Chapter 5

OPTIMAL TREATMENT STRATEGY USING SEMI-MARKOV DECISION PROCESS
5.1 Introduction

In some medical treatment, decision must be made sequential and in an uncertain environment. A physician determining a course of treatment must consider patient’s health as well as the best treatment decision in the future. Often decisions are to be taken in a dynamic environment. Physiological as well as physical changes in patients, may sometime contribute to the changes of the environment. Uncertain environment arises mainly due to patients respond differently even to same treatment for a disease.

Physicians always need to make subjective judgement about the treatment strategies. However a mathematical decision model that provide insight into the nature of optimal decision can aid the treatment. Markov decision processes (MDPs)are appropriate technique useful in class of problems involving complex, stochastic and dynamic decisions for which it can find optimal solutions. The goal of a MDP is to provide a optimal policy which is a decision strategy to optimize a particular criterion such as maximizing total discounted reward or minimizing the total discounted cost.

MDPs are a general framework for modeling dynamic systems under uncertainty. It binds previous, current and future treatment decision through the proper definition of patient’s states defined as variables that contain the relevant information for making future decisions. The treatment model evolves in the following manner. The condition or state of patient is observed (or partially observed), an action is taken, a cost is incurred (or a reward is received) and the patient get into a new state according to a known probability distribution. The state variable defined so that given current state
of patient, the future transitions and rewards are independent of the past. It is the standard assumption of a Markov Process

The broad classes of MDPs are Finite Horizon MDPs and Infinite Horizon MDPs. The number \( N \) of decision epochs is finite in the former and it goes to infinite in the latter. For a finite horizon model optimal policy for both the average reward per state and the total reward criterion are equivalent. Infinite horizon models requires a large amount of data hence it is assumed that the data are time homogeneous. As a result the states of infinite horizon MDP must be carefully defined to ensure that the state transition of patients are stationary. When the data are time dependent the time homogeneity assumption can be satisfied by properly augmenting the state definition with the time at which the transition occurs

In most of the medical investigation, state of a patients is decided in the light of a series of medical tests which are subjected to test errors. A modified MDPs called Partially Observed MDPS (POMDP) have been developed to deal the data with imperfect information (Lovejoy (1991), White et al.(1989)) In these models it is assumed that uncertainty exist, in patient’s transition and the state he/she truly occupies. Therefore the objective is to find an optimal policy based on the observation of the patient and the previous decision rule applied.

In MDP models the treatment decision are taken at each of a sequence of unit time intervals or fixed epochs and the sojourn time in states has no effect on rewards or incurring costs for patient. However in health-care and other application, decision are
taken over continuous time intervals such as varying treatment can be administered. The sojourn time in states may depend on the duration of his/her current health status. The MDP models might not be suitable to model such disease progression instead Semi-Markov Decision Process (SMDP) models are more appropriate. In SMDP models allow patients’ state transition to occur in continuous time and allow to assume any probability distribution for sojourn time in a state.

For a range of multi-state diseases in which the data arises as transition-times and states, Markov and Semi-Markov process are appropriate models. Historically it has been difficult to adopt realistic models for biomedical applications since the likelihood turns out to be prohibitively complicated. With the development of computational methods and aids, the problem of tractability can be overcome. Kay(1986) introduced a k-state Markov model for continuous time processes for analyzing cancer markers. Similar models have been applied to AIDS (Longini et al(1989)), heart transplantation (Sharples(1993)), diabetes (Anderson(1988), Marshall and Jones (1995)), infectious diseases , dementia etc. Extension to these models have been further extended to include fixed time and time-varying covariate information (see for eg: Anderson et al (1991), Gauvreau et al(1994)). Methods for estimation of transition rates are generally numerically based and have usual maximum likelihood sampling schemes such as Metroplois-Hasting method(Sharples (1993), Prevost et al.(1998), Richard et al.(1993), or have used population-based approaches akin to weighted least squares( see Chen et al. 1996). These approaches have all concentrated on continuous-time Markov processes usually due to unequally spaced observation times.
In this chapter, we consider a complex survival model that lives in a randomly changing environment which affect model parameters. The term ‘environment’ is used in the generic sense so that it represents any set of conditions that affect the stochastic structure of the model investigated. The concept of ‘environment’ process in one form or another, has been used in the literature for various purpose.

The use of environmental process to modulate the deterministic and stochastic parameters of Operation Research models can be seen in reliability, inventory and queueing applications. One may refer to Ozekici and Soyer(2003) for an expository coverage in reliability theory. The problem of optimal replacement of a semi-Markov system under semi-Markov environment is studied by Hu and Yue(2003). Ozekici(1996), Ozekici and Parler(1999) discuss other applications in inventory and queueing. A comprehensive discussion on Markov modulated queueing system can be found in Prabu and Zhu(1988).

Although the literature cited above illustrate the use of random environment in reliability, inventory and queueing model, the concept is of paramount interest in survival analysis. It is generally assumed that a patient stays in a given fixed environment. The probability law of his ageing and death process there remain intact throughout his useful life. The life duration and corresponding hazard rate is taken to be the one obtained through statistical life testing procedures that are believed to be under ideal conditions. There has been growing interest in the recent years in lifetime models under random environment.
This is necessitated by the fact that the subject/patient often lives in varying environments during which they are subjected to varying environment conditions with significant effects on performance/health status. During a treatment period whole environment of the patient may change due to occurrence of other contagious diseases, hypertension, high blood pressure, cardiac problems, severe climatic/seasonal changes or adopting entirely new treatment strategy on medical team’s advice. When environment changes, the state of patient also changes. The deterioration and failure process therefore depends on the environment. This makes it crucial to identify an optimal treatment strategy especially for a range of multi-state disease processes.

The remainder of this chapter is organized as follows. In section 2 prognosis of three diseases are discussed. In section 3 the model is presented in the frame of MDP. Section 4 deals with the problem of optimal control limit policies. Section 5 addresses a special case of Markov environment in which computationally feasible solution is arrived at. A numerical example is provided in the next section to illustrated the methodology. Followed by a discussion in the final section.

To begin with we shall examine some situations where we observe the progression of disease through stages.
5.2 Prognosis of Some Diseases

Alzheimer’s Disease

It is a progressive, degenerative disease that destroys vital brain cells. As each area of the brain is affected, certain functions or abilities can be lost. The losses affect the individual’s ability to think, to remember, to understand and to make decisions. In addition to affecting a person’s mental abilities, Alzheimer’s disease affects moods and emotions. Along with loss of abilities, changes in behaviour occur. Gradually, independence disappears. The progression of Alzheimer’s disease varies from person to person and can span three to twenty years (the average length of the disease is between eight and twelve years). The progression can be described as a series of stages, providing a guide to the pattern of the disease, which can help when making care decisions. One staging system explains the disease in three stages: early, middle and late. Another staging system, often used by medical professionals, is the Global Deterioration Scale (also called the Reisberg Scale). This scale divides the disease into seven stages.

SmallPox

Highly contagious disease Smallpox was the most feared epidemic for centuries. Now it is eradicated from the most part of the globe by vaccination. Progression of smallpox has several medically well defined stages of disease.
<table>
<thead>
<tr>
<th>Stages</th>
<th>Duration</th>
<th>Contagious?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Incubation period</td>
<td>7 to 17 days</td>
<td>Not Contagious</td>
</tr>
<tr>
<td>2. Initial symptoms (Prodome)</td>
<td>2 to 4 days</td>
<td>Possibly contagious</td>
</tr>
<tr>
<td>3. Early rash</td>
<td>About 4 days</td>
<td>Highly contagious</td>
</tr>
<tr>
<td>4. Pustular rash</td>
<td>About 5 days</td>
<td>Contagious</td>
</tr>
<tr>
<td>5. Pustules and scabs</td>
<td>About 5 days</td>
<td>Contagious</td>
</tr>
<tr>
<td>6. Resolving scabs</td>
<td>About 6 days</td>
<td>Contagious</td>
</tr>
<tr>
<td>7. Scabs resolved</td>
<td></td>
<td>Not contagious</td>
</tr>
</tbody>
</table>

**Incubation period:** Exposure to the virus is followed by an incubation period during which people do not have any symptoms and they may feel fine. The average incubation period is 12 to 14 days after exposure to the virus but it can range from 7 to 17 days. At this time the infected person is not contagious.

**Initial Symptoms - Prodome:** The first symptoms of smallpox include: Fever (38 degrees Celcius), malaise, head and body aches, and sometimes vomiting. At this time people are normally prostrate.

**Early rash:** The smallpox rash has a characteristic centrifugal distribution. The rash emerges first as small red spots on the tongue and in the mouth. These develop into sores and then the sores break open which releases large amounts of the virus into the mouth and throat. When this happens the person becomes contagious. About the same time that the sores break open a rash appears on the skin. It starts on the face, spreads to the arms and legs and then to the hands and feet. The rash will usually spread to all parts of the body within 24 hours. As the rash appears the fever may fall and the person may feel a bit better. The rash becomes raised bumps by the third day of the rash. By the forth day the bumps fill with a thick opaque fluid and often have a depression in the centre. The fever will often increase again at
this time and may remain high until scabs have formed over the bumps.

**Pustular rash:** The bumps become pustules, which are sharply raised and usually round and firm to the touch.

**Pustules and scabs:** The pustules begin to form a crust and then a scab. By the end of the second week after the rash has appeared most of the sores will have scabbed over.

**Resolving Scabs:** The scabs will begin to fall off but will often leave marks on the skin that will become pitted scars. Most of the scabs will have fallen off three weeks after the rash first appeared. The person is contagious to others until all of the scabs have fallen off. Scabs Resolved: Once the scabs have fallen off the person is no longer contagious.

**Liver Disease**

The prognosis of liver disease has various stages namely, Inflammation, Fibrosis, Cirrhosis, Liver cancer and Liver failure. Healthy liver helps fight infections and cleans our blood. It also helps digest food and stores energy for when we need it. It has the amazing ability to grow back, or regenerate, when it is damaged. Anything that keeps the liver from doing its job - or from growing back after injury - may put our life in danger. Whether the liver is infected with a virus, injured by chemicals, or under attack from your own immune system, the basic danger is the same - that
the liver will become so damaged that it can no longer function properly.

**Stage 1 . Inflammation:** In the early stage of any liver disease, liver may become inflamed. It may become tender and enlarged. Inflammation shows that the body is trying to fight an infection or heal an injury. But if the inflammation continues over time, it can start to hurt liver permanently. When most other parts of our body become inflamed, we can feel it - the area becomes hot and painful. But an inflamed liver may cause no discomfort at all. If the liver disease is diagnosed and treated successfully at this stage, the inflammation may go away.

**Stage 2 . Fibrosis:** If left untreated, the inflamed liver will start to scar. As excess scar tissue grows, it replaces healthy liver tissue. This process is called fibrosis. (Scar tissue is a kind of fibrous tissue.) Scar tissue cannot do the work that healthy liver tissue can. Moreover, scar tissue can keep blood from flowing through the liver. As more scar tissue builds up, liver may not work as well as it once did. Or, the healthy part of liver has to work harder to make up for the scarred part. If the liver disease is diagnosed and treated successfully at this stage, there’s still a chance that the liver can heal itself over time.

**Stage 3 . Cirrhosis:** But if left untreated, liver may become so seriously scarred that it can no longer heal itself. This stage - when the damage cannot be reversed - is called cirrhosis. Once one been diagnosed with cirrhosis, treatment will focus on keeping his/her condition from getting worse. It may be possible to stop or slow the liver damage. It is important to protect the healthy liver tissue that have
left. Cirrhosis can lead to a number of complications, including liver cancer. In some people, the symptoms of cirrhosis may be the first signs of liver disease.

**Stage 4 . Liver cancer:** Cancer that starts in the liver is called primary liver cancer. Cirrhosis and hepatitis B are leading risk factors for primary liver cancer. But cancer can develop in the liver at any stage in the progression of liver disease.

**Stage 5 . Liver failure** Liver failure means that your liver is losing or has lost all of its function. It is a life-threatening condition that demands urgent medical care. The first symptoms of liver failure are often nausea, loss of appetite, fatigue, and diarrhea. Because these symptoms can have any number of causes, it may be hard to tell that the liver is failing. But as liver failure progresses, the symptoms become more serious. The patient may become confused and disoriented, and extremely sleep. There is a risk of coma and death. Immediate treatment is needed. The medical team will try to save whatever part of the liver that still works. If this is not possible, the only option may be a liver transplant. When liver failure occurs as a result of cirrhosis, it usually means that the liver has been failing gradually for some time, possibly for years. This is called chronic liver failure. Chronic liver failure can also be caused by malnutrition. More rarely, liver failure can occur suddenly, in as little as 48 hours. This is called acute liver failure and is usually a reaction to poisoning or a medication overdose.
5.3 Semi-Markov Model for Treatment Strategy

we shall state the general form of the model that represents the foregoing situation may be stated as follows:

1. The patient is in a semi-Markov environment \(\{(J_n, L_n), n \geq 0\}\) on a set \(K\) of countable environment states, where \(J_n\) is the state of the environment immediately after its \(n\)th transition epoch \(T_n\), and \(0 = T_0 < T_1 < T_2 < \cdots\). \(L_n\) is the time duration of the patient in the state \(J_n\). Let the state’s kernel be

\[
G_{kk}(t) = Pr(T_{n-1} - T_n \leq t, J_{n+1} = k'/J_n = k)
\]

and let \(\psi_{kk'} = G_{kk'}(\infty)\) and \(G_k(t) = \sum_{k' \in K} G_{kk'}(t)\).

2. During an environment state \(k\), the patient goes through several states of disease according to a semi-Markov process with a kernel \(\{P_{ij}(t), i, j \in S\}\) and a set \(S = \{0, 1, 2, \ldots\}\) of countable states, where the state 0 represents disease-free state, and states 1, 2, \ldots represent the different adverse disease states of the patient and the bigger the value, the more serious is the condition.

Let \(P^k_{ij}(\infty)\), \(T^k_{ij}(t) = P^k_{ij}(t)/P_{ij}^k\) and \(T^k_i(t) = \sum_{j \in S} P^k_{ij}(t)\).

3. Suppose that the patient is in environment \(k\), then one of the following two actions can be chosen if his state transfers to \(i\):

(a) Continue the present treatment strategy (denoted by \(C\)) with a cost rate \(h^k(i)\);

(b) Initiating a rejuvenating treatment strategy (denoted by \(R\)) like chemother-
apy, radiation, surgery, organ transplant and/or admission in ICU etc., with a cost rate \( c^k(i) \), and the time of action \( R \) is assumed to be a random variable with probability distribution function \( F^k(t) \), and the state after rejuvenation will be 0, the disease-free state.

4. When the environment state changes from \( k \) to \( k' \) if action \( C \) or \( R \) is chosen then the patient’s state will change immediately according to a probability \( q_{ij}^k \) and an instantaneous cost \( R^k(i, C) \) occurs; while if action \( R \) is chosen then within no time it completed and an instantaneous cost \( R^k(i, R) \) occurs.

5. The objective is to minimize the expected discounted total costs with discount factor \( \alpha > 0 \).

The above treatment strategy can be modeled by a semi-Markov decision process (SMDP) in a semi- Markov environment, presented and studied by Hu (1997), as follows.

During the environment state \( k \), i.e., \( J_n = k \) for some \( n \geq 0 \), it can be modeled by the following SMDP:

\[
SMDP_k := \{ S, A, p^k(j|i, a), T^k(.|i, a, j), r^k(i, a, j, u) \}
\]

where \( S \) is the state space and \( A = \{ C, R \} \) is the action set. The transition probability \( p^k \), the distribution function \( T^k \) of the transition time, and the one step cost function
are given, respectively by
\[ P^k(j|i, C) = P^k_{ij}, \quad P^k(j|i, R) = \delta_{j0} \]
\[ T^k(t|i, C, j) = T^k_{ij}(t), \quad T^k(t|i, R, 0) = F^k(t) \]
\[ r^k(i, C, j, u) = h^k(i) \int_0^u e^{-\alpha t} dt = h^k(i) \alpha^{-1}(1 - e^{-\alpha t}) \]
\[ r^k(i, R, j, u) = C^k(i) \int_0^u e^{-\alpha t} dt = C^k(i) \alpha^{-1}(1 - e^{-\alpha t}) \] (5.3.2)
\[ \delta_{j0} = \begin{cases} 1 & \text{if } j = 0, \\ 0 & \text{otherwise} \end{cases} \]

For SMDP in a semi-Markov environment, when the environment state changes from \( k \) to \( k' \) i.e., at \( L_{n+1} \) for some \( n \geq 0 \) with \( J_n = k \) and \( J_{n+1} = k' \), the patient’s state changes immediately to \( j \) with a probability \( q(j|i, a, k, k') \) if the patient’s state is \( i \) at \( L_{n+1} - 0 \) and the last action taken before \( L_{n+1} \) is ’a’, and in the same time, an instantaneous cost \( R^k(i, a) \) occurs, where
\[ q(j|i, C, k, k') = q^k_{ij} \]
\[ q(j|i, R, k, k') = \delta_{j0} \]

To simplify notations, for \( k \in K \) and \( s, t \geq 0 \), we let
\[ h^k(s, t) = \sum_{k'} Pr(T_{n+1} - T_n > t, J_{n+1} = k' / J_n = k) \int_0^t e^{-\alpha u} du \]
\[ + \int_{s+}^{s+t} \sum_k Pr(T_{n+1} - T_n \leq u, J_{n+1} = k' / J_n = k) \int_0^{u-s} e^{-\alpha l} dl \]
\[ = \alpha^{-1}(1 - e^{-\alpha t})[1 - G_k(t + s)] + \alpha^{-1} \int_{s+}^{s+t} (1 - e^{-\alpha(u-s)}) dG_k(u), \]
\[ g^{k,k'}(s, t) = \int_{s+}^{s+t} e^{-\alpha(u-s)} dG_k(u), \] (5.3.3)
\[ g^k(s, t) = \sum_{k' \in K} g^{k,k'}(s, t) = \int_{s+}^{s+t} e^{-\alpha(u-s)} dG_k(u), \]
Let \( x = (k, s, i) \in \Omega = \{(k, s, i) : k \geq 0, s \geq 0, i \in S\} \) be the mathematical state which means that the environment is in state \( k \) just since time \( s \) ago and the patient’s state just transfers to \( i \). Then, we define Expected discounted cost occurring when the state \( x \) is reached and action \( 'a' \) is taken, the \( r(x,a) \)

\[
r(x, C) = h^k(i) \int_0^\infty h^k(s, t) \, dT^k_i(t) + R^k(i, C) \int_0^\infty g^k(s, t) \, dT^k_i(t) \\
r(x, R) = c^k(i) \int_0^\infty h^k(s, t) \, dF^k(t) + R^k(i, R) \int_0^\infty g^k(s, t) \, dF^k(t) \tag{5.3.4}
\]

and \( \beta(x, a, k') \) is corresponding to a discount factor depending on the state \( x \), when the action is \( 'a' \) and the next environment state \( k' \).

\[
\beta(x, C, k') = \int_0^\infty g^{k'}(s, t) \, dT^k_i(t) \\
\beta(x, R, k') = \int_0^\infty g^{k'}(s, t) \, dF^k(t)
\]

Now, it follows that \( V^*(x) \), the minimal expected discounted total cost starting from the initial state \( x \), is the minimal nonnegative solution of the following optimality equation

\[
V^*(x) = \min \{V^*(x, C), \ V^*(x, R)\} \tag{5.3.5}
\]

where, \( x = (k, s, i) \in \Omega \)

\[
V^*(x, C) = r(x, C) + \sum_{k' \in K} \beta(x, C, k') \sum_{j \in S} q_{ij}^k V^*(k', 0, j) \\
+ \sum_{j \in S} p_{ij}^k \int_0^\infty e^{-\alpha t} V^*(k, s + t, j) \, dT^k_{ij}(t) \tag{5.3.6}
\]

\[
V^*(x, R) = r(x, R) + \sum_{k' \in K} \beta(x, R, k') V^*(k', 0, 0) \\
+ \int_0^\infty e^{-\alpha t} V^*(k, s + t, 0) \, dF^k(t)
\]
are respectively the discounted total cost if action C or R is used in the first horizon with the mathematical state x and then the optimal policy is used in the remaining horizons.

### 5.4 Optimal Control Limit Policies

From the standard results in discrete time Markov decision processes DTMDP, Eqn. (5.3.5) can be considered as an optimality equation for an adequate DTMDP with state space Ω. Thus we can consider its n-horizon problem with the optimality equation

\[
V^*_n(x) = \min\{V^*_n(x, C), V^*_n(x, R)\}, \text{ where } x = (k, s, i) \in \Omega  \tag{5.4.1}
\]

where \(V^*_n(x)\) is the optimal value from state x for n horizons problem, while

\[
\begin{align*}
V^*_n(x, C) &= r(x, C) + \sum_{k \in K} \beta(x, C, k') \sum_{j \in S} q_{ij}^k V^*_{n-1}(k', 0, j) \\
&\quad + \sum_{j \in S} P_{ij} \int_0^\infty e^{-\alpha t} V^*_{n-1}(k, s + t, j) \, dT_{ij}^k(t) \\
V^*_n(x, R) &= r(x, R) + \sum_{k \in K} \beta(x, R, k') V^*_{n-1}(k', 0, 0) \\
&\quad + \int_0^\infty e^{-\alpha t} V^*_{n-1}(k, s + t, 0) \, dF^k(t)
\end{align*}
\tag{5.4.2}
\]

are the values from state x in n horizons if action C or R is used respectively in the first horizon and then an optimal policy in the remaining horizons. The initial conditions are

\[
V^*_0(x, C) = V^*_0(x, R) = 0
\]

131
Let \( v_n(x) = V^*_n(x, C) - V^*_n(x, R), \quad v(x) = V^*(x, C) - V^*(x, R) \) \}

\( x = (k, s, i) \in \Omega \)

then it follows from the standard theory of DTMDP that

\[
\lim_{n \to \infty} V^*_n(x, a) = V^*(x, a), \quad a = C, R
\]

(5.4.3)

\[
\lim_{n \to \infty} v_n(x) = v(x)
\]

while the optimal policies can be depicted as \( f^*_n(x) = C \iff v_n(x) < 0, f^*(x) = C \iff v(x) < 0 \). So, \((f^*_N, f^*_N, \ldots, f^*_0)\) is optimal for N-horizon problem; and \( f^* \) is optimal for the infinite horizon discounted criterion. A concept of stochastic order between two distribution functions is needed. For two distribution functions F and G, F is said to be smaller stochastically than G, denoted by \( F \preceq G \), if \( F(t) \geq G(t) \) for each \( t \)

We have the following familiar result on stochastic order.

For two distribution functions F and G, \( F \preceq G \) if and only if \( \int_{-\infty}^{\infty} f(t) dF(t) \leq \int_{-\infty}^{\infty} f(t) dG(t) \) for each nondecreasing function \( f \).

To obtain some properties of the optimal policies, we introduce the following assumption.

ASSUMPTIONS A

For each \( k \in K \),

(A.1) \( \sum_{j=m}^{\infty} q^k_{ij} \) nondecreasing in \( i \) for each \( m \geq 0 \);

(A.2) \( h^k(i), c^k(i), R^k(i, C) \) and \( R^k(i, R) \) are all nonnegative and nondecreasing in \( i \) ;
(A.3) both \( h^k(i) - c^k(i) \) and \( R^k(i, C) - R^k(i, R) \) are nondecreasing in \( i \);

(A.4) \( F \preceq T^k_0 \preceq T^k_1 \preceq T^k_2 \ldots \), i.e., \( T^k_i \) is stochastically nondecreasing in \( i \) and \( F^k(.) \) is the smallest;

(A.5) \( \int_0^\infty e^{-at} \sum_{j \in S} V(t, j)p^k_{ij}dT^k_{ij}(t) \) is nondecreasing in \( i \) if \( V(t, j) \) is nonnegative and nondecreasing in \( j \) for each \( t \geq 0 \).

As usual, Assumption (A.1) that \( \sum_{j=m}^\infty q^k_{ij} \) is nondecreasing in \( i \) for \( m > 0 \) means that more serious is the condition of patient, faster the critical state reached by the environment change. Assumption (A.3) that \( h^k(i) - c^k(i) \) is nondecreasing in \( i \) indicates that the treatment cost increases faster than the rejuvenating cost as the increasing of serious condition of patient, and similar for \( R^k(i, C) - R^k(i, R) \). In fact, Assumptions A.1), (A.2) and (A.3) are those in the literature for the discrete time model, while Assumption (A.4) is given for the continuous time case here. Assumption (A.4) means that the sojourn time in a state for patient is nondecreasing as seriousness of his condition increases, and the rejuvenating time is smaller than the sojourn time in any state. Assumption (A.5) follows that if \( T^k_{ij}(t) \) is absolutely continuous with probability density function \( t^k_{ij}(t) \), and \( \sum_{j=m}^\infty p^k_{ij}t^k_{ij}(t) \) is nondecreasing in \( i \) for each \( t \geq 0, m \geq 0 \), which is similar as (A.1). It is easy to see that the latter two conditions are involved in the state definition.

By the earlier stated result on stochastic order and Assumption A, it is easy to see that the \( g_{kk'}(t) \) is nondecreasing in \( i \) which implies that \( \beta(x, C, k') \) is nondecreasing in \( i \).
We have the following well known result on transition probabilities

Let \( r_{ij} \) be a transition probability matrix, then the following two are equivalent:

(1) For each \( m \geq 0 \), \( \sum_{j=m}^{\infty} r_{ij} \) is nondecreasing in \( i \);

(2) For each nonnegative and nondecreasing function \( h(j) \), \( \sum_{j=0}^{\infty} r_{ij} h(j) \) is nondecreasing in \( i \).

Then by using the induction method it can be shown from the above result and Assumption A that all of \( V_n^*(x, C) \), \( V_n^*(x, R) \) and \( V^*(x) \) are nondecreasing in \( i \).

Now for each \( k \in K \) and \( s \geq 0 \)

\[
\begin{align*}
  r(x, C) - r(x, R) &= h^k(i) \int_0^\infty h^k(s, t) \left[ dT^k_i(t) - dF^k(t) \right] \\
  &\quad + [h^k(i) - c^k(i)] \int_0^\infty h^k(s, t) \ dF^k(t) \\
  &\quad + R^k(i, C) \int_0^\infty g^k(s, t) \left[ dT^k_i(t) - dF^k(t) \right] \\
  &\quad + [R^k(i, C) - R^k(i, R)] \int_0^\infty g^k(s, t) \ dF^k(t)
\end{align*}
\]

is also nondecreasing in \( i \) due to Assumption A, result on stochastic order and the fact that both \( h^k(s, t) \) and \( g^k(s, t) \) are nondecreasing in \( t \) for each \( k \in K, s \geq 0 \). It should be noted that the latter two terms in \( V_n^*(x, R) \) of Eqn. (5.4.2) are independent of \( i \). So \( v_n(x) \) is nondecreasing in \( i \) and thus the following result

**Theorem 5.4.1** Under Assumption A, both \( V_n^*(k, s, i) \) and \( v_n(k, s, i) \) are nondecreas-
ing in i for each n ≥ 0, k ∈ K, s ≥ 0, so

\[ V^*_n(k, s, i) = V^*_n(k, s, i, c), 0 \leq i \leq i^*_n(k, s) \]

\[ = V^*_n(k, s, i, R), i \geq i^*_n(k, s) \]  \hspace{1cm} (5.4.4)

where \( i^*_n(k, s) := \min(i|v_n(k, s, i) \geq 0) \).

Similarly, both \( V^*(k, s, i) \) and \( v(k, s, i) \) are also nondecreasing in i and

\[ V^*(k, s, i) = V^*(k, s, i, c), 0 \leq i \leq i^*(k, s) \]

\[ = V^*(k, s, i, R), i \geq i^*(k, s) \]  \hspace{1cm} (5.4.5)

where \( i^*(k, s) := \min(i|v(k, s, i) \geq 0) \).

The above Theorem states that there exists a state limit \( i^*(k, s) \) just since time s ago for each \( k \in K \) and \( s \geq 0 \) such that if the patient enters a state \( i \) while the environment is in state \( k \), then the optimal action is to replace the patient by a new one if and only if the deteriorative degree of the patient is over the limit \( i^*(k, s) \), i.e., \( i \leq i^*(k, s) \). Such a policy is called a control limit policy. So the Theorem shows that there exists optimal control limit policies for both finite and infinite-horizon problems. In the next section, we will discuss a special case of Markov environment and get some better results.
5.5 Markov Environment-A Special Case.

In this section, we consider that the environment is Markov as follows:

\[ G_{kk'}(t) = \psi_{kk'} G_k(t), \quad G_k(t) = 1 - e^{-\lambda_k t}, t \geq 0, k \text{ and } k' \in K. \quad (5.5.1) \]

In this case, it will be shown that the variable \( s \) in state \( x = (k, s, i) \) can be deleted.

Let

\[ t_k^F = \int_0^\infty [1 - e^{-(\lambda_k + \alpha) t}] dF_k(t) \]
\[ t_{ij}^k = \int_0^\infty [1 - e^{-(\lambda_k + \alpha) t}] dT_{ij}^k(t) \]
\[ t_i^k = \sum_{j \in S} P_{ij}^k \lambda_k = 1 - t_i^F, \quad \alpha_{ij}^k = 1 - t_{ij}^k, \quad \alpha_i^k = 1 - t_i^k \quad (5.5.2) \]

where \( F_k(t) \) and \( T_{ij}^k(t) \) are defined in Section 1. Then it can be calculated, due to Eqn. (5.3.4), that

\[ r(x, C) = r'(k, i, C) e^{-(\lambda_k s)} = \frac{t_i^F}{\lambda_k + \alpha} [h_i^k(i) + \lambda_k R_i^k(i, C)] e^{-(\lambda_k s)} \]
\[ r(x, R) = r'(k, i, R) e^{-(\lambda_k s)} = \frac{t_i^F}{\lambda_k + \alpha} [c_i^k(i) + \lambda_k R_i^k(i, R)] e^{-(\lambda_k s)} \quad (5.5.3) \]
\[ \beta(x, C, k') = \frac{\lambda_k t_i^F}{\lambda_k + \alpha} \psi_{kk'} e^{-\lambda_k s} \]
\[ \beta(x, R, k') = \frac{\lambda_k t_i^F}{\lambda_k + \alpha} \psi_{kk'} e^{-\lambda_k s} \]

Based on Eqn. (5.5.3), it can be shown that \( e^{\lambda_k s} V^*(k, s, i) \) and therefore \( e^{\lambda_k s} V^*(k, s, i, C) \), \( e^{\lambda_k s} V^*(k, s, i, R) \) are independent of \( s \) and thus

\[ e^{\lambda_k s} V^*(k, s, i) = V^*(k, 0, i) \]
\[ e^{\lambda_k s} V^*(k, s, i, C) = V^*(k, 0, i, C) \]
\[ e^{\lambda_k s} V^*(k, s, i, R) = V^*(k, 0, i, R) \]
We denote by
\[ V^*(k, i) := V^*(k, 0, i), \quad V^*(k, i, C) := V^*(k, 0, i, C), \quad V^*(k, i, R) := V^*(k, 0, i, R) \]
and
\[ v(k, i) = V^*(k, i, C) - V^*(k, i, R) \]

Then \( V^*(k, i) \) is the minimal nonnegative solution of the following optimality equation
\[ V^*(k, i) = \min \{ V^*(k, i, C), V^*(k, i, R) \} \] (5.5.4)
with corresponding
\[ V^*(k, i, C) = r'(k, i, C) + \frac{\lambda_k t_k^i}{\lambda_k + \alpha} \sum_{k' \in K} \sum_{j \in S} q_{k'}^j V^*(k', j) + \sum_{j \in S} p_0^j \alpha_{ij}^k V^*(k, j) \]
\[ V^*(k, i, R) = r'(k, i, R) + \frac{\lambda_k t_k^i}{\lambda_k + \alpha} \sum_{k' \in K} \psi_{kk'} V^*(k', 0) + \alpha_k^i V^*(k, 0) \] (5.5.5)

Now, the problem is simplified by deleting the time variable \( s \), and we can solve for \( V^*(k, i) \) only. From the standard results in DTMDP, Eqn. (5.5.4) can also be considered as the optimality equation of an adequate defined DTMDP with state space \( S' = \{(k, i) : k \in K, i \in S\} \) and action set \( A = \{C, R\} \).

In the case of Markov environment, Assumptions (A.4) and (A.5) can be replaced, respectively, by the following weaker ones:

(A.4') \( t_k^i \leq t_0^k \leq t_1^k \leq t_2^k \leq \ldots \) for each \( k \in K \);

(A.5') \( \sum_{j=m}^{\infty} p_{ij}^k \alpha_{ij}^k \) is nondecreasing in \( i \) for each \( k \in K \) and \( m \geq 0 \).

The following corollary can be proved exactly as that of Theorem 5.4.1 from the above discussions

137
Corollary 5.5.1 For the Markov environment case, suppose that Assumptions (A.1), (A.2), (A.3), (A.4'), (A.5') hold, then \( v(k, i) = V^*(k, i, C) - V^*(k, i, R) \) is nondecreasing in \( i \) and

\[
V^*(k, i) = V^*(k, i, O), \quad i < i^*(k)
\]

\[
= V^*(k, i, R), \quad i \geq i^*(k)
\]

(5.5.6)

where \( i^*(k, i) = \min\{i|v(k, i) \geq 0\} \)

The corollary says that the state limit is also independent of the time variable \( s \), that is, \( i^*(k, s) = i^*(k) \).

Remark 5.5.1 (1) If \( t_F^k \leq t_0^k \) is not true, then it can be shown similarly that Corollary 5.5.1 holds in \( i \geq J_k := \min\{i|t_i^k \geq t_F^k\} \) and \( i^*(k) \) should be redefined by \( i^*(k) := \min\{i \geq J_k|v(k, i) \geq 0\} \). In this case, the optimal policy is to operate if \( J_k \leq i < i^*(k) \) and to replace if \( i \geq i^*(k) \), while it is not known what optimal action is when \( 0 \leq i < J_k \). We call such a policy an extended control limit policy.

(2) Due to the expressions of \( r(x, a) \) of Eqn. 5.5.3, one can know that both the optimal value and the optimal policies depend on \( T_{ij}^k(t) \) only through \( t_{ij}^k \). This is to say that the model with a Markov environment is robust with respect to the distribution function \( T_{ij}^k(t) \) of the time of state transition for the patient.

Moreover, if

\[
P_{ij}^k(t) = P_{ij}^k T_i^k(t), \quad \forall i, j, k
\]
k then, we can assume that the patient is Markov, i.e.,

\[ T_i^k(t) = 1 - e^{-\mu_i^k t} \]

where \( t_i^k \) and \( t_i^k \) are determined by each other with

\[
\begin{align*}
 t_i^k &= \frac{\lambda_k + \alpha}{\lambda_k + \alpha + \mu_i^k} \\
 \mu_i^k &= (\lambda_k + \alpha) \frac{1 - t_i^k}{t_i^k}
\end{align*}
\]

In Assumption (A.4'), \( t_i^k \) is nondecreasing in \( i \), so \( \sum_{j=m}^{\infty} P_{ij}^k \alpha_{ij}^k = \sum_{j=m}^{\infty} P_{ij}^k (1 - t_{ij}^k) \) may not be nondecreasing. The following lemma gives a sufficient condition for it.

**Lemma 5.5.1** Suppose that \( T_{ij}^k(t) = T_i^k(t) \) for all \( i, j \in S, k \in K \) and \( m \geq 0 \), then \( \sum_{j=m}^{\infty} P_{ij}^k \alpha_{ij}^k \) is nondecreasing in \( i \) if and only if

\[
\frac{t_{i+1}^k - t_i^k}{1 - t_i^k} \leq \frac{\sum_{j=m}^{\infty} P_{i+1,j}^k - \sum_{j=m}^{\infty} P_{ij}^k}{\sum_{j=m}^{\infty} P_{ij}^k}
\]  

Eqn. (5.5.7) means that the increasing speed of \( \sum_{j=m}^{\infty} P_{ij}^k \) in \( i \) for each \( m \geq 0 \) is larger than or equals to the decreasing speed of \( (1 - t_i^k) \).

Proof: It follows the given condition that

\[
t_{ij}^k = t_i^k, \quad \alpha_{ij}^k = \alpha_i^k = 1 - t_i^k, \quad \sum_{j=m}^{\infty} P_{ij}^k \alpha_{ij}^k = (1 - t_i^k) \sum_{j=m}^{\infty} P_{ij}^k
\]

It is obvious that for two nonnegative functions \( h(i) \) and \( g(i) \), if \( h(i) \) is non-increasing while \( g(i) \) is nondecreasing, then \( h(i), g(i) \) is nondecreasing if and only if

\[
\frac{h(i)}{h(i+1)} \leq \frac{g(i+1)}{g(i)} \quad \text{or} \quad \frac{h(i) - h(i+1)}{h(i+1)} \leq \frac{g(i+1) - g(i)}{g(i)}
\]
which immediately implies the Lemma 4.1. The optimal policies \( f^*_n \) and \( f^* \) are characterized by \( i^*_n(k) \) and \( i^*(k) \), respectively. We have the following result about the upper bound of these numbers, which is useful for the state reduction problem discussed below.

**Lemma 5.5.2** Under the conditions given in Corollary 5.5.1, if \( t^k_0 = t^k_F \), then \( i^*_n(k) \leq i^*_0(k) := \min \{ i | \Delta r(k, i) \geq 0 \} \) and \( i^*(k) \leq i^*_0(k) \) where \( \Delta r(k, i) = r^*(k, i, C) - r^*(k, i, R) \).

Proof: If \( t^k_0 = t^k_F \), then \( \alpha^k_0 = \alpha^k_F \). So it follows from Theorem 5.4.1 that

\[
\sum_{j \in S} q^k_{ij} V_n(k', j) - t^k_F V_n(k', 0) \geq \sum_{j \in S} q^k_{ij} V_n(k, 0) - t^k_F V_n(k, 0) = 0
\]

\[
\sum_{j \in S} p^k_{ij} \alpha^k_{ij} V_n(k, j) - \alpha^k_F V_n(k, 0) \geq \sum_{j \in S} p^k_{ij} \alpha^k_0 V_n(k, j) - \alpha^k_F V_n(k, 0)
\]

\[
= \alpha^k_0 V_n(k, 0) - \alpha^k_F V_n(k, 0) = 0
\]

So, we can get that \( v_n(k, i) \geq \Delta r(k, i) \), which implies the lemma immediately. Assumption (A.3) is about the cost rate, we now replace it by a new one about the expected total cost in a state. (A.3') for each \( k \in K \), both \( h^k(i)t^k_i - c^k(i)t^k_F \) and \( R^k(i, C)t^k_i - R^k(i, R)t^k_F \) are nondecreasing in \( i \). Here, \( h^k(i)t^k_i \) and \( c^k(i)t^k_F \) are respectively the expected treatment and rejuvenating treatment costs in state \( i \) when the environment state is \( k \). So the nondecreasingness of \( h^k(i)t^k_i - c^k(i)t^k_F \) means that the expected treatment cost increases faster than the expected rejuvenating treatment cost as the patient’s state increases. The nondecreasingness of \( R^k(i, C)t^k_i - R^k(i, R)t^k_F \) has a similar meaning.
Theorem 5.5.1 Under Assumptions (A.1), (A.2), (A.3), (A.4′) and (A.5′), for each \( k \in K \) and \( n \geq 1 \), \( v_n(k,i) := V_n^*(k,i,C) - V_n^*(k,i,R) \) is nondecreasing in \( i \), so \( v_n(k,i) < 0 \) iff \( i < i_n^*(k) := \min\{i|v_n(k,i) \geq 0\} \); moreover, \( v(k,i) := V^*(k,i,C) - V^*(k,i,R) \) is also nondecreasing in \( i \), and \( v(k,i) < 0 \) iff \( i < i^*(k) := \min\{i|v(k,i) \geq 0\} \). Thus, there exist optimal control limit policies.

Proof: It should be noted first that under the given conditions,

\[(\lambda_k + \alpha)\Delta r(k,i) = [h^k(i)t^k_i - c^k(i)t_F^k] + \lambda_k[R^k(i,C)t^k_i - R^k(i,R)t_F^k] \]

is nondecreasing in \( i \). Then the theorem can be proved exactly as that of Theorem 5.4.1.

Remark 5.5.2 The above theorem shows the existence of optimal control limit policies whose state limit \( i^*(k) \) depends only on the environment state \( k \). Thus, the Markov environment case is more simpler than the semi-Markov environment case.

Now we reduce the number of states of the patient under the Markov environment (see Eqn.( 5.5.1)). First, we suppose that

\[i_n^*(k) \leq j(k), \quad n \geq 0, \quad k \in K \tag{5.5.8}\]

for some \( j(k) \), where \( i_n^*(k) \) is defined in Theorem 5.5.1.

By Theorem 5.5.1, we have

\[V^*(k,i) = V^*(k,i,R) = r'(k,i,R) + V_0(k), \quad i \geq j(k), \quad k \in K \tag{5.5.9}\]
Where

\[ V_0(k) = \frac{\lambda_k t_k^k}{\lambda_k + \alpha} \sum_{k' \in K} \psi_{kk'} V^*(k', 0) + \alpha V^*(k, 0). \]

Thus one can get by Eqn. (5.5.4) for \( i \geq 0 \) as follows:

\[
V^*(k, i, C) = r'(k, i, C) + \frac{\lambda_k t_k^k}{\lambda_k + \alpha} \sum_{k' \in K} \psi_{kk'} \left\{ \sum_{j=0}^{j(k')-1} q_{ij}^k V^*(k', j) + \sum_{j=j(k')}^{\infty} q_{ij}^k [r'(k, i, R) + V_0(k)] \right\}
\]

\[
= r'(k, i, C) + \frac{\lambda_k t_k^k}{\lambda_k + \alpha} \sum_{k' \in K} \psi_{kk'} \left\{ \sum_{j=0}^{j(k')-1} q_{ij}^k [r'(k', j, R) - r'(k', j(k'), R)] + \sum_{j=j(k')}^{\infty} p_{ij}^k \alpha^k \right\}
\]

\[
= r'(k, i, C) + \frac{\lambda_k t_k^k}{\lambda_k + \alpha} \sum_{k' \in K} \psi_{kk'} \left\{ \sum_{j=0}^{j(k')-1} q_{ij}^k [r'(k', j, R) - r'(k', j(k'), R)] + \sum_{j=j(k')}^{\infty} p_{ij}^k \alpha^k \right\}
\]

We define that

\[
\tilde{q}_{ij}^{kk'} = \begin{cases} q_{ij}^k, & \text{When } j < j(k') \\ \sum_{j=j(k')}^{\infty} q_{ij}^k, & \text{When } j = j(k') \end{cases}
\]

\[
\tilde{p}_{ij}^k = \begin{cases} p_{ij}^k, & \text{When } j < j(k') \\ \sum_{j=j(k')}^{\infty} p_{ij}^k, & \text{When } j = j(k') \end{cases}
\]

\[
\tilde{T}_{ij}^k(t) = \begin{cases} T_{ij}^k(t), & \text{When } j < j(k') \\ \sum_{j=j(k')}^{\infty} T_{ij}^k(t)/p_{ij}^k, & \text{When } j = j(k') \end{cases}
\]

Thus

\[
\sum_{j=j(k)}^{\infty} p_{ij}^k \alpha_{ij}^k = \frac{\tilde{T}_{ij}^k(t)}{\tilde{p}_{ij,k(k)}^k \alpha_{ij,k(k)}}
\]
Where $\tilde{\alpha}_{i,j(k)}^k$ is defined as $\alpha_{i,j(k)}^k$ with $T_{ij}^k(t)$ being replaced by $\tilde{T}_{ij}^k(t)$. Let

$$\tilde{h}^k(i) = h^k(i) + \lambda_k \sum_{k' \in K} \psi_{kk'} \sum_{j=0}^{\infty} q_{ij}^k \frac{t_{F}^{k'}}{\lambda_{k'}} [c^{k'}(j) - c^{k'}(j(k'))]$$

$$+ (t_{i}^{k})^{-1} t_{F}^{k} \sum_{j=0}^{\infty} p_{ij}^k \alpha_{ij}^k [c^{k'}(j) - c^{k'}(j(k'))]$$

$$\tilde{R}^k(i, C) = R^k(i, C) + \sum_{k' \in K} \psi_{kk'} \sum_{j=0}^{\infty} q_{ij}^k \frac{t_{F}^{k'}}{\lambda_{k'}} + \lambda_{k'} [R^{k'}(j, R) - R^{k'}(j(k'), R)]$$

$$+ (t_{i}^{k})^{-1} t_{F}^{k} \sum_{j=0}^{\infty} p_{ij}^k \alpha_{ij}^k [R^{k'}(j, R) - R^{k'}(j(k'), R)]$$

$$\tilde{r}(k, i, C) = \frac{t_{i}^{k}}{\lambda_{k} + \alpha} \tilde{h}^k(i) + \lambda_k R^k(i, C)$$

It is easy to see that $\tilde{r}(k, i, C)$ is still nondecreasing in $i$ for each $k$ under Assumption A. Then for $i \geq 0$,

$$V^*(k, i, C) = \tilde{r}(k, i, C) + \frac{\lambda_k t_{i}^{k}}{\lambda_{k'} \alpha} \sum_{k' \in K} \psi_{kk'} \sum_{j=0}^{j(k')} q_{ij}^{kk'} V^*(k', j) + \sum_{j=0}^{j(k)} \tilde{p}_{ij}^k \tilde{\alpha}_{ij}^k V^*(k, j).$$

(5.5.12)

Now, we construct a new rejuvenating model (NRM), which is similar as the original rejuvenating model (ORM) excepts that

1. the state set of the patient in environment $k$ is $S_k = \{0, 1, ..., j(k)\}$ for $k \in K$;
2. the parameters $p_{ij}^k$, $T_{ij}^k(t)$, $q_{ij}^k$, $h^k(i)$ and $R^k(i, C)$ are replaced by $\tilde{p}_{ij}^k$, $\tilde{T}_{ij}^k(t)$, $\tilde{q}_{ij}^k$, $\tilde{h}^k(i)$ and $\tilde{R}^k(i, C)$, respectively, which are defined the above;
3. the patient must be given rejuvenating treatment in state $j(k)$ during environment state $k$ (due to Eqn. (5.5.8)).
From the above discussions, we know that the NRM and the ORM are equivalent under the meanings that the optimal objective values are identical and their optimality equations are equivalent for both the finite- and infinite-horizon problems. So their optimal policies are identical. The difference between them is that the number of patient's states is finite for NRM. Certainly, the problem with finite states is simpler than that with infinite states, e.g., the computation for the case of finite states is feasible while that for the case of infinite states should be approximated.

When \( j(k) \leq j^* \) for some \( j^* \), it can take the state set as \( S_k = \{0, 1, \ldots, j^*\} \), which is irrespective of \( k \). In the remaining of this section, we consider two further special cases. The first is that the state of patient itself is Markov, i.e.,

\[
T^k_{ij}(t) = 1 - e^{-\mu_{k,i} t}, \quad F^k(t) = 1 - e^{-\mu_F t} \quad (5.5.13)
\]

then \( \alpha^k_{ij} = \frac{\mu_{k,i}}{\lambda_k + \mu_{k,i} + \alpha} \), \( \beta^k_{ij} = \frac{\lambda_k + \alpha}{\lambda_k + \mu_{k,i} + \alpha} \).

The second further special case is that the environment is a Poisson process with rate \( \psi \), i.e., the Markov environment (see Eqn.( 5.5.1) with

\[
\psi_{k,k'} = 1 \quad G_k(t) = 1 - e^{-\lambda t}, \quad t \geq 0, \quad k \in K. \quad (5.5.14)
\]

Moreover, it is assumed that each adverse factor increases the degree of the seriousness in the condition of the patient with a probability distribution \( \{q_j, j \geq 0\} \) as follows:

\[
d^k_{ij} = 0 \quad \text{for} \ j < 0 \quad \text{and} \quad d^k_{ij} = q_{j-1} \quad \text{for} \ j \geq i \quad (5.5.15)
\]

Furthermore, all \( p^k_{ij}, T^k_{ij}, h^k(i), c^k(i), R^k(i, C) \) and \( R^k(i, R) \) are independent of \( k \) and will be denoted by \( p_{ij}, T_{ij}(t) \) and so on, by only deleting \( k \) in the original notations.
Then $t^k_{ij}, t^k_i, \alpha^k_{ij}, \alpha^k_i, t^k_F, \alpha^k_F$ are also independent of $k$ and will be denoted by $t_{ij}, t_i$ and so on.

Under these conditions, it can be shown that $V^*(k,i)$ and therefore $V^*(k,i,C)$ and $V^*(k,i,R)$ are independent of $k$. So $i^*(k) = i^*$ is also independent of $k$.

### 5.6 Numerical Example

Consider a numerical example where the environment is a Markov process having two states with parameters as follows

$$\psi_{kk'} = \begin{pmatrix} 0.64 & 0.36 \\ 0.57 & 0.43 \end{pmatrix}, \quad \lambda_1 = 0.076, \quad \lambda_2 = 0.093$$

and the state transition probabilities for the system are

$$P^1_{ij} = \begin{pmatrix} 0.68 & 0.26 & 0.06 & 0.0 & 0.0 \\ 0.04 & 0.68 & 0.22 & 0.00 & 0.0 \\ 0.00 & 0.07 & 0.65 & 0.28 & 0.00 \\ 0.00 & 0.00 & 0.03 & 0.65 & 0.32 \\ 0.00 & 0.00 & 0.00 & 0.00 & 1.0 \end{pmatrix}, \quad P^2_{ij} = \begin{pmatrix} 0.83 & 0.15 & 0.02 & 0.0 & 0.0 \\ 0.03 & 0.72 & 0.18 & 0.07 & 0.0 \\ 0.00 & 0.04 & 0.67 & 0.18 & 0.11 \\ 0.00 & 0.00 & 0.05 & 0.61 & 0.34 \\ 0.00 & 0.00 & 0.00 & 0.00 & 1.0 \end{pmatrix}$$

while two probability systems caused by the environment changes are

$$q^1_{ij} = \begin{pmatrix} 0.58 & 0.28 & 0.14 & 0.00 & 0.0 \\ 0.06 & 0.48 & 0.28 & 0.18 & 0.0 \\ 0.00 & 0.04 & 0.41 & 0.27 & 0.18 \\ 0.00 & 0.00 & 0.00 & 0.42 & 0.58 \\ 0.00 & 0.00 & 0.00 & 0.00 & 1.0 \end{pmatrix}, \quad q^2_{ij} = \begin{pmatrix} 0.64 & 0.26 & 0.10 & 0.0 & 0.0 \\ 0.06 & 0.55 & 0.27 & 0.12 & 0.0 \\ 0.00 & 0.06 & 0.47 & 0.29 & 0.18 \\ 0.00 & 0.00 & 0.00 & 0.48 & 0.52 \\ 0.00 & 0.00 & 0.00 & 0.00 & 1.0 \end{pmatrix}$$

The cost rate functions are as follows:

$$h^1(i) = 18 + 2i, \quad c^1(i) = 46 + i, \quad R^1(i,C) = 55 + i, \quad R^1(i,R) = 0$$

$$h^2(i) = 15 + 2i, \quad c^2(i) = 41 + i, \quad R^2(i,C) = 50 + i, \quad R^2(i,R) = 0$$
Now, it is assumed that the continuous discount factor is \( \alpha = 0.45 \), and

\[
(t_1^1, t_2^1, t_3^1, t_4^1) = (0.46, 0.73, 0.74, 0.76, 0.78, 0.80),
\]

\[
(t_1^2, t_2^2, t_3^2, t_4^2) = (0.57, 0.78, 0.79, 0.80, 0.82, 0.86).
\]

Thus for \( i=0,1,2,3,4 \)

\[
r'(1, i, C) = \frac{t_1^i}{\lambda_1 + \alpha} (21.496 + 2.076i),
\]

\[
r'(2, i, C) = \frac{t_1^i}{\lambda_1 + \alpha} (18.731 + 2.093i),
\]

\[
r'(1, i, R) = \frac{t_F^i}{\lambda_1 + \alpha} (46 + i),
\]

\[
r'(2, i, R) = \frac{t_F^i}{\lambda_1 + \alpha} (41 + i),
\]

Now we compute the finite-horizon optimal values \( V_n(k, i) \) iteratively by

\[
V_{n+1}(k, i, C) = r'(k, i, C) + \frac{\lambda k t_k^i}{\lambda_k + \alpha} \sum_{k' \in K} \sum_{j \in S} \psi_{kk'} q_{ij} V_n(k', j) + \sum_{j \in S} p_{ik}^j (1 - t_j^k) V_n(k, j),
\]

\[
V_{n+1}(k, i, R) = r'(k, i, R) + \frac{\lambda k t_F^i}{\lambda_k + \alpha} \sum_{k' \in K} \sum_{j \in S} \psi_{kk'} V_n(k', 0) + (1 - t_F^i) V_n(k, 0),
\]

\[
V_{n+1}(k, i) = \min \{ V_{n+1}(k, i, C) - V_{n+1}(k, i, R) \}
\]

for \( n \geq 0 \) with \( V_0(k, i, C) = V_0(k, i, R) = 0 \), \( \forall k, i \) \hspace{1cm} (5.6.1)

The numerical results are shown in Table 1 and Table 2 when \( n=26 \)

\[
|V_{n+1}(k, i, a) - V_n(k, i, a)| \leq 0.01 \text{ for all } (k, i, a),
\]

so we take the optimal value \( V^*(k, i) = V_{27}(k, i) \). Now \( v(k, i) \) is shown in the last line of Table (5.1) and thus the optimal limits for both the environments 1 and 2 is 3.

Namely,

\[
v^*(1) = 3, \quad v^*(2) = 3.
\]

The optimal policy in this example, is to initiating a rejuvenating treatment strategy if and only if the state of the patient reaches or exceeds 3 in both the environments 1 and 2.
### Table 5.1 Computed Values of $V_n(k, i)$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$V_n(1, i)$ (Environment 1)</th>
<th>$V_n(2, i)$ (Environment 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$i = 0$</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>129.69</td>
<td>144.16</td>
</tr>
<tr>
<td>2</td>
<td>225.48</td>
<td>247.43</td>
</tr>
<tr>
<td>3</td>
<td>297.28</td>
<td>324.65</td>
</tr>
<tr>
<td>4</td>
<td>355.19</td>
<td>384.76</td>
</tr>
<tr>
<td>5</td>
<td>398.77</td>
<td>429.02</td>
</tr>
<tr>
<td>6</td>
<td>430.30</td>
<td>459.77</td>
</tr>
<tr>
<td>7</td>
<td>452.28</td>
<td>480.17</td>
</tr>
<tr>
<td>8</td>
<td>467.18</td>
<td>493.55</td>
</tr>
<tr>
<td>9</td>
<td>477.14</td>
<td>502.36</td>
</tr>
<tr>
<td>10</td>
<td>483.76</td>
<td>508.19</td>
</tr>
<tr>
<td>11</td>
<td>488.15</td>
<td>512.05</td>
</tr>
<tr>
<td>12</td>
<td>491.07</td>
<td>514.62</td>
</tr>
<tr>
<td>13</td>
<td>493.01</td>
<td>516.32</td>
</tr>
<tr>
<td>14</td>
<td>494.30</td>
<td>517.45</td>
</tr>
<tr>
<td>15</td>
<td>495.15</td>
<td>518.20</td>
</tr>
<tr>
<td>16</td>
<td>495.72</td>
<td>518.70</td>
</tr>
<tr>
<td>17</td>
<td>496.10</td>
<td>519.03</td>
</tr>
<tr>
<td>18</td>
<td>496.35</td>
<td>519.25</td>
</tr>
<tr>
<td>19</td>
<td>496.51</td>
<td>519.40</td>
</tr>
<tr>
<td>20</td>
<td>496.62</td>
<td>519.50</td>
</tr>
<tr>
<td>21</td>
<td>496.70</td>
<td>519.56</td>
</tr>
<tr>
<td>22</td>
<td>496.74</td>
<td>519.61</td>
</tr>
<tr>
<td>23</td>
<td>496.78</td>
<td>519.63</td>
</tr>
<tr>
<td>24</td>
<td>496.80</td>
<td>519.65</td>
</tr>
<tr>
<td>25</td>
<td>496.81</td>
<td>519.67</td>
</tr>
<tr>
<td>26</td>
<td>496.82</td>
<td>519.67</td>
</tr>
</tbody>
</table>

$v(k, i)$: $-87.54 \text{ to } -16.40$

### 5.7 Discussion

In this chapter, we studied an optimal treatment strategy of a patient in a semi-Markov environment. It considered the performance/health status both by the patient
itself in a semi-Markov setup and by the influence of the environment to the patient. For both the finite- and infinite-horizon discounted criterions, it was shown that there exist optimal control limit policies. A special case for a Markov environment was discussed. When the control limits are bounded for each environment state, the countable states of the patient was simplified equivalently to a finite one. Finally, a numerical example was illustrated to prove the correctness and validity of the analysis.