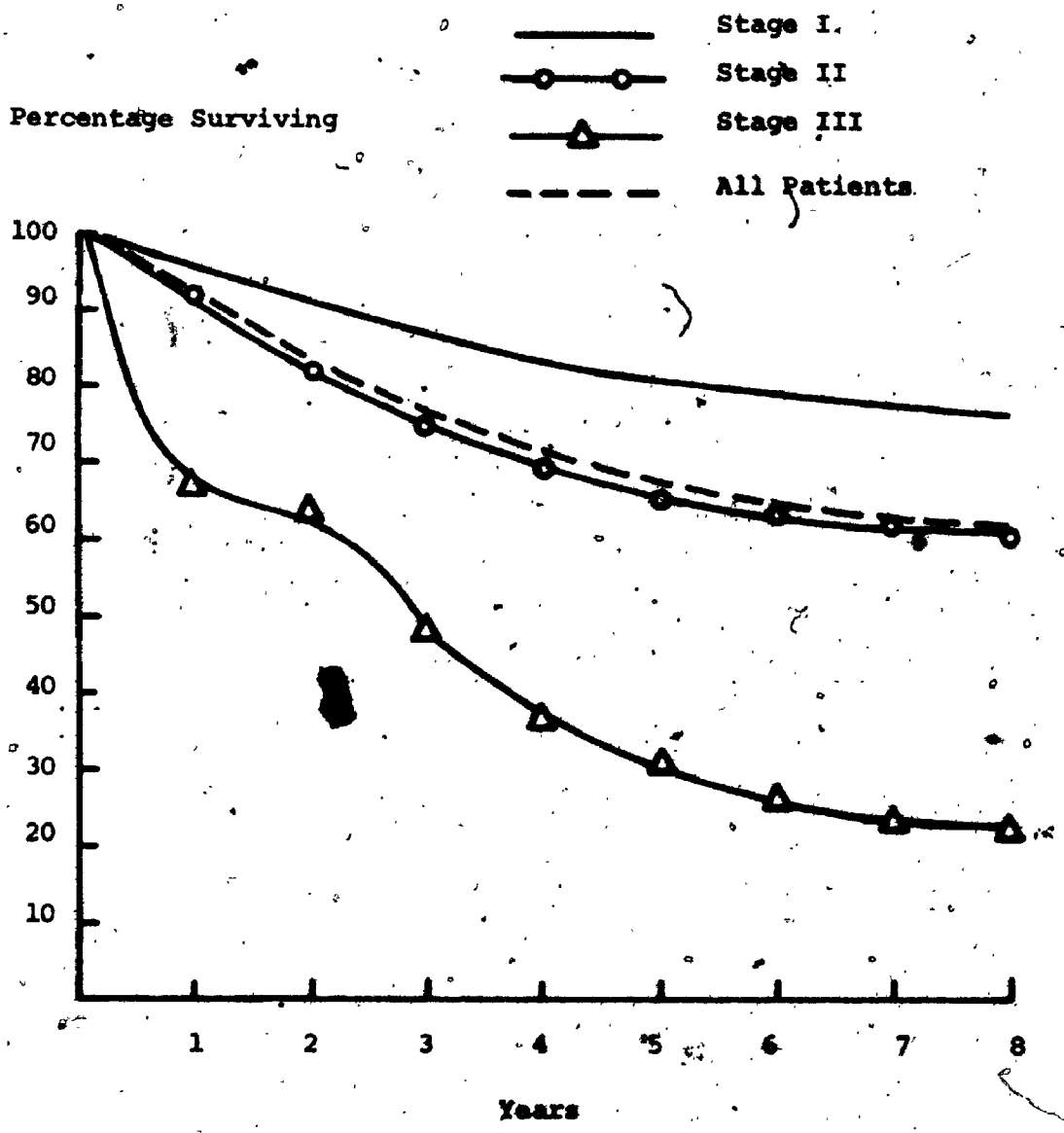


Figure 5.3.3

Survival Curves Estimated by the Semi-Markov Model from the Time of Recovery According to Stage of Disease



5.4 The Effect of Parameters change on Survival

A Semi-Markov model can also be used to study the impact of change in any or a combination of parameters. Clinicians and health planner may be interested in estimating the T-year survival rate if the probability of dying under-treatment could be reduced by certain percentage points. This model can provide an estimate immediately without conducting a long term trial. Of course, this approach does not replace the actual observed fact which will be available several years hence; but simply allows one to make intelligent decisions.

The estimates are subject to error. It is important to investigate the tolerance of the model with respect to these errors in the estimation of parameters. This is important since large fluctuations in the final results with minor changes in parameters may require the estimation of these parameters with great degree of precision. That may mean large data sets and consequently high cost.

Though these two questions raised above are conceptually distinct, but the methodological approach is exactly identical. Certain changes in parameters are hypothesized and resulting survival rates at distinct future time points are analysed. This approach will be demonstrated by changing the probability of dying, relapse rate and duration of stay in recovery state. In fact, any combination of change could have been used but to keep the discussion simple and to the point only the above three will be considered.

Change in Probability of Dying Due to Cancer of the Cervix During Treatment

Let us suppose that it is possible to develop a treatment for cancer of the cervix which may reduce the probability of dying during treatment by 5 percent. The model is then used to estimate the effect of this change on the survival rates. In order to use the model, one has to adjust the other probabilities in the transition probability matrix. In our case, the transition probability from under-treatment to recovery is increased by 5 percent. The estimates of this change on the survival rates for all patients are given in table 5.4.1.

Table 5.4.1

Survival Rates Based on Semi-Markov Model for All Patients With Changes in Probability of Dying Due to Cancer of Cervix

Years after treatment	No change	5 percent decrease in probability of dying due to cancer of the cervix
1	87.0	91.5
2	76.1	83.1
3	68.7	76.7
4	63.5	72.3
5	59.9	69.2
6	57.4	67.2
7	55.7	65.8
8	54.5	64.8

The table 5.4.1 shows that the survival rates are increased by almost 10 percent after 5 years. Thus, an investigator can weigh the importance of this 10 percent increase in survival against the cost of reducing the mortality during treatment.

Changes in Relapse Rate

The table 5.4.2 presents the survival rate when the probability of relapse is reduced at an interval of 2 percent to a maximum of 10 percent.

Table 5.4.2

Survival Rates Based on the Semi-Markov Model for Cancer of the Cervix Patients with Changes in Relapse Probability

Years after treatment	Decrease in relapse probability by					
	No change	2%	4%	6%	8%	10%
1	87.0	87.2	87.4	87.6	87.9	88.1
2	76.1	77.1	77.7	78.3	78.9	79.6
3	68.7	69.7	70.7	71.7	72.7	73.6
4	63.5	64.7	65.9	67.1	68.4	69.6
5	59.9	61.3	62.6	64.1	65.5	66.8
6	57.4	59.9	60.5	61.9	63.5	65.0
7	55.7	57.4	59.0	60.6	62.2	63.8
8	54.5	56.3	58.0	59.7	61.3	62.9

The decrease in transition probability from recovery to relapse given that the patient was in recovery state by 10 percent in table 5.4.2 shows that the survival rate increases about 7 percent at 5 years and 8.4 percent at 8 years. However, there is not much change in the survival rates during the first two years. The survival rates change very slowly with the decrease in the probability of relapse. The changes lie between 0.2 to 2 percent for each 2 percent reduction in the relapse probability. It seems that in order to improve survival rate by a noticeable margin, a substantial reduction in relapse probability will have to be achieved. In other words, it can be interpreted that if the error in estimation of relapse rate is small then the predicted results of the model will not fluctuate to any large extent.

Change in Expected Length of Stay in Recovery State

The other possible variation in the parameters of the model could be the expected length of stay in recovery state before relapse. The effect of these changes on the survival rates are presented in table 5.4.3.

Table 5.4.3

Survival Rates Based on the Semi-Markov Model for All Cancer of Cervix Patients With Changes in the average Length of State in Recovery Before Relapse

Years after treatment	Changes in average length of stay in recovery before relapse				
	No change	Decrease by 5 months	decrease by 10 months	increase by 5 months	increase by 10 months
1	87.0	86.2	84.9	87.6	88.1
2	76.1	74.5	71.9	77.9	79.2
3	68.7	66.4	63.5	70.8	72.4
4	63.5	61.1	58.6	65.5	67.3
5	59.9	57.8	55.8	61.7	63.4
6	57.4	55.8	54.3	59.1	60.6
7	55.7	54.5	53.5	57.1	58.5
8	54.[illegible]	53.7	53.0	55.7	56.9

The results in the above table show that the increase in the average length of stay in the state of recovery by 10 months increases the survival rate by 3.5 percent at 5 years and 2.4 percent at 8 years; while the survival rate is decreased by 4.1 percent at 5 years and 1.5 percent at 8 years when the duration of stay in the recovery state is decreased by 10 months. With the increase or

decrease in duration of the recovery state there has been some increase or decrease in the survival rates at future time points. However, these changes have been very modest. Again it will appear that even as large a change of about 25 percent in duration of stay in recovery does not appreciably alter the probability of survival after 5 years of follow-up. Similarly, conclusion can also be drawn that to change the survival pattern by changing the length of recovery state, will require a substantial increase or decrease in this parameter.

5.5 Discussion

It was demonstrated that Semi-Markov model can also be used to provide estimates and information on aspects other than survival alone. It can be used to estimate future survival patterns from the time of treatment, recovery or relapse. Expected duration of stay in various states of the disease can also be easily estimated.

The model was used to study the effect of relapse on the subsequent survival. It was estimated that 14.6 percent of the patients who relapsed survived 2 years from that time, whereas 76.1 percent of the patients survived from the time of treatment. Obviously, the prognosis of the patients who relapse is bleak. The model also provides estimates of the probability of surviving from the time patient goes under-treatment and when she recovers. These may be important in case there is a substantial mortality during treatment.

Expected duration of stay in recovery state were estimated for the three stages of the disease. In agreement with clinical judgement, the estimated duration was highest for stage I and lowest for stage III. This provides an alternative measure to compare the efficacy of two possible treatments. The treatment for which the expected duration of recovery state is longer, other measures remaining constant, will logically be considered to be superior.

Finally, the model provides a useful tool for exploring the impact of changes in the parameters. It was shown that a 5 percent decrease in the mortality due to cancer of the cervix during treatment increased the survival after 5 years by 10 percent. The survival rates did not change appreciably with the decrease in the probability of relapse or with substantial changes in the duration of recovery. This type of analysis allows the clinician to anticipate the future outcome of changes in treatment policies. Without the model, he has no other recourse but to wait for several years to see the results to come.

CHAPTER 6

SUMMARY AND CONCLUSION

Survivorship studies in chronic diseases have relied heavily on life table methods. This in general required a long follow-up of patients, depending upon the time period for which the rates were needed. Investigators could not predict future survivorship from a limited follow-up. Boag (1949) in his attempt to predict the survivorship used a lognormal model to describe the mortality experience of 'uncured' patients. This method required iterative procedures to estimate the parameters which in some instances did not converge. Moreover, the model did not provide a good fit to the mortality data as observed by Berg (1965), Gehan (1969) and Campos (1971).

Fix and Neyman (1951) used a Markov process to develop a technique for predicting mortality for cancer patients. They assumed that the transition probability from any state S_i to any other state S_j depends only on the state S_i. Their model does not account for the duration of stay in a given state. This time dimension appears to be an important factor in the total prognosis of cancer. The principles of the Semi-Markov process provide a handy

approach to deal with this problem. Based on these principles, a 5 states Semi-Markov model for the prognosis of cancer of the cervix was proposed. The five states of the model are:

S_1 Under-treatment;

S_2 Recovery;

S_3 Relapse;

S_4 Death due to cancer of the cervix; and

S_5 Death due to other causes or lost to follow-up.

The length of stay distribution in various transient states were assumed to follow a geometrical distribution depending on the state occupied as well as on the next state entered.

In order to investigate the validity of the model, the data from Victoria Hospital Cancer Clinic, London, Ontario was collected for cancer of the cervix patients. There were 101 patients. The performance of the model was first tested on all patients. Later on, the data was classified according to the stage of the disease and different age-groups. The parameters of the model (i.e., Transition probability matrix and Holding-time matrix) were estimated by the maximum likelihood method based on data for the first five years and the predictions were made for 8 years. These predicted rates were compared with the Markov model and observed rates. A brief summary for these comparisons in different situations at 8-years is described as follows:

	Survival percentage at 8-years by Semi-Markov model	Survival percentage at 8-years by Markov model	Observed survival percentage
All Patients	54.5	36.7	55.5
Stage I	70.7	52.3	69.4
Stage II	58.4	41.4	59.4
Stage III	14.5	5.7	13.6
< 40 Years	65.2	47.9	65.3
40-60 Years	58.5	39.9	59.5
> 60 Years	33.4	25.4	29.6

The above results indicate that the Semi-Markov model provides a closer fit to the observed data; while the survival rates predicted by Markov model underestimate the observed rates in each of these cases.

The model was further tested on another set of data on cancer of the cervix available from Sorensen (1958) to validate the applicability of this approach to a more general situation. The predicted survival rates at ten years based on Semi-Markov and Markov model were 41.1 and 33.3 percent respectively; whereas the observed rate was 42.3 percent. Again, in this case, the Semi-Markov model provided better estimates of survival rates than those obtained from the Markov model.

Thus, the Semi-Markov model gave a better fit to the observed data than the one provided by the Markov approach. It was observed in chapter 4 that the Markov model had a tendency to

underestimate the survival rates and this underestimation increases rapidly beyond 5 years. This underestimation was probably due to the assumption that the length of stay distribution in any transient state before transitting to any other state is a geometrical distribution. This implies a constant force of mortality. This assumption is not realistic since the force of mortality decreases as the survival time increases. The Semi-Markov model does not require this assumption. Infact, it assumes that the survival distribution in a given state depends also on the next state visited.

In addition to being able to describe the survival pattern of the observed data more closely, it provides a tool for estimating the length of stay in the transient states. The length of stay is an important dimension of the total prognosis which along with the survival rates could be used for comparing the efficacy of different treatments. The quantitative structure of this model could be utilized to calculate the survival from the time of recovery, relapse or any other specified state since these survival probabilities were already available in the system during the estimation procedure. Thus, the Semi-Markov model provides a comprehensive and compact way to estimate these survival probabilities; whereas in the life-table one has to construct different life-tables for different starting points.

The other use of the model could be in the planning purposes. The proposed model can be used to consider the effect of changes in the parameters on the prognosis of the disease.

Several illustrations of this aspect of the model are provided in chapter 5. For instance, it was shown that a decrease of 5% in the probability of dying during treatment increases the survival probability by 10 percent. However, decrease or increase in duration of about 10 months did not effect the survival to any great extent. This information is vital to any decision maker. The model provides an explicit and logical tool for the analysis and evaluation of changes in the system. A better insight into the prognosis of the disease without waiting for the results for a long time is achieved. This, of course, does not eliminate the need for actual observation over a period of time. It, however, provides an objective assessment of the possible outcome of various alternative policies for the decision maker to evaluate.

In summary, it can be concluded that the Semi-Markov model provides an adequate description for the prognosis of cancer of the cervix. It is superior to the Markov model which underestimates the survival probabilities.

Suggestions for Further Work

The model presented in this thesis suggests at least some directions in which future research might prove useful. One of these is exploration of the applicability of similar models to other types of cancer. The other is the exploration of more complex models, such as models in which the assumption of a closed population is relaxed, i.e., the population state is an open system. In modeling a closed system, a patient cohort is assumed to follow a Semi-Markov

process while adding new patients into the system means that each time a patient enters the recovery system, another Semi-Markov process begins. Combining these many Semi-Markov processes, it is possible to calculate the proportion of patients in various states of the recovery system.

APPENDIX 1

Form Used In Data Collection

Case No.	Year	Age	Sex	Stage of disease	Type of treatment	Observed State	Observed Time period
						Under-treatment Recovery Relapse Death due to disease Death due to other causes or lost to follow-up	

APPENDIX 2

Interval Transition Probability Matrices $\Phi(t)$, at $t = 12, 24, 36, 48, 60, 72, 84$, and 96 months for Various Stages of the Disease

Table 1

Interval Transition Probabilities for Stage I

$$\Phi(12) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.00738 & 0.85789 & 0.03615 & 0.09856 \\ 0 & 0.92916 & 0.04077 & 0.03006 \\ 0 & 0.18769 & 0.21519 & 0.59711 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(24) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.00065 & 0.81076 & 0.03723 & 0.15135 \\ 0 & 0.87952 & 0.03991 & 0.08057 \\ 0 & 0.17729 & 0.06196 & 0.76075 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(36) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.00006 & 0.77613 & 0.03012 & 0.19369 \\ 0 & 0.84305 & 0.03191 & 0.12504 \\ 0 & 0.16963 & 0.02059 & 0.80978 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(48) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.00001 & 0.75069 & 0.02285 & 0.22646 \\ 0 & 0.81625 & 0.02411 & 0.15964 \\ 0 & 0.16399 & 0.00867 & 0.82734 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$\Phi(72)$ =

	S_1	S_2	S_3	S_4
S_1	0	0.73198	0.01698	0.25104
S_2	0	0.79655	0.01789	0.18556
S_3	0	0.15986	0.00469	0.83545
S_4	0	0	0	1.00000

$\Phi(84)$ =

	S_1	S_2	S_3	S_4
S_1	0	0.70814	0.00922	0.28264
S_2	0	0.77144	0.00971	0.21885
S_3	0	0.15458	0.00210	0.84331
S_4	0	0	0	1.00000

$\Phi(96)$ =

	S_1	S_2	S_3	S_4
S_1	0	0.70072	0.00677	0.29250
S_2	0	0.76362	0.00712	0.22924
S_3	0	0.15294	0.00152	0.84554
S_4	0	0	0	1.00000

Note: There are no deaths due to other causes or lost to follow-up; thus there is no state S_5 for stage I cases.

Table 2

Interval Transition Probabilities for Stage II

$\phi(12)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.01527	0.72435	0.06687	0.08682	0.10669
S_2	0	0.74528	0.07042	0.07164	0.11266
S_3	0	0.08064	0.16624	0.74328	0.00984
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\phi(24)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00536	0.63489	0.05603	0.19292	0.11079
S_2	0	0.65676	0.05665	0.17061	0.11599
S_3	0	0.07659	0.03308	0.87782	0.01251
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\phi(36)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00188	0.57934	0.03829	0.26854	0.11195
S_2	0	0.60118	0.03842	0.24327	0.11713
S_3	0	0.06959	0.00948	0.90801	0.01293
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\phi(48)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00066	0.54352	0.02509	0.31805	0.11269
S_2	0	0.56533	0.02513	0.29168	0.11787
S_3	0	0.06466	0.00419	0.91810	0.01305
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(60) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0.00023	0.52037	0.01627	0.34997	0.11316
S_2	0	0.54216	0.01629	0.32321	0.11835
S_3	0	0.06143	0.00239	0.92306	0.01312
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(72) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0.00008	0.50541	0.01053	0.37051	0.11347
S_2	0	0.52718	0.01054	0.34363	0.11865
S_3	0	0.05935	0.00149	0.92600	0.01317
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(84) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0.00001	0.49574	0.00680	0.38375	0.11367
S_2	0	0.51750	0.00681	0.35683	0.11885
S_3	0	0.05799	0.00095	0.92785	0.01319
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(96) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0.00001	0.48949	0.00439	0.39230	0.11379
S_2	0	0.51125	0.00440	0.36536	0.11898
S_3	0	0.05712	0.00061	0.92905	0.01321
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

Table 3

Interval Transition Probabilities for Stage III

$\Phi(12)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.07098	0.46710	0.12841	0.32362	0.00988
S_2	0	0.65442	0.19424	0.13738	0.01596
S_3	0	0	0.30629	0.69371	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(24)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.01596	0.32149	0.11907	0.52471	0.01875
S_2	0	0.45339	0.16874	0.34925	0.02861
S_3	0	0	0.09381	0.90618	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(36)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00359	0.23633	0.08204	0.65225	0.02579
S_2	0	0.33572	0.11447	0.51117	0.03864
S_3	0	0	0.02873	0.97127	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(48)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00081	0.18610	0.05117	0.73056	0.03136
S_2	0	0.26625	0.07094	0.61622	0.04659
S_3	0	0	0.00880	0.99119	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(60)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00018	0.15616	0.03056	0.77732	0.03578
S_2	0	0.22479	0.04223	0.68009	0.05289
S_3	0	0	0.00269	0.99973	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(72)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00004	0.13808	0.01786	0.80473	0.03928
S_2	0	0.19970	0.02465	0.71776	0.05788
S_3	0	0	0.00083	0.99917	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(84)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00001	0.12696	0.01033	0.82063	0.04206
S_2	0	0.18426	0.01425	0.73964	0.06185
S_3	0	0	0.00025	0.99975	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(96)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.12300	0.00594	0.82979	0.04426
S_2	0	0.17456	0.00819	0.75226	0.06498
S_3	0	0	0.00008	0.99992	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

Table 4

Interval Transition Probabilities for All Patients

$\Phi(12)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.02498	0.76624	0.06663	0.12895	0.01321
S_2	0	0.84682	0.07647	0.06034	0.01637
S_3	0	0.09087	0.23083	0.67708	0.00123
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(24)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00573	0.67406	0.06287	0.23260	0.02474
S_2	0	0.74764	0.06875	0.15466	0.02895
S_3	0	0.08167	0.06247	0.85320	0.00265
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(36)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00131	0.61168	0.04638	0.30702	0.03359
S_2	0	0.68046	0.05014	0.23078	0.03861
S_3	0	0.07369	0.01943	0.90311	0.00376
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(48)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00030	0.56928	0.03199	0.35804	0.04039
S_2	0	0.63478	0.03445	0.28475	0.04602
S_3	0	0.06824	0.00746	0.91968	0.00461
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(60)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00007	0.54038	0.02157	0.39239	0.04558
S_2	0	0.60365	0.02319	0.32147	0.05169
S_3	0	0.06451	0.00362	0.92661	0.00526
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(72)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00002	0.52064	0.01442	0.41536	0.04956
S_2	0	0.58237	0.01550	0.34611	0.05602
S_3	0	0.06197	0.00207	0.93019	0.00576
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(84)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00001	0.50712	0.00961	0.43068	0.05259
S_2	0	0.56779	0.01033	0.36255	0.05932
S_3	0	0.06023	0.00129	0.93233	0.00614
S_4	0	0	0	1.00000	0.
S_5	0	0	0	0	1.00000

$\Phi(96)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.49782	0.00640	0.44088	0.05489
S_2	0	0.55776	0.00688	0.37351	0.06184
S_3	0	0.05904	0.00084	0.93369	0.00643
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

APPENDIX 3

Interval Transition Probability Matrices, $\Phi(t)$

at t = 12, 24, 36, 48, 60, 72, 84

and 96 Months Corresponding

to Different Age-groups

Table 1

Interval Transition Probabilities for Age-Group

Less Than 40 Years

$$\Phi(12) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.01638 & 0.81974 & 0.05221 & 0.11157 \\ 0 & 0.88718 & 0.05821 & 0.05462 \\ 0 & 0.09548 & 0.18722 & 0.71730 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(24) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.00309 & 0.75532 & 0.04616 & 0.19543 \\ 0 & 0.82076 & 0.04877 & 0.13047 \\ 0 & 0.11382 & 0.04188 & 0.84430 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(36) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.00058 & 0.71510 & 0.03265 & 0.25167 \\ 0 & 0.77906 & 0.03413 & 0.18681 \\ 0 & 0.11373 & 0.01322 & 0.87305 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(48) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.00011 & 0.68915 & 0.02199 & 0.28874 \\ 0 & 0.75208 & 0.02294 & 0.22498 \\ 0 & 0.11034 & 0.60609 & 0.88357 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(60) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.00002 & 0.67216 & 0.01463 & 0.31318 \\ 0 & 0.73439 & 0.01525 & 0.25035 \\ 0 & 0.10718 & 0.00353 & 0.88929 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(72) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.00001 & 0.66097 & 0.00969 & 0.32933 \\ 0 & 0.72274 & 0.01010 & 0.26715 \\ 0 & 0.10484 & 0.00223 & 0.89292 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(84) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.65357 & 0.00642 & 0.34000 \\ 0 & 0.71505 & 0.00669 & 0.27827 \\ 0 & 0.10323 & 0.00146 & 0.89531 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(96) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.64869 & 0.00425 & 0.34707 \\ 0 & 0.70996 & 0.00443 & 0.28562 \\ 0 & 0.10215 & 0.00096 & 0.89689 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

Note: Since there are no deaths due to other causes for age-group less than 40 years; thus there is no state S_5 in this case.

Table 2

Interval Transition Probabilities for Age-Group

Between 40-60 Years

$$\Phi(12) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0.02508 & 0.80177 & 0.04264 & 0.12216 & 0.00834 \\ 0 & 0.85327 & 0.04501 & 0.09188 & 0.00985 \\ 0 & 0.12432 & 0.04514 & 0.82932 & 0.00123 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(24) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0.00865 & 0.71207 & 0.02978 & 0.23394 & 0.01555 \\ 0 & 0.76049 & 0.03080 & 0.19128 & 0.01743 \\ 0 & 0.11023 & 0.00641 & 0.88101 & 0.00235 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(36) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0.00298 & 0.65355 & 0.01924 & 0.30313 & 0.02108 \\ 0 & 0.69997 & 0.01987 & 0.25692 & 0.02324 \\ 0 & 0.10103 & 0.00310 & 0.89265 & 0.00321 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$\Phi(48) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0.00103 & 0.61529 & 0.01236 & 0.34599 & 0.02532 \\ 0 & 0.66039 & 0.01276 & 0.29916 & 0.02769 \\ 0 & 0.09502 & 0.00195 & 0.89916 & 0.00387 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$\Phi(60)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00035	0.59022	0.00794	0.37294	0.02855
S_2	0	0.63443	0.00819	0.32629	0.03108
S_3	0	0.09108	0.00125	0.90329	0.00438
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(72)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00012	0.57373	0.00509	0.39005	0.03100
S_2	0	0.61736	0.00526	0.34372	0.03366
S_3	0	0.08849	0.00080	0.90595	0.00476
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(84)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00004	0.56286	0.00327	0.40096	0.03287
S_2	0	0.60609	0.00338	0.35491	0.03561
S_3	0	0.08678	0.00052	0.90765	0.00505
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(96)$ =

	S_1	S_2	S_3	S_4	S_5
S_1	0.00001	0.55565	0.00210	0.40795	0.03428
S_2	0	0.59864	0.00217	0.36209	0.03709
S_3	0	0.08565	0.00033	0.90875	0.00527
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

Table 3

Interval Transition Probabilities for Age-Group Greater than 60 Years

$\Phi(12)$ =

	S_1	S_2	S_3	S_4	S_4
S_1	0.02254	0.59942	0.09957	0.23198	0.04649
S_2	0	0.74359	0.13350	0.05754	0.06536
S_3	0	0	0.46370	0.53629	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(24)$ =

S_1	0.00234	0.46466	0.11415	0.33488	0.08397
S_2	0	0.57940	0.14439	0.16469	0.11151
S_3	0	0	0.21502	0.78498	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(36)$ =

S_1	0.00024	0.37809	0.09494	0.41630	0.11043
S_2	0	0.47389	0.11793	0.26411	0.14407
S_3	0	0	0.09971	0.90029	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(48)$ =

S_1	0.00003	0.32227	0.06998	0.47852	0.12910
S_2	0	0.40583	0.08618	0.34093	0.16706
S_3	0	0	0.04623	0.95377	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$\Phi(60) =$

	s_1	s_2	s_3	s_4	s_5
s_1	0.00001	0.28614	0.04849	0.52309	0.14228
s_2	0	0.36176	0.05943	0.39554	0.18328
s_3	0	0	0.02144	0.97856	0
s_4	0	0	0	1.00000	0
s_5	0	0	0	0	1.00000

$\Phi(72) =$

	s_1	s_2	s_3	s_4	s_5
s_1	0	0.24562	0.03841	0.56438	0.15158
s_2	0	0.31035	0.04734	0.44757	0.19474
s_3	0	0	0.00698	0.99302	0
s_4	0	0	0	1.00000	0
s_5	0	0	0	0	1.00000

$\Phi(84) =$

	s_1	s_2	s_3	s_4	s_5
s_1	0	0.22381	0.02726	0.59078	0.15815
s_2	0	0.28355	0.03355	0.48008	0.20282
s_3	0	0	0.00305	0.99695	0
s_4	0	0	0	1.00000	0
s_5	0	0	0	0	1.00000

$\Phi(96) =$

	s_1	s_2	s_3	s_4	s_5
s_1	0	0.20863	0.01915	0.60943	0.16278
s_2	0	0.26489	0.02355	0.50302	0.20853
s_3	0	0	0.00133	0.99866	0
s_4	0	0	0	1.00000	0
s_5	0	0	0	0	1.00000

APPENDIX 4

Transition Probability Matrix (P_o) Corresponding to Transient States S_1, S_2, & S_3, Inverse of The Matrix ($I-P_o$) and Diagonal Matrix (M) of Mean Waiting Time

Table 1

Transition Probability Matrix for Transient States

Stage		S_1	S_2	S_3
I	S_1	0	0.9167	0
	S_2	0	0.6970	0.3030
	S_3	0	0.2000	0
II	S_1	0	0.9565	0
	S_2	0	0.4771	0.4091
	S_3	0	0.1111	0
III	S_1	0	0.6842	0
	S_2	0	0.1538	0.7692
	S_3	0	0	0
I+ II + III	S_1	0	0.8911	0
	S_2	0	0.5111	0.4222
	S_3	0	0.1033	0

Table 2

Stage	Inverse of the Matrix $(I - P_o)$			
		S_1	S_2	S_3
I	S_1	1	3.7818	1.1459
	S_2	0	4.1254	0.1249
	S_3	0	0.8251	0.1249
II	S_1	1	2.0034	0.8196
	S_2	0	2.0945	0.8568
	S_3	0	0.2327	1.0952
III	S_1	1	0.8086	0.6219
	S_2	0	1.1818	0.9090
	S_3	0	0	1.0000
I + II + III	S_1	1	2.0049	0.8465
	S_2	0	2.2500	0.9499
	S_3	0	0.2369	1.1000

Table 3

Stage	Diagonal Matrix (M)

I

	S_1	S_2	S_3
S_1	1.7024	0	0
S_2	0	53.012	0
S_3	0	0	7.8980

II

	S_1	S_2	S_3
S_1	1.6589	0	0
S_2	0	39.818	0
S_3	0	0	7.0678

III

	S_1	S_2	S_3
S_1	3.5622	0	0
S_2	0	30.120	0
S_3	0	0	10.6500

I + II + III

	S_1	S_2	S_3
S_1	2.0748	0	0
S_2	0	45,564	0
S_3	0	0	8.5409

APPENDIX 5

t-step Transition Probability Matrices P(t), for t = 1, 2, 3, 4, 5, 6, 7, 8 Years Corresponding to Various Stages of the Disease and Different Age-groups

Table 1

P(t) for Stage I

$$P(1) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.00002 & 0.84836 & 0.04253 & 0.10909 \\ 0 & 0.91464 & 0.04928 & 0.03608 \\ 0 & 0.15229 & 0.20242 & 0.64528 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(2) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.78244 & 0.05042 & 0.16714 \\ 0 & 0.84406 & 0.05505 & 0.10088 \\ 0 & 0.17013 & 0.04848 & 0.78139 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(3) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.72333 & 0.04877 & 0.22790 \\ 00. & 0.78039 & 0.05274 & 0.16686 \\ 0 & 0.16298 & 0.01819 & 0.81882 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(4) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.66901 & 0.04552 & 0.28547 \\ 0 & 0.72181 & 0.04914 & 0.22905 \\ 0 & 0.15185 & 0.01172 & 0.83643 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(5) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \\ \begin{bmatrix} 0 & 0.61883 & 0.04218 & 0.33898 \\ 0 & 0.66768 & 0.04552 & 0.28680 \\ 0 & 0.14067 & 0.00986 & 0.84947 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(6) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \\ \begin{bmatrix} 0 & 0.57243 & 0.03904 & 0.38853 \\ 0 & 0.61761 & 0.04212 & 0.34027 \\ 0 & 0.13016 & 0.00893 & 0.86091 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(7) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \\ \begin{bmatrix} 0 & 0.52951 & 0.03611 & 0.43438 \\ 0 & 0.57131 & 0.03896 & 0.38973 \\ 0 & 0.12041 & 0.00822 & 0.87137 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(8) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \\ \begin{bmatrix} 0 & 0.48981 & 0.03341 & 0.47678 \\ 0 & 0.52847 & 0.03604 & 0.43549 \\ 0 & 0.11138 & 0.00759 & 0.88102 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

Table 2

P(t) for Stage II

$$P(1) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0.00002 & 0.81174 & 0.08484 & 0.07115 & 0.03225 \\ 0 & 0.82702 & 0.09671 & 0.03759 & 0.03867 \\ 0 & 0.05649 & 0.44218 & 0.50002 & 0.00131 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(2) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.67613 & 0.11602 & 0.14409 & 0.06375 \\ 0 & 0.68942 & 0.12275 & 0.11705 & 0.07077 \\ 0 & 0.07169 & 0.20099 & 0.72325 & 0.00407 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(3) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.56572 & 0.11669 & 0.22753 & 0.09005 \\ 0 & 0.57709 & 0.12096 & 0.20435 & 0.09759 \\ 0 & 0.07065 & 0.09581 & 0.82644 & 0.00711 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(4) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.47445 & 0.10631 & 0.30715 & 0.11208 \\ 0 & 0.48409 & 0.10929 & 0.28653 & 0.12007 \\ 0 & 0.06384 & 0.04919 & 0.87700 & 0.00996 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$P(5) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.39839	0.09289	0.37815	0.13057
S_2	0	0.40653	0.09515	0.35939	0.13893
S_3	0	0.05557	0.02793	0.90400	0.01249
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(6) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.33472	0.07961	0.43958	0.14609
S_2	0	0.34158	0.08139	0.42225	0.15478
S_3	0	0.04754	0.01772	0.92006	0.01468
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(7) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.28132	0.06757	0.49197	0.15914
S_2	0	0.28709	0.06903	0.47579	0.16809
S_3	0	0.03032	0.01245	0.93071	0.01654
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(8) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.23645	0.05709	0.53633	0.17011
S_2	0	0.24133	0.05829	0.52109	0.17929
S_3	0	0.03404	0.00939	0.93844	0.01812
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

Table 3

P(t) for Stage III

$$P(1) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0.01912 & 0.47107 & 0.12185 & 0.36975 & 0.01820 \\ 0 & 0.60809 & 0.20903 & 0.14721 & 0.03566 \\ 0 & 0 & 0.30629 & 0.69371 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(2) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0.00037 & 0.29546 & 0.13812 & 0.53070 & 0.03535 \\ 0 & 0.36978 & 0.19113 & 0.38174 & 0.05735 \\ 0 & 0 & 0.09381 & 0.90619 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(3) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0.00001 & 0.17984 & 0.10411 & 0.67015 & 0.04589 \\ 0 & 0.22486 & 0.13584 & 0.56877 & 0.07053 \\ 0 & 0 & 0.02873 & 0.97127 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(4) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.10936 & 0.06948 & 0.76885 & 0.05231 \\ 0 & 0.13674 & 0.08861 & 0.69610 & 0.07855 \\ 0 & 0 & 0.00881 & 0.99119 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$P(5) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.06650	0.04414	0.83315	0.05621
S_2	0	0.08315	0.05572	0.77770	0.08343
S_3	0	0	0.00269	0.99730	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(6) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.04044	0.02742	0.87356	0.05858
S_2	0	0.05056	0.03449	0.82859	0.08639
S_3	0	0	0.00083	0.99917	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(7) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.02459	0.01685	0.89854	0.06002
S_2	0	0.03075	0.02112	0.85994	0.08819
S_3	0	0	0.00025	0.99975	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(8) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.01495	0.01030	0.91385	0.06089
S_2	0	0.01869	0.01289	0.87911	0.08929
S_3	0	0	0.00008	0.99992	0
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

Table 4

P(t) for All Patients

$$P(1) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0.00029 & 0.77671 & 0.05464 & 0.14799 & 0.02036 \\ 0 & 0.84850 & 0.06543 & 0.05889 & 0.02717 \\ 0 & 0.07863 & 0.18076 & 0.73918 & 0.00144 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(2) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.66357 & 0.06072 & 0.23417 & 0.04155 \\ 0 & 0.72510 & 0.06735 & 0.15723 & 0.05032 \\ 0 & 0.08093 & 0.03782 & 0.87742 & 0.00383 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(3) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.56781 & 0.05439 & 0.31813 & 0.05967 \\ 0 & 0.62055 & 0.05962 & 0.24972 & 0.07012 \\ 0 & 0.07164 & 0.01213 & 0.91014 & -0.00608 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(4) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.48607 & 0.04699 & 0.39177 & 0.07518 \\ 0 & 0.53122 & 0.05138 & 0.33033 & 0.08707 \\ 0 & 0.06174 & 0.00688 & 0.92333 & 0.00805 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$P(5) \doteq$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.41612	0.04029	0.45513	0.08845
S_2	0	0.45478	0.04405	0.39959	0.10158
S_3	0	0.05293	0.00528	0.93205	0.00974
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(6) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.35625	0.03451	0.50942	0.09982
S_2	0	0.38935	0.03772	0.45894	0.11399
S_3	0	0.04533	0.00442	0.93907	0.01118
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(7) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.30499	0.02955	0.55591	0.10956
S_2	0	0.33333	0.03229	0.50975	0.12463
S_3	0	0.03881	0.00376	0.94501	0.01242
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(8) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.26111	0.02509	0.59572	0.11788
S_2	0	0.28537	0.02765	0.55325	0.13373
S_3	0	0.03322	0.00322	0.95008	0.01348
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

Table 5

P(t) for Age-group Less Than 40 Years

$$P(1) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0.00023 & 0.83894 & 0.04765 & 0.11317 \\ 0 & 0.90447 & 0.05675 & 0.03878 \\ 0 & 0.10275 & 0.24195 & 0.65529 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(2) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.76389 & 0.05915 & 0.17696 \\ 0 & 0.82389 & 0.06506 & 0.11104 \\ 0 & 0.11780 & 0.06437 & 0.81783 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(3) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.69699 & 0.05766 & 0.24535 \\ 0 & 0.75187 & 0.06249 & 0.18563 \\ 0 & 0.11316 & 0.02226 & 0.86458 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(4) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.63633 & 0.05351 & 0.31016 \\ 0 & 0.68647 & 0.05779 & 0.25574 \\ 0 & 0.10464 & 0.01181 & 0.88355 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(5) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.58104 & 0.04906 & 0.36990 \\ 0 & 0.62683 & 0.05294 & 0.32023 \\ 0 & 0.09586 & 0.00879 & 0.89535 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(6) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.53057 & 0.04485 & 0.42458 \\ 0 & 0.57238 & 0.04838 & 0.37923 \\ 0 & 0.57238 & 0.04838 & 0.37923 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(7) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.48449 & 0.04096 & 0.47455 \\ 0 & 0.52267 & 0.04419 & 0.43314 \\ 0 & 0.08081 & 0.00680 & 0.91319 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(8) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 0.44245 & 0.03741 & 0.52014 \\ 0 & 0.47728 & 0.04035 & 0.48237 \\ 0 & 0.07306 & 0.00618 & 0.92076 \\ 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

Table 6

P(t) for Age-group Between 40-60 Years

$$P(1) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0.00017 & 0.80879 & 0.08092 & 0.09944 & 0.01068 \\ 0 & 0.85041 & 0.09683 & 0.03918 & 0.01358 \\ 0 & 0.07717 & 0.41630 & 0.50591 & 0.00062 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(2) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.69419 & 0.11202 & 0.17208 & 0.02171 \\ 0 & 0.73067 & 0.12266 & 0.12149 & 0.02518 \\ 0 & 0.09775 & 0.18078 & 0.71954 & 0.00193 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(3) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.59899 & 0.11385 & 0.25595 & 0.03121 \\ 0 & 0.63084 & 0.12182 & 0.21217 & 0.03518 \\ 0 & 0.09708 & 0.08473 & 0.81483 & 0.00337 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(4) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.51817 & 0.10539 & 0.33702 & 0.03941 \\ 0 & 0.54587 & 0.11179 & 0.29851 & 0.04382 \\ 0 & 0.08909 & 0.04467 & 0.86149 & 0.00474 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$P(5) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.44879	0.09405	0.41064	0.04651
S_2	0	0.47284	0.09940	0.37646	0.05130
S_3	0	0.07921	0.02722	0.88759	0.00598
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(6) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.38892	0.08261	0.47581	0.05266
S_2	0	0.40978	0.08717	0.44527	0.05778
S_3	0	0.06947	0.01900	0.90446	0.00707
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(7) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.33712	0.07205	0.53284	0.05799
S_2	0	0.35521	0.07597	0.50542	0.06340
S_3	0	0.06054	0.01464	0.91679	0.00802
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

$P(8) =$

	S_1	S_2	S_3	S_4	S_5
S_1	0	0.29225	0.06264	0.58250	0.06261
S_2	0	0.30793	0.06602	0.55777	0.06827
S_3	0	0.05261	0.01196	0.92658	0.00886
S_4	0	0	0	1.00000	0
S_5	0	0	0	0	1.00000

Table 7

P(t) for Age-group Greater Than 60 Years

$$P(1) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0.00094 & 0.60871 & 0.09558 & 0.24149 & 0.05328 \\ 0 & 0.73377 & 0.13963 & 0.04468 & 0.08192 \\ 0 & 0 & 0.54609 & 0.45391 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(2) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.44723 & 0.13728 & 0.31229 & 0.10319 \\ 0 & 0.53843 & 0.17871 & 0.14085 & 0.14202 \\ 0 & 0 & 0.29822 & 0.70178 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(3) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.32817 & 0.13741 & 0.39459 & 0.13983 \\ 0 & 0.39508 & 0.17277 & 0.24602 & 0.18613 \\ 0 & 0 & 0.16285 & 0.83715 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(4) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.24080 & 0.12086 & 0.47163 & 0.16671 \\ 0 & 0.28990 & 0.14951 & 0.34209 & 0.21849 \\ 0 & 0 & 0.08893 & 0.91107 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(5) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.17669 & 0.09963 & 0.53725 & 0.18643 \\ 0 & 0.21272 & 0.12213 & 0.42291 & 0.24224 \\ 0 & 0 & 0.04856 & 0.95144 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(6) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.12965 & 0.07907 & 0.59036 & 0.20091 \\ 0 & 0.15609 & 0.09639 & 0.48785 & 0.25966 \\ 0 & 0 & 0.02652 & 0.97348 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(7) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.09514 & 0.06128 & 0.63205 & 0.21153 \\ 0 & 0.11454 & 0.07444 & 0.53858 & 0.27245 \\ 0 & 0 & 0.01448 & 0.98552 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

$$P(8) = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.06981 & 0.04675 & 0.66412 & 0.21932 \\ 0 & 0.08404 & 0.05664 & 0.57748 & 0.28183 \\ 0 & 0 & 0.00791 & 0.99209 & 0 \\ 0 & 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0 & 1.00000 \end{bmatrix} \end{array}$$

REFERENCES

Alling, D.W. (1958), The After-History of Pulmonary Tuberculosis, A stochastic Model, Biometrics 14, 527-547.

Anderson, T.W. and Goodman, L.A. (1957), Statistical Inference about Markov Chains, Ann. Math. Stat., 28, 89-110.

Bailey, N.T.J. (1957), The Mathematical Theory of Epidemics, Griffen, London.

Bailey, N.T.J. (1964), The Element of Stochastic Processes, John Wiley and Sons, Inc.

Berg, J.W. (1965), The Distribution of Cancer Deaths in Time, Brit. J. Cancer 19, 695-711.

Berkson, J. and Gage, R.P. (1952), Survival Curve for Cancer Patients Following Treatment, Journal of the American Statistical Association, vol. 47, 501-515.

Boaw, J.W. (1948), The Presentation and Analysis of the Results of Radiotherapy, Brit. J. Radiol. 21, 128-138.

Boag, J.W. (1949), Maximum Likelihood Estimates of the Proportion of Patients Cured by Cancer Therapy, Journal of the Royal Statistical Society, Vol. B 11, 15-53.

Campos, J.L. (1971), Mortality Trends in Carcinoma of the Cervix Uteri, J. Chron. Dis. 24, 701-709.

Cane, V. (1959), Behaviour Sequences as Semi-Markov Chains, J.R. Stat. Ser. B, 21, 36-58.

Chatfield, C. (1973), Statistical Inference Regarding Markov Chain Model, Applied Stat. 22, 7-20.

Chiang, C.L. (1968), Introduction to Stochastic Processes in Biostatistics, John Wiley & Sons, New York.

Chiang, C.L. (1971), A Stochastic Model of Human Fertility, Biometrics, 27, 345-356.

Chung, K.L. (1967), Markov Chains with Stationary Transition Probabilities, Springer-Verlag, New York.

Cutler, S.J. and Ederer, F. (1958), Maximum Utilization of the Life Table Method in Analyzing Survival, J. Chron. Dis., 8, 699-713.

Darling, D.A. and Kac, M. (1957), An Occupation Times for Markov Processes, Trans. Amer. Math. Soc., 84, 444-458.

Doob. J.L. (1953), Stochastic Processes, John Wiley and Sons, New York.

Falk, G. (1971), Average Length of Stay in Long-Term Institutions, Health Services Research, 6, 251-255.

Feller, W. (1957), An Introduction to Probability Theory and its Applications, Vol I & II, John Wiley and Sons, New York.

Fix, E. and Neyman, J. (1951), A Simple Stochastic Model of Recovery, Relapse, Death and Loss of Patients, Human Biology, 23, 205-241.

Fleiss, J.L. (1973), Statistical Methods for Rates and Proportion, John Wiley and Sons.

Fetter, R.B. and Thompson, J.D. (1969), A Decision Model for the Design & Operation of a Progressive Patient Care Hosp., Med. Care, 2, 450-62.

Friedell, G.Sherman, J. and Sommers, S. (1961), Growth Curves of Human Cancer Transplants During Exp. Chemotherapy, Cancer 12, 1117-21.

Gehan, E.A. (1969), Estimating Survival Function from the Life Table, J. Chr. Dis., Vol. 21, 629-644.

Gustafson, D.H. (1968), Length of Stay: Prediction and Explanation, Health Services Research 3, 12-13.

Hanson, B.L. (1973), A Statistical Model for Length of Stay in a Mental Hospital, Health Services Research, 8, 37-45.

Howard, R.A. (1971), Dynamic Prob. Systems, Vol 1 & 2, Wiley.

Jain, R.K. (1975), use of Canonical Transition Matrix in the study of Mortality among Cancer Patients, Contributed Paper at the 1975 Annual Meeting of Statistical Science Association of Canada.

Kac, M. (1957), A Class of Lim. Theo., Trans. Amer. Math., 84, 459-71.

Kaplan, H.S. (1962), The Radical Radiotherapy of Localized Hodgkin's Disease, Radiology, 78, 553-61.

Karlin, S. (1969), A First Course in Stoch. Proc., Academic Press.

Kemeny, J.G. And Snell, L. (1960), Finite Markov Chain, Van Nostrand.

Kesten, H. (1962), Occupation Times for Markov And Semi-Markov Chains, Trans. Amer. Math. Soc., 103, 82-112.

Kolmogorov, A.N. (1950), Foundations of the Theory of Probability, Chelsa Press, New York.

Kurtz, T.G. (1971), Comparison of Semi-Markov and Markov Processes, The Annal of Mathematical Statistics, Vol. 42, No. 3, 991-1002.

Levy, P. (1954), Systems Semi-Markoviens a plus une infinite denombrable d'etats possibles, Proc. Internat. Congr. Math., Amsterdam, Vol. 2, 294; Processus Semi-Markoviens, ibid., Vol. 3. 416-426.

Lindell, A. (1952), Carcinoma of the Uterine Cervix: Incidence and Influence of Age, Acta. Radiol. Supp. 92.

Marshall, A.W. and Goldhamer, H. (1955), An Application of Markov Processes to the study of the Epidemiology of Mental Disease, J. Amer. Stat. Assoc., 50, 99-129.

Meier, P. (1955), Note on Estimation in a Markov Process with constant Transition Rates, Human Biology, 27, 121-124.

Myers, M.H., Axtell, L.M. and Zelen, M. (1966), The Use of Prognostic Factors in Predicting Survival for Breast Cancer Patients, J. Chron. Dis. 19, 923-933.

Navarro. V. (1969), Planning Personal Health Services: A Markovian Model, Medical Care, Vol. VII, No. 3, 242-249.

Nolan. J.F. and Dusault, L. (1951), Radiation Treatment of Cancer of the Cervix Uteri, Amer. J. Roentgenol 65, 709-714.

Parzen, E.B. (1962), Stochastic Processes, Holden-Day.

Perrin, E.B. and Sheps, M.C. (1964), Human Reproduction: A Stochastic Process, Biometrics, 20, 28-45.

Pyke, R. (1958), On Renewal Processes Related to Type I and Type II Counter Models, Ann. Math. Stat. 29, 737-754.

Pyke, R. (1961), Markov Renewal Processes: Definitions and Preliminary Properties, Ann. Math. Stat., 32, 1231- 1242.

Pyke, R. (1961), Markov Renewal Processes with Finitely many States, Ann. Math. Stat., 32, 1243-1259.

Ross, S.M. (1970), Applied Probability Models with Optimization Applications, San Franciso: Holden-Day.

Sheps, M.C. and Perrin, E.B. (1966), Further Results from a Human Fertility Model with a Variety of Pregnancy Outcomes, Human Biology, 38, 180-193.

Sheps, M.C. and Menken, J.A. (1971), A Model for Studying Birth Rates Given Time Dependent Changes in Reproduction Parameters, Biometrics, 27, 325-343.

Smith, W.L. (1955), Regenerative Stochastic Processes, Proc. Roy. Soc., Ser. A, Vol. 232, 6-31.

Sorensen, B. (1958), Late Results of Radium Therapy in Cervical Carcinoma, Acta. Radiol. Supp., 169.

Staff, P.J. and Vagholkar, M.K. (1971), Stationary Distributions of Open Markov Processes in Discrete Time with Application to Hospital Planning, J. Appl. Prob., 8, 668-680.

Sverdrup, E. (1965), Estimates and Test Procedures in Connection with Stochastic Models of Deaths, Recoveries and Transfers Between Different States of Health, Skand. Aktuar.,46, 184-211.

Watterson, G. (1961), Markov Chains with Absorbing States: A Genetic Example, Ann. Math Stat., 33, 939-957.

Zahl, S. (1955), A Markov Process Model for Follow-up Studies, Human Biology, 27, 90-120.