

1973

# Aspects of the convergence of Bayes policies and posterior distributions for Markov decision processes

Mohamed Fathi Ahmed El-Sabbagh  
*Iowa State University*

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>

 Part of the [Statistics and Probability Commons](#)

## Recommended Citation

El-Sabbagh, Mohamed Fathi Ahmed, "Aspects of the convergence of Bayes policies and posterior distributions for Markov decision processes" (1973). *Retrospective Theses and Dissertations*. 5003.  
<https://lib.dr.iastate.edu/rtd/5003>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact [digirep@iastate.edu](mailto:digirep@iastate.edu).

## INFORMATION TO USERS

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

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

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.
2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.
4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.
5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

**Xerox University Microfilms**

300 North Zeeb Road  
Ann Arbor, Michigan 48106

74-531

EL-SABBAGH, Mohamed Fathi Ahmed, 1938-  
ASPECTS OF THE CONVERGENCE OF BAYES POLICIES  
AND POSTERIOR DISTRIBUTIONS FOR MARKOV  
DECISION PROCESSES.

Iowa State University, Ph.D., 1973  
Statistics

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.

Aspects of the convergence of Bayes policies and  
posterior distributions for Markov decision processes

by

Mohamed Fathi Ahmed El-Sabbagh

A Dissertation Submitted to the  
Graduate Faculty in Partial Fulfillment of  
The Requirements for the Degree of  
DOCTOR OF PHILOSOPHY

Departments: Industrial Engineering  
Statistics

Co-Majors: Industrial Engineering  
Statistics (Operations Research)

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Departments

Signature was redacted for privacy.

For ~~the~~ Graduate College

Iowa State University  
Ames, Iowa

1973

## TABLE OF CONTENTS

	Page
I. INTRODUCTION AND REVIEW OF LITERATURE	1
II. LARGE DEVIATIONS FOR MARKOV CHAINS	17
A. Introduction	17
B. Primitive Matrices	17
C. (2 x 2) Primitive Matrices	19
D. A Theorem Concerning Special Family of (2 x 2) Primitive Matrices	25
E. The Moment Generating Function for the (2 x 2) Case	30
F. Lower Bound for the Probability of a Large Deviation	33
G. Upper Bound for the Probability of a Large Deviation	40
H. The Main Theorem	42
III. ALMOST-SURE POSTERIOR CONVERGENCE AND ITS RATE	43
A. Introduction	43
B. Setting of the Problem in the (N x K x L)- Case	43
C. Posterior Convergence; The Special Case (N, K, L) = (2, 2, 2)	55
D. Rate of Almost-Sure Convergence of the Posterior Distribution	61
IV. POLICY CONVERGENCE	75

V. SUMMARY AND SUGGESTIONS FOR FUTURE STUDY	83
A. Problem Structure	83
B. Findings	83
C. Suggestions for Future Study	84
VI. LITERATURE CITED	86
VII. ACKNOWLEDGMENTS	91

## LIST OF DIAGRAMS

	Page
Diagram 1. Non-Bayesian optimal Markov decision process	8
Diagram 2. A Bayesian Markov decision process	12
Diagram 3. Part of the decision tree for n stages	45
Diagram 4. The boundary $\mathcal{L}$ : $\mathcal{L}_1 - \mathcal{L}_2 - \mathcal{L}_3$	68

## LIST OF TABLES

	Page
Table 1. The set $\{\psi_i^k\}$	51



## I. INTRODUCTION AND REVIEW OF LITERATURE

A stochastic process is a process which proceeds in time governed by probabilistic laws; it is a collection of random variables  $\{X_t\}$  where the index  $t$  is a member of a given set  $\{t\}$ , and  $X_t$  represents a characteristic under study. Often  $t$  symbolizes time and  $\{t\}$  is taken to be the nonnegative integers  $\{0,1,2,\dots\}$ . The unit involved can be any predetermined time interval which determines when the random variable  $X_t$  is to be measured or observed. This interval can be of any length, e.g., one minute, three days, etc. A stochastic process of this structure is often referred to as a time series. For example  $\{X_0, X_1, X_2, \dots\}$  can be thought of as the collection of queue lengths in a queuing system measured at different points in time  $t_0, t_1, t_2, \dots$ .

Note that  $X_t$  might be a vector; indeed so might  $t$ , when it does not denote time; however we restrict attention to the scalar case. The possible values or materializations of the random variables  $X_t$ , which will be frequently referred to as the "state space", need not be finite or even countable. An example of a continuum of possible outcomes is when the  $X_t$  are normally distributed. Note that the index  $t$  also might assume a continuum of values. In any case the collection  $\{t\}$  will be referred to as the "time domain".

Throughout the following chapters, this thesis will consider a time domain  $\{t_0, t_1, t_2, \dots, t_r, \dots\}$  where

$$t_{r+1} = t_r + \Delta t, \quad \Delta t > 0$$

being the time unit. At these times the system will be in one of a finite number  $N$  of mutually exclusive and exhaustive categories or states, i.e.,  $X_t \in \{1, 2, \dots, N\}$ . A specific example of the general situation just outlined, in addition to the queuing example above, is an inventory system whose level is reviewed only at discrete equally spaced intervals of time, e.g., each week.

In order to be able to obtain useful analytic results about the system, an assumption regarding the joint distribution of  $X_0, X_1, X_2, \dots$  is necessary. A widely used assumption which leads to tractable analytic manipulation is that the stochastic process has the Markovian property. A stochastic process  $\{X_t\}$  with countable time domain  $\{0, 1, 2, \dots\}$  and finite state space  $\{0, 1, 2, \dots, N\}$  is called a Markov chain if;

$$P\{X_{t+1}=j/X_0=m_0, \dots, X_{t-1}=m_{t-1}, X_t=i\} = P\{X_{t+1}=j/X_t=i\} \quad (1.1)$$

for  $t=\{0, 1, 2, \dots\}$  and every sequence  $m_0, m_1, \dots, m_{t-1}, i, j$  with elements in  $\{1, 2, \dots, N\}$ . In other words, the Markovian property signifies that the conditional probability of

the system (or  $X_t$ ) being in any future state given a particular past state history and the present state does not depend functionally on the past state history. The conditional probabilities mentioned above are called one-step transition probabilities. They are referred to as stationary if

$$P\{X_{t+1}=j/X_t=i\} = P\{X_1=j/X_0=i\} \quad (1.2)$$

for all  $t=0,1,2,\dots$  and all  $i,j \in \{1,2,\dots,N\}$  and are then usually represented by  $p_{ij}$ . In such a stationary situation the set of all the one-step transition probabilities is conveniently represented by a square matrix called the transition probability matrix  $P$ :

$$P \equiv [p_{ij}] \equiv \begin{matrix} & \begin{matrix} 1 & 2 & \dots & j & \dots & N \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ \cdot \\ \cdot \\ \cdot \\ i \\ \cdot \\ \cdot \\ \cdot \\ N \end{matrix} & \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1j} & \dots & p_{1N} \\ p_{21} & p_{22} & \dots & p_{2j} & \dots & p_{2N} \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ p_{i1} & p_{i2} & \dots & p_{ij} & \dots & p_{iN} \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ p_{N1} & p_{N2} & \dots & p_{Nj} & \dots & p_{NN} \end{bmatrix} \end{matrix} \quad (1.3)$$

The  $N$  rows and columns correspond to the  $N$  states the system

can occupy. The elements in row  $i$ ,  $i=1,2,\dots,N$  are the probabilities of one-step transition to state  $j$ ,  $j=1,2,\dots,N$ , conditioning on being in state  $i$ , and thus

$$0 \leq p_{ij} \leq 1 \quad \text{for all } i,j=1,2,\dots,N$$

and

$$\sum_{j=1}^N p_{ij} = 1 \quad \text{for all } i=1,2,\dots,N \quad . \quad (1.4)$$

Markov chains were first introduced by Markov in 1907. In the 1930's, Kolmogorov extended the mathematical theory to include the case of chains with an infinite number of states, i.e., infinite "state domain". By 1950 it was well recognized that the Markov chain principle provides a mathematical model with much potential for applications. The mathematical analysis of Markovian systems and their applications to physical as well as social sciences is a subject in itself about which several books have been written. Examples of specific applications are found in Derman and Lieberman (1967) on joint replacement and stocking problems, in Klein (1966) on production scheduling and control and maintenance-replacement decision problems, and in Herniter and Magee (1961) on customer behavior and marketing policies.

Howard and Derman introduced Markov decision processes which involve repetitive decision making in a Markov chain

environment. In that context, Howard (1960) has introduced the notion of returns or rewards in a Markov chain setting. Corresponding to the transition matrix  $P$ , he considers a matrix  $R$  of transition rewards:

$$R \equiv [r_{ij}] \equiv \begin{matrix} & \begin{matrix} 1 & 2 & \dots & j & \dots & N \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ \vdots \\ i \\ \vdots \\ N \end{matrix} & \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1j} & \dots & r_{1N} \\ r_{21} & r_{22} & \dots & r_{2j} & \dots & r_{2N} \\ \vdots & \vdots & & \vdots & & \vdots \\ r_{i1} & r_{i2} & \dots & r_{ij} & \dots & r_{iN} \\ \vdots & \vdots & & \vdots & & \vdots \\ r_{N1} & r_{N2} & \dots & r_{Nj} & \dots & r_{NN} \end{bmatrix} \end{matrix} \quad (1.5)$$

which specifies a return  $r_{ij}$  for a system transition from state  $i$  to state  $j$ .

In the case of a finite horizon Markovian process (finite time domain  $\{t\}$  or number of stages), the expected value of the system (expected total reward) can be calculated if we know both the transition probability matrix  $P$  and the associated reward matrix  $R$ . In this context, Howard (1960) introduced a recurrence relation for the expected total reward in the next  $n$  transitions,  $v_i(n)$ , given that the system is currently in state  $i$ :

$$v_i(n) = \sum_{j=1}^n p_{ij} [r_{ij} + v_j(n-1)] \quad , \quad i=1,2,\dots,N$$

$$n=1,2,3,\dots \quad (1.6)$$

In the infinite horizon case, a discount factor  $\beta$  may be introduced in order to be able to deal with finite total rewards. The discount factor  $\beta$ ,  $0 < \beta < 1$ , will discount the reward received at the  $h+1$ <sup>st</sup> transition by  $\beta^h$  to bring it down to present worth. It is still relevant, in the infinite horizon case, to consider finite horizons as stepping stones, but the recurrence relations (Equation 1.6) are changed slightly to allow for the discount factor  $\beta$ :

$$v_i(n) = \sum_{j=1}^n p_{ij} [r_{ij} + \beta v_j(n-1)] \quad , \quad i=1,2,\dots,N$$

$$n=1,2,\dots \quad (1.7)$$

The discount factor  $\beta$  can be interpreted in two different but related ways. First it has the psychological meaning that a reward received immediately has a greater intrinsic value than the same reward received at a future date (which suggests that the same reward received in the infinite future has no value). The other interpretation arises from the engineering economist's point of view, who would consider the discount factor  $\beta^h$  as the present worth factor  $(1+i)^{-h}$ , where  $i$  is the effective interest rate for the period of time between two successive transitions. In this context one may then appropriately speak of the present

worth of the sum of all future expected rewards.

Decision making was introduced into the multistage Markovian models by permitting a choice among several transition and reward matrices. In other words, the decision maker is confronted with  $K_i$  choices whenever he is in state  $i$ , each of these  $K_i$  choices leading to its corresponding  $1 \times N$  vector of transition probabilities and a corresponding  $1 \times N$  vector of transition rewards. Without loss of generality it will be assumed in this thesis that  $K_i = K$  not dependent on  $i$ . Note that, in this context of decision alternatives, a vector of transition rewards from state  $i$  to state  $j$  under decision  $k \in (1, 2, \dots, K)$ , should be thought of as a vector of "net" rewards taking into account the cost of moving from  $i$  to  $j$  under  $k$ . In the above decision making context, the details of a one-step transition may be described by the following diagram (Diagram 1).

At this point it may be appropriate to define the matrix

$$\mathcal{P} \equiv K \cdot N \quad \begin{array}{c} N \\ \boxed{\begin{array}{c} [p_{ij}^1] \\ \text{---} \text{---} \text{---} \\ [p_{ij}^2] \\ \text{---} \text{---} \text{---} \\ \vdots \\ \text{---} \text{---} \text{---} \\ [p_{ij}^K] \end{array}} \end{array} \quad (1.8)$$

$i, j: 1, 2, \dots, N$   
 $k: 1, 2, \dots, K$

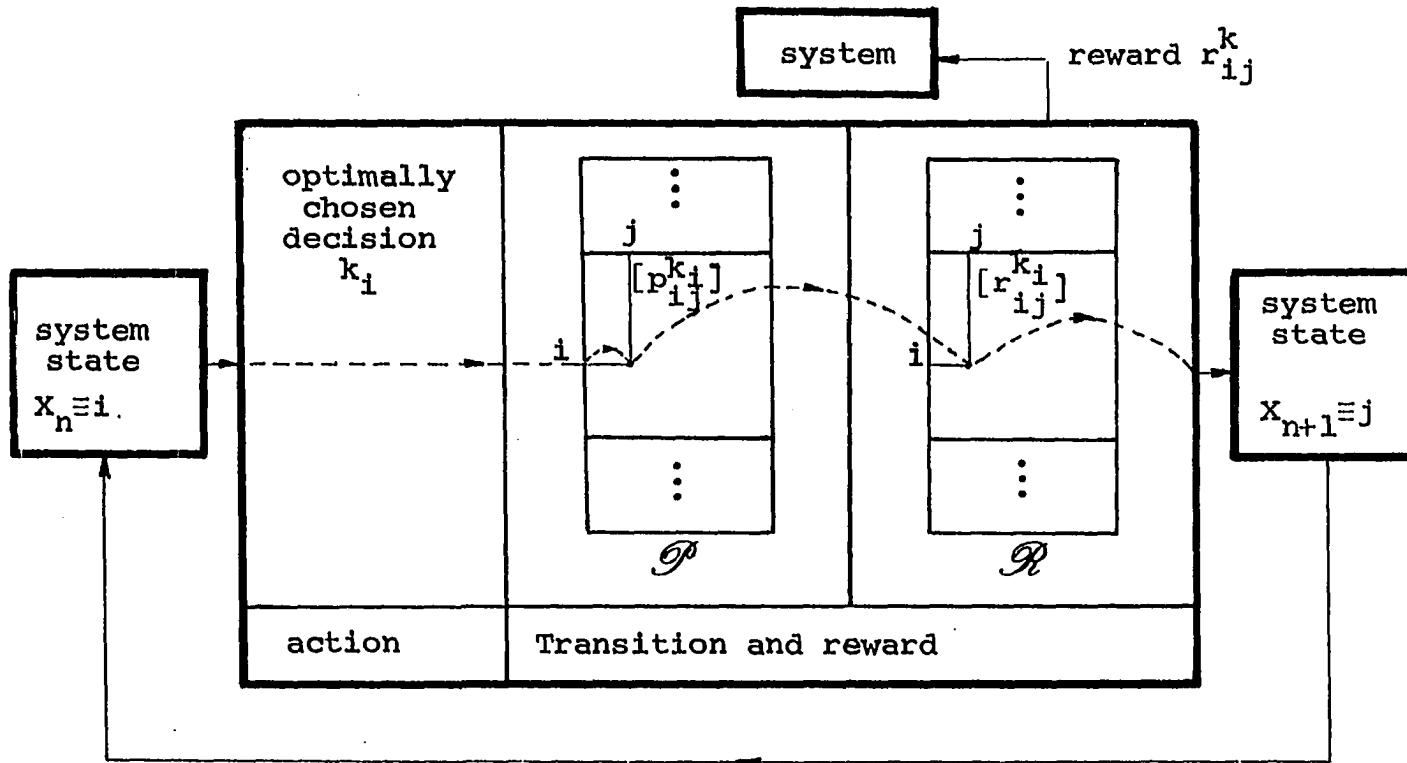


Diagram 1. Non-Bayesian optimal Markov decision process



as a generalized stochastic matrix since the word stochastic matrix will be reserved for an  $(N \times N)$  square matrix. Also the corresponding  $(K \times N \times N)$  reward matrix will be called the generalized reward matrix,  $\mathcal{R}$ .

Adopting a (stationary policy) in a Markovian decision process is to decide on which alternative action  $k$  to choose in state  $i$ . Howard (1960) has contributed to this problem by developing a procedure (the policy iteration method), to determine choices that maximize total expected reward. The method applies in particular to the discounted infinite horizon case. An alternate solution method, linked to standard dynamic programming approaches, Blackwell (1962, 1965), is based on the recursion system

$$v_i(n) = \max_k [q_i^k + \beta \sum_{j=1}^N p_{ij}^k v_j(n-1)] \quad i=1,2,\dots,N \quad (1.9)$$

where

$$q_i^k = \sum_{j=1}^N p_{ij}^k r_{ij}^k$$

is the expected immediate one-step reward and  $v_i(n)$  denotes the maximum total expected reward obtainable in  $n$  stages when starting in state  $i$ . In this context Howard (1960) has demonstrated the existence of a stationary optimal policy for the infinite horizon problem.

In most of the writing mentioned so far, the transition

matrix  $P$  (without alternatives) or  $\mathcal{P}$  (with alternatives) is assumed to be known with certainty. This assumption may not be justified, especially when dealing with a new system. The Bayesian approach attempts to circumvent this assumption, and considers a set of generalized stochastic matrices  $\{\mathcal{P}\}$ , finite or infinite, rather than a single matrix  $\mathcal{P}$ . The Bayesian decision maker starts with prior (raw) knowledge about, which matrix  $\mathcal{P}^* \in \{\mathcal{P}\}$  governs the system. This prior knowledge is described by a "prior" distribution (a preliminary probability distribution function  $h(\mathcal{P})$  over the set  $\{\mathcal{P}\}$ ). The prior distribution function may be discrete or continuous, depending on the set  $\{\mathcal{P}\}$ . The decision maker updates or improves his knowledge about the prior distribution with each transition, as he observes the successive random variables  $X_t$ . This improved knowledge is quantified by a "posterior" distribution that replaces the original prior. Since the outcomes  $X_t$  are actually generated according to the matrix  $\mathcal{P}^*$ , it is likely that the decision maker's updated posterior distribution  $h(\mathcal{P} | X_t, k)$  will place more probability mass near  $\mathcal{P}^*$ . Hopefully after a reasonably large number of transitions  $n^*$ , there will be probability mass near unity near  $\mathcal{P}^*$  and the optimal decision will be identical to the optimal decisions used by the decision maker who knows  $\mathcal{P}^*$ . In particular, the adopted policy will become stationary, in the sense that alternatives chosen in given states

will be the same for all  $n \geq n^*$ . Related problems in the firming up of information have been treated from several points of view; for example in Savage (1954), Blackwell and Dubins (1962), Doob (1953), and Rose (1971). The details of a one-step transition in the Bayes case may be described by the following diagram (Diagram 2).

Recent applications of Bayesian decision theory to Markov chains include studies, done at the Massachusetts Institute of Technology by Silver (1963) on chains with uncertain transition probabilities. Gonzales-Zubita and Miller (1965) uses mainly experimental and heuristic approaches to the problem and relies upon simulation rather than analysis.

Martin (1967) also considered the matter of selection of prior distribution functions, for the unknown matrix of transition. Related to the notion of natural conjugacy he discussed the idea of a family of distributions closed under consecutive sampling, i.e., a family such that the decision maker's revision of the prior will produce a posterior which belongs to a given family of distributions.

The problem of choosing a sequence of policies which maximizes the total expected discounted reward over an infinite period of time, was formulated by Martin (1967) in terms of a set of functional equations, once again linked to dynamic programming. The uniqueness and existence of

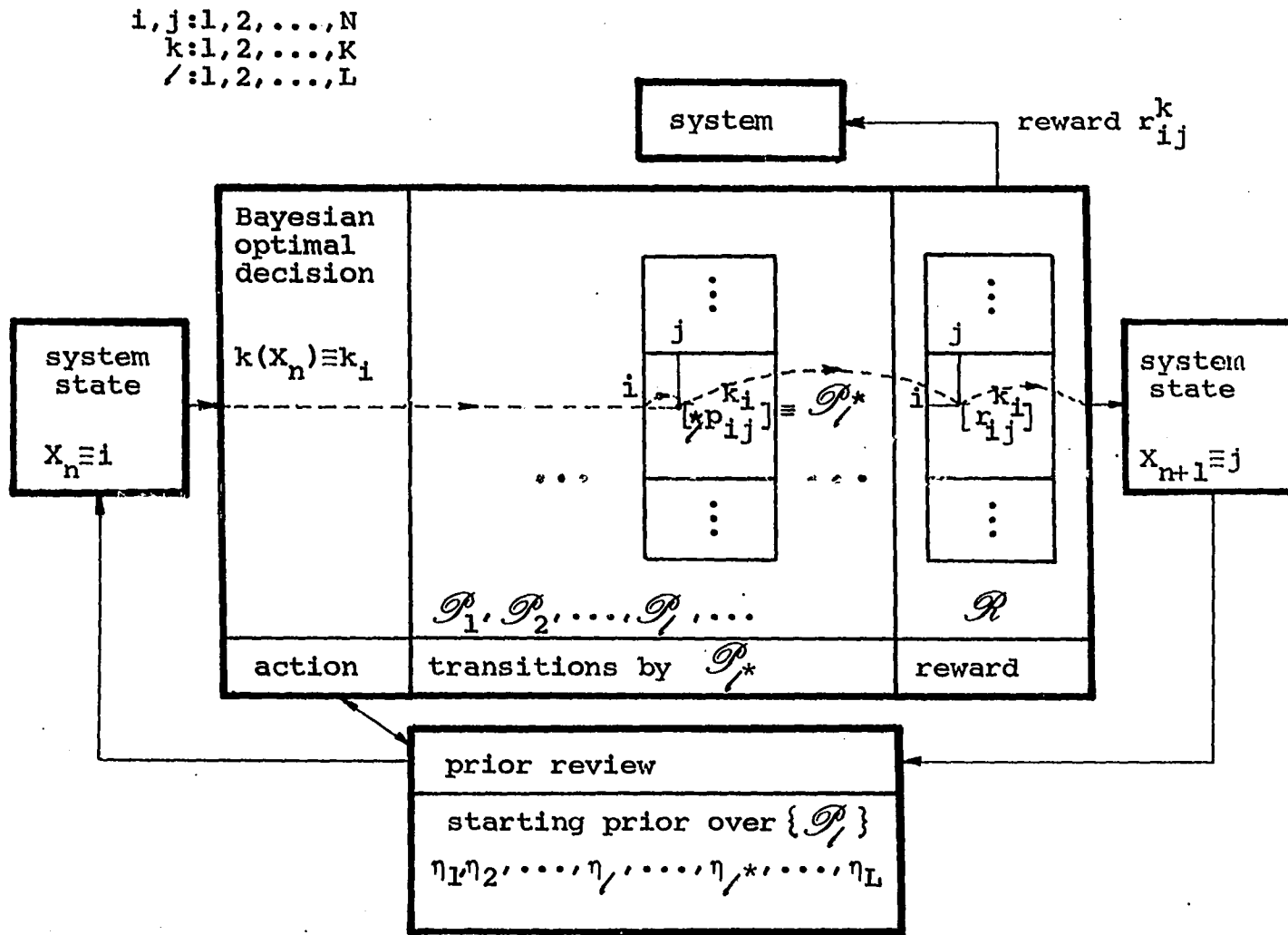


Diagram 2. A Bayesian Markov decision process

solutions of such equations were discussed, together with a successive approximation method, converging monotonically to solution. The above functional relations mentioned may be summarized as follows.

$$v_i(h) = \max_{1 \leq k \leq K} \left\{ \bar{q}_i^k + \beta \sum_{j=1}^N \bar{p}_{ij}^k(h) v_j(T_{ij}^k(h)) \right\}$$

$$i=1, 2, \dots, N$$

$$0 \leq \beta < 1 \quad (1.10)$$

where

$$\bar{p}_{ij}^k(h) = \sum_{\mathcal{P}} p_{ij}^k h(\mathcal{P})$$

is the marginal prior expectation of  $p_{ij}^k$ ,  $h$  is the current distribution over the states,  $T_{ij}^k(h)$  is the corresponding posterior and

$$\bar{q}_i^k(h) = \sum_{j=1}^N \bar{p}_{ij}^k(h) r_{ij}^k \quad \begin{array}{l} k=1, 2, \dots, K, \\ i=1, 2, \dots, N \end{array}$$

denote the mean one-step transition reward when the system is in state  $i$  and alternative  $k$  is used.

Generally both the transition matrix  $\mathcal{P}$  and the reward matrix  $\mathcal{R}$  can be considered random. Wolf (1970) considered the case where the  $\mathcal{P}$  matrix is known with certainty and the reward matrix follow a prior probability distribution. In particular he considered the rewards to be random, distributed

in accordance with a distribution belonging to a given class, and prior as well as posterior distributions pertain to this class. An upper and lower bound for the maximum total expected reward were developed as well as a method for finding an optimum policy.

This thesis considers a Markov decision process with alternative actions and uncertain transition probabilities. The system undergoing the above process is assumed to have a finite state space  $\mathcal{N}$  which consists of  $N$  states, with  $K$  alternative actions available to the decision maker in each state, and an infinite operating time domain. A system review (sampling) takes place at fixed points in time  $t_0, t_1, \dots, t_n, \dots$  with a prespecified constant time interval  $\Delta t = t_{n+1} - t_n, n=0,1,2,\dots$  where  $n$  denotes the  $n$ th sampling stage. A discounted reward structure is in effect with the  $n$ th reward  $r_{ij}^k$ , collected immediately after the  $n$ th transition occurs between states  $i$  and  $j$  under alternate  $k$ , discounted by a discount factor  $\beta^n$ .  $\beta, 0 < \beta < 1$ , is the discount rate.

The system is assumed to be governed by a generalized stochastic  $(K \cdot N \times N)$  matrix  $\mathcal{P}_l^*$ , a member of a finite set of  $L$  generalized stochastic matrices  $\{\mathcal{P}_l\}$ . The decision maker's uncertainty is introduced by a prior probability distribution over the set  $\{\mathcal{P}_l\}$ , reflecting his guess about the probabilities:  $\Pr\{\mathcal{P}_l \equiv \mathcal{P}_l^*\}, l=1,2,\dots,L$ .

A primary concern of this thesis is the convergence of the posterior distribution to a degenerate probability distribution with probability mass one at the matrix  $\mathcal{P}_1^*$ , and the effect of this convergence on the system's total discounted reward.

The approach adopted is to analyze a related "mixed" random walk, whose step sizes are functions of the transition probabilities of the set  $\{\mathcal{P}_i\}$ , and whose partial sums  $\mathcal{S}_n$  uniquely determine the posterior development. The random walk passes through two regions, the first related to a certain "state-nonstationary" phase of the policy, the second related to a certain "state-stationary" phase.

Chapter II is concerned with the task of bounding the probability that the posterior maintains at least a certain convergence rate, assuming state-stationarity. The idea is to use the moment-generating function approach for large deviations, explored for example by Bahadur and Rao (1960) and Chanda (1972), to the matrix-iteration expression for the moment-generating functions of a cumulative sum of scalar functions of Markov chain transitions, as it appears in Montroll (1947).

Based on Chapter II, Chapter III develops an upper bound for the rate of posterior convergence, with a certain assumption made about the process which insures a state-stationary phase for the related random walk. The procedure is to use the results of Dubins and Savage (1965a) to follow

the early course of the posterior in the state-nonstationary phase, then the results of Chapter II to follow the eventual state-stationary course. In section B, the random walk formulation is developed for the general case  $(N, K, L)$ , followed in sections C and D by the development of an upper bound for the convergence rate, in the special case  $(N = 2, K = 2, L = 2)$ .

Let  $\mathcal{D}_2$  be a Bayesian decision-maker with a prior assigning some weight to  $\mathcal{P}_1^*$ , the actual generalized matrix  $\mathcal{P}_1^*$  governing the evolution of the process, and let  $\mathcal{D}_1$  be a decision maker who knows  $\mathcal{P}_1^*$  and acts optimally accordingly. Chapter IV discussed the almost-sure relative near-equality for large  $\beta$  of the total discounted rewards earned by  $\mathcal{D}_2$  and  $\mathcal{D}_1$ . Two further assumptions are made here, the first of which essentially concerns independence from  $\beta$  of the state-stationary phase of the process.

Chapter V summarizes the results developed in the previous chapters and suggests possible considerations for future studies.



## II. LARGE DEVIATIONS FOR MARKOV CHAINS

### A. Introduction

This chapter explores the large deviations of cumulative sums of scalar functions of transitions of Markov chains. The term "large deviations" is used here in the sense that it is used for example in Bahadur and Rao (1960) or in Chanda (1972). The general approach will be to wed the moment-generating function approach explored by the above two authors to the matrix-iteration expression for the moment-generating functions of cumulative sums of scalar functions of Markov chain transitions, as it appears for example in Montroll (1947).

### B. Primitive Matrices

A nonnegative square matrix

$$\mathcal{B} \equiv [b_{ij}] \equiv \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1s} \\ b_{21} & b_{22} & \cdots & b_{2s} \\ \vdots & & & \vdots \\ b_{s1} & b_{s2} & \cdots & b_{ss} \end{bmatrix} \quad (2.1)$$

is such that  $b_{ij} \geq 0$  for all  $i, j=1, 2, \dots, s$ . An important property of nonnegative square matrices is demonstrated by the theorem of Perron and Forbenius (Cox and Miller, 1965), which states that a nonnegative square matrix has a maximal nonnegative

characteristic value which is not exceeded in absolute value by any other characteristic value, and corresponding to which there is a nonnegative characteristic vector.

A nonnegative irreducible matrix is one that cannot by suitable permutations applied to both rows and columns, be put into the form

$$\mathcal{B} \equiv \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix} \quad (2.2)$$

where the  $B_{11}$ ,  $B_{22}$  are square matrices.

Some properties of a nonnegative irreducible matrix are as follows:  $\mathcal{B}$  has a simple real positive characteristic root  $\lambda_1$  with the following properties;

(1) Corresponding to  $\lambda_1$  there is a positive right characteristic vector  $X$ , i.e., there exists a vector  $X > 0$  such that:

$$\mathcal{B}X = \lambda_1 X \quad (2.3)$$

(2) If  $\lambda'$  is any other characteristic root of  $\mathcal{B}$  then

$$|\lambda'| \leq \lambda_1 \quad ; \quad (2.4)$$

(3)  $\lambda_1$  is simple root of the determinantal equation

$$|\lambda I - \mathcal{B}| = 0 \quad (2.5)$$

(4)  $\lambda_1 \leq \max_i (\sum_j b_{ij})$ ,  $\lambda_1 \leq \max_j (\sum_i b_{ij})$ , i.e.

$$\lambda_1 \leq \min[\max_i(\sum_j b_{ij}), \max_j(\sum_i b_{ij})] \quad (2.6)$$

(5) The facts stated under (1) also apply to a positive left characteristic vector, as can easily be verified as follows: consider a nonnegative irreducible matrix  $\mathcal{B}$ , with maximal characteristic root  $\lambda_1$  then  $\mathcal{B}'$  is also nonnegative irreducible and has the same maximal characteristic root  $\lambda_1$ , as implied by Equation 2.5. Hence according to (1) there exists a vector  $\tilde{X} > 0$ , such that

$$\mathcal{B}' \tilde{X} = \lambda_1 \tilde{X} \quad , \quad (2.7)$$

i.e., a vector  $\tilde{X}'$  such that

$$\tilde{X}' \mathcal{B} = \lambda_1 \tilde{X}' \quad . \quad (2.8)$$

A nonnegative irreducible primitive (i.e., primitive) matrix is a nonnegative irreducible matrix such that  $\lambda_1$  is the only root of modulus  $|\lambda_1|$ . It can be shown, Gantmacher (1960) that, alternatively, a primitive matrix is a nonnegative matrix one of whose powers is a positive matrix.

### C. (2 x 2) Primitive Matrices

Consider a primitive matrix  $M$

$$M \equiv \begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad . \quad (2.9)$$

In view of the characterization of primitivity at the end of the last section, there can be at most one zero among the numbers A, B, C and D, and both B and C must be positive.

The characteristic roots of the matrix M are obtained by solving the determinantal equation

$$|M - \lambda I| = 0 \quad (2.10)$$

where I is the identity matrix of size 2 and  $\lambda$  is a scalar.

Equation 2.10 is equivalent to

$$\begin{vmatrix} (A-\lambda) & B \\ C & (D-\lambda) \end{vmatrix} = 0$$

$$(A-\lambda)(D-\lambda) - BC = 0$$

$$\lambda^2 - \lambda(A+D) + (AD-BC) = 0 \quad , \quad (2.11)$$

a second degree equation in  $\lambda$ , which has a solution

$$\begin{aligned} \lambda_1 &= \frac{(A+D) + \sqrt{(A+D)^2 - 4(AD-BC)}}{2} \\ &= \frac{(A+D) + \sqrt{A^2 + 2AD + D^2 - 4AD + 4BC}}{2} \\ &= \frac{(A+D) + \sqrt{(A-D)^2 + 4BC}}{2} \end{aligned} \quad (2.12)$$

and similarly

$$\lambda_2 = \frac{(A+D) - \sqrt{(A-D)^2 + 4BC}}{2} \quad (2.13)$$

where  $\lambda_1, \lambda_2$  are the characteristic roots of the matrix M.

Now let  $X = (x_1, x_2)$  and  $Y = (y_1, y_2)$  be the corresponding normalized right characteristic vectors for  $\lambda_1, \lambda_2$  respectively. The elements of the characteristic vector X are determined by the solution of the following set of equations

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda_1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (2.14)$$

or equivalently

$$Ax_1 + Bx_2 = \lambda_1 x_1 \quad (2.15)$$

$$Cx_1 + Dx_2 = \lambda_1 x_2 \quad (2.16)$$

together with the normalizing condition

$$x_1^2 + x_2^2 = 1 \quad (2.17)$$

Using Equations 2.15, 2.17 we get

$$x_2 = x_1 \frac{(\lambda_1 - A)}{B} \quad (2.18)$$

and

$$x_1^2 + x_1^2 \left( \frac{\lambda_1 - A}{B} \right)^2 = 1$$

↓

$$x_1^2 = \frac{B^2}{B^2 + (\lambda_1 - A)^2}$$

↓

$$x_1 = \frac{B}{\sqrt{B^2 + (\lambda_1 - A)^2}} \quad (2.19)$$

and consequently

$$\begin{aligned} x_2 &= \frac{B}{\sqrt{B^2 + (\lambda_1 - A)^2}} \cdot \left( \frac{\lambda_1 - A}{B} \right) \\ &= \frac{(\lambda_1 - A)}{\sqrt{B^2 + (\lambda_1 - A)^2}} \end{aligned} \quad (2.20)$$

By the same procedure used above the elements  $y_1, y_2$  of the other characteristic vector  $Y$  are determined by the solution of the set of equations

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \lambda_2 \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad (2.21)$$

or equivalently

$$AY_1 + BY_2 = \lambda_2 Y_1 \quad (2.22)$$

$$CY_1 + BY_2 = \lambda_2 Y_2 \quad (2.23)$$

together with the normalizing condition

$$y_1^2 + y_2^2 = 1 \quad (2.24)$$

The solution of Equations 2.23 and 2.24 gives

$$y_1 = \frac{(\lambda_2 - D)}{\sqrt{C^2 + (\lambda_2 - D)^2}} \quad (2.25)$$

and

$$y_2 = \frac{C}{\sqrt{C^2 + (\lambda_2 - D)^2}} \quad (2.26)$$

Hence the two characteristic roots are

$$\lambda_1 = \frac{(A+D) + \sqrt{(A-D)^2 + 4BC}}{2} \quad (2.27)$$

$$\lambda_2 = \frac{(A+D) - \sqrt{(A-D)^2 + 4BC}}{2} \quad (2.28)$$

and the corresponding normalized right characteristic vectors are

$$X = \left\{ \frac{2B}{\sqrt{4B^2 + [(D-A) + \sqrt{(A-D)^2 + 4BC}]^2}} \right. \\ \left. \frac{(D-A) + \sqrt{(A-D)^2 + 4BC}}{\sqrt{4B^2 + [(D-A) + \sqrt{(A-D)^2 + 4BC}]^2}} \right\} \quad (2.29)$$

$$Y = \left\{ \frac{(A-D) - \sqrt{(A-D)^2 + 4BC}}{\sqrt{4C^2 + [(A-D) - \sqrt{(A-D)^2 + 4BC}]^2}} \right. \\ \left. \frac{2C}{\sqrt{4C^2 + [(A-D) - \sqrt{(A-D)^2 + 4BC}]^2}} \right\} \quad (2.30)$$

If now one defines

$$V = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \quad (2.31)$$

Then  $V$  can be shown to have an inverse  $W$  by noting that  $V$  has a positive determinant under the assumptions considered about the matrix  $M$  at the beginning of Section C of this chapter together with the values of  $X$ ,  $Y$  in Equations 2.29, 2.30. Then one can write

$$M \cdot V = V \Lambda$$

and

$$M = V \Lambda V^{-1} \\ = V \Lambda W \quad (2.32)$$



where

$$\Lambda \equiv \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} .$$

D. A Theorem Concerning A Special Family  
of (2 x 2) Primitive Matrices

Consider the stochastic matrix M(0)

$$M(0) \equiv \begin{bmatrix} \theta_1 & (1-\theta_1) \\ (1-\theta_2) & \theta_2 \end{bmatrix} \quad (2.33)$$

where  $0 \leq \theta_1, \theta_2 < 1$  and  $\theta_1, \theta_2$  are not both zero. Consider also the matrix family M(t),

$$M(t) \equiv \begin{bmatrix} \theta_1 e^{at} & (1-\theta_1) e^{bt} \\ (1-\theta_2) e^{ct} & \theta_2 e^{dt} \end{bmatrix}, \quad 0 \leq t < +\infty \quad (2.34)$$

where a, b, c and d are arbitrary real numbers.

Theorem 2.1: Let  $\lambda_1(t)$  be the dominant characteristic root for M(t) and let  $\pi = (\pi_1, \pi_2)$  be the stationary probabilities vector for M(0). Let also

$$m = \pi_1 \cdot [a \cdot \theta_1 + b(1-\theta_1)] + \pi_2 \cdot [c(1-\theta_2) + d \cdot \theta_2] \quad (2.35)$$

If

$$m < Q < \max[a, b, c, d] \quad (2.36)$$

then there is a  $\tau$  such that

$$1) \quad e^{-Qt} \lambda_1(t) < 1 \quad \text{at } t=\tau \quad (2.37)$$

$$2) \quad \frac{d}{dt}[e^{-Qt} \lambda_1(t)] = 0 \quad \text{at } t=\tau \quad (2.38)$$

Proof

(1) The stochastic matrix  $M(0)$  has a stationary probability vector  $\pi$ , with elements determined by the solution of the set of equations:

$$\pi[I - M(0)] = 0$$

↓

$$[\pi_1, \pi_2] \begin{bmatrix} 1-\theta_1 & -(1-\theta_1) \\ -(1-\theta_2) & (1-\theta_2) \end{bmatrix} = 0$$

↓

$$\pi_1 \cdot (1-\theta_1) - \pi_2 \cdot (1-\theta_2) = 0 \quad (2.39)$$

$$-\pi_1(1-\theta_1) + \pi_2 \cdot (1-\theta_2) = 0 \quad (2.40)$$

together with the conditions

$$\pi_1 + \pi_2 = 1 \quad , \quad (2.41)$$

$$\pi_1, \pi_2 > 0 \quad . \quad (2.42)$$

Using Equations 2.39 and the conditions 2.41 and 2.42 one obtains

$$\pi \equiv (\pi_1, \pi_2) = \left[ \frac{(1-\theta_2)}{(1-\theta_1) + (1-\theta_2)}, \frac{(1-\theta_1)}{(1-\theta_1) + (1-\theta_2)} \right] . \quad (2.43)$$

Now, in view of Equation 2.27, the dominant characteristic root of the matrix  $M(t)$  is given by

$$\lambda_1(t) = \frac{\theta_1 e^{at} + \theta_2 e^{dt}}{2} + \sqrt{\left[ \frac{\theta_1 e^{at} - \theta_2 e^{dt}}{2} \right]^2 + (1-\theta_1)(1-\theta_2) \cdot e^{(b+c)t}} . \quad (2.44)$$

With reference to Equation 2.44

$$e^{-Qt} \lambda_1(t) = \frac{\theta_1 e^{(a-Q)t} + \theta_2 e^{(d-Q)t}}{2} + \sqrt{\left[ \frac{\theta_1 e^{(a-Q)t} - \theta_2 e^{(d-Q)t}}{2} \right]^2 + (1-\theta_1)(1-\theta_2) e^{[(b-Q)+(c-Q)] \cdot t}} \quad (2.45)$$

One can note the following properties of the quantity

$$e^{-Qt} \lambda_1(t)$$

a)  $e^{-Qt} \lambda_1(t) \Big|_{t=0}$  is the dominant root of a stochastic matrix, and hence is equal to unity. This may be seen as well by direct substitution in Equation 2.45

b)  $e^{-Qt} \lambda_1(t)$  is easily shown to be differentiable on the interval  $[0, \infty)$ ; in fact, with reference to Equation 2.45

$$\begin{aligned} & \frac{d}{dt} [e^{-Qt} \cdot \lambda(t)] \\ &= [\theta_1(a-Q) \cdot e^{(a-Q)t} + \theta_2(d-Q) e^{(d-Q)t}] / 2 \\ & \quad + \frac{1}{2} \{ [\theta_1 e^{(a-Q)t} - \theta_2 e^{(d-Q)t}]^2 / 4 \\ & \quad + (1-\theta_1)(1-\theta_2) e^{[(b-Q)+(c-Q)] \cdot t} \}^{-1/2} \\ & \quad \cdot \{ [\theta_1 e^{(a-Q)t} - \theta_2 e^{(d-Q)t}] \\ & \quad \cdot \left[ \frac{\theta_1(a-Q)e^{(a-Q)t} - \theta_2(d-Q)e^{(d-Q)t}}{2} \right] \\ & \quad + (1-\theta_1)(1-\theta_2) [(b-Q)+(c-Q)] \cdot e^{[(b-Q)+(c-Q)]t} \} \end{aligned} \quad (2.46)$$

- c)  $[e^{-Qt} \lambda(t)]$  has a negative derivative at  $t=0$   
since Equation 2.46 reduces to

$$\begin{aligned}
 & \left. \frac{d}{dt} [e^{-Qt} \cdot \lambda(t)] \right|_{t=0} \\
 &= [\theta_1(a-Q) + \theta_2(d-Q)]/2 \\
 & \quad + \frac{1}{2} \{ [\theta_1 - \theta_2]^2/4 + (1-\theta_1)(1-\theta_2) \}^{-1/2} \\
 & \quad \cdot \{ [\theta_1 - \theta_2] \left[ \frac{\theta_1(a-Q) - \theta_2(d-Q)}{2} \right] \\
 & \quad + (1-\theta_1)(1-\theta_2) [(b-Q)+(c-Q)] \} \\
 &= [\theta_1(a-Q) + \theta_2(d-Q)]/2 + \frac{1}{2[(1-\theta_1)+(1-\theta_2)]} \\
 & \quad \cdot \{ (\theta_1 - \theta_2) [\theta_1(a-Q) - \theta_2(d-Q)] \\
 & \quad + 2(1-\theta_1)(1-\theta_2) [(b-Q)+(c-Q)] \} \\
 &= \frac{1}{(1-\theta_1)+(1-\theta_2)} \{ [a_1\theta_1(1-\theta_2) + (b+c)(1-\theta_1)(1-\theta_2) \\
 & \quad + d\theta_2(1-\theta_1)] - Q[\theta_1(1-\theta_2) + \theta_2(1-\theta_1) + 2(1-\theta_1)(1-\theta_2)] \} \\
 &= m - Q \tag{2.47}
 \end{aligned}$$

which is a negative quantity in view of condition 2.36.

- d) It is also noted that  $e^{-Qt} \lambda_1(t)$  tends to  $\infty$  as  $t$  goes to infinity under inequality 2.36, in view of

Equation 2.45.

And now one can conclude:

- i)  $e^{-Qt}\lambda_1(t)$  is continuous, so by a), c) and d),  
there is a "first"  $T > 0$  with

$$e^{-QT}\lambda_1(T) = 1 \quad . \quad (2.48)$$

- ii)  $e^{-Qt}\lambda_1(t)$  is continuous on  $[0, T]$  and hence  
achieves its infimum,

$$\inf_{[0, T]} [e^{-Qt}\lambda_1(t)] \text{ at some } \tau, 0 \leq \tau \leq T \quad . \quad (2.49)$$

- iii)  $e^{-Q\tau}\lambda_1(\tau) < 1$  (2.50)

by reference to i), ii), a) and c) above so that

- iv)  $0 < \tau < T$  .

Hence, in view of iv), ii) and b),

- v)  $\frac{d}{dt}[e^{-Qt}\lambda_1(t)] = 0$  at  $t = \tau$  (2.51)

### E. The Moment Generating Function for the (2 x 2) Case

Consider a two-state Markov chain on two states 1 and  
2 with primitive transition matrix

$$M(0) = \begin{bmatrix} \theta_1 & 1-\theta_1 \\ 1-\theta_2 & \theta_2 \end{bmatrix}, \quad 0 \leq \theta_1, \theta_2 < 1; \\ \theta_1, \theta_2 \text{ not both zero.} \quad (2.52)$$

Suppose that the initial state of the Markov chain is  $i_0$  (=1 or 2).

Consider a random walk with partial sums

$$\mathcal{S}_{i_0, n}^* = \sum_{v=1}^n Z_v \quad (2.53)$$

defined on the Markov chain, with the  $v$ 'th step  $Z_v$  is the real number  $\alpha, \beta, \gamma$  or  $\delta$  depending on whether the  $v$ 'th transition of the Markov chain is  $1 \rightarrow 1, 1 \rightarrow 2, 2 \rightarrow 1$  or  $2 \rightarrow 2$ .

It is verified by direct computation that the moment generating function of  $\mathcal{S}_{i_0, n}^*$  is given by

$$\psi_n(t) = I_0 [M(t)]^{nJ} \quad ; \quad t \text{ real} \quad (2.54)$$

where  $I_0$  is a  $1 \times 2$  vector with unity at  $i_0$  and zero elsewhere, where  $J$  is a  $2 \times 1$  vector of one's and where the matrix  $M(t)$  is given by

$$M(t) = \begin{bmatrix} \theta_1 e^{\alpha t} & (1-\theta_1) e^{\beta t} \\ (1-\theta_2) e^{\gamma t} & \theta_2 e^{\delta t} \end{bmatrix} \quad (2.55)$$

Such expressions for the moment-generating functions of Markov chain-related random walks seem to have appeared first in Montroll (1947).

Since  $M(0)$  is primitive by assumption, so will be  $M(t)$ , and one may apply the facts deduced in Section C to write  $M(t)$  in the form

$$M(t) = V(t) \Lambda(t) W(t) \quad , \quad (2.56)$$

where  $W(t)$ ,  $\Lambda(t)$  and  $V(t)$  are, respectively, the matrices  $W$ ,  $\Lambda$  and  $V$  of Section C, with the substitutions

$$\begin{aligned} A &= \theta_1 e^{\alpha t} \\ B &= (1-\theta_1) e^{\beta t} \\ C &= (1-\theta_2) e^{\gamma t} \\ D &= \theta_2 e^{\delta t} \quad . \end{aligned} \quad (2.57)$$

Since  $W(t) = V(t)^{-1}$  it follows, in view of Equations 2.54 and 2.56, that

$$\psi_n(t) = I_0 V(t) [\Lambda(t)]^n W(t) J \quad , \quad (2.58)$$

which may be written in the form

$$\Delta(t) \cdot [\lambda_1(t)]^n \cdot \left\{ 1 + \left[ \frac{\lambda_2(t)}{\lambda_1(t)} \right]^n \cdot \delta(t) \right\} \quad (2.59)$$



where  $\lambda_1(t)$ ,  $\lambda_2(t)$ ,  $\Delta(t)$  and  $\delta(t)$  are analytic in  $t$ ,  $\lambda_2(t)/\lambda_1(t)$  is less than one for all  $t$  and is uniformly bounded away from one in any finite  $t$  interval, and where  $\lambda_1(t)$  is given by the expression

$$\lambda_1(t) = \frac{e^{\alpha t} + e^{\delta t}}{2} + \sqrt{\left[\frac{\theta_1 e^{\alpha t} - \theta_2 e^{\delta t}}{2}\right]^2 + (1-\theta_1)(1-\theta_2) \cdot e^{(\beta+\gamma)t}} \quad (2.60)$$

#### F. Lower Bound for the Probability of a Large Deviation

Consider the random walk of Section E, with the four possible steps denoted by  $a$ ,  $b$ ,  $c$  and  $d$  rather than by  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$ . Denote the corresponding partial sums by  $\mathcal{S}_{i_0, n}$ , rather than by  $\mathcal{S}_{i_0, n}^*$ .

We are interested in the excess probability

$$\Pr\{\mathcal{S}_{i_0, n} \geq n \cdot Q\} \quad (2.61)$$

for some real number  $Q$ .

If we now equate  $\alpha$  to  $(a-Q)$ ,  $\beta$  to  $(b-Q)$ , etc., this excess probability may be written in terms of the random walk of Section E as

$$\Pr\{\mathcal{S}_{i_0, n}^* \geq 0\} = \int_{x=0}^{\infty} dF_n(x) \quad (2.62)$$

where  $F_n(x)$  is the cdf of  $\mathcal{S}_{i_0, n}^*$ . Suppose now that

$$m < Q < \max(a, b, c, d) \quad (2.63)$$

where

$$m = \left\{ \left[ \frac{(1-\theta_2)}{(1-\theta_1)+(1-\theta_2)} \right] [\theta_1 a + (1-\theta_1) b] \right. \\ \left. + \left[ \frac{(1-\theta_1)}{(1-\theta_1)+(1-\theta_2)} \right] [(1-\theta_2) c + \theta_2 d] \right\} \quad (2.64)$$

Then in view of the theorem of Section D, there exists a real number  $\tau$ ,  $0 < \tau < +\infty$ , such that

$$(1) \quad \lambda_1(\tau) < 1 \quad (2.65)$$

$$(2) \quad \lambda_1'(\tau) = 0 \quad (2.66)$$

where  $\lambda_1(t)$  is the function (2.60), with  $a-Q$  substituted for  $\alpha$ ,  $b-Q$  for  $\beta$ , etc., i.e., the function (2.45) of Section D.

Using this  $\tau$ , define a new cdf  $G_n(x)$  by

$$dG_n(x) = \exp(\tau x) dF_n(x) / \psi_n(\tau) \quad , \quad (2.67)$$

where as defined in Section E,  $\psi_n(t)$  is the moment-generating function of  $\mathcal{S}_{i_0, n}^* = \mathcal{S}_{i_0, n} - nQ$ . Then in analogy to the development in Bahadur and Rao (1960) and Chanda (1972),

$$\Pr\{\mathcal{S}_{i_0, n} \geq nQ\} / \psi_n(\tau) = \Pr\{\mathcal{S}_{i_0, n}^* \geq 0\} / \psi_n(\tau)$$

$$= \int_{x=0}^{\infty} dF_n(x) / \psi_n(\tau)$$

$$= \int_{x=0}^{\infty} \exp(-\tau x) dG_n(x)$$

$$= \int_{x=0}^{\infty} \exp(-\tau x) dG_n(x)$$

$$= \int_{y=0}^{\infty} \exp(-\tau \sqrt{n} y) dG_n(\sqrt{n} y)$$

$$= e^{-\sqrt{n}\tau y} G_n(\sqrt{n}y) \Big|_0^{\infty} - \int_{y=0}^{\infty} (-\sqrt{n}\tau) e^{-\sqrt{n}\tau y} G_n(\sqrt{n}y) dy$$

$$= -G_n(0) + \int_{y=0}^{\infty} \sqrt{n}\tau e^{-\sqrt{n}\tau y} G_n(\sqrt{n}y) dy$$

$$= \sqrt{n}\tau \cdot \int_{y=0}^{\infty} e^{-\sqrt{n}\tau y} [G_n(\sqrt{n}y) - G_n(0)] dy$$

$$\geq \sqrt{n}\tau \int_{y=\varepsilon}^{\infty} e^{-\sqrt{n}\tau y} [G_n(\sqrt{n}y) - G_n(0)] dy$$

$$\begin{aligned}
&\geq \sqrt{n\tau} \cdot [G_n(\sqrt{n}\epsilon) - G_n(0)] \cdot \int_{y=\epsilon}^{\infty} e^{-\sqrt{n\tau}y} dy \\
&= [G_n(\sqrt{n}\epsilon) - G_n(0)] \cdot e^{-\sqrt{n\tau}\epsilon} \quad . \quad (2.68)
\end{aligned}$$

Thus

$$e^{\sqrt{n\tau}t} \cdot \Pr\{\mathcal{S}_{i_0, n} \geq nQ\} / \psi_n(\tau) \geq G_n(\sqrt{n}\epsilon) - G_n(0) \quad . \quad (2.69)$$

In order to obtain the lower bound we seek, it remains for us to examine  $\psi_n(\tau)$  and  $G_n(\sqrt{n}\epsilon) - G_n(0)$ .

We first examine  $\psi_n(\tau)$ : As defined earlier  $\psi_n(t)$  is the moment-generating function of the partial sum  $\mathcal{S}_{i_0, n}^*$  for the random walk with steps

$$\begin{aligned}
\alpha &= a-Q \\
\beta &= b-Q \\
\gamma &= c-Q \\
\delta &= d-Q \quad , \quad (2.70)
\end{aligned}$$

based on the Markov chain with transition matrix  $M(0)$ , starting in state  $i_0$ . Hence, in view of the calculations of Section E, and in particular Equation 2.59,

$$\psi_n(\tau) = \Delta(\tau) \cdot [\lambda_1(\tau)]^n \cdot [1 + \rho(\tau)^n \cdot \delta(\tau)] \quad (2.71)$$

where

$$\rho(\tau) = \lambda_2(\tau)/\lambda_1(\tau) < 1 \quad .$$

Moreover, in view of property (1), Equation 2.65, of  $\tau$  given earlier in this section,

$$\lambda_1(\tau) < 1 \quad .$$

We next examine  $G_n(\sqrt{n}\epsilon) - G_n(0)$ , again in analogy with the work of Bahadur and Rao (1960) and Chanda (1972). Recall the definition of  $G_n(x)$  in Equation 2.67. The moment-generating function corresponding to  $G_n(x)$  is

$$\begin{aligned} \Phi_n(t) &= \int_{-\infty}^{\infty} \exp(tx) \cdot \exp(\tau x) \cdot dF_n(x) / \psi_n(\tau) \\ &= \psi_n(t+\tau) / \psi_n(\tau) \quad , \end{aligned} \quad (2.72)$$

so that the moment-generating function  $\Theta_n(t)$  corresponding to the cdf  $G_n(\sqrt{n}x)$  is

$$\psi_n(\tau+t/\sqrt{n}) / \psi_n(\tau) \quad , \quad (2.73)$$

and in view of Equation 2.71

$$\begin{aligned} \Theta_n(t) &= \frac{\Delta(\tau+t/\sqrt{n})}{\Delta(\tau)} \cdot \left[ \frac{\lambda_1(\tau+t/\sqrt{n})}{\lambda_1(\tau)} \right]^n \\ &\quad \cdot \left[ \frac{1 + (\rho(\tau+t/\sqrt{n}))^n \delta(\tau+t/\sqrt{n})}{1 + (\rho(\tau))^n \delta(\tau)} \right]^n \quad . \end{aligned} \quad (2.74)$$

Now since  $\Delta(t)$  and  $\delta(t)$  are continuous, and  $\rho(t)$  is less than one for all  $t$  and is uniformly bounded away from one in any finite neighborhood of  $\tau$ , one may write

$$\theta_n(t) = \left[ \frac{\lambda_1(\tau+t/\sqrt{n})}{\lambda_1(\tau)} \right]^n \cdot \eta_n(t) \quad , \quad (2.75)$$

where as  $n$  increases,

$$\eta_n(t) \rightarrow 1 \quad (2.76)$$

for any value of  $t$ .

As for the first factor of the expression, since  $\lambda_1(t)$  is analytic one may expand  $\lambda_1(t)$  around  $t=\tau$ , obtaining

$$\lambda_1(\tau+t/\sqrt{n}) = \lambda_1(\tau) + \frac{t}{\sqrt{n}} \cdot \lambda_1'(\tau) + \frac{t^2}{2n} \cdot \lambda_1''(\tau) + O_n(t) \quad (2.77)$$

where  $n \cdot O_n(t)$  tends to zero as  $n$  increases for all  $t$ .

But, in view of property (2), Equation 2.66, given earlier in this section,

$$\lambda_1'(\tau) = 0 \quad ,$$

so that

$$\frac{\lambda_1(\tau+t/\sqrt{n})}{\lambda_1(\tau)} = 1 + \frac{t^2 \cdot \lambda_1''(\tau)}{2n\lambda_1(\tau)} + O_n(t) \quad (2.78)$$

and as  $n$  increases.

$$\left[ \frac{\lambda_1(\tau + t/\sqrt{n})}{\lambda_1(\tau)} \right]^n \rightarrow \exp[t^2 \lambda_1''(\tau) / 2\lambda_1(\tau)] \quad (2.79)$$

Now if we assume that  $\lambda_1''(\tau) > 0$ , then the right-hand side is the moment generating function of the normal distribution with mean zero and variance  $\lambda_1''(\tau) / \lambda_1(\tau)$ .

So in view of Equation 2.75, 2.76 and 2.79,  $\theta_n(t)$  tends with  $n$  to the moment-generating function of the normal  $N[0, \lambda_1''(\tau) / \lambda_1(\tau)]$ , which implies that

$$G_n(\sqrt{nx}) \rightarrow \Phi[x \cdot \sigma(\tau)] \quad (2.80)$$

where

$$\sigma(\tau) = \sqrt{\lambda_1(\tau) / \lambda_1''(\tau)} \quad .$$

One now has all the ingredients needed. Combining Equation 2.54 and 2.68 we have

$$\frac{e^{\sqrt{nt}\varepsilon} \cdot \Pr\{\mathcal{S}_{i_0, n} \geq nQ\}}{[\lambda_1(\tau)]^n} \geq \Delta(\tau) \cdot [1 + (\rho(\tau))^n \cdot \delta(\varepsilon)] \quad .$$

$$[G_n(\sqrt{n\varepsilon}) - G_n(0)] \quad , \quad (2.81)$$

and using Equation 2.80 and the fact that  $\rho(\tau) < 1$ ,

$$\liminf \frac{e^{\sqrt{n}\tau\varepsilon} \cdot \Pr\{\mathcal{S}_{i_0, n} \geq nQ\}}{[\lambda_1(\tau)]^n} \geq \Delta(\tau) \cdot [\Phi(\varepsilon\sigma(\tau)) - \Phi(0)]$$

$$\equiv B(\varepsilon) \quad . \quad (2.82)$$

In other words,

$$e^{\sqrt{n}\tau\varepsilon} \cdot \Pr\{\mathcal{S}_{i_0, n} \geq nQ\} / [\lambda_1(\tau)]^n$$

eventually is not below  $B(\varepsilon)$ , or

$$\sqrt{n}\tau\varepsilon + \ln \Pr\{\mathcal{S}_{i_0, n} \geq nQ\} - n \ln \lambda_1(\tau)$$

eventually is not below  $B(\varepsilon)$ , or, for any  $\varepsilon$ , there is a  $n(\varepsilon)$  such that

$$\frac{1}{n} \ln \Pr\{\mathcal{S}_{i_0, n} \geq nQ\} \geq \lambda_1(\tau) - \varepsilon \quad \text{for all } n > n(\varepsilon)$$

$$(2.83)$$

### G. Upper Bound for the Probability of a Large Deviation

An upper bound analogous to the lower bound derived above is readily established as well, again arguing in analogy to Bahadur and Rao (1960) and Chanda (1972):

$$\Pr\{\mathcal{S}_{i_0, n} \geq nQ\} = \Pr\{\mathcal{S}_{i_0, n}^* \geq 0\}$$



$$\begin{aligned}
&= \Pr\{\mathcal{S}_{i_0, n}^* \geq 0\} \\
&= \Pr\{e^{\mathcal{S}_{i_0, n}^*} \geq 1\} \\
&\leq E[e^{\mathcal{S}_{i_0, n}^*}] \\
&\equiv \psi_n(\tau) \quad . \qquad (2.84)
\end{aligned}$$

The inequality follows from the fact that the expectation of a nonnegative random variable  $X$  equals the area of the region  $\Omega$  in the plane determined by

$$F_X(x) \leq y \leq 1 \quad \text{and} \quad x \geq 0 \quad ,$$

whereas  $\Pr\{X \geq 1\}$  equals the area of the subregion of  $\Omega$  determined by  $F_X(1) \leq y \leq 1$  and  $0 \leq x \leq 1$ .

Hence using Equation 2.71 of the previous section, one finds that

$$\frac{\Pr\{\mathcal{S}_{i_0, n} \geq nQ\}}{[\lambda_1(\tau)]^n} \leq \Delta(\tau) \cdot [1 + \rho(\tau)^n \cdot \delta(\tau)] \quad , \quad (2.85)$$

or

$$\overline{\lim} \frac{\Pr\{\mathcal{S}_{i_0, n} \geq nQ\}}{[\lambda_1(\tau)]^n} \leq \Delta(\tau) \quad , \quad (2.86)$$

or for any  $\varepsilon$  there is a  $n(\varepsilon)$  such that

$$\frac{1}{n} \ln \Pr\{\mathcal{S}'_{i_0, n} \geq nQ\} \leq \ln \lambda_1(\tau) + \varepsilon, \quad n > n(\varepsilon) \quad . \quad (2.87)$$

#### H. The Main Theorem

We recapitulate the development of this chapter in one theorem:

**Theorem 2.2:** Consider a two-state Markov chain with transition matrix (Equation 2.52). Consider a random walk  $(n, \mathcal{S}'_{i_0, n})$  defined on this Markov chain which starts at state  $i_0$  and takes steps a, b, c or d depending on which of the four possible transitions occurs. Let  $Q$  be a real number satisfying Equation 2.63 and define the function  $\lambda_1(t)$  as in Equation 2.60. Then there is a positive number  $\tau$ ,  $\lambda_1(\tau) < 1$ , with the following property.

If  $\lambda_1'(\tau) > 0$ , then, given  $t$ , there exists  $n(\varepsilon)$  such that, for all  $n > n(t)$ ,  $\ln \lambda_1(\tau) - \varepsilon \leq \frac{1}{n} \ln \Pr\{\mathcal{S}'_{i_0, n} \geq nQ\} \leq \ln \lambda_1(\tau) + \varepsilon$ .

In view of known analogues of Equation 2.58 for general primitive Markov kernels, for example in Harris (1964), and Debreu and Herstein (1957), we expect general versions of this theorem to hold as well.

### III. ALMOST-SURE POSTERIOR CONVERGENCE AND ITS RATE

#### A. Introduction

This chapter explores the rate of almost-sure convergence of the posterior distribution when the decision maker uses a Bayes strategy based on a prior that puts some weight on the true state of nature  $\mathcal{P}_j^*$ . The two main assumptions are (1) that  $\mathcal{P}_j^*$  is positive and (2) that there is a "prior" interval about unity where the Bayes strategy is state-stationary. Certain other assumptions also are made, but are verified for a certain low-dimensional case in sections C, D, and F.

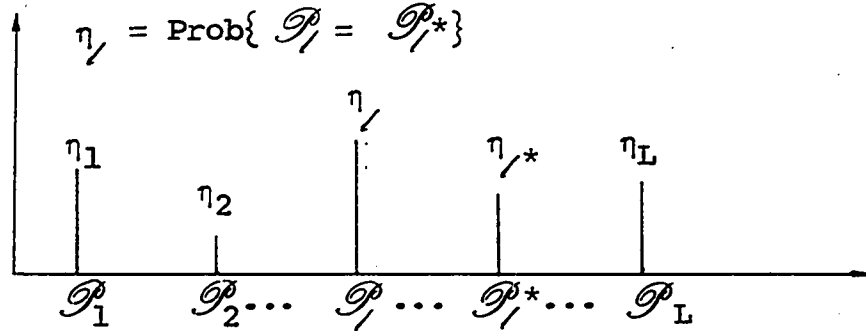
The general idea is to follow the early course of the posterior by means of a result of Dubins and Savage (1965a), and to then follow its eventual "state-stationary course" by means of the material of Chapter II.

#### B. Setting of the Problem in the (NXKXL)-Case

The model will involve:

1. A finite state space with  $N$  states  $1, 2, \dots, N$ .
2. A finite number of actions  $\{1, 2, \dots, K\}$  available for the decision maker in each state.
3. A prior knowledge about the transition matrix used by the process, in the form of a finite set of  $(N.K \times N)$  generalized stochastic matrices  $\{\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_j, \dots, \mathcal{P}_j^*, \dots, \mathcal{P}_L\}$  together with a prior

probability distribution  $\{\eta_1, \eta_2, \dots, \eta_j, \dots, \eta_{j^*}, \dots, \eta_L\}$ , with  $\mathcal{P}^*$  the matrix in fact governing the process. This may be represented as follows



$$\text{and } \sum_{\text{all } j} \eta_j = 1.$$

4. A  $(N \times K \times N)$  reward matrix  $\mathcal{R}$ .

The decision maker observes the process every time unit  $\{0, 1, 2, \dots\}$ , makes his decisions and collects rewards. This can be represented by a decision tree, with part of it shown in Diagram 3.

Let us define

$X_t$  = The state of the system at time  $t$  (or at the beginning of period number  $t+1$ )

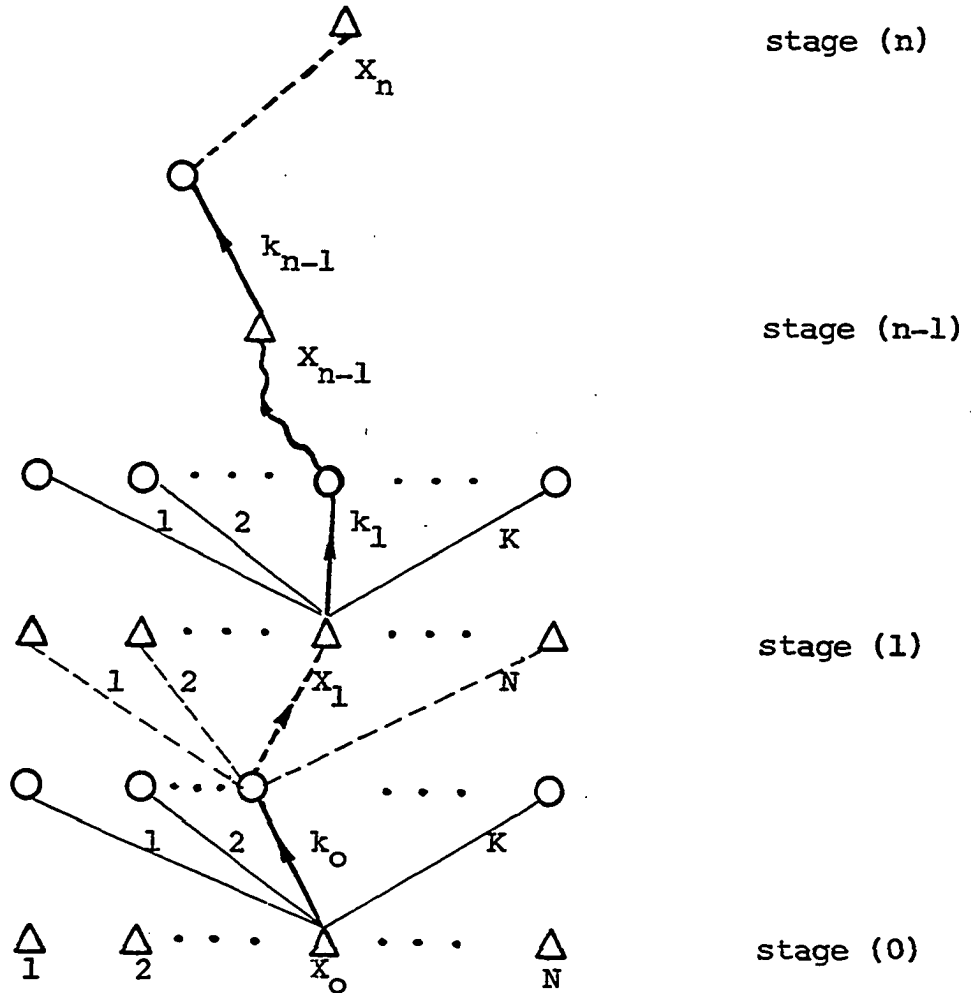
$$X_t \in \{1, 2, \dots, N\}$$

$$t \in \{0, 1, 2, \dots\} .$$

$k_t$  = Decision taken at time  $t$  or at the beginning of period number  $t+1$

$$k_t \in \{1, 2, \dots, K\}$$

$$t \in \{0, 1, 2, \dots\} .$$



$\triangle$  decision and reward nodes  
 $\circ$  chance nodes  
 $\longrightarrow$  decision made  
 $\dashrightarrow$  state materialized  
 reward  $r_{x_{h-1}x_h}^{k_{h-1}}$  are collected immediately after the transition  
 $x_{h-1} - x_h, h:0,1,2,\dots,n$

Diagram 3. Part of the decision tree for n stages

After  $n$  periods of times we will have a system state history

$$X_n = \{x_0, x_1, x_2, \dots, x_n\} .$$

The likelihood of that state history will depend on  $\mathcal{P}$  and we obtain

$$\begin{aligned} (x_0, x_1, x_2, \dots, x_n | \mathcal{P}) &= \overset{k_0}{p_{x_0, x_1}} \cdot \overset{k_1}{p_{x_1, x_2}} \cdots \overset{k_t}{p_{x_t, x_{t+1}}} \\ &\cdots \overset{k_{n-1}}{p_{x_{n-1}, x_n}} \end{aligned} \quad (3.1)$$

where

$$\overset{k_t}{p_{x_t, x_{t+1}}}$$

is the probability that the system is in state  $x_{t+1}$  at time  $t+1$ , given that  $\mathcal{P}$  underlies the process, and that the system is in state  $x_t$  and alternative  $k_t$  is chosen at time  $t$ .

Let us also define the transition count matrix

$\mathcal{N} =$

$n_{11}^1$	$n_{12}^1$	...	$n_{1N}^1$
$n_{21}^1$	$n_{22}^1$	...	$n_{2N}^1$
·		$n_{ij}^1$	·
·			·
$n_{N1}^1$	$n_{N2}^1$	...	$n_{NN}^1$
·			
·			
·			
$n_{11}^k$	$n_{12}^k$	...	$n_{1N}^k$
$n_{21}^k$	$n_{22}^k$	...	$n_{2N}^k$
·		$n_{ij}^k$	·
·			·
$n_{N1}^k$	$n_{N2}^k$	...	$n_{NN}^k$
·			
·			
·			
$n_{11}^K$	$n_{12}^K$	...	$n_{1N}^K$
$n_{21}^K$	$n_{22}^K$	...	$n_{2N}^K$
·		$n_{ij}^K$	·
·			·
$n_{N1}^K$	$n_{N2}^K$	...	$n_{NN}^K$

action 1 is taken

action k is taken

action K is taken

(3.2)

where

$n_{ij}^k$  = the number of transitions observed from state  $i$  to state  $j$  when alternative  $k$  is chosen in state  $i$ .

Note that

$$\sum_{k=1}^K \sum_{j=1}^N \sum_{i=1}^N n_{ij}^k = n .$$

After the decision maker observes the process for  $n$  periods of time, he should be able to adjust his prior knowledge according to the resulting state history,  $(x_0, x_1, x_2, \dots, x_n)$ . The resulting posterior knowledge is quantified by a new probability distribution  $\{\eta'_1, \eta'_2, \dots, \eta'_j, \dots, \eta'_{j^*}, \dots, \eta'_L\}$ .

The object of this chapter is the study of the rate of convergence of this posterior distribution to  $\{0, 0, \dots, 0, \dots, 1, \dots, 0\}$  where unity occurs at position  $j^*$ .

The (posterior) probability distribution  $\{\eta'_1, \eta'_2, \dots, \eta'_j, \dots, \eta'_{j^*}, \dots, \eta'_L\}$  is given by

$$\eta'_j = \frac{\eta_j \mathcal{L}(x_0, x_1, x_2, \dots, x_n | \mathcal{P}_j)}{\sum_{\text{all } l} \eta_l \mathcal{L}(x_0, x_1, x_2, \dots, x_n | \mathcal{P}_l)} ; j = 1, 2, \dots, L$$

(3.3)

and in particular,



$$\begin{aligned}
\eta'_{/*} &= \eta_{/*} \mathcal{L}(x_0, x_1, x_2, \dots, x_n / \mathcal{P}_{/*}) / [\eta_1 \mathcal{L}(x_0, x_1, x_2, \dots, x_n / \mathcal{P}_1) \\
&+ \dots + \eta_{/} \mathcal{L}(x_0, x_1, x_2, \dots, x_n / \mathcal{P}) + \dots \\
&+ \eta_L \mathcal{L}(x_0, x_1, x_2, \dots, x_n / \mathcal{P}_L)] \\
&= \eta_{/*} (*\overset{k_0}{p}_{x_0, x_1} \dots *\overset{k_{n-1}}{p}_{x_{n-1}, x_n}) / [\eta_1 (1^{\overset{k_0}{p}_{x_0, x_1}} \dots 1^{\overset{k_{n-1}}{p}_{x_{n-1}, x_n}}) \\
&+ \dots + \eta_{/} (\overset{k_0}{p}_{x_0, x_1} \dots \overset{k_{n-1}}{p}_{x_{n-1}, x_n}) + \dots \\
&+ \eta_{/*} (*\overset{k_0}{p}_{x_0, x_1} \dots *\overset{k_{n-1}}{p}_{x_{n-1}, x_n}) + \dots \\
&+ \eta_L (L^{\overset{k_0}{p}_{x_0, x_1}} \dots L^{\overset{k_{n-1}}{p}_{x_{n-1}, x_n}})] \quad \dots \quad (3.4)
\end{aligned}$$

$$\begin{aligned}
&\equiv \eta_{/*} / \{ \eta_1 [ \left( \frac{1^{\overset{k_0}{p}}}{*\overset{k_0}{p}} \right)_{x_0, x_1} \dots \left( \frac{1^{\overset{k_{n-1}}{p}}}{*\overset{k_{n-1}}{p}} \right)_{x_{n-1}, x_n} ] + \dots \\
&+ \eta_{/} [ \left( \frac{1^{\overset{k_0}{p}}}{*\overset{k_0}{p}} \right)_{x_0, x_1} \dots \left( \frac{1^{\overset{k_{n-1}}{p}}}{*\overset{k_{n-1}}{p}} \right)_{x_{n-1}, x_n} ] + \dots + \eta_{*} + \dots \\
&+ \eta_L [ \left( \frac{L^{\overset{k_0}{p}}}{*\overset{k_0}{p}} \right)_{x_0, x_1} \dots \left( \frac{L^{\overset{k_{n-1}}{p}}}{*\overset{k_{n-1}}{p}} \right)_{x_{n-1}, x_n} ] \} \quad \dots \quad (3.5)
\end{aligned}$$

where we have assumed  $\psi_{ij}^k > 0$ , all  $i, j, k$ , and for example

$$\left( \frac{\binom{p}{*p}}{\binom{p}{*p}} \right)_{x_i, x_{i+1}}^{k_0} \equiv \left( \frac{\binom{p}{x_i, x_{i+1}}^{k_0}}{\binom{p}{*p}} \right)_{x_i, x_{i+1}}.$$

Now let us look at the quantity

$$\begin{aligned} & \eta_{/} \left[ \left( \frac{\binom{p}{*p}}{\binom{p}{*p}} \right)_{x_0, x_1}^{k_0} \cdots \left( \frac{\binom{p}{*p}}{\binom{p}{*p}} \right)_{x_i, x_{i+1}}^{k_i} \cdots \left( \frac{\binom{p}{*p}}{\binom{p}{*p}} \right)_{x_{n-1}, x_n}^{k_{n-1}} \right] \\ &= \eta_{/} \left\{ \prod_{i,j,k} \left[ \left( \frac{\binom{p}{*p}}{\binom{p}{*p}} \right)_{i,j}^k \right]^{n_{ij}^k} \right\} \\ &= \eta_{/} \exp \left[ \sum_{i,j,k} n_{ij}^k \ln \left( \frac{\binom{p}{*p}}{\binom{p}{*p}} \right)_{i,j}^k \right] \quad \dots \quad (3.6) \end{aligned}$$

where, as defined above,  $n_{ij}^k$  is the number of transitions from state  $i$  to state  $j$  under alternative  $k$ .

The exponent of the quantity (2.6) can be viewed as

$$\sum_{\alpha=1}^n \psi_{\alpha}^{Z_{\alpha}}$$

where the joint distribution of the  $Z_{\alpha}$  is such that the conditional distribution of any  $Z_{\alpha}$ , given all previous  $Z$ 's, is one of the K.N distributions  $\psi_i^k$  given in Table 1.

Table 1. The set  $\{\psi_i^k\}$ 

Distribution Name	Domain of Distribution	Probabilities
$\psi_1^1$	$\ln\left(\frac{p}{*p}\right)_{11}^1 \equiv a_{11}^1$ $\vdots$ $\ln\left(\frac{p}{*p}\right)_{1N}^1 \equiv a_{1N}^1$	$*p_{11}^1$ $\vdots$ $*p_{1N}^1$
$\vdots$	$\vdots$	$\vdots$
$\psi_i^k$	$\ln\left(\frac{p}{*p}\right)_{i1}^k \equiv a_{i1}^k$ $\vdots$ $\ln\left(\frac{p}{*p}\right)_{iN}^k \equiv a_{iN}^k$	$*p_{i1}^k$ $\vdots$ $*p_{iN}^k$
$\vdots$	$\vdots$	$\vdots$
$\psi_N^K$	$\ln\left(\frac{p}{*p}\right)_{N1}^K \equiv a_{N1}^K$ $\vdots$ $\ln\left(\frac{p}{*p}\right)_{NN}^K \equiv a_{NN}^K$	$*p_{N1}^K$ $\vdots$ $*p_{NN}^K$

By the above explanation and in reference to Table 1, the quantity

$$\sum_{\alpha=1}^n Z_{\alpha}$$

can be looked at as the sum of  $n$  steps of a "mixed" random walk featuring a set of  $(K.N)$  conditional distributions governing the steps of the walk:

$$\psi = \begin{bmatrix} \psi_1^1 \\ \cdot \\ \cdot \\ \psi_i^k \\ \cdot \\ \cdot \\ \psi_N^K \end{bmatrix} \quad (3.7)$$

(N.K) x 1

involving probabilities taken from  $\mathcal{P}_i^*$  and step sizes taken from the following matrix  $\mathcal{A}$ :

$$\mathcal{A}_i = \begin{array}{c}
 \begin{array}{cccc}
 1 & /a_{11}^1 & /a_{12}^1 & \dots & /a_{1N}^1 \\
 2 & /a_{21}^1 & /a_{22}^1 & \dots & /a_{2N}^1 \\
 \vdots & \vdots & & /a_{ij}^1 & \vdots \\
 N & /a_{N1}^1 & /a_{N2}^1 & \dots & /a_{NN}^1
 \end{array} & 1 \\
 \vdots & & & & \\
 \begin{array}{cccc}
 1 & /a_{11}^k & /a_{12}^k & \dots & /a_{1N}^k \\
 2 & /a_{21}^k & /a_{22}^k & \dots & /a_{2N}^k \\
 \vdots & \vdots & & /a_{ij}^k & \vdots \\
 N & /a_{N1}^k & /a_{N2}^k & \dots & /a_{NN}^k
 \end{array} & k \\
 \vdots & & & & \\
 \begin{array}{cccc}
 1 & /a_{11}^K & /a_{12}^K & \dots & /a_{1N}^K \\
 2 & /a_{21}^K & /a_{22}^K & \dots & /a_{2N}^K \\
 \vdots & \vdots & & /a_{ij}^K & \vdots \\
 N & /a_{N1}^K & /a_{N2}^K & \dots & /a_{NN}^K
 \end{array} & K \\
 \dots & & & & \dots & (3.8)
 \end{array}$$

The conditional expectation of  $Z_\alpha$ , given that it is governed by the conditional distribution  $\psi_i^k$ , is:

$$E(Z_\alpha / Z_\alpha \in \psi_i^k) = \sum_{j=1}^N (a_{ij}^k * p_{ij}^k) \dots (3.9)$$

and can be shown to be negative, and so also the expectation of the sum

$$\sum_{\alpha=1}^n \gamma_{\alpha}$$

The proof will be left to the next section when the special case ( $N=2, K=2, L=2$ ), is considered.

Equation 3.5 may now be written as

$$\eta'_{\gamma^*} = \frac{\eta_{\gamma^*}}{\eta_1 \cdot \exp\left(\sum_{\alpha=1}^n \gamma_{\alpha}\right) + \dots + \eta_{\gamma} \cdot \exp\left(\sum_{\alpha=1}^n \gamma_{\alpha}\right) + \dots + \eta_{\gamma^*} + \dots + \eta_L \cdot \exp\left(\sum_{\alpha=1}^n \gamma_{\alpha}\right)}$$

(3.10)

It can be shown that each of the sums

$$\sum_{i=1}^n \gamma_{\alpha}$$

(3.11)

$\forall \gamma$  other than  $\gamma^*$ , will tend to  $(-\infty)$  as  $n$  tends to  $\infty$ , and thus

$$\lim_{n \rightarrow \infty} \eta'_{\gamma^*} \rightarrow 1$$

$$\uparrow \downarrow$$

$$\lim_{n \rightarrow \infty} \eta'_{\gamma} \rightarrow 0, \quad \forall \gamma \text{ other than } \gamma^*. \quad (3.12)$$

This will also be considered below in the above special case.

## C. Posterior Convergence; The Special

Case  $(N,K,L) = (2,2,2)$ 

Consider without loss of generality the special case with  $N=2$ , i.e., when the state space consists only of the two states  $\underline{1}$  and  $\underline{2}$ ; with  $K=2$ , i.e., where there are two alternative actions in each state, and with  $L=2$ , i.e., where there are only two generalized stochastic matrices to be considered,  $\overline{\Pi}$  and  $\mathcal{P}$ , with corresponding prior distribution  $\eta, (1-\eta)$ .

The decision maker's prior knowledge about the process is that there are two generalized matrices,  $\overline{\Pi}, \mathcal{P}$ , of which one will be used by the process. However our analysis proceeds under the assumption that one of the two, say  $\overline{\Pi}$ , underlies the process, which fact is denoted by  $\mathcal{P}^* = \overline{\Pi}$ , in accordance with previous notation. Finally we assume that the decision maker is a Bayesian, putting prior  $\eta$  on

$$\mathcal{P}^* \equiv \overline{\Pi} \equiv \begin{bmatrix} \pi_{11}^1 & \pi_{12}^1 \\ \pi_{21}^1 & \pi_{22}^1 \\ \pi_{11}^2 & \pi_{12}^2 \\ \pi_{21}^2 & \pi_{22}^2 \end{bmatrix} \quad \begin{array}{l} k=1 \\ \\ k=2 \end{array} \quad (3.13)$$

and prior  $1-\eta$  on

$$\mathcal{P} = \begin{bmatrix} p_{11}^1 & p_{12}^1 \\ p_{21}^1 & p_{22}^1 \\ p_{11}^2 & p_{12}^2 \\ p_{21}^2 & p_{22}^2 \end{bmatrix} \quad \begin{array}{l} k=1 \\ \\ k=2 \end{array} \quad (3.14)$$

Again, let the state history be  $X_{0,n} = (x_0, x_1, x_2, \dots, x_n)$  and let

$$\mathcal{N} = \begin{bmatrix} n_{11}^1 & n_{12}^1 \\ n_{21}^1 & n_{22}^1 \\ n_{11}^2 & n_{12}^2 \\ n_{21}^2 & n_{22}^2 \end{bmatrix} \quad (3.15)$$

be the corresponding transition count matrix. Then the likelihood of the sample state history can be written

$$\mathcal{L}(x_0, x_1, x_2, \dots, x_n / \bar{\Pi}) = (\pi_{11}^1)^{n_{11}^1} \cdot (\pi_{12}^1)^{n_{12}^1} \dots (\pi_{22}^2)^{n_{22}^2}$$

and

$$\mathcal{L}(x_0, x_1, x_2, \dots, x_n / \mathcal{P}) = (p_{11}^1)^{n_{11}^1} \cdot (p_{12}^1)^{n_{12}^1} \dots (p_{22}^2)^{n_{22}^2}$$

and the posterior probabilities of  $\bar{\Pi}$  and  $\mathcal{P}$  given the state



history  $x_{0,n}$  are

$$\begin{aligned}
 \eta' &= \frac{\eta \cdot \mathcal{L}(x_0, x_1, x_2, \dots, x_n / \mathcal{P})}{\eta \cdot \mathcal{L}(x_0, x_1, x_2, \dots, x_n / \mathcal{P}) + (1-\eta) \cdot \mathcal{L}(x_0, x_1, x_2, \dots, x_n / \mathcal{Q})} \\
 &= \frac{\eta \cdot \prod_{i,j,k=1}^2 (\pi_{ij}^k)^{n_{ij}^k}}{\eta \cdot \prod_{i,j,k=1}^2 (\pi_{ij}^k)^{n_{ij}^k} + (1-\eta) \cdot \prod_{i,j,k=1}^2 (p_{ij}^k)^{n_{ij}^k}} \\
 &= \frac{\eta}{\eta + (1-\eta) \prod_{i,j,k=1}^2 \left( \frac{p_{ij}^k}{\pi_{ij}^k} \right)^{n_{ij}^k}} \\
 &= \frac{\eta}{\eta + (1-\eta) \exp \left[ \sum_{i,j,k=1}^2 \ln \left( \frac{p_{ij}^k}{\pi_{ij}^k} \right)^{n_{ij}^k} \right]} \\
 &= \frac{\eta}{\eta + (1-\eta) \exp \left[ \sum_{\alpha=1}^n Z_{\alpha} \right]} \tag{3.16}
 \end{aligned}$$

and

$$(1-\eta)' = 1-\eta' \quad , \tag{3.17}$$

where, as before, the quantity

$$\sum_{\alpha=1}^n Z_{\alpha}$$

will be considered as the sum of  $n$  steps of a "mixed" random walk governed by four conditional distributions  $\psi_i^k$ , with probabilities provided by the four rows of  $\overline{\Pi}$ , and domains (step-size domains) provided by the four rows of the step-size matrix  $\mathcal{A}$ ,

$$\mathcal{A} \equiv \begin{bmatrix} \ln\left(\frac{p_{11}^1}{\pi_{11}^1}\right) & \ln\left(\frac{p_{12}^1}{\pi_{12}^1}\right) \\ \ln\left(\frac{p_{21}^1}{\pi_{21}^1}\right) & \ln\left(\frac{p_{22}^1}{\pi_{22}^1}\right) \\ \ln\left(\frac{p_{11}^2}{\pi_{11}^2}\right) & \ln\left(\frac{p_{12}^2}{\pi_{12}^2}\right) \\ \ln\left(\frac{p_{21}^2}{\pi_{21}^2}\right) & \ln\left(\frac{p_{22}^2}{\pi_{22}^2}\right) \end{bmatrix} \equiv \begin{bmatrix} a_{11}^1 & a_{12}^1 \\ a_{21}^1 & a_{22}^1 \\ a_{11}^2 & a_{12}^2 \\ a_{21}^2 & a_{22}^2 \end{bmatrix} \quad (3.18)$$

Thus, recalling earlier notation (Equation 3.9),

$$\begin{aligned} E[Z_\alpha / Z_\alpha \varepsilon \psi_1^1] \\ = \sum_{j=1}^2 \ln\left(\frac{p_{1j}^1}{\pi_{1j}^1}\right) \cdot \pi_{1j}^1 \\ \vdots \end{aligned}$$

$$\begin{aligned}
E[Z_\alpha / Z_\alpha \varepsilon \psi_2^2] \\
= \sum_{j=1}^2 \ln\left(\frac{p_{2j}^2}{\pi_{2j}^2}\right) \cdot \pi_{2j}^2 \quad . \quad (3.19)
\end{aligned}$$

It may now be observed that

$$E[Z_\alpha / Z_\alpha \varepsilon \psi_i^k] \quad , \quad i, k=1, 2$$

is a continuous function of  $p_{il}^k$ , attains a maximum value equal to zero when  $p_{il}^k = \pi_{il}^k$ , and is negative otherwise. In other words its maximum will occur when the matrices  $\mathcal{P}$  and  $\overline{\Pi}$  have identical elements. The above fact, related to likelihood maximization, may be proved as follows. Let

$$\begin{aligned}
\mu_i^k &= E[Z_\alpha / Z_\alpha \varepsilon \psi_i^k] \\
&= \ln\left(\frac{p_{i1}^k}{\pi_{i1}^k}\right) \cdot \pi_{i1}^k + \ln\left(\frac{p_{i2}^k}{\pi_{i2}^k}\right) \cdot \pi_{i2}^k \\
&= \ln\left(\frac{p_{i1}^k}{\pi_{i1}^k}\right) \cdot \pi_{i1}^k + \ln\left(\frac{1-p_{i1}^k}{1-\pi_{i1}^k}\right) \cdot (1-\pi_{i1}^k) \quad , \quad 0 < \pi_{i1}^k < 1 \quad . \\
& \quad \quad \quad (3.20)
\end{aligned}$$

Then, for  $0 < p_{i1}^k < 1$ ,

$$\frac{d(\mu_i^k)}{dp_{il}^k} = \left(\frac{\pi_{il}^k}{p_{il}^k}\right) - \left(\frac{1-\pi_{il}^k}{1-p_{il}^k}\right) \quad (3.21)$$

$$= 0$$

↓

$$p_{il}^k = \pi_{il}^k$$

↓

$$\mu_i^k = 0 \quad \text{at} \quad p_{il}^k = \pi_{il}^k \quad (3.22)$$

and

$$\begin{aligned} \frac{d^2(\mu_i^k)}{(dp_{il}^k)^2} &= \frac{d}{dp_{il}^k} \left( \frac{d(\mu_i^k)}{dp_{il}^k} \right) \\ &= \frac{d}{dp_{il}^k} \left[ \left( \frac{\pi_{il}^k}{p_{il}^k} \right) - \left( \frac{1-\pi_{il}^k}{1-p_{il}^k} \right) \right] \\ &= -\pi_{il}^k \cdot \frac{1}{(p_{il}^k)^2} - (1-\pi_{il}^k) \cdot \frac{1}{(1-p_{il}^k)^2} \quad (3.23) \end{aligned}$$

Thus the R.H.S. of Equation 3.20 has a unique maximum at  $p_{il}^k = \pi_{il}^k$  in the open interval  $0 < p_{il}^k < 1$ , where in fact it is zero. Since, in addition, Equation 3.20 tends to  $-\infty$  for  $p_{il}^k$  tending to zero or one, it follows that Equation 3.20 is nonpositive in the corresponding closed interval, and zero only at  $p_{il}^k = \pi_{il}^k$ . Kullback (1959) has studied general

versions of such information inequalities. Since the value  $\mu_i^k$  has this characteristic then the sum

$$\sum_{\alpha=1}^n Z_{\alpha}$$

has a nonpositive expectation and

$$n \cdot \{\min_{i,k} [\mu_i^k]\} \leq E \left[ \sum_{\alpha=1}^n Z_{\alpha} \right] \leq n \cdot \{\max_{i,k} [\mu_i^k]\} \quad . \quad (3.24)$$

#### D. Rate of Almost-Sure Convergence of the Posterior Distribution

Consider the "mixed" random walk

$$\sum_{\alpha=1}^n Z_{\alpha}$$

underlying the decision process that has been introduced in Section C. The evolution of this random walk involves not only its partial sums

$$\sum_{\alpha=1}^n Z_{\alpha}$$

on the real line, but also the states 1 and 2. Hence a more detailed description of the progress of the walk in the first  $n$  steps involves both the partial sum achieved, call it

$$\mathcal{I}_n = \sum_{\alpha=1}^n z_{\alpha} ,$$

and the current state, call it  $i_n$ . This more detailed vector description  $(\mathcal{I}_n, i_n)$  will be of use in this section.

The purpose of this section is to show that the sum  $\mathcal{I}_n$  tends to  $(-\infty)$  almost surely as  $n$  approaches  $\infty$ , and hence that the posterior  $\eta'$  tends to unity; in particular the rate of this convergence will be examined. Note that the sum  $\mathcal{I}_n$  determines the posterior (Equation 3.16) and consequently the decision maker's choice among the available alternatives, as indicated in Equation 1.10.

As indicated in Chapter I, it will now be assumed that there is a posterior neighborhood of unity, i.e., a neighborhood  $\mathcal{J} = [\eta'_0, 1]$  for the posterior  $\eta'$ , where the Bayes policy is state stationary. This means that, whenever the posterior is in  $\mathcal{J}$ , a particular decision  $k(1)$  is called for by the Bayes strategy when the process is in state  $\underline{1}$ , and a particular decision  $k(2)$  is called for by the Bayes strategy when the process is in state  $\underline{2}$ . In view of the correspondence between the partial sum  $\mathcal{I}_n$  and the posterior  $\eta'$  (Equation 3.16) this assumption can be reformulated as follows: there is a value  $\mathcal{J}^* < 0$  with the property that, given  $(\mathcal{I}_n, i_n)$  with  $\mathcal{I}_n < \mathcal{J}^*$ , the next step  $z_{n+1}$  materializes in accordance with one specific square transition matrix and

a corresponding conditional probability distribution. In other words, conditionally on any  $(\mathcal{I}_n, i_n)$  with  $\mathcal{I}_n \leq \mathcal{I}^*$ , the probability of any event  $\mathcal{E}$  involving the further vector states of the random walk may be computed as if the Markov chain underlying the walk were stationary (indeed as if  $(\mathcal{I}_n, i_n)$  were a stationary Markov process), as long as  $\mathcal{E}$  belongs to the event

$$(\mathcal{I}_{n+1}, \mathcal{I}_{n+2}, \dots \leq \mathcal{I}^*) \quad . \quad (3.25)$$

The  $(2 \times 2)$  transition matrices associated with the stationary situation mentioned above are denoted by  $\pi_s$  and  $P_s$  with corresponding posterior probabilities  $\eta'$  and  $(1-\eta')$  respectively where

$$\pi_s \equiv \begin{bmatrix} \pi_{11}^{k(1)} & \pi_{12}^{k(1)} \\ \pi_{21}^{k(2)} & \pi_{22}^{k(2)} \end{bmatrix} \quad (3.26)$$

and

$$P_s \equiv \begin{bmatrix} p_{11}^{k(1)} & p_{12}^{k(1)} \\ p_{21}^{k(2)} & p_{22}^{k(2)} \end{bmatrix} \quad . \quad (3.27)$$

The corresponding step size matrix is

$$A_s \equiv [a_{ij}^{k(s)}] \equiv \begin{bmatrix} a_{11}^{k(1)} & a_{12}^{k(1)} \\ a_{21}^{k(2)} & a_{22}^{k(2)} \end{bmatrix} \equiv \begin{bmatrix} \ln\left(\frac{p_{11}^{k(1)}}{\pi_{11}^{k(1)}}\right) & \ln\left(\frac{p_{11}^{k(1)}}{\pi_{11}^{k(1)}}\right) \\ \ln\left(\frac{p_{12}^{k(2)}}{\pi_{12}^{k(2)}}\right) & \ln\left(\frac{p_{22}^{k(2)}}{\pi_{22}^{k(2)}}\right) \end{bmatrix} \quad (3.28)$$

The main idea of the procedure adopted in this section is to exhibit a boundary  $\mathcal{L}$ , composed of three segments  $\mathcal{L}_1$ ,  $\mathcal{L}_2$  and  $\mathcal{L}_3$  with the property that  $\mathcal{I}_n$  stays below  $\mathcal{L}$  with probability near 1. This boundary leads directly to the main theorem of the chapter which states, essentially, that with probability near 1,  $\eta'$  approaches unity at an exponential rate determined by the "eventual" expected size of the steps  $Z_\alpha$ .

The construction of the first part of the boundary ( $\mathcal{L}_1$ ) is based on a theorem of Dubins and Savage (1965a, 1965b) which is also applied in an article by Dubins and Freedman (1965). Following will be a restatement of the Dubins-Savage theorem without proof and with appropriate changes in wording to suit the present exposition.

**Theorem 3.1:** Let  $Z_1, Z_2, \dots, Z_\alpha, \dots, Z_n$  be a real valued stochastic process. Let  $\mu_\alpha$  be the conditional expectation of  $Z_\alpha$  given the past and  $V_\alpha$  the conditional variance of  $Z_\alpha$  given the past. Suppose that for every  $\alpha$ ,  $\mu_\alpha$  is finite almost surely. (No assumption is needed about  $V_\alpha$ .) Let



$a_1, a_2$  be positive numbers. Then the probability that, for all  $n$ ,

$$(Z_1 + Z_2 + \dots + Z_n) \leq a_1 + (\mu_1 + \mu_2 + \dots + \mu_n) + a_2(V_1 + V_2 + \dots + V_n)$$

is greater than or equal to  $\frac{a_1 a_2}{1 + a_1 a_2}$  . (3.29)

Using the notation of Section C of this chapter let

$$\tilde{\mu} = \max_{i,k} \{\mu_i^k\} \quad i=1,2;k=1,2 \quad (3.30)$$

be the maximum of the expectations of the conditional distributions  $\psi_i^k$ . One may note that Equation 3.30 implies, viewing our random walk as the process of the theorem, that

$$\tilde{\mu} \geq \max_{\alpha} \{\mu_{\alpha}\} \quad \alpha=1,2,\dots,n;\forall n \quad (3.31)$$

Let also

$$\tilde{V} = \max_{i,k} \{\text{Var}(Z_{\alpha}) | Z_{\alpha} \in \psi_i^k\} \quad , \quad i=1,2;k=1,2 \quad (3.32)$$

be the maximum of the variances of the conditional distributions  $\psi_i^k$ . Note also that Equation 3.32 implies that

$$\tilde{V} \geq \max_{\alpha} \{V_{\alpha}\} \quad \alpha=1,2,\dots,n;\forall n \quad (3.33)$$

The inequality in Equation 3.29 then implies that

$$\Pr\left\{ \sum_{\alpha=1}^n Z_{\alpha} \leq a_1 + n(\tilde{\mu} + a_2 \tilde{V}); \text{ all } n \right\} \geq \frac{a_1 a_2}{1 + a_1 a_2} \quad (3.34)$$

Of special interest for us are choices for  $a_1$  and  $a_2$  such that  $a_2$  is small enough that  $\tilde{\mu} + a_2\tilde{\nu}$  is negative, and  $a_1$  is large enough so that

$$\frac{a_1 a_2}{1 + a_1 a_2}$$

is arbitrary large.

Given such values for  $a_2$  and  $a_1$  (i.e.,  $a_2$  "small enough" and  $a_1$  "large enough"), the linear expression  $a_1 + n(\tilde{\mu} + a_2\tilde{\nu})$  provides us with the portion  $\mathcal{L}_1$  of our boundary, as is shown in Diagram 4.

The following remarks may be made concerning  $\mathcal{L}_1$ :  $\mathcal{L}_1$  is by itself a possible boundary providing a rate-of-convergence upper bound arbitrarily near  $\tilde{\mu}$  for  $\mathcal{I}_n$ , which in turn provides an exponential rate-of-convergence upper bound for  $\eta'$ . However we choose to utilize  $\mathcal{L}_1$  only to take us through the "nonstationary" region of our random walk, relying on the further construction of the boundary portion  $\mathcal{L}_3$  to provide us with a sharper rate-of-convergence upper bound.

For the construction of the next part  $\mathcal{L}_2$  of  $\mathcal{L}$  we consider the maximum step size  $a^*$ , which must be positive, of the four "stationary" steps  $a_{ij}^{k(s)}$  discussed earlier in this section (Equation 3.28):

$$a^* = \max\left\{\ln\left(\frac{p_{11}^{k(1)}}{\pi_{11}^{k(1)}}\right), \ln\left(\frac{p_{12}^{k(1)}}{\pi_{12}^{k(1)}}\right), \ln\left(\frac{p_{21}^{k(2)}}{\pi_{21}^{k(2)}}\right), \ln\left(\frac{p_{22}^{k(2)}}{\pi_{22}^{k(2)}}\right)\right\}. \quad (3.35)$$

We then fix on some large number  $M$ , to be specified later on, and extend the linear boundary segment  $\mathcal{L}_1$  to the highest point that (1) is at least  $Ma^*$  below the level  $\mathcal{S}$  and (2) has an integer abscissa, see Diagram 4. The abscissa of the point  $0_1$ , the sampling state  $n_s$ , is then given by the equation

$$n_s = \left[ \frac{\mathcal{S}^* - Ma^* - a_1}{\tilde{\mu} + a_2\tilde{v}} \right] + 1 \quad (3.36)$$

where  $[ ]$  denotes "largest integer less than".

The point  $0_1 : [n_s, a_1 + n_s(\tilde{\mu} + a_2\tilde{v})]$  actually is the "right" end point of the boundary segment  $\mathcal{L}_1$ , as well as the "left" end point of the boundary segment  $\mathcal{L}_2$ . The "right" end point of the segment  $\mathcal{L}_2$  will be located at the point

$$0_1^* : [n_s + M, a_1 + n_s \cdot (\tilde{\mu} + a_2\tilde{v}) + Ma^*] \quad , \quad (3.37)$$

which according to the value of  $n_s$  in Equation 3.36 will be in the stationary region, i.e., below or at the level  $\mathcal{S}^*$ . With the above locations of  $0_1, 0_1^*$ , the boundary portion  $\mathcal{L}_2$  of  $\mathcal{L}$  will be a line segment  $0_1-0_1^*$  with positive slope  $a^*$  defined by the Equation 3.35. One may further notice that a random walk proceeding only by steps of magnitudes  $a_{ij}^{k(s)}$  which starts at  $0_1$  cannot reach the point  $0_1^*$  unless it proceeds exclusively by  $M$  steps of size  $a^*$ .

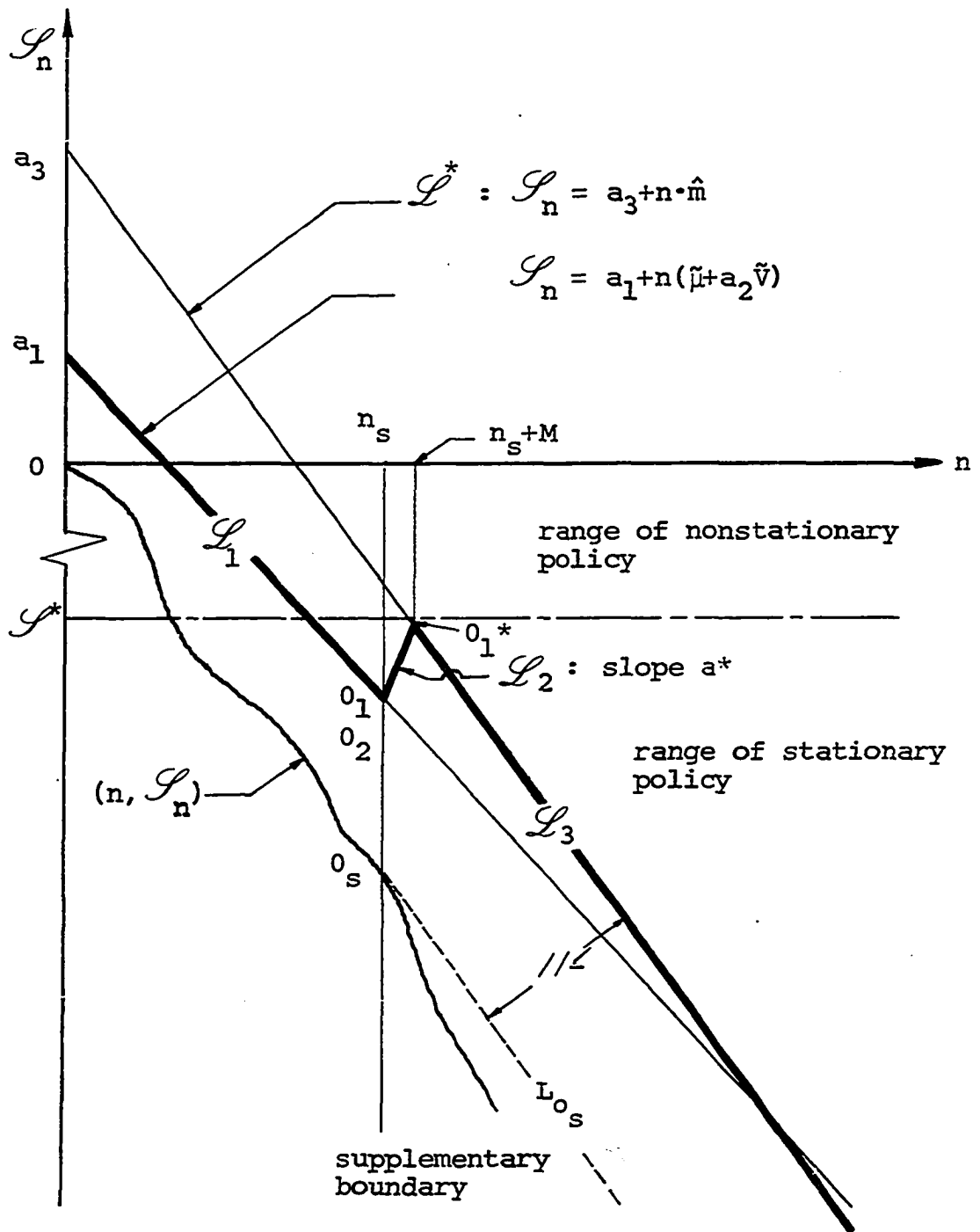


Diagram 4. The boundary  $L : L_1 - L_2 - L_3$

As for the third segment  $\mathcal{L}_3$ , let  $m$  be defined as in Equation 2.35 with the following identifications:

$$(1) \begin{bmatrix} \theta_1 & (1-\theta_1) \\ (1-\theta_2) & \theta_2 \end{bmatrix} = \begin{bmatrix} \pi_{11}^{k(1)} & \pi_{12}^{k(1)} \\ \pi_{21}^{k(2)} & \pi_{22}^{k(2)} \end{bmatrix} \quad (3.38)$$

(2) The step sizes  $a$ ,  $b$ ,  $c$  and  $d$  are replaced by the elements of the matrix of Equation 3.28.

In other words let  $m$  be defined by

$$\begin{aligned} m = & \frac{[1-\pi_{22}^{k(2)}]}{[1-\pi_{11}^{k(1)}]+[1-\pi_{22}^{k(2)}]} \left\{ \left[ \ln\left(\frac{p_{11}^{k(1)}}{\pi_{11}^{k(1)}}\right) \right] \cdot \pi_{11}^{k(1)} \right. \\ & \left. + \left[ \ln\left(\frac{p_{12}^{k(1)}}{\pi_{12}^{k(1)}}\right) \right] \cdot \pi_{12}^{k(1)} \right\} \\ & + \frac{[1-\pi_{11}^{k(1)}]}{[1-\pi_{11}^{k(1)}]+[1-\pi_{22}^{k(2)}]} \left\{ \left[ \ln\left(\frac{p_{21}^{k(2)}}{\pi_{21}^{k(2)}}\right) \right] \cdot \pi_{21}^{k(2)} \right. \\ & \left. + \left[ \ln\left(\frac{p_{22}^{k(2)}}{\pi_{22}^{k(2)}}\right) \right] \cdot \pi_{22}^{k(2)} \right\} \quad (3.39) \end{aligned}$$

The third boundary segment  $\mathcal{L}_3$  is another linear segment, defined as follows in terms of  $m$ : it starts at  $0_1^*$  and extends downward with a negative slope  $\hat{m} = m + \varepsilon$ . This segment of the boundary may be looked at as a portion of a

line given by the following equation:

$$\mathcal{I}_n = a_3 + n \cdot \hat{m} \quad (3.40)$$

as demonstrated in Diagram 4. One would notice that, with reference to the coordinates of the point  $0_1^*$  given in Equation 3.37 together with the equation of the line in 3.40, the value of  $a_3$  is given by:

$$a_3 = a_1 + n_s(\tilde{\mu} + a_2\tilde{V}) + Ma^* - (n_s + M) \cdot \hat{m} \quad (3.41)$$

which is greater than  $a_1$ . It is clear that this "modified" boundary  $\mathcal{L}^*$  described by Equation 3.40 will also be an upper bound for  $\mathcal{I}_n, \Psi_n$  with probability near 1 if  $\mathcal{L}$  is so.

We are now ready to estimate the probability that no partial sum  $\mathcal{I}_n$  exceeds the boundary  $\mathcal{L}: \mathcal{L}_1 - \mathcal{L}_2 - \mathcal{L}_3$  composed of the successive linear segments  $\mathcal{L}_1, \mathcal{L}_2,$  and  $\mathcal{L}_3$ .

Let  $\mathcal{I}_s$  be a possible value of  $\mathcal{I}_n$  at  $n = n_s$ , given that  $\mathcal{I}_n$  has not exceeded  $\mathcal{L}_1$  in the previous  $n_s$  steps. Consider now the conditioning event that  $(\mathcal{I}_n, i_n) = (\mathcal{I}_s, i_s)$  at  $n = n_s$ , i.e., being at the point  $0_s, s:1,2,\dots,$  with coordinates  $(n_s, \mathcal{I}_s)$  while the process is in state  $i_s, i_s:1,2$ . In keeping with remarks made earlier in this section, the probability that, during its progress beyond  $(n_s, \mathcal{I}_s), \mathcal{I}_n$  never exceeds  $\mathcal{L}_2 - \mathcal{L}_3$  may be computed or

bounded as if the stationary matrix (Equation 3.26) underlies the walk, with transition-dependent step sizes given by Equation 3.28. Under this simpler model the material of Chapter II can be brought to bear, and we may claim, for  $n > n_s$ , that there exists a  $\rho < 1$  such that

$$\Pr\{\mathcal{I}_n - \mathcal{I}_s > \hat{m}(n-n_s) \mid (\mathcal{I}_n, i_n) = (\mathcal{I}_s, i_s) \\ \text{at } n = n_s\} \leq \rho^{(n-n_s)} . \quad (3.42)$$

Since the above is an excess probability for the supplementary boundary segment  $L_{O_s}$  starting at  $(n_s, \mathcal{I}_s)$  and falling off with slope  $\hat{m}$ , and since  $L_{O_s}$  is below  $\mathcal{L}_2 - \mathcal{L}_3$ , it follows that for  $n > n_s$ , also

$$\Pr\{\mathcal{I}_n \text{ exceeds } \mathcal{L}_2 - \mathcal{L}_3 \mid (\mathcal{I}_n, i_n) = (\mathcal{I}_s, i_s) \\ \text{at } n = n_s\} \leq \rho^{(n-n_s)} . \quad (3.43)$$

But now  $\mathcal{L}_2 - \mathcal{L}_3$  was so constructed that, for  $n - n_s < M$ , the left hand side of Equation 3.42 is in fact zero. Hence

$$\Pr\{\mathcal{I}_n \text{ exceeds } \mathcal{L}_2 - \mathcal{L}_3 \text{ for some } n > n_s \mid \\ (\mathcal{I}_n, i_n) = (\mathcal{I}_s, i_s) \text{ at } n = n_s\} \\ \leq 0 + 0 + \dots + \rho^M + \rho^{M+1} + \rho^{M+2} + \dots + \rho^{n-n_s}$$

$$\begin{aligned}
& + \rho^{n-n_s+1} + \rho^{n-n_s+2} + \dots \\
= & \sum_{(n-n_s)=M}^{\infty} \rho^{(n-n_s)} \\
= & \frac{\rho^M}{(1-\rho)} \tag{3.44}
\end{aligned}$$

Putting this bound together with the facts deduced earlier for  $\mathcal{L}_1$  and conditioning on the process starting at the sure initial vector state  $(0, i_0)$  we find:

$$\begin{aligned}
& \Pr\{\mathcal{I}_n \text{ does not exceed } \mathcal{L}_1 - \mathcal{L}_2 - \mathcal{L}_3 \text{ for any } n\} \\
= & \sum_{(\mathcal{I}_s, i_s)} \Pr\{\mathcal{I}_n \text{ does not exceed } \mathcal{L}_1 \text{ for any } n \leq n_s \\
& \text{and } (\mathcal{I}_n, i_n) = (\mathcal{I}_s, i_s) \text{ at } n = n_s\} \cdot \\
& \Pr\{\mathcal{I}_n \text{ does not exceed } \mathcal{L}_2 - \mathcal{L}_3 \text{ for any } n > n_s | \\
& \mathcal{I}_n \text{ does not exceed } \mathcal{L}_1 \text{ for any } n \leq n_s \text{ and} \\
& (\mathcal{I}_n, i_n) = (\mathcal{I}_s, i_s) \text{ at } n = n_s\} \quad . \quad . \tag{3.45}
\end{aligned}$$

Now the Markovian nature of  $(\mathcal{I}_n, i_n)$  allows us to write the following:

$$\begin{aligned}
& \Pr\{\mathcal{I}_n \text{ does not exceed } \mathcal{L}_2 - \mathcal{L}_3 \text{ for any } n > n_s | \\
& \mathcal{I}_n \text{ does not exceed } \mathcal{L}_1 \text{ for any } n \leq n_s \text{ and}
\end{aligned}$$



$$\begin{aligned}
& (\mathcal{I}_n, i_n) = (\mathcal{I}_s, i_s) \text{ at } n = n_s \} \\
& = 1 - \Pr\{\mathcal{I}_n \text{ exceeds } \mathcal{L}_2 - \mathcal{L}_3 \text{ for some } n > n_s | \\
& (\mathcal{I}_n, i_n) = (\mathcal{I}_s, i_s) \text{ at } n = n_s \} \\
& \geq [1 - \frac{\rho^M}{(1-\rho)}] \tag{3.46}
\end{aligned}$$

in view of 3.44. Then, in view of 3.45, 3.46,

$$\begin{aligned}
& \Pr\{\mathcal{I}_n \text{ does not exceed } \mathcal{L}_1 - \mathcal{L}_2 - \mathcal{L}_3 \text{ for any } n\} \\
& \geq (\mathcal{I}_s, i_s)^\Sigma \Pr\{\mathcal{I}_n \text{ does not exceed } \mathcal{L}_1 \text{ for any } n \leq n_s \\
& \text{and } (\mathcal{I}_n, i_n) = (\mathcal{I}_s, i_s) \text{ at } n = n_s\} \cdot [1 - \frac{\rho^M}{(1-\rho)}] \tag{3.47}
\end{aligned}$$

$$\begin{aligned}
& = [1 - \frac{\rho^M}{(1-\rho)}] \cdot \Pr\{\mathcal{I}_n \text{ does not exceed } \mathcal{L}_1 \text{ for any} \\
& n \leq n_s\} \\
& \geq [1 - \frac{\rho^M}{(1-\rho)}] \cdot \Pr\{\mathcal{I}_n \text{ does not exceed } \mathcal{L}_1 \text{ for any } n\} \\
& \geq [1 - \frac{\rho^M}{(1-\rho)}] \cdot [\frac{a_1 a_2}{1 + a_1 a_2}] \equiv 1 - \delta . \tag{3.48}
\end{aligned}$$

Equation 3.48 implies that the partial sum  $\mathcal{I}_n$  tends to

$-\infty$  as  $n$  tends to  $\infty$  with probability at least  $(1 - \delta)$ .

By appropriately choosing the values of  $M$ ,  $a_1$  and  $a_2$ , the quantity  $(1 - \delta)$  can be made as close to one as we please.

With reference to the "modified" boundary  $\mathcal{L}$  given by Equation 3.40 and illustrated in Diagram 4, a convenient summary of all the above can be provided by the following theorem:

Theorem 3.2: Given  $\varepsilon$  and  $\delta$ , there exists an  $A(\varepsilon, \delta)$  such that

$$\Pr\{\mathcal{L}_n \leq n(m+\varepsilon) + A(\varepsilon, \delta), \text{ all } n\} \geq 1 - \delta \quad . \quad (3.49)$$

In fact  $A(\varepsilon, \delta)$  is just the quantity  $a_3$  of Equation 3.41 for  $a_1$  and  $M$  chosen sufficiently large.

Corollary 3.1: Suppose  $\eta > 0$ ; given  $\varepsilon$  and  $\delta$  there exists  $B(\varepsilon, \delta)$  such that

$$\Pr\{1 - \eta'_n \leq B(\varepsilon, \delta) (e^{m+\varepsilon})^n, \text{ all } n\} \geq 1 - \delta \quad . \quad (3.50)$$

The corollary follows from the theorem and relation 3.16 between the posterior  $\eta'_n$  and the partial sum  $\mathcal{L}_n$ .

## IV. POLICY CONVERGENCE

This chapter considers a sense in which the Bayesian decision maker's total infinite-horizon reward nearly equals the total reward of the decision maker who known the true state of nature, in a sense much like that usually associated with one and two-armed bandit problems. The plan here is to combine our knowledge regarding the behavior of the posterior distribution with facts concerning the recurrence of the situation in which two decision makers find themselves in the same process state. These last facts come from the following lemma.

Lemma: Consider a Markov chain  $X_n$  over a finite state space  $\mathcal{N}$  composed of  $N$  states, governed by a positive stochastic transition matrix. Consider as well a second process  $Y_n$  on the state space  $\mathcal{N}$  of the process  $X_n$ , such that (Assumption A)  $\Pr\{X_{m+i} \neq Y_{m+i}; i:1, \dots, M | Y_{m+i} = Y_{m+i}; i:1, \dots, M\} \leq \Pr\{X_{m+i} \neq Y_{m+i}; i:1, \dots, M\}$ . Then

$$\Pr\{X_n = Y_n \text{ i.o.}\} = 1 \quad . \quad (4.1)$$

Proof 4.1: Let  $[p_{ij}]$  denote the transition matrix  $X_n$  and define

$$\varepsilon \equiv \min_{i,j} p_{ij} \quad . \quad (4.2)$$

Let  $(y_{m+1}, y_{m+2}, \dots, y_{m+M})$  be a point in the  $M$ -fold Cartesian

product  $(\mathcal{N} \times \dots \times \mathcal{N})$ . Then it is clear that

$$\Pr\{X_{m+1} \neq Y_{m+1}, X_{m+2} \neq Y_{m+2}, \dots, X_{m+M} \neq Y_{m+M}\} \leq (1-\varepsilon)^M . \quad (4.3)$$

Hence, letting  $y$ ,  $Y$  and  $X$  denote the vectors  $(y_{m+1}, y_{m+2}, \dots, y_{m+M})$ ,  $(Y_{m+1}, Y_{m+2}, \dots, Y_{m+M})$  and  $(X_{m+1}, X_{m+2}, \dots, X_{m+M})$  respectively, one finds

$$\Pr\{X \neq Y\} \leq \sum_Y \Pr\{X \neq y\} \cdot \Pr\{Y = y\} \leq (1-\varepsilon)^M . \quad (4.4)$$

Hence

$$\begin{aligned} & \Pr\{X_{m+v} \neq Y_{m+v}; \forall v: 1, 2, \dots\} \\ & \leq \Pr\{X_{m+v} \neq Y_{m+v}, 1 \leq v \leq M\} \leq (1-\varepsilon)^M \end{aligned} \quad (4.5)$$

in view of Equation 4.4.

In other words, choosing  $M$  arbitrarily large, one concludes that

$$\Pr\{X_{m+v} \neq Y_{m+v}; \forall v: 1, 2, \dots\} = 0 \quad (4.6)$$

or

$$\Pr\{X_{m+v} = Y_{m+v}; \text{some } v: 1, 2, \dots\} = 1 . \quad (4.7)$$

Now let us define the events  $\mathcal{G}_m$  and  $\mathcal{E}_n$  as follows

$$\mathcal{G}_m \equiv \{X_{m+v} = Y_{m+v}; \text{some } v: 1, 2, \dots\} \quad (4.8)$$

and

$$\mathcal{E}_n \equiv \{X_n = Y_n\} \quad . \quad (4.9)$$

Then Equation 4.7 implies that

$$\Pr\left\{\bigcap_{m=1}^{\infty} \mathcal{E}_m\right\} = \Pr\left\{\bigcap_{m=1}^{\infty} \bigcup_{v=1}^{\infty} \mathcal{E}_{m+v}\right\} = 1 \quad (4.10)$$

or equivalently

$$\Pr\left\{\mathcal{E}_n \text{ i.o.}\right\} = 1 \quad . \quad (4.11)$$

We now intend to show that, in a certain sense, a Bayesian decision-maker is as well-off in terms of total discounted rewards as a decision maker who knows the true state of nature, if the discount factor  $\beta$  is large.

Consider then the special case of a Markov decision situation with two possible states of nature, involving two possible generalized transition matrices  $\overline{\Pi}$  and  $\mathcal{P}$ , a certain discount factor  $\beta$ , and a certain reward matrix  $\mathcal{R}$ .

Consider in this context a first decision-maker  $\mathcal{D}_1$  who knows that  $\overline{\Pi}$  applies. In accordance with the work of Howard (1960), that decision maker will possess a stationary optimal policy which specifies a certain action alternative  $k(i)$  whenever in state  $i$ . Accordingly the successive states  $Y_n$  that  $\mathcal{D}_1$  occupies when acting optimally will

be generated by a Markov chain with a square transition matrix  $[\pi_{ij}^{k(i)}]$ .

Consider as well a second decision maker  $\mathcal{D}_2$ , occupying successive states  $X_n$ , who does not know that the true state of nature is  $\bar{\Pi}$ , but rather proceeds as a Bayesian optimizer using a sequential Bayes strategy with prior probabilities  $\eta$  and  $(1-\eta)$  on  $\bar{\Pi}$  and  $\mathcal{P}$  respectively with  $\eta > 0$ .

Suppose that the situation is such that the "state-stationary assumption" of Chapter III is satisfied. In other words suppose that there exists an interval  $\mathcal{I}:[\eta_0, 1]$  about unity with the property that the Bayes strategy is state-stationary for  $\eta' \in \mathcal{I}$ .

We note, since  $\mathcal{D}_1$ 's policy is in fact a Bayes policy for  $\eta=1$ , that the "state-stationary assumption" in fact implies that  $\mathcal{D}_1$  and  $\mathcal{D}_2$  will be operating under the same state-stationary policy whenever  $\mathcal{D}_2$ 's posterior  $\eta'$  is in  $\mathcal{I}$ , i.e., that, whenever  $\eta'$  is in  $\mathcal{I}$ , they both will be using the same policy  $k(i)$  and consequently will both be subject to state transitions in accordance with the stochastic matrix  $[\pi_{ij}^{k(i)}]$ .

Hence it follows from the work of Chapter III that, if  $\bar{\Pi} > 0$ , then with probability 1,  $\mathcal{D}_1$  and  $\mathcal{D}_2$  eventually will operate under the same "state-stationary" policy  $k(i)$ . At the same time, in view of the lemma at the beginning of this section,  $\mathcal{D}_1$  and  $\mathcal{D}_2$  will find themselves in the same

state infinity often with probability 1 under Assumption A.

Hence, putting together the two claims of the last paragraph, we conclude that, with probability 1, there will be a finite time period (or number of sampling stages)  $n^*$  with the following properties:

1. The decision makers  $\mathcal{D}_1, \mathcal{D}_2$  find themselves in the same process state at  $n^*$ .
2. From  $n^*$  onward  $\mathcal{D}_1$  and  $\mathcal{D}_2$  follow the same "state-stationary" policy  $k(i)$ .

These two properties of  $n^*$  can be made to reflect the near-equivalence of the two strategies in several ways. For example, suppose for purposes of illustration that  $\mathcal{N}$  consists of only two states (i.e.,  $N=2$ ) and that there are two available alternative actions in each state (i.e.,  $K=2$ ). Consider a measure space  $\mathcal{M}$  whose elements are quadruple sequences

$$\omega: X_1, Y_1, Z_1, V_1; X_2, Y_2, Z_2, V_2; X_3, Y_3, Z_3, V_3; \dots \quad (4.12)$$

of ones and twos, with probability assigned to the usual corresponding  $\sigma$ -algebra of  $\omega$ -sets assuming independence among all sequence elements, with

$$\Pr\{X_i = 1\} = 1 - \Pr\{X_i = 2\} = \pi_{11}^1$$

$$\Pr\{Y_i = 1\} = 1 - \Pr\{Y_i = 2\} = \pi_{11}^2$$

$$\begin{aligned} \Pr\{Z_i = 1\} &= 1 - \Pr\{Z_i = 2\} = \pi_{21}^1 \\ \Pr\{V_i = 1\} &= 1 - \Pr\{V_i = 2\} = \pi_{21}^2 \end{aligned} \quad (4.13)$$

This measure space yields a measure space for pairs of state-histories arising under any policy pair, in particular the optimum-policy pair for the two above-mentioned decision makers  $\mathcal{D}_1$  and  $\mathcal{D}_2$ .

In addition, this measure space of course gives correct state-history probabilities for individual policies. For example consider the policy  $\mathcal{H}$  that recommends the second alternative in the first state and the first alternative in the second state. Then, assuming that the process starts in the first state, then

$$\begin{aligned} &\Pr_{\mathcal{H}}\{1;2;1;\dots\} \\ &= \Pr\{\text{all } \omega \text{ leading to } (1;2;1;\dots) \text{ under } \mathcal{H}\} \\ &= \Pr\{\text{all } \omega \text{ of the form } (\text{unspecified}, 2, \text{unspecified}, \\ &\quad \text{unspecified}; \text{unspecified}, \text{unspecified}, 1, \text{unspecified}; \\ &\quad \dots)\} \end{aligned} \quad (4.14)$$

= the summation of the probabilities of all  $\omega$ 's with 2 in the second location and 1 in the seventh location  
 $= \pi_{12}^2 \pi_{21}^1$  .



Assume now that Assumption A applies for our measure space  $\mathcal{M}$ . Then the  $n^*$  of page 79 exists with probability 1. In other words there is an  $\omega$ -set of probability one all of whose members possess an  $n^*(\omega)$  satisfying properties 1 and 2 on page 79. But, 1 and 2 imply in  $\mathcal{M}$  that the state and decision histories for  $\mathcal{D}_1$  and  $\mathcal{D}_2$  will coincide from  $n^*(\omega)$  onward. In other words there will be a set of  $\omega$ 's of probability 1 for which there is an  $n^*(\omega)$  such that the transition rewards for  $\mathcal{D}_1$  and  $\mathcal{D}_2$  coincide from  $n^*(\omega)$  onward.

Now let us make the further assumption that the rewards are bounded away from zero:

$$0 < r \leq r_{ij}^k \leq R, \quad i, j, k: 1, 2 \quad (4.15)$$

where

$$R = \max_{i, j, k} r_{ij}^k$$

$$r = \min_{i, j, k} r_{ij}^k .$$

Then, in view of the comments following Equation 4.14, with probability 1 there is an  $n^*(\omega)$  such that

$$\frac{|\mathcal{R}_1(\beta) - \mathcal{R}_2(\beta)|}{\mathcal{R}_1(\beta)} \leq \frac{\sum_{k=0}^{n^*-1} \beta^k R}{\sum_{k=1}^{\infty} \beta^k r} \quad (4.16)$$

where  $\mathcal{R}_i(\beta)$  is the total discounted infinite horizon reward gained by  $\mathcal{D}_i$ .

Finally, if we assume that  $n^*$  may be chosen independently of  $\beta$  for  $\beta$  in some interval  $(\beta_0, 1)$  near 1, we have that

$$\Pr\left\{\lim_{\beta \rightarrow 1^-} \frac{|\mathcal{R}_1(\beta) - \mathcal{R}_2(\beta)|}{\mathcal{R}_1(\beta)} = 0\right\} = 1 \quad (4.17)$$

## V. SUMMARY AND SUGGESTIONS FOR FUTURE STUDY

### A. Problem Structure

Considered is a system governed by a "mixed" Markov chain with uncertain transition probabilities. The system can be exclusively in one of a finite number of states,  $N$ , with the provision of a finite number  $K$  of possible alternative actions in each state. An infinite operating time domain is considered with a discount factor  $0 < \beta < 1$  such that a reward  $r_{ij}^k$  received immediately after the  $n$ th transition, is discounted by  $\beta^n$ .

Two decision makers are considered; one is a Bayesian optimizer  $\mathcal{D}_1$  who considers a prior probability distribution over a finite set of generalized stochastic ( $K \cdot N \times N$ ) matrices  $\{\mathcal{P}_i\}$  which assigns some weight to the actual matrix  $\mathcal{P}_i^*$  governing the process; the other decision maker  $\mathcal{D}_2$  knows the matrix  $\mathcal{P}_i^*$  and acts optimally accordingly.

### B. Findings

A summary of the finding, about the above-described system is given below:

1. As the Bayesian decision maker  $\mathcal{D}_1$  proceeds in time, his posterior converges almost surely to a distribution degenerate at  $\mathcal{P}_i^*$ , provided the latter is positive. The topic of large deviations of Markov chains is introduced,

and the probability that, the posterior maintains, at least a certain exponential convergence rate in a certain "state-stationary" region is thereby estimated. Based on this development, a boundary  $\mathcal{L}: \mathcal{L}_1 - \mathcal{L}_2 - \mathcal{L}_3$  is constructed for a related "mixed" random walk which, under an additional assumption, provides a rate-of-convergence bound for the posterior.

The detailed analysis is carried out just for the special case ( $N=2, K=2, L=2$ ).

2. The total discounted reward earned by the system under the supervision of a Bayesian decision-maker is almost surely, for  $\beta \rightarrow 1-$ , relatively equal to that earned by another decision maker who knows  $\mathcal{P}_1^*$  and acts optimally accordingly.

### C. Suggestions for Future Study

The following are some suggestions for future studies. Some of these suggestions concern strengthening some of the assumptions made in this thesis; others concern extensions of this analysis which were not pursued due to time limitations.

1. Development of tighter bounds for the posterior convergence in the "state-nonstationary" phase; possibly through a study of the large deviation principle.

2. Study of the large deviation principle for Markov

chains under milder assumption about the transition matrices involved.

3. Confirming the assumption concerning the existence of a "state-stationary" phase.

4. Study of the dependency on the discount factor  $\beta$  of the posterior development.

5. Free Chapter IV of its several assumptions, possibly by changing the sense of the near equality of the total rewards to  $\mathcal{D}_1$  and  $\mathcal{D}_2$ .

## VI. LITERATURE CITED

- Anderson, T. W. and Goodman, L. A. 1957. Statistical inference about Markov chains. *Annals of Mathematical Statistics* 28: 89-110.
- Bahadur, R. R. and Rao, R. R. 1960. On deviation of the sample mean. *Annals of Mathematical Statistics* 31: 1015-1027.
- Barber, M. N. and Ninham, B. W. 1970. *Random and restricted walks theory and application*. New York, New York, Gordon and Breach Science Publishers.
- Bellman, R. 1957. *Dynamic programming*. Princeton, New Jersey, Princeton University Press.
- Bellman, R. 1957. A Markovian decision process. *Journal of Mathematics and Mechanics* 6: 679-684.
- Bickel, P. J. and Yahav, Y. A. 1967. Asymptotically pointwise optimization procedure in sequential analysis. *Proc. Fifth Berkeley Symp. Math. Statist. Probability* 1: 401-413.
- Billingsley, P. 1961a. *Statistical inference for Markov processes*. Chicago, Illinois, University of Chicago Press.
- Billingsley, P. 1961b. Statistical methods in Markov chains. *Annals of Mathematical Statistics* 32: 12-40.
- Blackwell, D. 1962. Discrete dynamic programming. *Annals of Mathematical Statistics* 33: 719-726.
- Blackwell, D. 1965. Discounted dynamic programming. *Annals of Mathematical Statistics* 36: 226-335.
- Blackwell, D. and Dubins, L. 1962. Mergine of opinions with increasing information. *Annals of Mathematical Statistics* 33: 882-886.
- Blackwell, D. and Girshick, M. A. 1954. *Theory of games and statistical decisions*. New York, New York, John Wiley and Sons, Inc.
- Block, H. W. 1971. Accuracy of convergence of sums of dependent random variables with variances not necessarily finite. *Annals of Mathematical Statistics* 42: 2134-2138.

- Box, G. E. P. and Hill, W. J. 1967. Discrimination among mechanistic models. *Technometrics* 9: 57-71
- Bradt, R. N. and Karlin, S. 1956. On the design and comparison of certain dichotomous experiments. *Annals of Mathematical Statistics* 27: 390-409.
- Bradt, R. N., Johnson, S. M. and Korlins, S. 1956. On sequential designs for maximizing the sum of  $n$  observations. *Annals of Mathematical Statistics* 27: 1060-1074.
- Breiman, L. 1968. *Probability*. Reading, Massachusetts, Addison-Wesley Publishing Company.
- Chanda, K. C. 1972. On estimation of tail end probabilities of the sample mean for linear stochastic processes. *Annals of Mathematical Statistics* 43, No. 5: 1680-1686.
- Chung, K. L. 1967. *Markov chains with stationary transition probabilities*. 2nd ed. New York, New York, Springer-Verlag.
- Conn, P. S. 1969. Asymptotic properties of sequences of positive kernels. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University.
- Cox, D. R. and Miller, H. D. 1965. *The theory of stochastic processes*. New York, New York, John Wiley and Sons, Inc.
- Debreu, G. and Herstein, I. N. 1957. Non-negative square matrices. *Econometrica* 21: 597-607.
- Derman, C. 1963. On sequential decisions and Markov chains. *Management Science* 9: 16-24.
- Derman, C. and Lieberman, G. J. 1967. A Markovian decision process for a joint replacement and stocking problem. *Management Science* 13: 607-617.
- Dobrushin, R. L. 1956. Central limit theorem for non-stationary Markov chains. I. *Theory of Probability and Its Applications* 1: 64-80.
- Doob, J. L. 1953. *Stochastic processes*. New York, New York, John Wiley and Sons, Inc.
- Dubins, L. E. and Freedman, D. A. 1965. A sharper form of the Borel-Cantelli lemma and the strong law. *Annals of Mathematical Statistics* 36: 800-807.

- Dubins, L. E. and Savage, L. J. 1965a. A Tchebycheff-like inequality for stochastic processes. Proc. Nat. Acad. Sci. U.S.A. 53: 274-275.
- Dubins, L. E. and Savage, L. J. 1965b. How to gamble if you must. New York, McGraw-Hill.
- Dykin, E. B. and Uspenskii, V. A. 1963. Random walks (translated from Russian). Boston, Mass., D. C. Heath and Company.
- Farrell, R. 1967. Weak limits of sequences of Bayes procedure in estimation theory. Proc. Fifth Berkeley Symp. Math. Statist. Probability 1: 83-111.
- Feldman, Dorian. 1962. Contributions to the "two-armed bandit" problems. Annals of Mathematical Statistics 33: 847-656.
- Feller, W. c1967. An introduction to probability theory and its application, Vol. 1. New York, New York, John Wiley and Sons, Inc.
- Freedman, D. A. 1965. Asymptotic behavior of Bayes estimates in discrete case. II. Annals of Mathematical Statistics 36: 454-456.
- Gantmacher, F. R. 1960. The theory of matrices. II. New York, New York, Chelsea Publishing Company.
- Gnedenko, B. V. 1968. The theory of probability (translated from Russian). 4th ed. New York, New York, Chelsea Publishing Company.
- Gonzales-Zubita, R. and Miller, R. L. March, 1965. Markovian decision process with uncertain transition probabilities. Technical Report No. 11. Research in the control of complex systems. Cambridge, Massachusetts, Operations Research Center, Massachusetts Institute of Technology.
- Harris, T. E. 1964. Theory of branching processes. New York, New York, Prentice-Hall, Inc.
- Herniter, J. D. and Magee, J. F. 1961. Customer behavior as a Markov process. Operation Research 9: 105-122.
- Howard, R. A. c1960. Dynamic programming and Markov processes. Cambridge, Massachusetts, The M.I.T. Press.



- Ifram, A. F. 1965. On the asymptotic behavior of densities with application to sequential analysis. *Annals of Mathematical Statistics* 37: 615-637.
- Klein, M. 1966. Markovian decision models for reject allowance problems. *Management Science* 12: 349-358.
- Kullback, S. 1959. *Information theory and statistics*. New York, New York, John Wiley and Sons, Inc.
- Lee, T. C., Judge, G. G. and Takayama, T. 1965. On estimating the transition probabilities of a Markov process. *Journal of Farm Economics* 47: 742-762.
- Lee, T. C., Mudge, G. G. and Zallner, A. 1968. Maximum likelihood and Bayes estimates of transition probabilities. *Journal of the American Statistical Association* 63: 1162-1179.
- Losifeseu, M. and Theodoreseu, R. 1969. *Random processes and learning* (translated from German). New York, New York, Springer-Verlag.
- Manne, A. S. 1960. Linear programming and sequential decisions. *Management Science* 6: 259-267.
- Martin, J. J. c1967. *Bayesian decision problems and Markov chains*. New York, New York, John Wiley and Sons, Inc.
- Montroll, E. W. 1947. On the theory of Markov chains. *Annals of Mathematical Statistics* 18: 18-36.
- Moran, P. A. 1968. *An introduction to probability theory*. Oxford, England, Oxford University Press.
- Nemhauser, G. L. c1966. *Introduction to dynamic programming*. New York, New York, John Wiley and Sons, Inc.
- Pierce, D. A. and Folks, J. L. 1969. Sensitivity of Bayes procedure to the prior distribution. *Operations Research* 17: 334-350.
- Raiffa, H. and Schlaifer, R. 1961. *Applied statistical decision theory*. Boston, Massachusetts, Graduate School of Business Administration, Harvard University.
- Robbins, H. 1952. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society* 58: 527-535.

- Romanovsky, V. I. 1970. Discrete Markov chains. Groningen, The Netherlands, Wollers-Noordhoff Publishing.
- Rose, J. S. 1971. Markov decision processes under uncertainty. Unpublished Ph.D. thesis. Evanston, Illinois, Library, Northwestern University.
- Rosen, B. 1967. On the central limit theorem for sums of dependent random variables. Zeit Wahrscheinlichkeitstheorie 7: 48-82.
- Rosenblatt, M. 1962. Random processes. New York, New York, Oxford University Press.
- Ryzin, J. V. 1966. Bayes risk consistency of classification procedures using density estimation. Sankayā, Series A 28: 261-270.
- Savage, L. J. 1954. The foundation of statistics. New York, New York, John Wiley and Sons, Inc.
- Silver, E. A. August, 1963. Markovian decision process with uncertain transition probabilities or rewards. Technical Report No. 1, Research in the control of complex systems. Cambridge, Massachusetts, Operations Research Center, Massachusetts Institute of Technology.
- Spitzer, F. 1964. Principles of random walks. Princeton, New Jersey, D. Van Nostrand Company, Inc.
- Wagner, H. M. 1960. On the optimality of pure strategies. Management Science 6: 268-269.
- Wald, A. 1947. Sequential analysis. New York, New York, John Wiley and Sons, Inc.
- Walker, A. M. 1969. On the asymptotic behavior of posterior distributions. Journal of the Royal Statistical Society, Series B 31: 80-88.
- Walter, A. R. 1966. A local limit theorem for a certain class of random walks. Annals of Mathematical Statistics 37: 855-869.
- Wertherill, G. R. 1961. Bayesian sequential analysis. Biometrika 48: 281-292.
- Wolf, F. K. 1970. Markovian decision processes with uncertain rewards. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University.

## VII. ACKNOWLEDGMENTS

I would like to thank Assuit University, Egypt—A.I.D. Program, U.S.A. for their support and for getting this graduate study started. I would like also to thank the Department of Industrial Engineering and the Department of Statistics at Iowa State University, Ames, Iowa, for their financial and moral support during this period of my graduate study. I would like to express my appreciation to my advisor (in the Industrial Engineering Department) Dr. A. C. Kleinschmidt for his valuable guidance and support. Special acknowledgment is due my advisor (in the Departments of Statistics and Industrial Engineering) Dr. H. T. David for his excellent guidance, encouragement and endless giving of his valuable time and counsel throughout the preparation of this dissertation. I would like also to extend my appreciation to the other members of my committee, Professor J. K. Walkup (Head of the Industrial Engineering Department), Dr. R. W. Mensing (Department of Statistics), Dr. R. S. Dañiya (Department of Mathematics), for generously giving their time and support. The author would like also to thank the alternate committee member, Dr. J. C. Even (Department of Industrial Engineering) for his valuable comments.

Appreciation is also due my family and relatives in Egypt, whose constant encouragement and prayers, gave me an infinite source of energy. A special acknowledgment to my

wife Susanna, for her patience, help and understanding.

Finally, my appreciation to Mrs. Marlys Phipps, for the excellent typing of this thesis.