AD-A046 158 CENTER FOR NAVAL ANALYSES ARLINGTON VA OPERATIONS EVA--ETC F/G 12/1
SEMI-MARKOV PROCESSES: AN INTRODUCTION,(U)
JUL 77 W R NUNN, A M DESIDERIO
UNCLASSIFIED CRC-335 N00014-76-C-0001
NL

| OF |
AD
A046 158

END
DATE
FILMED
12-77
DDC

(14) CRC-335

(12) B.S.

(6) # SEMI-MARKOV PROCESSES: AN INTRODUCTION .

CENTER FOR NAVAL ANALYSES
1401 Wilson Boulevard
Arlington, Virginia 22209

Operations Evaluation Group
Naval Warfare Analysis Group

(10) By Walter R./Nunn Anthony M./Desiderio

(11) July 1977

(12) 30 p.

(15) N0014-76-C-0001

Prepared for:

OFFICE OF NAVAL RESEARCH
Department of the Navy
Arlington, Virginia 22217

OFFICE OF THE CHIEF OF NAVAL OPERATIONS (Op 03)
Department of the Navy
Washington , D.C. 20350

270 700

02 033500.00

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CRC 335 ✓	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Semi-Markov Processes: An Introduction		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Walter R. Nunn, Anthony M. Desiderio		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0001 ✓
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Naval Analyses 1401 Wilson Boulevard Arlington, Virginia 22209		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlington, Virginia 22217		12. REPORT DATE July 1977 13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Office of the Chief of Naval Operations (Op03) Department of the Navy Washington, D.C. 20350		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This Research Contribution does not necessarily represent the opinion of the Department of the Navy.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

algorithms, Markov processes, probability, stochastic processes, transitions

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This Research Contribution describes the basic methodology for analyzing semi-Markov processes and demonstrates how to solve the necessary equation numerically. Two "mechanisms" for producing semi-Markov processes are described and examples presented. Algorithms needed to analyze semi-Markov (continuous- and discrete-time) processes are derived.

150

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601 |

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Center for Naval Analyses
an affiliate of the University of Rochester

1401 Wilson Boulevard Arlington, Virginia 22209 703/524-9400

An Equal Opportunity Employer

30 September 1977

MEMORANDUM FOR DISTRIBUTION LIST

Subj: Center for Naval Analyses Research Contribution 335

Encl: (1) CRC 335, "Semi-Markov Processes: An Introduction," by Walter R. Nunn and Anthony M. Desiderio July 1977

1. Enclosure (1) is forwarded as a matter of possible interest.

2. Research Contributions are distributed for their potential value in other studies and analyses. They do not necessarily represent the opinion of the Department of the Navy.

3. The enclosure is approved for public release.

PHIL E. DePOY
Director
Operations Evaluation Group

Distribution List:
Reverse page

Subj: Center for Naval Analyses Research Contribution 335

DISTRIBUTION LIST

Department of the Navy

SNDL Part II:

A1	ASST SECNAV (R&D)
A2A	ONR
E3A	NRL
FF38	USNA, NIMITZ LIBRARY
FKA6A2	NAVWPNSCEN
FKR3C	NAVAIRTESTCEN
FKR4A	PACMISTESTCEN
OpNav:	Op-09BH
	Op-96
	Op-03

Other

Director, Defense Research and Engineering
Ass't SecDef, Program Analysis & Evaluation (2 copies)
Defense Documentation Center (12 copies)
Aerospace Corporation
Analytic Services, Inc.
Applied Research Laboratory
Institute for Defense Analyses
Jet Propulsion Laboratory
Lincoln Laboratory
Mathematics Research Center
The Rand Corporation
Stanford Research Institute
TRW Defense and Space Systems Group
National Academy of Sciences
University of Rochester (2 copies)

TABLE OF CONTENTS

INTRODUCTION

The idea of a semi-Markov process was proposed in 1954 (references 1 and 2). The semi-Markov process is similar to the Markov process in that both processes are described by a set of states whose transitions are governed by a transition probability matrix. The semi-Markov process, however, differs from the Markov process in that the times between transitions may be random variables. Further, the amount of time spent in any state after entering it is a random variable described by a probability density that can be a function of both the state of occupancy and the states to which transitions can occur.

The statistical time behavior of the semi-Markov process is described by a set of linear integral equations. The solution of this set of equations yields the probabilities that the process occupies the states of the system as a function of time. In many practical cases where the semi-Markov process is complex or involves many states, an analytic solution is difficult to obtain and numerical procedures must be used.

The purposes of this Research Contribution are (1) to acquaint the analyst with the semi-Markov process, (2) to illustrate that use of the semi-Markov process is common in systems and operations analysis, and (3) to show how numerical solution of the associated equations can be accomplished.

The general formulation of the semi-Markov process as developed in references 3 and 4 is discussed informally. A more formal treatment, however, appears in references 5 or 6. Numerical techniques are emphasized since algorithms for analyzing semi-Markov processes are not well developed or well known. Algorithms for both continuous-time and discrete-time processes are derived, and their numerical accuracies are discussed. Two examples are worked out in detail.

GENERAL FORMULATION OF A SEMI-MARKOV PROCESS

BASIC FEATURES

Let N be the number of states of a semi-Markov process. Let $h_{ij}(t)dt$ be the probability of a transition to state j between t and $t+dt$ given that the last transition was to state i at time zero. These holding time functions $h_{ij}(.)$ are elements of the transition matrix $H(.)$, i.e.,

$$H(t) = \{h_{ij}(t)\} \qquad 1 \le i, j \le N$$

and must satisfy

$$\int_0^\infty dt \sum_{j=1}^{N} h_{ij}(t) = 1, \qquad 1 \le i \le N \quad . \tag{1}$$

The functions $h_{ij}(.)$ are "dishonest" density functions in that they need not integrate to unity, *although they are otherwise* like density functions. Equation (1) expresses the intuitive requirement that there is unit probability that the system will be in one of the N states of the system at some point in the future, given start in i .

When the system makes a transition from state i to state j , the transition is said to be real if $j \neq i$ and virtual if $j = i$. Virtual transitions are represented in the transition matrix by nonzero diagonal elements. This formulation allows for both kinds of transitions.

Let $w_i(t)dt$ be the probability of leaving state i between t and $t+dt$. The quantity $w_i(t)$ is called the unconditional waiting time density and is expressed in terms of the $h_{ij}(.)$ by:

$$w_i(t) = \sum_j h_{ij}(t) \quad . \tag{2}$$

GENERATING A SEMI-MARKOV PROCESS

There are two general types of random mechanisms that give rise to semi-Markov processes. In the first type of mechanism, the process is envisioned, after having entered state i , first to make a random draw to determine the successor state j , and then to make another random draw to determine the amount of time it will stay in state i before going to state j . The outcome of the first draw is based on the set of transition probabilities p_{ij} , $1 \le i, j \le N$. The probability p_{ij} is the conditional probability that the next

transition is to state j, given that the last transition was to state i. The outcome of the second draw is based on the functions $f_{ij}(.)$, the holding time density functions for duration in state i given that the next transition is to state j. Then:

$$p_{ij} = \int_0^\infty h_{ij}(t)\, dt$$

and

$$f_{ij}(t) = h_{ij}(t)/p_{ij} \quad .$$

It is clear from the definition that $f_{ij}(.)$ integrates to unity; from equation (1), it is seen that:

$$\sum_j p_{ij} = 1 , \qquad 1 \le i \le N \quad .$$

In the second type of mechanism the process, after having entered state i, makes N random draws -- one draw from each of the N holding time densities $f_{ij}(.)$. It then determines the successor state and length of time in state i from the smallest draw. Thus, if the results of the N draws are the values t_{ij}, $1 \le j \le N$, and if $t_{ik} = \underset{j}{MIN}(t_{ij})$, then the next transition is to state k and the length of time the process holds in state i before going to k is t_{ik}.

An example of the first type of mechanism is the problem of the whimsical travelling salesman. The salesman randomly decides which city he will visit next, given that he is in a particular city (decision based on a random draw and the p_{ij}'s). Then the amount of time it takes the salesman to travel to the next city is based on a random draw from the density of travel times to that city.

An example of the second type of mechanism is the reliability of a system. If a system is composed of N critical subsystems, each with known failure time densities, then the path by which the system fails and the time to failure is determined by the smallest draw from the failure time densities.

The holding time probabilities $h_{ij}(t)$ for the two types of mechanisms are expressed by:

$$h_{ij}(t) = \begin{cases} p_{ij}\, f_{ij}(t) & \text{type 1 mechanism} \\ f_{ij}(t) \prod_{k \neq j} (1 - F_{ik}(t)) & \text{type 2 mechanism} \end{cases} \qquad (3)$$

where

$$F_{ik}(t) = \int_0^t f_{ik}(t)\, dt \quad .$$

For the type 2 mechanism, $\prod_{k \neq j} (1 - F_{ik}(t))$ is the probability that the process has not made a transition to a state other than j by time t. Further, $f_{ij}(t)dt$ is the probability that the transition to j will occur in the next instant of time between t and $t+dt$.

It is highly important in treating a semi-Markov process that the process be correctly identified as either type 1 or 2. An example of its importance is when there is a system composed of two independent, identical subsystems each of which has an exponential failure time density $(1/r)e^{-t/r}$. Then the mean time between failure is r if the system is type 1, but $1/2\ r$ if the system is type 2.

In certain cases, a system may be composed of some states that obey a type 1 transition mechanism, and others that obey a type 2 mechanism. There is no difficulty in modeling such a system, providing that $h_{ij}(t)$ is specified properly for all states.

Reflection shows that the type 2 mechanism is common and is capable of considerable generalization, e.g., the successor state may be chosen as some function of the N draws, not necessarily the state yielding the smallest draw. The possibility of anything but a type 1 mechanism is seldom mentioned in the literature, and the quantities p_{ij} and $f_{ij}(.)$ are usually treated as fundamental, with $h_{ij}(.)$ defined in terms of these. Here, $h_{ij}(.)$ has been taken as fundamental to show that it can be expressed in various ways according to the way in which the semi-Markov process is generated. Other mechanisms for choosing time in state and successor state can be found and give rise to still other forms for $h_{ij}(.)$. Reference 4 discusses the case in which the choice of successor state is probabilistic, conditional on the value of the waiting time in the current state.

If it is not known a priori that the system under consideration is of type 1, type 2, or otherwise, it may be possible in some cases to construct $h_{ij}(t)$ directly from data. When sufficient data is available, a convenient representation of $h_{ij}(t)$ is $w_i(t)\, \bar{p}_{ij}(t)$, where $w_i(t)$ is the waiting time density for transition out of state i, and $\bar{p}_{ij}(t)$ is the probability

that the system will make the next transition to j, given time t and current state i. This latter probability may be easy to estimate from the data.

THE SYSTEM EQUATIONS

This subsection derives the equations obeyed by the quantities which are usually of primary interest to an analyst. No attempt at completeness is made; equations for other quantities of possible interest are found in references 4, 5, and 6.

Let $\phi_{ij}(t)$ be the probability that the system is in state j at time t, given that it entered state i at time zero. The quantities $\phi_{ij}(t)$ are the state probabilities of the semi-Markov process. Knowing the values of these quantities for all time t is equivalent to knowing the probabilistic time behaviour of the system. The governing equation for these probabilities is now derived. A system starting in state i can be in state j at time t in the following ways:

1. $i=j$
   a. System never leaves state i, or
   b. System leaves state i but returns to i by time t.
2. $i \neq j$
   a. System leaves state i and manages to reach j by time t.

The set of equations that describes these events is:

$$\phi_{ij}(t) = \delta_{ij} W_i(t) + \sum_k \int_0^t d\tau \, h_{ik}(\tau) \, \phi_{kj}(t-\tau) \tag{4}$$

where

$$t \geq 0, \quad 1 \leq i, j \leq N, \quad \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}, \quad \text{and} \quad W_i(t) \quad \text{is given by:}$$

$$W_i(t) = 1 - \int_0^t w_i(t)dt \, . \tag{5}$$

The first term on the right-hand side of equation (4) is nonzero when $i=j$, in which case it represents the probability that the system did not leave its starting state i by time t. In the second term, the quantity $h_{ik}(\tau)d\tau$ is the probability that the system will make a transition to state k between times τ and $\tau+d\tau$. This is multiplied by $\phi_{kj}(t-\tau)$, the probability that the system will proceed to state j in the remaining time $t-\tau$, having entered state k. The product is integrated over all τ between 0 and t and summed over all states k.

When $N = 1$, equation (4) reduces to a single (renewal) equation; for $N > 1$, it is called a "Markov renewal equation" (references 5 and 6). Many quantities of interest in the analysis of semi-Markov processes obey a Markov renewal equation; this Research Contribution restricts its attention to such quantities.

When the semi-Markov process has few states and the densities are simple functions, it is sometimes possible to take the Laplace or geometric transform of equation (4), invert a matrix and take the inverse transform to obtain $\phi_{ij}(t)$, as described in reference 3. Transform techniques are usually difficult to apply and a direct numerical evaluation of equation (4) is needed. In the continuous-time case, there is little literature about the numerical solution of equation (4). Appendix A derives an algorithm for solving equation (4) in this case. In the discrete-time case, the situation is somewhat improved; equation (4) can be used almost "as is" to produce $\phi_{ij}(t)$ in a recursive manner. This is also discussed in reference 4. Appendix B derives an algorithm for solving equation (4), which has excellent numerical accuracy characteristics.

Absorbing States

How absorbing states are treated in this methodology is now discussed. An absorbing state is a state that is never left once it is entered; all other states are called transient states. The set of transient states is denoted T.

In brief, the effect of absorbing states is to simplify the solution of equation (4). When i is an absorbing state, it is seen that:

$$\phi_{ij}(t) = \delta_{ij}$$

for all t and all j. Hence, it is necessary to solve equation (4) only when i is a transient state. Also, when j is also a transient state, further simplification results from noting that the summation in equation (4) must be made only over the transient states since $\phi_{kj}(t-\tau)$ is zero whenever k is an absorbing state. With these modifications, the methodology for solving equation (4) remains effective regardless of the presence or absence of absorbing states.

Frequently, the analyst is particularly interested in the case where i is a transient state and j is an absorbing state. In this case, $\phi_{ij}(.)$ is a cumulative distribution function, obtained by solving equation (4) as written. However, when N is large, it may be computationally advantageous to note that equation (4) can be rewritten in this case as:

$$\phi_{ij}(t) = \int_0^t d\tau\, h_{ij}(\tau) + \sum_{k \in T} \int_0^t d\tau\, h_{ik}(\tau)\, \phi_{kj}(t-\tau) \quad , \tag{6}$$

where the summation is over just the transient states. This is of the same form as equation (4), but the summation is over fewer states.

Sometimes, the quantity of interest is the density function $\phi^*_{ij}(.)$ corresponding to the distribution function $\phi_{ij}(.)$. While the density function can be gotten by numerically differentiating $\phi_{ij}(.)$, it will usually be better practice to differentiate both sides of equation (6) and obtain the following equation for $\phi^*_{ij}(.)$:

$$\phi^*_{ij}(t) = h_{ij}(t) + \sum_{k \in T} \int_0^t d\tau \, h_{ik}(\tau) \, \phi^*_{kj}(t-\tau) \quad . \tag{7}$$

First Passage Time Distributions

Sometimes, the quantity of main interest is the distribution of the "first passage time," i.e., the time to the first visit of state j given start in state i . A straightforward way to calculate this distribution is to make state j absorbing, then apply the above methodology to calculate $\phi_{ij}(.)$ or $\phi^*_{ij}(.)$.

Expected Time in State

Sometimes the quantity of interest is the expected amount of time spent in a given state in a given time interval $[0, T]$. The random variable X_{ij} is defined equal to one if the process is in state j at time t , given start in state i at time zero, and equal to zero otherwise. Then the total time spent in state j is:

$$\int_0^T X_{ij}(t) \, dt$$

and the expected time in state j is:

$$\int_0^T E\left[X_{ij}(t)\right] dt = \int_0^T \phi_{ij}(t) \, dt \quad . \tag{8}$$

Hence, the desired expectation is gotten by solving equation (4) for $\phi_{ij}(.)$, then integrating it.

A CONTINUOUS-TIME APPLICATION

A simple stochastic process will now be discussed. Also, the equations developed in the preceding section will be applied to produce the numerical solution of the probabilistic time behaviour of the process. The numerical solution will then be compared with the analytic solution.

The problem to be considered is of an aircraft searching for a submerged submarine in the ocean. The aircraft drops sensors to detect the submarine. The aircraft is assumed to detect the submarine at a constant rate λ_2 , where λ_2 is the reciprocal of the mean time to detect. After the aircraft detects the submarine, it drops additional sensors to localize (refine the location) the submarine. Localization is assumed to take place at a constant rate λ_1 , where λ_1 is then reciprocal of the mean time to localize. Once the aircraft has localized the submarine, the mission is considered successful and the aircraft returns to base.

While the aircraft is attempting localization, however, the submarine may move out of detection range of the sensors and the aircraft may lose contact. The aircraft is assumed to lose contact on the submarine while attempting to localize it at a constant rate λ_3 , where λ_3 is the reciprocal of the mean time to lose contact. If the aircraft loses contact on the submarine, additional sensors are laid to redetect the submarine. The redetection rate is assumed to be equal to the detection rate λ_2 . Given that the aircraft can search for a long time and has an inexhaustible supply of sensors, then the probabilities of the aircraft being in the search, detection, and localization states as a function of time must be determined.

The elementary process just described is a three-state system consisting of two transient states and one absorbing state (figure 1). The arrows in the figure indicate the direction of transitions from each of the system.

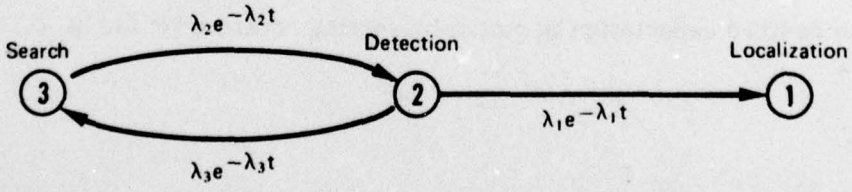

FIG. 1: THREE-STATE STOCHASTIC PROCESS

State 1 (localization) is the absorbing state because once entered it is never left, while states 2 and 3 (detection and search, respectively) are transient states. The holding time densities between states of the system are exponential and are indicated in the figure.

The system in figure 1 is assumed to make transitions according to a type 2 mechanism. The localization state and the search state compete for transitions from the detection state on a time basis. Using equation (3), the probability transition matrix is:

$$H(t) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} \delta(t-\infty) & 0 & 0 \\ \lambda_1 e^{-(\lambda_1+\lambda_3)t} & 0 & \lambda_3 e^{-(\lambda_1+\lambda_3)t} \\ 0 & \lambda_2 e^{-\lambda_2 t} & 0 \end{pmatrix} \end{array} \tag{9}$$

where $\delta(t-\infty)$ is a formalism to mean that state 1 is never left, once entered (except at $t=\infty$) . From equation (5), the $W(t)$ matrix is given by:

$$W(t) = \begin{pmatrix} 1 \\ e^{-(\lambda_1+\lambda_3)t} \\ e^{-\lambda_2 t} \end{pmatrix} \tag{10}$$

Using equations (9) and (10) in the recursive equation (A-5) of appendix A, $\Phi(t)$ is produced numerically. Figures 2 and 3 present the solution $\Phi(t)$ for $\lambda_1=1/24$, $\lambda_2=1/12$, $\lambda_3 = 1/18$, and $\Delta t=0.5$. Note that $\Phi_{11}=1$, $\Phi_{12}=\Phi_{13}=0$ for all time because state 1 is the absorbing state. Figure 2 displays the solution if the system initially started in state 3. The probability of being in the absorbing state Φ_{31} increases rapidly and then asymptotes to 1. The probability of being in state 3, given that the process started there Φ_{33} , rapidly decays to zero. The probability of being in state 2 given that the system started in state 3 Φ_{32} , builds up from 0 at time zero to a maximum of .39 at time 14 and then decays to zero. Figure 2 shows similar behaviour when the system starts in state 2.

Table 1 compares the numerical solution of equation (A-5) for Φ_{31} and Φ_{21} , the absorption probabilities, with a numerical evaluation of the analytic solution of the system of differential equations:

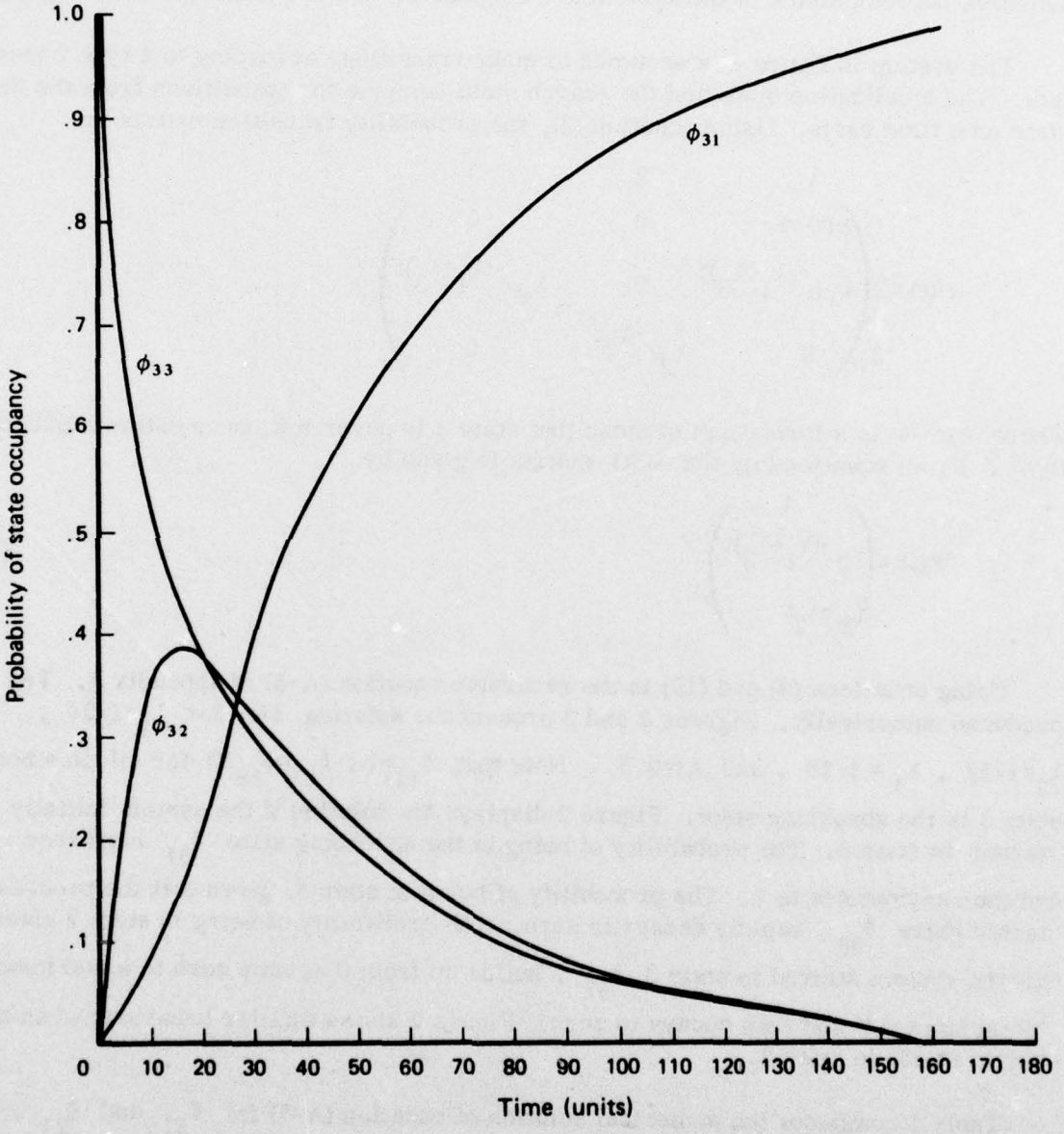

FIG. 2: EXAMPLE SOLUTION FOR SYSTEM STARTING IN STATE 3

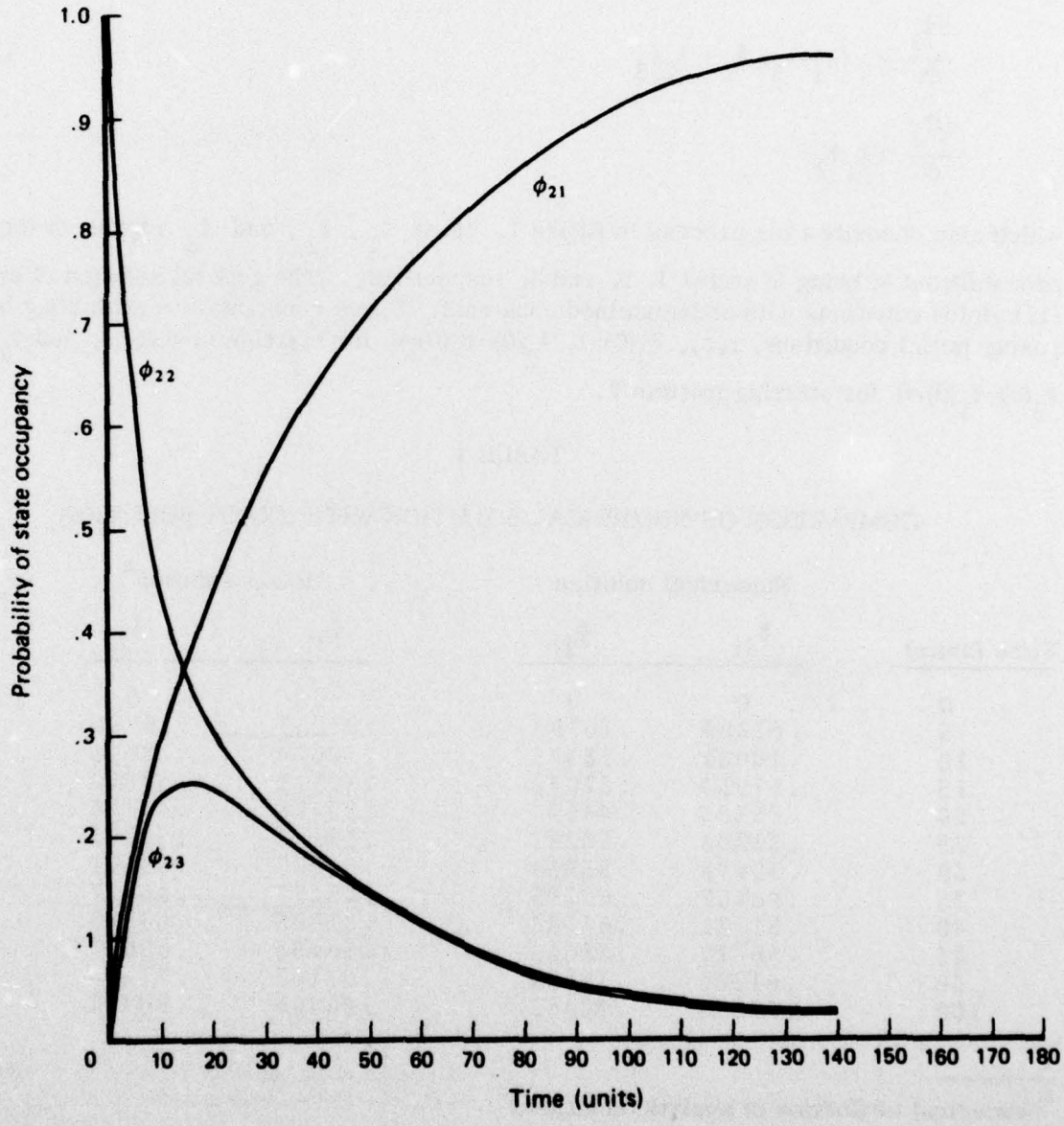

FIG. 3: EXAMPLE SOLUTION FOR SYSTEM STARTING IN STATE 2

$$\frac{d\Phi_3}{dt} = \lambda_2\Phi_3 + \lambda_3\Phi_2$$

$$\frac{d\Phi_2}{dt} = -(\lambda_1+\lambda_3)\,\Phi_2 + \lambda_2\Phi_3 \qquad (11)$$

$$\frac{d\Phi_1}{dt} = \lambda_1\Phi_2$$

which also describes the process in figure 1. Here Φ_1, Φ_2, and Φ_3 represent the probabilities of being in states 1, 2, and 3, respectively. The general solution of equation (11) yields equations with undetermined constants. These constants are evaluating by imposing initial conditions, i.e., $\Phi_3(0)=1$, $\Phi_2(0)=\Phi_1(0)=0$ for starting in state 3, and $\Phi_2(0)=1$, $\Phi_3(0)=\Phi_1(0)=0$ for starting in state 2.

TABLE 1

COMPARISON OF NUMERICAL SOLUTION WITH EXACT SOLUTION

	Numerical solution		Exact solution[a]	
Time (units)	Φ_{31}	Φ_{21}	Φ_{31}	Φ_{21}
0	0	0	0	0
5	.03264	.16792	.03261	.16786
10	.10081	.28325	.10074	.28315
15	.17953	.37072	.17941	.37059
20	.25805	.44194	.25788	.44176
25	.33203	.50251	.33182	.50231
30	.39998	.55534	.39973	.55511
35	.46162	.60203	.46133	.60177
40	.51721	.64358	.51688	.64329
45	.56719	.68069	.56683	.68037
50	.61206	.71389	.61167	.71355
100	.87058	.90462	.86996	.90411

[a] Numerical evaluation of analytic solution.

As can be seen in table 1 the difference between the numerical solution and the exact solution is less than 1 part of 10^3. A smaller step size would have yielded better accuracy but required much more computation time. The numerical error inherent in the trapezoidal integration rule is given by equation (A-3) as:

$$\frac{(\Delta x)^3 f''(\xi)}{12} .$$

The relative error is given by:

$$R = \frac{(\Delta x)^3 f''(\xi)}{12 f(\xi)}$$

where f is the function being integrated. In our example the most rapidly varying function is $e^{-\lambda_2 t}$; therefore, the relative error is:

$$R \approx \frac{(.5)^3 \lambda^2 e^{-\lambda_2 t}}{12\, e^{-\lambda_2 t}} = \frac{(.5)^3 (1/12)^2}{12} = 7.23 \times 10^{-5}$$

or less than 1 part in 10^4 .

A DISCRETE-TIME APPLICATION

This section models a dogfight between "Snoopy" and the "Red Baron" as a semi-Markov process in discrete time to obtain the probability that Snoopy wins as a function of time. The numerical solution is then compared with the analytic solution.

The dogfight is thought of as being in one of five possible states at any time: (1) win for Snoopy, (2) offensive position for Snoopy, (3) neutral, (4) defensive position for Snoopy, and (5) loss for Snoopy. The transitions between states are governed by the transition matrix below:

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ .6 & 0 & .4 & 0 & 0 \\ 0 & .6 & 0 & .4 & 0 \\ 0 & 0 & .6 & 0 & .4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Modeling the dogfight as a type 1 mechanism seems most sensible, and the holding time densities are defined by:

$$f_{ij}(n) = p_i q_i^{n-1}, \quad n=1,2,\ldots \text{ for } i=2,3,4, \text{ with } p_2=p_4=.2,\ p_3=.1,\ q_i=1-p_i \ .$$

States 1 and 5 are absorbing states so that their holding time densities need not be specified. This choice of $f_{ij}(.)$ is made primarily for analytic convenience and these functions do not represent realistic holding time distributions in actual dogfights.

Of interest here is the probability that Snoopy wins, as a function of time. To make matters concrete, it is assumed that the Red Baron ambushes Snoopy (the engagement starts in state 4), and also that the fight breaks off (due to fuel exhaustion) after 2 minutes if no one wins before that time.

The discrete-time version of equation (7) is wanted; specifically, $\phi^*_{41}(n)$ is desired since state 1 is the absorbing state corresponding to a win by Snoopy, and the fight starts in state 4. Also, $\Delta t = 1$ second, so that $t_n = n$, $n=0,\ldots,120$. Then:

$$\phi^*_{i1}(n) = h_{i1}(n) + \sum_{\tau=1}^{n} \sum_{k=2}^{4} h_{ik}(\tau)\, \phi^*_{k1}(n-\tau) \qquad (12)$$

just as equation (B-1) in appendix B was gotten from equation (4). To get ϕ^*_{41}, it is necessary to solve for all ϕ^*_{i1}, i=2, 3, 4. To get the final form of the recursion scheme, equation (12) is rewritten in matrix/vector form as:

$$\underline{\phi}^*(n) = \underline{h}(n) + \sum_{\tau=1}^{n} G(\tau)\,\underline{\phi}^*(n-\tau) \qquad (13)$$

where $\underline{\phi}^*(n)$ is the 3x1 vector with components $\phi^*_{i1}(n)$, i=2, 3, 4; $\underline{h}(n)$ is the corresponding vector with components $h_{i1}(n)$, and G is the 3x3 submatrix of H, consisting only of the transient states 2, 3, and 4. Equation (13) is a discrete form of equation (7). Alternatively, ϕ^*_{41} could be gotten by solving equation (B-1) for ϕ_{41} and "differentiating" to get ϕ^*_{41}. But this way is computationally superior in the same way that equation (7) is superior to equation (6) when absorbing states are involved.

Equation (13) can be solved using the z-transform (reference 4), and the probability of a win by Snoopy given the start in state 4 is:

$$\frac{.27}{11}(.96)^{n-1} + \frac{.72}{11}(.74)^{n-1} - .09(.8)^{n-1}, \quad n \geq 1 .$$

A comparison of the exact values and the values gotten by recursive solution using equation (13) is given in table 2.

There is agreement, in the table, through the first nine digits. However, it is possible that the numerical solution is more accurate than the "exact" solution, depending on how well the CNA computer raises numbers to large powers.

Figure 4 gives the cumulative probability of a win for Snoopy as a function of the length of the dogfight for all three starting states, 2, 3, and 4. The probability of a loss for Snoopy could be calculated similarly by changing the vector $\underline{h}(n)$ to the vector with components $h_{15}(n)$. The exchange ratio as a function of time can be gotten by dividing the probability of a win by the probability of a loss.

TABLE 2

COMPARISON OF EXACT VALUES AND VALUES FROM RECURSIVE SOLUTION

Time (sec)	Numerical solution	Exact[a] solution
0	0	0
10	$9.274338537E^{-3}$	$9.27433854E^{-3}$
20	$1.021862002E^{-2}$	$1.021862002E^{-2}$
30	$7.384697292E^{-3}$	$7.384697295E^{-3}$
40	$4.980723436E^{-3}$	$4.980723439E^{-3}$
50	$3.319363533E^{-3}$	$3.319363535E^{-3}$
60	$2.207700531E^{-3}$	$2.207700533E^{-3}$
70	$1.467846696E^{-3}$	$1.467846698E^{-3}$
80	$9.758826697E^{-4}$	$9.75882671E^{-4}$
90	$6.487997539E^{-4}$	$6.487997548E^{-4}$
100	$4.313433694E^{-4}$	$4.313433701E^{-4}$
110	$2.867711621E^{-4}$	$2.867711626E^{-4}$
120	$1.90654829E^{-4}$	$1.906548293E^{-4}$

[a]Numerical evaluation of analytic solution.

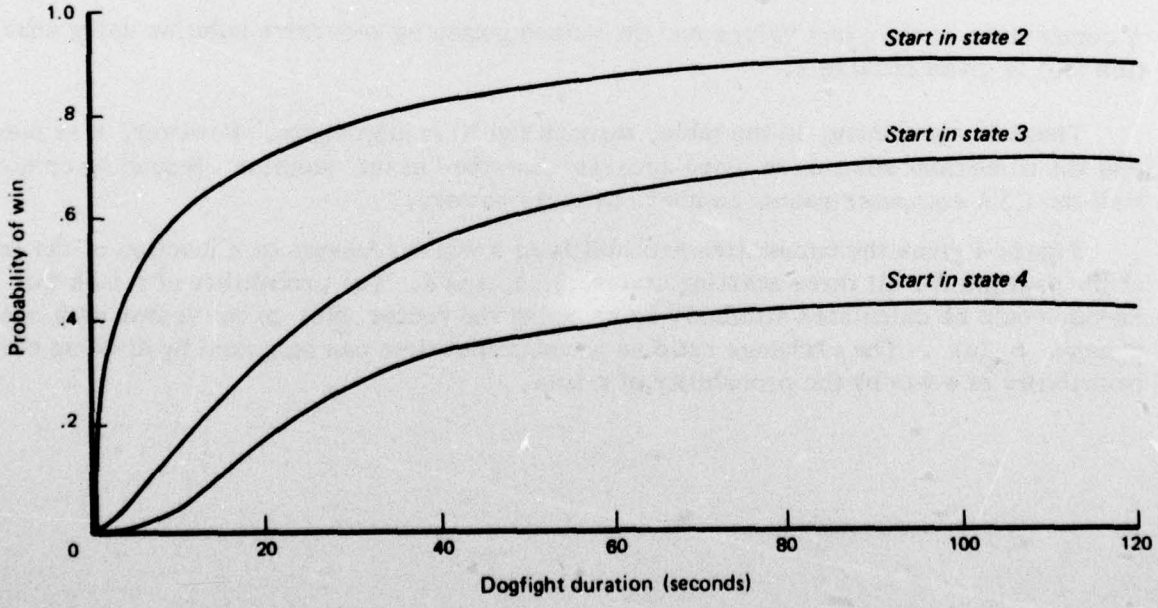

FIG. 4: PROBABILITY OF WIN FOR SNOOPY vs. DOGFIGHT DURATION

DISCUSSION OF ALGORITHMS

In the discrete-time case, algorithms for the solution of equation (4) are easy to write. The algorithm in appendix B is fast running and suitable for time-sharing usage. The answers are exact except for truncation and roundoff errors, both of which accumulate slowly.

In the continuous-time case, the situation is not as good. An algorithm based on the trapezoidal rule is given in appendix A. It is easy to program, but requires a small step size for accuracy. The running time for the algorithm is proportional to n^2, where n is the number of subdivisions of the time axis, so that small step sizes imply long running times. Appendix A also contains an algorithm based on Simpson's rule. This algorithm gives about three more digits of accuracy than the trapezoidal algorithm for the same step size for most applications; however, the Simpson algorithm is more complicated than the trapezoidal algorithm and is inherently slower running. All of the above algorithms can be programmed in various ways, e.g., to emphasize speed or to minimize storage.

Relatively little literature exists concerning the numerical solution to equation (4). Many authors suggest the use of transform methods, involving a numerical inversion of the transform. This is a nontrivial operation in most cases. The idea of a direct numerical attack on equation (4) is not new; reference 7, for example, outlines an algorithmic procedure very similar to that used in appendix A. although no discussion of numerical accuracy or running time is given. Reference 8 describes a solution based on expansion in power series, resulting in a simple recursive scheme for determining the unknown coefficients. The technique must be tailored to the specific distributions in use, however. Reference 9 discusses a rather elaborate scheme using spline functions and Lobatto integration to do calculations similar to those needed for equation (4). But it is not clear that such a complex technique is needed for routine solution of equation (4).

Reference 10 gives an annotated bibliography on the computational aspects of probabilistic modeling. In it will be found very few references to the computational problems involved in solving Markov renewal equations (equation (4)). It seems clear that considerably more algorithmic research is needed before such equations can be routinely solved in a fast and accurate manner.

REFERENCES

1. Levy, P., "Systems Semi-Markoviens Ayant au plus une Inifinite Denombrable d'Etats Possibles," Proc. Intern. Congr. Math., Amsterdam, The Netherlands, vol. 2, p. 294, 1954; "Processus Semi-Markoviens," vol, 3, pp. 416-426, 1954

2. Smith, W. L., "Regenerative Stochastic Processes," Proc. Roy. Soc (GB), series A, vol. 232, pp. 6-31, 1955

3. Howard, R. A., "System Analysis of Semi-Markov Processes," IEEE Transactions on Military Electronics, vol. MIL-8, no. 2, pp. 114-124, Apr 1964

4. Howard, R. A., "Dynamic Probabilistic Systems, Vol. II: Semi-Markov and Decision Processes," John Wiley and Sons, Inc., New York, 1971

5. Ross, S. M., "Applied Probability Models with Optimization Applications," Holden-Day, Inc., San Francisco, California, 1970

6. Cinlar, E., "Introduction to Stochastic Processes," Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1975

7. Smith, W. L., "Renewal Theory and Its Ramifications," J. Roy. Stat. Soc., series B, vol. 20, pp. 243-302, 1958

8. Smith, W. L. and Leadbetter, M. R., "On the Renewal Function for the Weibull Distribution," Technometrics, vol. 5, no. 3, pp. 393-396, 1963

9. Cleroux, R. and McConalogue, D. J., "A Numerical Algorithm for Recursively Defined Convolution Integrals Involving Distribution Functions," Management Science, vol. 22, no. 10, pp. 1138-1146, 1976

10. Neuts, M. F., "An Annotated Bibliography on Computational Probability," J. Comp. & Applied Math., vol. 3, no. 1, pp. 53-74, 1977

11. Abramowitz, M. and Stegun, I. A., "Handbook of Mathematical Functions," National Bureau of Standards Applied Mathematical Series 55, Jun 1964

APPENDIX A

CONTINUOUS-TIME SOLUTION FOR THE SYSTEM EQUATIONS

APPENDIX A

CONTINUOUS-TIME SOLUTION FOR THE SYSTEM EQUATIONS

In this appendix, two algorithms are derived for solving equation (4) (from main text) by numerical integration. First, an algorithm is derived based on the trapezoidal rule; then an algorithm is derived based on Simpson's rule.

TRAPEZOIDAL RULE

Equation (4) is solved by finding $\phi_{ij}(t)$ on a set of equally spaced points in time. The time points are denoted $t_0, t_1, \ldots, t_m$. The step size, i.e., the interval between the time points, is denoted Δt. The time points can be represented as:

$$t_n = \Delta t \cdot n, \quad 0 \le n \le m \quad .$$

With $\phi_{ij}(0) = \delta_{ij}$, the solution for t_n, $n > 0$, is now desired.

For convenience, equation (4) is rewritten in matrix notation as:

$$\phi(t_n) = W(t_n) + \int_0^{t_n} d\tau\, H(\tau)\, \phi(t_n - \tau) \qquad \text{(A-1)}$$

where

$$\phi(t_n) = \{\phi_{ij}(t_n)\}$$
$$H(t_n) = \{h_{ij}(t_n)\} \qquad 1 \le i, j \le N$$

and $W(t_n)$ is the diagonal matrix with elements $W_i(t_n)$ down the diagonal and zeros elsewhere. It is observed that when $\tau = 0$, the argument of ϕ in the integrand is the same as that of ϕ on the left of the equation. To capitalize on this, equation (A-1) is rewritten as follows:

$$\phi(t_n) = W(t_n) + \int_0^{t_1} d\tau\, H(\tau)\, \phi(t_n - \tau) + \int_{t_1}^{t_n} d\tau\, H(\tau)\, \phi(t_n - \tau) \quad . \qquad \text{(A-2)}$$

Equation (A-2) is evaluated using the trapezoidal integration rule (reference 11), i.e.,

$$\int_{x_1}^{x_2} f(y)\,dy = \frac{\Delta x}{2}\left[f(x_1) + f(x_2)\right] - \frac{(\Delta x)^3}{12} f''(\xi)\ ,\ (x_1 < \xi < x_2) \tag{A-3}$$

where x_1 and x_2 are separated by an interval Δx, $f(.)$ is the function to be integrated, and the last term on the right is the error term. Applying equation (A-3) without the error term produces:

$$\phi(t_n) = W(t_n) + \frac{\Delta t}{2}\left[H(0)\phi(t_n) + H(t_1)\phi(t_n - t_1)\right] + \Delta t \sum_{k=1}^{n} H(t_k)\phi(t_n - t_k) - \frac{\Delta t}{2}\left[H(t_1)\phi(t_n - t_1) + H(t_n)\phi(0)\right] \tag{A-4}$$

where $H(0) = H(t_0)$ has been used. Solving for $\phi(t_n)$ produces:

$$\phi(t_n) = \left[I - \frac{\Delta t}{2} H(0)\right]^{-1} \times \left[W(t_n) + \Delta t \sum_{k=1}^{n} H(t_k)\phi(t_n - t_k) - \frac{\Delta t}{2} H(t_n)\phi(0)\right]. \tag{A-5}$$

This equation is the basic recursive scheme used numerically to produce the solution of equation (4) on the points t_n, $1 \le n \le m$. The solution is started with $\phi(0) = W(0) = I$, and the solution for $\phi(t_1)$ is gotten by applying equation (A-5), etc. In this way, the solution of equation (4) for any finite t_n is produced.

The matrix $\left[I - \frac{\Delta t}{2} H(0)\right]$ in equation (A-5) need be inverted only once at the start of the numerical scheme. If the transition matrix $H(t)$ contains only density functions that are zero at time zero, no inversion is required.

The transition matrix should contain only density functions that are continuous in the time interval of interest. Discontinuous density functions may cause large numerical inaccuracies in the numerical scheme.

At each step, the solution of equation (A-5) must satisfy the following condition:

$$\sum_{j=1}^{N} \phi_{ij}(t_n) = 1 , \quad \text{for } 1 \le i \le N \quad . \tag{A-6}$$

This condition serves to check on both the numerical accuracy and stability of the solution.

The relative numerical accuracy of the iterative scheme is one order of magnitude less than the numerical accuracy of the trapezoidal integration rule. This occurs because of the compounding of numerical errors in the iterative scheme. The relative error for the trapezoidal rule is given by:

$$R = \frac{(\Delta x)^3 f''(\xi)}{12 f(\xi)}$$

where $f(.)$ is the function being integrated.

SIMPSON'S RULE

In this subsection, Simpson rules are applied to equation (A-1). Simpson's rule and Simpson's 3/8 rule (reference 11) are given by:

$$\int_{t_0}^{t_2} f(t)dt = \frac{\Delta t}{3} (f_0 + 4f_1 + f_2) - \frac{(\Delta t)^5 f^{(4)}(\xi)}{90} \tag{A-7}$$

and (A-8)

$$\int_{t_0}^{t_3} f(t)dt = \frac{3\Delta t}{8} (f_0 + 3f_1 + 3f_2 + f_3) - \frac{3(\Delta t)^5 f^{(4)}(\xi)}{80}$$

respectively, where Δt is the step size and $f_i = f(t_i)$. For larger intervals of integration, equation (A-7) is applied repeatedly to obtain:

$$\int_{t_0}^{t_n} f(t)dt \approx \frac{\Delta t}{3} (f_0 + 4f_1 + 2f_2 + 4f_3 + \ldots + 4f_{n-1} + f_n) \quad . \tag{A-9}$$

The right-hand side is denoted by $S(t_0, t_n)$. Equation (A-9) is also known as Simpson's rule. Note that n must be even and greater than or equal to two.

These rules are applied for the numerical solution of equation (A-1), and it will be seen that the cases "n even" and "n odd" must be handled differently.

Case I: n Even, $n \geq 2$

The integral in equation (A-1) is split at the point t_2, obtaining:

$$\phi(t_n) = W(t_n) + \frac{\Delta t}{3}\left[H(0)\phi(t_n) + 4H(t_1)\phi(t_{n-1}) + H(t_2)\phi(t_{n-2})\right] + S(t_2, t_n)$$

or (A-10)

$$\phi(t_n) = \left[I - \frac{\Delta t}{3} H(0)\right]^{-1}\left\{W(t_n) + \frac{\Delta t}{3}\left[4H(t_1)\phi(t_{n-1}) + H(t_2)\phi(t_{n-2})\right] + S(t_2, t_n)\right\} .$$

Case II: n Odd, $n \geq 3$

The integral in equation (A-1) is split at t_3, obtaining:

$$\phi(t_n) = W(t_n) + \frac{3\Delta t}{8}\left[H(0)\phi(t_n) + 3H(t_1)\phi(t_{n-1}) + 3H(t_2)\phi(t_{n-2}) + H(t_3)\phi(t_{n-3})\right] + S(t_3, t_n)$$

or

$$\phi(t_n) = \left[I - \frac{3\Delta t}{8} H(0)\right]^{-1}\left\{W(t_n) + \frac{3\Delta t}{8}\left[3H(t_1)\phi(t_{n-1}) + 3H(t_2)\phi(t_{n-2}) + H(t_3)\phi(t_{n-3})\right] + S(t_3, t_n)\right\} . \quad \text{(A-11)}$$

Equations (A-10) and (A-11) are the counterpart of equation (A-5). It should be noted that there is no way to get $\phi(t_1)$ by these formulas. This value must be input independently, if the analyst wishes to get the increased accuracy that Simpson's rule provides over the trapezoidal rule. Indeed, as the analyst passes to more accurate numerical integration schemes, he will have to input increasingly more points to begin the calculation. In practice, this creates no problem -- the analyst might use something like the trapezoidal rule to get these initial points, then shift to the more accurate scheme.

APPENDIX B

DISCRETE-TIME SOLUTION FOR THE SYSTEM EQUATIONS

APPENDIX B

DISCRETE-TIME SOLUTION FOR THE SYSTEM EQUATIONS

When the holding time variables are discrete and take their values on the lattice points $t_n = \Delta t \cdot n$, $0 \le n \le m$, as in appendix A, then equation (4) can be solved with a simple recursive algorithm which has excellent numerical error characteristics. The algorithm is derived in reference 4, but is rederived here for completeness.

In the discrete case, equation (A-1) in appendix A reduces to:

$$\phi(t_n) = W(t_n) + \sum_{\tau=t_0}^{t_n} H(\tau)\, \phi(t_n - \tau) \quad .$$

However, it has been required that the holding time densities $h_{ij}(.)$ have no impulse component at the origin, so that in the discrete case $H(0)$ is the zero matrix. Then:

$$\phi(t_n) = W(t_n) + \sum_{\tau=t_1}^{t_n} H(\tau)\, \phi(t_n - \tau) \quad . \qquad \text{(B-1)}$$

It is also seen that $\phi(t_n)$ has been expressed in terms of $\phi(t_k)$, $0 \le k \le n-1$. Starting with $\phi(t_0) = W(t_0)$, equation (B-1) permits the recursive computation of (t_n) .

Equation (B-1) is remarkably well suited to numerical computation since (1) there are no matrix inversions, (2) there are no subtractions or divisions, and (3) there are no scaling problems since all the quantities involved are probabilities and hence all scaled between zero and one. It may therefore be anticipated that roundoff and truncation errors will accumulate very slowly, even in problems involving large dimensions.