LA-UR-14-22145
Approved for public release; distribution is unlimited.

Title: First Passage Moments of Finite-State Semi-Markov Processes

Author(s): Warr, Richard
Cordeiro, James

Intended for: Report

Issued: 2014-03-31

First Passage Moments of Finite-State Semi-Markov Processes

Richard L. Warr*and James D. Cordeiro*

March 12, 2014

Abstract

In this paper, we discuss the computation of first-passage moments of a regular time-homogeneous semi-Markov process (SMP) with a finite state space to certain of its states that possess the property of universal accessibility (UA). A UA state is one which is accessible from any other state of the SMP, but which may or may not connect back to one or more other states. An important characteristic of UA is that it is the state-level version of the oft-invoked process-level property of irreducibility. We adapt existing results for irreducible SMPs to the derivation of an analytical matrix expression for the first passage moments to a single UA state of the SMP. In addition, consistent point estimators for these first passage moments, together with relevant R code, are provided.

KEY WORDS: First Passage Distributions; Markov Renewal Process; Spectral Radius; Statistical Flowgraph Model; Universally Accessible
AMS MSC 2010: 60K15; 60G10; 62M05

1 Introduction

Since the seminal works of Levy [17, 18] and Smith [26], semi-Markov processes (SMPs) have been utilized as a framework for a wide variety of applications within the scientific literature. Much of the interest is due to the

*Department of Mathematics and Statistics, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, USA. The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

fact that the SMP relaxes the assumption of exponential sojourn times, yet retains a measure of the tractability of classical continuous time Markov chains. An area of study most frequently associated with SMPs is that of survival analysis and reliability, for which the definitive reference is [3], and which has been continued by the likes of [1, 5, 19] and others. Of special note are the areas of semi-Markov decision processes and PH-distributions [11, 28], often used in reliability, but which also appear in the context of SMP first passage moments, as in [29]. Other areas that have seen the application of SMP models are DNA analysis [2], queueing theory [13, 21], finance [9], artificial intelligence [28], and transportation [4, 16], to name but a few.

In this article, we will show the existence of and then derive the moments of first passage to states of a SMP with a finite state space that have the property of being accessible from every other state. We call this property *universal accessibility* (UA) and note that it can be likened to a state-level version of the property of irreducibility. This comes as a consequence of the fact that, as we will later show, UA of every state is a necessary and sufficient condition for irreducibility to hold. In this sense, then, UA of a subset of states of a SMP may be considered a natural relaxation of the property of irreducibility, which has been the standard assumption in the work of all researchers dating from Pyke [22, 23] onwards. Rather than being a simple generalization, we will show here that UA is, in fact, a *minimal* condition required for the existence of finite moments of first passage. This demonstration requires an application of the Perron-Frobenius theorem generalized to reducible matrices (and hence reducible processes) in order to arrive at the existence of the required matrix inverse. For further details on the Perron-Frobenius theorem, and spectral theory in general, see [7]. Although the proof of invertibility is somewhat convoluted, one gains the advantage of being able to consider only those first passage moments to a given universally accessible state, thus reducing the dimensionality of the problem. In addition, the expression that we derive does not suffer the presence of inverses of singular matrix terms. Contrast this to the situation of Hunter [8] and later researchers, whose expressions for first passage moments involved a noninvertible matrix, thus requiring a generalized inverse approach. Another significant advantage is that we are able to discard the somewhat strong assumptions of positive recurrence, and thus irreducibility, thereby increasing the class of SMPs for which a unified analytical approach to computing the first passage moments is available.

Explicit time-domain formulas for the first two moments of the first passage distribution of a irreducible ergodic SMP with a finite state space have long been known. Pyke [23] inverted Laplace-Stieltjes transform matrices under restrictive non-singularity conditions in order to derive the first and second moments. Hunter [8] repeated this analysis by means of Markov renewal theory, and then solved for the matrix of first passage moments M of the SMP through multiplication of the matrix $I - P$ by its generalized inverse, where P is the transition probability matrix of the embedded discrete time Markov chain. Although the role of the fundamental matrix of the embedded DTMC in solving the problem of finding the first passage moments was recognized since at least Kemeny and Snell [10], it was Hunter [8] that recognized its fundamental importance by proving that the fundamental matrix is a generalized inverse for $I - P$. Some years later, Yao [27] was able to use the generalized inverse to find *all* moments of first passage. Zhang and Hou [29] likewise employed the generalized inverse method in order to derive exact first passage moments for SMPs with phase- (PH-)distributed sojourn times between states, thus capitalizing on the robust interest in the reliability community for these somewhat exponential-like statistical distributions. All of these previous investigations assumed irreducibility, and are thus useful background for, though not directly applicable to the type of reducible process that we consider here.

The remainder of the paper will proceed as follows. In Section 2, we define notation, terminology, and assumptions that guide the remainder of the discourse. In section 3, we introduce the notion of universal accessibility, as well as a result that explains its relationship to irreducibility. We then present the main result in Section 4, which is the derivation of the formula for the first passage moments under the condition of universal accessibility. Finally, in Section 5, we present a method for estimating the first passage moments of SMPs and a brief example.

2 Notation and Basic Definitions

In this section we introduce the the notation used in this paper. A boldface symbol without indices refers to a matrix (e.g., $\mathbf{F}(t)$ is a matrix with elements $F_{ij}(t)$ in the ith row and jth column). We will sometimes drop the function argument for simplicity's sake; e.g., $\mathbf{F} = \mathbf{F}(t)$. In the usual way, we define the Dirac-δ function as

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

Additionally, we will specify that the m-dimensional square matrices $\boldsymbol{I}$ and $\boldsymbol{J}$ denote the identity matrix and the matrix whose entries consist of ones, respectively. Finally, the matrix binary operator '$\circ$' denotes Hadamard (element-wise) multiplication; i.e.

$$\left[\mathbf{A} \circ \mathbf{B}\right]_{ij} = A_{ij} B_{ij}.$$

We now define a regular time-homogeneous SMP $\{Z(t) : t \geq 0\}$ with a finite state space $\mathcal{S} = \{1, 2, \ldots, m\}$. Note that the assumption of regularity implies that the process may transition only a finite number of times in a finite time interval with probability 1. Let S_k, $k = 0, 1, 2, \ldots$ be the transitional epochs of the SMP and let $Z_k = Z(S_k)$. We define the kernel matrix $\boldsymbol{Q}(x) = [Q_{ij}(x)]$ of the SMP as

$$\begin{aligned} Q_{ij}(x) &= \mathrm{P}\{Z_{k+1} = j,\, S_{k+1} - S_k \leq x \,|\, Z_k = i\} \\ &= \mathrm{P}\{Z_1 = j,\, S_1 \leq x \,|\, Z_0 = i\}, \end{aligned}$$

which are the joint probabilities of waiting times and transitions from state $i \in \mathcal{S}$ to state $j \in \mathcal{S}$. The transition matrix of the embedded discrete time Markov chain (DTMC) is thus given by $\boldsymbol{p} = \boldsymbol{Q}(\infty)$. In addition, we define the matrix of distribution functions $\boldsymbol{F}(x) = [F_{ij}(x)]$ of the sojourn times in state i, given that the process transitions to state j as

$$F_{ij}(x) = \mathrm{P}\{S_1 \leq x \,|\, Z_0 = i,\, Z_1 = j\}, \tag{1}$$

with associated rth moments

$$\boldsymbol{e}^{(r)} = [e_{ij}^{(r)}], \quad \boldsymbol{e} = [e_{ij}] = \boldsymbol{e}^{(1)}, \qquad r \geq 1$$

from which it immediately follows that

$$Q_{ij}(x) = p_{ij} F_{ij}(x), \tag{2}$$

or, alternatively, as the Hadamard matrix product $\boldsymbol{Q} = \boldsymbol{p} \circ \boldsymbol{F}$.

The similarity in behavior of an SMP to a Markov chain at transition epochs $\{S_k : k = 0, 1, 2, \ldots\}$ is due to the classification of these transitions as *Markov renewal epochs*. These are times at which the process in question possesses the Markov, or memoryless property:

$$\mathrm{P}\{Z_{k+1} = j \,|\, Z_k = i,\, Z_{k-1}, \ldots, Z_1, Z_0\} = \mathrm{P}\{Z_{k+1} = j \,|\, Z_k = i\}.$$

Define the random variable $N_k(t)$ to be the number of transitions (Markov renewals) of the SMP into state k up to and including $t \geq 0$ and let

$$\boldsymbol{N}(t) \equiv [N_k(t)]_{k \in \mathcal{S}}$$

be the vector consisting of the random counting variables $N_k(t)$. Also define the scalar random variable

$$N(t) \equiv \sum_{k=1}^{m} N_k(t)$$

to be the total number of transitions, or Markov renewals, of the SMP up to t. We thus obtain the relationship $Z(t) = Z_{N(t)}$, between the SMP $\{Z(t) : t \geq 0\}$ and its embedded DTMC $\{Z_k : k \geq 0\}$. The vector counting process $\{\boldsymbol{N}(t) : t \geq 0\}$ is known as the *Markov renewal process associated to the SMP* $\{Z(t) : t \geq 0\}$.

The state properties of the SMP such as irreducibility and recurrence may be elicited from the properties of its embedded DTMC $\{Z_n : n \geq 0\}$. We say that state j is *accessible* from state i ($i \to j$) if there is a nonzero probability that $\{Z_n\}$ may transition to state j in a finite number of steps, given that it begins in state i. Mathematically, this means that there is some $n \in \mathbb{Z}_+$ such that $p_{ij}^{(n)} > 0$, where

$$p_{ij}^{(n)} = \mathrm{P}\left\{Z_n = j \,|\, Z_0 = i\right\}.$$

The matrix $\boldsymbol{p}^{(n)} = [p_{ij}^{(n)}]$ is called the *nth-step transition probability matrix*. The ijth element of the matrix denotes the probability of the DTMC transitioning from state i to state j in n stages and can be computed using the identity $\boldsymbol{p}^{(n)} = \boldsymbol{p}^n$. On the other hand, we say that state j is *not accessible* from state i (denoted $i \nrightarrow j$) if $p_{ij}^{(n)} = 0$ for all n. There may also exist a state $0 \in \mathcal{S}$ known as an *absorbing state*, which is to say that, for any other state $j \in \mathcal{S}$, $0 \nrightarrow j$. In this case, the SMP, having transitioned to state 0, sojourns for an *infinite* amount of time in this state. Many applications in survival and reliability analysis may be modeled using stochastic processes with one or more absorbing states. Transitioning to an absorbing state is tantamount to death or complete failure in the original process.

If i and j are mutually accessible (that is, $i \to j$ and $j \to i$, otherwise denoted as $i \leftrightarrow j$), then they are said to *communicate*. Since communication fulfills the axioms of reflexivity, transitivity, and symmetry, it is an equivalence relation, and thus defines a partitioning of the state space $\mathcal{S}$ into various disjoint *communicating classes*. If $\mathcal{S}$ is itself comprised of a single communicating class, then the SMP is called *irreducible*; otherwise, it is known as *reducible*. On the other hand, a nonnegative $m \times m$ matrix $\boldsymbol{A} = [a_{ij}]$ is an *irreducible matrix* if, for each i and j, there exists some $0 < \eta < \infty$ such that the ijth element of $\boldsymbol{A}^\eta$ is greater than 0. The algebraic and probabilistic definitions of irreducibility coincide if the irreducible matrix is the transition probability matrix $\boldsymbol{p}$, for then the ijth element of $\boldsymbol{p}^{(\eta)} = \boldsymbol{p}^{\eta}$ is strictly positive if and only if j is accessible from state i in a finite number η of steps with nonzero probability. This last statement can be made precise by reference to the *digraph associated to* $\boldsymbol{A}$, denoted $\mathcal{G}(\boldsymbol{A})$. This is the digraph with vertices in the set $V(\mathcal{G}(\boldsymbol{A})) = \{1, 2, \ldots, m\}$ such that the directed arc, or edge, (i, j) exists if and only if $a_{ij} > 0$. $\mathcal{G}(\boldsymbol{A})$ is said to be *strongly connected* if, for each ordered pair $i, j \in V(\mathcal{G}(\boldsymbol{A}))$, there exists a (directed) path in $\mathcal{G}(\boldsymbol{A})$ from i to j. In either case of there being an edge or directed path from i to j, the implication is clearly $i \to j$. The final connection between irreducibility and connectedness is made in the following Proposition:

Proposition 2.1 *Let $\boldsymbol{A}$ be a nonnegative square matrix. $\boldsymbol{A}$ is irreducible if and only if $\mathcal{G}(\boldsymbol{A})$ is strongly connected.*

Proof See Shao [25].

We next address the first passage times of an SMP. To this end, define the random variable

$$T_j = \inf\{t \geq S_1 : Z(t) = j\},\ j \in \mathcal{S},$$

which represents the time of first passage from an initial state i to state j if $i \neq j$, and the time of first return to j otherwise. The distribution function $G_{ij}(t)$ of first passage, conditioned on being in the initial state $i \in \mathcal{S}$, is defined as

$$G_{ij}(t) = \mathrm{P}\left\{T_j \leq t \,|\, Z(0) = i\right\},$$

and for which the corresponding rth moments $\mu_{ij}^{(r)}$, $r \geq 1$, if they exist, are given by

$$\mu_{ij}^{(r)} = \mathrm{E}\left[T_j^r \,|\, Z(0) = i\right].$$

We thus define $\boldsymbol{G}(t)$ and $\boldsymbol{\mu}^{(r)} = \left[\mu_{ij}^{(r)}\right]$ to be the matrices of first passage distribution functions and moments.

As stated in Proposition 5.15 of [24, pg104] and Lemma 4.1 of [29], the moments of first passage for an irreducible SMP may be computed as the finite solution to the systems of equations given by

$$\mu_{ij}^{(1)} = \sum_{k=1}^{m} p_{ik} \left[\left(1 - \delta_{kj}\right) \mu_{kj} + \mu_{ik}\right] \tag{3}$$

$$\mu_{ij}^{(r)} = \sum_{k=1}^{m} p_{ik} \mu_{ik}^{(r)} + \sum_{s=1}^{r} \binom{r}{s} \left[\sum_{k \neq j} p_{ik} e_{ik}^{(r-s)} \mu_{kj}^{(s)}\right], \qquad r \geq 2. \tag{4}$$

Clearly, a necessary condition for $\mu_{ij}^{(r)} < \infty$ is that $i \to j$, which is certainly true if the SMP is irreducible. In contrast, we observe that $G_{ij}(\infty) < 1$ (and $\mu_{ij} = \infty$) might occur for a pair of states $i,\, j \in \mathcal{S}$ if $i \nrightarrow j$. As we will later show, (3) and (4) still hold under the somewhat weakened assumption of universal accessibility for the terminal state j.

The recurrence properties of a SMP may be explained in terms of the distribution of the first passage of a SMP from a given state $i \in \mathcal{S}$ back to itself, otherwise known as the time of (first) return to a state $i \in \mathcal{S}$. The crucial step is to define

$$f_{ii} = \mathrm{P}\left\{N(T_i) < \infty \,|\, Z_0 = i\right\},$$

which is the probability that the number of steps required for the embedded DTMC $\{Z_n : n \geq 0\}$ to return to state i is finite. If $f_{ii} < 1$, then the state $i \in \mathcal{S}$ is called *transient*; otherwise, it is known as *recurrent*. If, in addition to recurrence, we have $\mu_{ii} < \infty$, then the state is called *positive recurrent*. The SMP itself is deemed recurrent, transient, or positive recurrent as a *process* if the corresponding condition holds for every state $i \in \mathcal{S}$. For an irreducible SMP with a finite state space, it is well-known that the process is automatically positive recurrent. This is not true, in general, for a reducible process, but may be evaluated on a state-by-state basis.

The Perron-Frobenius theorem adapted to finite-dimensional irreducible and nonnegative matrices is very useful for characterizing the set of eigenvalues of such matrices. As we will see later, the theory may be (indirectly) extended to even reducible nonnegative matrices by leveraging their distinctive canonical form. Let $\boldsymbol{A} \in \mathbb{R}_+^{m \times m}$ for some positive integer m. We define the *spectrum* of $\boldsymbol{A}$, denoted $\sigma_{\boldsymbol{A}}$, to be the set of its eigenvalues. Its *spectral radius*, denoted $\rho(\boldsymbol{A})$, is given by

$$\rho(\boldsymbol{A}) = \max\{|\lambda| : \lambda \in \sigma_{\boldsymbol{A}}\} \in \mathbb{R}_+,$$

which indicates the maximum radius of the disc that contains $\sigma_{\boldsymbol{A}}$ in the complex plane. Of particular interest is the case of a finite-dimensional *stochastic matrix* $\boldsymbol{A}$, which is a nonnegative square matrix such that $\boldsymbol{A}\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ is a column vector of ones. Perron-Frobenius theory, via Proposition 2.4 for the reducible case, implies that the spectral radius is likewise an eigenvalue of $\boldsymbol{A}$, denoted the *Perron root of* $\boldsymbol{A}$. Stochastic matrices comprise the boundary of the unit ball $\mathcal{A} = \{\boldsymbol{A} \in \mathbb{R}_+^{m \times m} : ||\boldsymbol{A}||_\infty \leq 1\}$ of finite-dimensional nonnegative matrices in the normed linear space induced by the *infinity norm* $||\cdot||_\infty$, which is given by the maximum absolute row sum of $\boldsymbol{A} = [a_{ij}]$, or

$$||\boldsymbol{A}||_\infty = \max_i \sum_{j=1}^{n} |a_{ij}| = \max(\boldsymbol{A}\mathbf{1}).$$

As the next Proposition will show, we may classify certain elements of $\boldsymbol{A} \in \mathcal{A}$ with spectral radius $\rho(\boldsymbol{A}) < 1$ as *substochastic*, which is to say that $0 < \min(\boldsymbol{A}\mathbf{1}) < 1$.

Proposition 2.2 *Suppose that $\boldsymbol{A} \in \mathcal{A}$. If $\rho(\boldsymbol{A}) < 1$, then $\boldsymbol{A}$ is substochastic.*

Proof Clearly, since $\boldsymbol{A} \in \mathcal{A}$, it must be either stochastic or substochastic. Therefore the only thing that must be proved is that $\boldsymbol{A}$ is not stochastic. Assume $\boldsymbol{A}$ is stochastic; i.e. $\boldsymbol{A}\mathbf{1} = \mathbf{1}$. This implies 1 is an eigenvalue, which contradicts $\rho(\boldsymbol{A}) < 1$. Therefore, $\boldsymbol{A}$ must be substochastic.

For an irreducible nonnegative matrix $\boldsymbol{A}$, it is, in fact, sufficient for $\boldsymbol{A}$ to have a spectral radius that is strictly less than unity in order to be substochastic, as the next Proposition shows.

Proposition 2.3 *If $\boldsymbol{A} \in \mathbb{R}_+^{m\times m}$ is an irreducible substochastic matrix, then $\rho(\boldsymbol{A}) < 1$.*

Proof See Theorem 7 in [14].

For such reasons, among others, it is very convenient to work with irreducible processes. Results for irreducible matrices (processes) may still be applied to the reducible case via an important consequence of the mathematical notion of reducibility. From the definition given above for a reducible matrix $\boldsymbol{A}$, it can be shown that there exists a permutation matrix $\boldsymbol{P}$ such that $\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1}$ is in upper block triangular form as follows:

$$\boldsymbol{A} \sim \boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1} = \left(\begin{array}{cccc|cccc} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1r} & \boldsymbol{A}_{1,r+1} & \boldsymbol{A}_{1,r+2} & \cdots & \boldsymbol{A}_{1M} \\ \mathbf{0} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2r} & \boldsymbol{A}_{2,r+1} & \boldsymbol{A}_{2,r+2} & \cdots & \boldsymbol{A}_{2M} \\ \vdots & & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{A}_{rr} & \boldsymbol{A}_{r,r+1} & \boldsymbol{A}_{r,r+2} & \cdots & \boldsymbol{A}_{rM} \\ \hline \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{A}_{r+1,r+1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \boldsymbol{A}_{r+2,r+2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{A}_{MM} \end{array}\right) \tag{5}$$

This is known as the *canonical form* for a reducible matrix. The canonical form is not unique, meaning that there may be two or more permutation matrices $\boldsymbol{P}$ such that a matrix $\boldsymbol{PAP}^{-1}$ is in canonical form. Its utility derives from several highly useful properties, which we will now discuss. First, the canonical matrix has the property such that every block matrix on the diagonal, $\boldsymbol{A}_{\nu\nu}$, $\nu \in \{1, \ldots, M\}$, is either $[0]_{1\times 1}$ or is irreducible. Moreover, the eigenvalues of $\boldsymbol{A}$ are invariant under the permutation transformation $\boldsymbol{PAP}^{-1}$. From a stochastic process perspective, we observe that transforming a reducible stochastic matrix $\boldsymbol{P}$ into its canonical form is equivalent to relabeling the state space of an associated (reducible) DTMC. The states are organized in the canonical transition matrix $\boldsymbol{P}$ in such a way that, for some $r \in \mathbb{Z}_+$, the transient state transitions are represented within the diagonal blocks $\boldsymbol{A}_{\nu\nu}$ for $1 \leq \nu \leq r$, while the blocks $\boldsymbol{A}_{\nu\nu}$ for $r+1 \leq \nu \leq M$ represent recurrent-state transitions. Additionally, as shown in Equation 8.4.7 of [20], $\rho(\boldsymbol{A}_{\nu\nu}) < 1$ for all $1 \leq \nu \leq r$, which, by Proposition 2.3, further shows that the first r diagonal sub-blocks of $\boldsymbol{A}$ are substochastic. The following Proposition rounds out this list of useful properties by relating the spectral radius of the sub-blocks of the canonical matrix to that of the entire matrix.

Proposition 2.4 *Suppose $\boldsymbol{A} \in \mathbb{R}_+^{m\times m}$ is a reducible matrix in canonical form. Then $\rho(\boldsymbol{A}) = \max_\nu \rho(\boldsymbol{A}_{\nu\nu})$ for $1 \leq \nu \leq M$.*

Proof See Lemma 1 in [12, pg. 303] with an additional induction argument to get the result or as argued in [7, pg.115].

Table 1 summarizes the important notation that we will use throughout this paper.

3 Universal Accessibility

In this section, we introduce the property of universal accessibility of state $j \in \mathcal{S}$. As we will later demonstrate, universal accessibility is a sufficient condition for the existence of a well-defined first-passage moment to a given state of the SMP.

Definition Let $\mathcal{P}$ be a stochastic process with state space $\mathcal{S}$. State $j \in \mathcal{S}$ is said to be *universally accessible* (UA) if, for every state $i \in \mathcal{S}$, we have $i \to j$.

If a proper subset of $\mathcal{S}$ is UA, then it is clear that the SMP is reducible. On the other hand, if *every* state is UA, then all states must communicate, as the next Proposition asserts.

Table 1: List of important symbols and notation.

n	The number of states in the SMP
x	The sojourn time in a state
t	Calendar time, or the time since the process began
p_{ij}	The probability that the next state in the process is j, given the process entered state i
$F_{ij}(x)$	The CDF of the waiting time distribution in state i, given the next transition to is state j
$G_{ij}(t)$	The CDF of the first passage distribution from state i to state j
$\boldsymbol{I}$	The $m \times m$ identity matrix
$\boldsymbol{I}(-j)$	The $m \times m$ identity matrix with the j^{th} row and column set to 0
$\mathbf{1}$	The $m \times m$ matrix of all 1s
$\mathbf{A}_j$	The j^{th} column of matrix $\mathbf{A}$
$\mathbf{A} \circ \mathbf{B}$	The element-wise product of two matrices
$e_{ij}^{(r)}$	$\int x^r dF_{ij}$, $e_{ij} \equiv e_{ij}^{(1)}$
$\mu_{ij}^{(r)}$	$\int x^r dG_{ij}$, $\mu_{ij} \equiv \mu_{ij}^{(1)}$
$\rho(\mathbf{A})$	The spectral radius of the matrix $\mathbf{A}$
$\lVert\mathbf{A}\rVert_\infty$	The infinity norm of a matrix

Proposition 3.1 *A SMP $\{Z(t) : t \geq 0\}$ with state space $\mathcal{S}$ is irreducible if and only if every $j \in \mathcal{S}$ is UA.*

The property of a state being UA is, in a sense, the minimal requirement for the existence of all first-passage moments. In the next section, we demonstrate the sufficiency of this condition by means of the Perron-Frobenius theorem applied to the canonical form of the reducible transition probability matrix of the embedded DTMC.

4 Moments of first passage time distributions

In this section, we derive a formula for determining the first and higher moments of first passage times in *reducible* SMPs to special states j that are UA. We begin with a technical result that will be needed in the proof of Theorem 4.2 to demonstrate that the matrix formula for the moments of first passage to a UA state $j \in \mathcal{S}$ is well-defined. For notational convenience, define $\boldsymbol{I}(-j)$ to be the identity matrix with the jth diagonal element set to zero.

Lemma 4.1 *Let $\{Z(t) : t \geq 0\}$ be a SMP with finite state space $\mathcal{S}$ and embedded DTMC at transition epochs with transition probabilities contained within the (stochastic) matrix $\boldsymbol{p}$. Then the matrix $[\boldsymbol{I} - \boldsymbol{p}\boldsymbol{I}(-j)]$ is nonsingular if and only if state $j \in \mathcal{S}$ is universally accessible (UA).*

Proof We begin with the observation that, since $\boldsymbol{A} = [\boldsymbol{A}_{\nu\kappa}] = \boldsymbol{p}\boldsymbol{I}(-j)$ is formed by setting each element of the jth column of $\boldsymbol{p}$ to 0, we essentially remove all directed arcs (i, j) in the digraph $\mathcal{G}(\boldsymbol{p})$ for each $i \in V(\mathcal{G}(\boldsymbol{p}))$ in order to produce $\mathcal{G}(\boldsymbol{A})$. This means that $\mathcal{G}(\boldsymbol{A})$ cannot be strongly connected, and thus $\boldsymbol{A}$ must be reducible. We may therefore assume that $\boldsymbol{A}$ is in canonical form (5). Furthermore, because the j^{th} column is zero, we will assume without loss of generality that the canonical form of $\boldsymbol{A}$ corresponds to the particular ordering of the states in $\mathcal{S}$ in which state j is re-designated as state 1. We impose the same permutation and partitioning on $\boldsymbol{p} = [\boldsymbol{p}_{\nu\kappa}]$ so that

$$\boldsymbol{A}_{\nu\kappa} = \begin{cases} \boldsymbol{p}_{\nu\kappa} & \text{if } (\nu, \kappa) \in \{1, \ldots, M\} \times \{2, 3, \ldots, M\}, \\ \boldsymbol{0} & \text{if } (\nu, \kappa) \in \{1, \ldots, M\} \times \{1\}, \end{cases} \tag{6}$$

where, as in (5), M is the dimension of $\boldsymbol{A}$. Notice that since $\boldsymbol{p}$ may be irreducible, the above does not necessarily imply that $\boldsymbol{p}$ can be put in canonical form, but rather is element-wise equivalent to $\boldsymbol{A}$, save for the first column, which, unlike that of $\boldsymbol{A}$, may contain positive entries. Stated succinctly, we have that

$$\mathbf{0} = \boldsymbol{A}_{\nu 1} \leq \boldsymbol{p}_{\nu 1}, \qquad \nu = 1, \ldots, M.$$

Assume that $[\boldsymbol{I} - \boldsymbol{A}]$ is nonsingular, which directly implies that $1 \notin \sigma_{\boldsymbol{A}}$; that is, 1 is not an eigenvalue of $\boldsymbol{A}$. Since $\boldsymbol{p}$ is a row-stochastic matrix, and because of the equivalence given in (6), the Gerschgorin Circle Theorem (see [20, Eqn. 7.1.13]) indicates that the spectral radius $\delta = \rho(\boldsymbol{A}) \leq 1$. Furthermore, the nonnegativity of $\boldsymbol{A}$ permits the use of Equation 8.3.1 of [20] to then assert that the Perron root $0 \leq \delta \leq 1$ exists. However, since we have shown that $1 \notin \sigma_{\boldsymbol{A}}$, it must then be the case that $\delta < 1$. This implies by Proposition 2.4 that $\rho(\boldsymbol{A}_{\nu\nu}) < 1$ for all $\nu \in \{1, \ldots, M\}$ and hence, by Proposition 2.3, each diagonal block $\boldsymbol{A}_{\nu\nu}$, $\nu \in \{1, \ldots, M\}$ must be substochastic.

We now consider the νth diagonal block in the canonical form of $\boldsymbol{A}$, where $\nu \in \{2, \ldots, M\}$, and proceed to show that each state i associated to the vertex set $V(\mathcal{G}(\boldsymbol{A}_{\nu\nu}))$ can access state 1. Because $\boldsymbol{p}$ is a row-stochastic matrix and $A_{\nu\nu}$ is substochastic, either or both of the following may hold:

1. $\boldsymbol{p}_{\nu 1} \neq \mathbf{0}$, or

2. $\boldsymbol{A}_{\nu\kappa} \neq \mathbf{0}$ for some $\kappa > \nu$.

For 1), $\boldsymbol{p}_{\nu 1} \neq \mathbf{0}$ indicates the existence of states $i_\nu \in V(\mathcal{G}(\boldsymbol{A}_{\nu\nu}))$ (with $i_\nu = i$ possible, but not necessary) and $1 \in V(\mathcal{G}(\boldsymbol{A}_{11}))$ for which there is a directed arc $(i_\nu, 1)$. Moreover, the irreducibility of $A_{\nu\nu}$ gives a directed path from i to i_ν. We thus obtain

$$i \to i_\nu \to 1.$$

In other words, there is a directed path from i to 1.

If 2) holds, there exists a directed arc from some state $i_\nu \in V(\mathcal{G}(\boldsymbol{A}_{\nu\nu})$ (again, with the possibility that $i_\nu = i$) to a state $i_\kappa \in V(\mathcal{G}(\boldsymbol{A}_{\kappa\kappa}))$. From here, we are again confronted with choices 1) and 2). If 1) holds, then the previous argument gives us a directed path from i_κ to 1. Since the irreducibility of $\boldsymbol{A}_{\nu\nu}$ implies the existence of a path from i to i_ν, we have the accessibility chain

$$i \to i_\nu \to i_\kappa \to 1,$$

and we are done. Otherwise, we proceed to the next diagonal block following $\boldsymbol{A}_{\kappa\kappa}$ and continue until $\nu > r$. If $\nu > r$ then the process is in a state $i_\nu \in V(\mathcal{G}(\boldsymbol{A}_{\nu\nu}))$. The only choice here, due to the this block being substochastic, is 1); that is, $\boldsymbol{p}_{\nu 1} \neq \mathbf{0}$, for which we have already demonstrated the existence of the connection $i_\nu \to 1$. Each of the preceding paths may then be combined to form a single directed path from an arbitrarily selected $i \in V(\mathcal{G}(\boldsymbol{A}_{\nu\nu}))$ to 1 so that

$$i \to i_\nu \to i_\kappa \to \cdots \to i_M \to 1.$$

Thus, state 1 is UA.

For the reverse implication, we will assume that state 1 is UA, and then proceed to show that $[\boldsymbol{I} - \boldsymbol{A}]$ is nonsingular. The reducibility of $\boldsymbol{A}$ allows us to assume that it possesses canonical form and, furthermore, that each submatrix on the diagonal of the canonical matrix corresponding to $\boldsymbol{A}$ is irreducible or zero. Consider an arbitrary nonzero, and hence irreducible, diagonal submatrix $\boldsymbol{A}_{\nu\nu}$ for some $\nu \in \{2, \ldots, M\}$ (recall that $\boldsymbol{A}_{11} = \boldsymbol{0}$ by definition of $\boldsymbol{A}$). By the assumption that state 1 is UA, there must be a directed path from each state in the vertex set $V(\mathcal{G}(\boldsymbol{A}_{\nu\nu}))$ to 1, which in turn implies that $\boldsymbol{A}_{\nu\nu}$ is substochastic. By Proposition 2.3, $\rho(\boldsymbol{A}_{\nu\nu}) < 1$. Using this fact, and the fact that the spectral radii of the zero submatrix blocks are 0, we may invoke Proposition 2.4, to state that $\rho(\boldsymbol{A}) < 1$. Hence, $[\boldsymbol{I} - \boldsymbol{A}]$ is nonsingular, which completes the proof.

For the following main result, we will show, using Lemma 4.1, that state j being UA is sufficient to derive a closed-form analytical expression for the rth first passage moments $\boldsymbol{\mu}^{(r)} = \left[\mu_{ij}^{(r)}\right]$, for $r \geq 1$ and for any given state $i \in \mathcal{S}$.

Theorem 4.2 *Let $\{Z(t) : t \geq 0\}$ be a regular, time-homogeneous SMP with a finite state space $\mathcal{S}$. Further suppose that $j \in \mathcal{S}$ is UA. Then the rth moments of the first passage times from all states $i \in \mathcal{S}$ to state j contained in the m-vector ($m = |\mathcal{S}|$)*

$$\boldsymbol{\mu}_j^{(r)} = \left[\mu_{ij}^{(r)}\right]_{i=1}^{m}, \quad r \geq 1,$$

are solutions to the system of equations given by

$$\boldsymbol{\mu}_j \equiv \boldsymbol{\mu}_j^{(1)} = [\boldsymbol{I} - \boldsymbol{p}\boldsymbol{I}(-j)]^{-1} (\boldsymbol{p} \circ \boldsymbol{e})\boldsymbol{1}, \tag{7}$$

$$\begin{aligned}\boldsymbol{\mu}_j^{(r)} &= [\boldsymbol{I} - \boldsymbol{p}\boldsymbol{I}(-j)]^{-1} \\ &\quad \times \left[\left(\boldsymbol{p} \circ \boldsymbol{e}^{(r)}\right) \boldsymbol{1} + \sum_{s=1}^{r-1} \binom{r}{s} \left[\left(\boldsymbol{p} \circ \boldsymbol{e}^{(r-s)}\right) \left((\boldsymbol{J} - \boldsymbol{I})_j \circ \boldsymbol{\mu}_j^{(s)} \right) \right] \right], \quad \textit{if } r > 1,\end{aligned} \tag{8}$$

where $\boldsymbol{1}$ is a column vector of ones and the scalar entries $\mu_{ij}^{(1)}$ and $\mu_{ij}^{(r)}$ for $r \geq 2$ are defined as in (3) *and* (4)*, respectively.*

Proof We first show, using induction on the rth moment, $r \geq 1$, that the system of equations (3) and (4) give a valid relationship between the first-passage moments to a given state j that is UA. For the mean time of first passage given by the system (3), we observe the following at the first transition epoch S_1 of the SMP:

1. $i \nrightarrow k$ at $S_1 \Rightarrow$ the corresponding kth term drops out of the expression, and
2. $i \rightarrow k$ at $S_1 \Rightarrow \mu_{ik}$ and μ_{kj} are well-defined, the latter because j is UA.

We thus conclude that a first-step analysis founded upon the state of the SMP at the first transition epoch S_1 (c.f. Proposition 5.15 of [24, pg104]) still holds for a terminal UA state j. Next, for the induction step, we consider expression (4) for the $(r+1)$th moment, where $r \geq 1$. We likewise claim that the original renewal argument given in Lemma 4.1 of [29] for the derivation of (4) for the rth moments of first passage is valid. In order to see this, we rewrite, for $i \in \mathcal{S}$, expression (4) as

$$\mu_{ij}^{(r+1)} = \sum_{k=1}^{m} p_{ik} \left[\left(1 - \delta_{kj}\right) \mu_{kj}^{(r+1)} + \mu_{ik}^{(r+1)} \right] + M_r \tag{9}$$

where

$$M_r = \sum_{s=1}^{r} \binom{r}{s} \left[\sum_{k \neq j} p_{ik} e_{ik}^{(r-s)} \mu_{kj}^{(s)} \right] .$$

The inductive hypothesis and items 1) and 2) above guarantee that M_r is well-defined while the remainder of (9) is in exactly the same form as (3), which has just been shown to have a finite solution via the base step.

Thus, for arbitrary $i \neq j$, where $i \in \mathcal{S}$, we may transform (3) into the equivalent matrix expression

$$\boldsymbol{\mu} = [\mu_{ij}] = \boldsymbol{p}\big((\boldsymbol{J} - \boldsymbol{I}) \circ \boldsymbol{\mu} \big) + \big(\boldsymbol{p} \circ \boldsymbol{e}\big)\boldsymbol{J} \ .$$

In this form we are not able to solve directly for $\boldsymbol{\mu}$, but, under the assumption that j is a specific UA state in $\mathcal{S}$, it is possible to solve for the j^{th} column of $\boldsymbol{\mu}$, which we denote as $\boldsymbol{\mu}_j$. We then obtain,

$$\boldsymbol{\mu}_j = \boldsymbol{p} \left[\big((\boldsymbol{J} - \boldsymbol{I}) \circ \boldsymbol{\mu} \big) \right]_j + \big(\boldsymbol{p} \circ \boldsymbol{e}\big)\boldsymbol{1} \ .$$

Next, we isolate $(\boldsymbol{p} \circ \boldsymbol{e})\boldsymbol{1}$ so that

$$\boldsymbol{\mu}_j - \boldsymbol{p} \left[\big((\boldsymbol{J} - \boldsymbol{I}) \circ \boldsymbol{\mu} \big) \right]_j = \big(\boldsymbol{p} \circ \boldsymbol{e}\big)\boldsymbol{1} \ .$$

Factoring out $\boldsymbol{\mu}_j$ gives

$$[\boldsymbol{I}-\boldsymbol{p}\boldsymbol{I}(-j)]\,\boldsymbol{\mu}_j = \big(\boldsymbol{p}\circ\boldsymbol{e}\big)\boldsymbol{1}$$

which allows us to finally solve for $\boldsymbol{\mu}_j$ as

$$\boldsymbol{\mu}_j = [\boldsymbol{I}-\boldsymbol{p}\boldsymbol{I}(-j)]^{-1}\big(\boldsymbol{p}\circ\boldsymbol{e}\big)\boldsymbol{1}\ .$$

By Lemma 4.1, the matrix $\boldsymbol{I}-\boldsymbol{p}\boldsymbol{I}(-j)$ is nonsingular. This proves that (7) is, indeed, well-defined.

A general formula for the rth moment, where $r \geq 2$, is given in Lemma 4.1 of [29] as

$$\mu_{ij}^{(r)} = \sum_{k=1}^{n} p_{ik}e_{ik}^{(r)} + \sum_{s=1}^{r}\binom{r}{s}\left[\sum_{k\neq j} p_{ik}e_{ik}^{(r-s)}\mu_{kj}^{(s)}\right],$$

which is expressed in matrix notation as

$$\boldsymbol{\mu}^{(r)} = \left(\boldsymbol{p}\circ\boldsymbol{e}^{(r)}\right)\boldsymbol{J} + \sum_{s=1}^{r}\binom{r}{s}\left[\left(\boldsymbol{p}\circ\boldsymbol{e}^{(r-s)}\right)\left((\boldsymbol{J}-\boldsymbol{I})\circ\boldsymbol{\mu}^{(s)}\right)\right].$$

Solving for the j^{th} column gives

$$\boldsymbol{\mu}_j^{(r)} = \left(\boldsymbol{p}\circ\boldsymbol{e}^{(r)}\right)\boldsymbol{1} + \sum_{s=1}^{r}\binom{r}{s}\left[\left(\boldsymbol{p}\circ\boldsymbol{e}^{(r-s)}\right)\left((\boldsymbol{J}-\boldsymbol{I})\circ\boldsymbol{\mu}^{(s)}\right)_j\right].$$

Using $\boldsymbol{e}^{(0)} = \boldsymbol{J}$ (the identity under the Hadamard product), we extract the rth term of the summation to obtain

$$\boldsymbol{\mu}_j^{(r)} - \boldsymbol{p}\left[(\boldsymbol{J}-\boldsymbol{I})\circ\boldsymbol{\mu}^{(r)}\right]_j = \left(\boldsymbol{p}\circ\boldsymbol{e}^{(r)}\right)\boldsymbol{1} + \sum_{s=1}^{r-1}\binom{r}{s}\left[\left(\boldsymbol{p}\circ\boldsymbol{e}^{(r-s)}\right)\left((\boldsymbol{J}-\boldsymbol{I})\circ\boldsymbol{\mu}^{(s)}\right)_j\right].$$

We further observe that

$$\boldsymbol{\mu}_j^{(r)} - \boldsymbol{p}\left[(\boldsymbol{J}-\boldsymbol{I})\circ\boldsymbol{\mu}^{(r)}\right]_j = (\boldsymbol{I}-\boldsymbol{p}\boldsymbol{I}(-j))\,\boldsymbol{\mu}_j^{(r)},$$

which gives

$$(\boldsymbol{I}-\boldsymbol{p}\boldsymbol{I}(-j))\,\boldsymbol{\mu}_j^{(r)} = \left(\boldsymbol{p}\circ\boldsymbol{e}^{(r)}\right)\boldsymbol{1} + \sum_{s=1}^{r-1}\binom{r}{s}\left[\left(\boldsymbol{p}\circ\boldsymbol{e}^{(r-s)}\right)\left((\boldsymbol{J}-\boldsymbol{I})\circ\boldsymbol{\mu}^{(s)}\right)_j\right].$$

Finally, we solve for $\boldsymbol{\mu}_j^{(r)}$ to obtain

$$\boldsymbol{\mu}_j^{(r)} = [\boldsymbol{I} - \boldsymbol{p}\boldsymbol{I}(-j)]^{-1} \left[\left(\boldsymbol{p} \circ \boldsymbol{e}^{(r)}\right) \mathbf{1} + \sum_{s=1}^{r-1} \binom{r}{s} \left[\left(\boldsymbol{p} \circ \boldsymbol{e}^{(r-s)}\right) \left((\boldsymbol{J} - \boldsymbol{I})_j \circ \boldsymbol{\mu}_j^{(s)} \right) \right] \right].$$

As argued in the proof of formula, (7), the inverse $[\boldsymbol{I} - \boldsymbol{p}\boldsymbol{I}(-j)]^{-1}$ exists. Hence, (8) is likewise well-defined.

We next investigate some statistical aspects in using Theorem 4.2 to estimate the first passage moments to universally accessible states in a SMP.

5 Estimation

In this section we will derive consistent estimates for first passage moments in SMPs. Since the SMP $\{Z(t) : t \geq 0\}$ is time-homogeneous, we assume without loss of generality that $Z(0) = i \in \mathcal{S}$. If we observe the SMP for a period of time $T > 0$, then, for any $j \in \mathcal{S}$, we may then define the point estimators $\hat{p}_{ij}$ for the probability and $\hat{e}_{ij}^{(r)}$ for the rth moment of the sojourn time of the SMP as it transitions from i to j as:

$$\hat{p}_{ij} \equiv \frac{n_{ij}}{\sum_{k \in \mathcal{S}} n_{ik}}, \qquad \hat{e}_{ij}^{(r)} \equiv \frac{1}{n_{ij}} \sum_{K=1}^{n_{ij}} x_{ijK}^r, \qquad i, j \in \mathcal{S},$$

where

$n_{ij} \equiv$ Number of observed transitions from state i to state j by time T,

$x_{ijK} \equiv$ K^{th} observed sojourn time from state i to state j by time T.

We further assume T is large enough so that at least one transition from i to j has been observed; in other words, $n_{ij} \geq 1$. In order to make inferential hypotheses using these estimators, it is useful to first show that they are consistent. A point estimator $\hat{\theta}_n$ is said to be *consistent* if it converges in probability to the true population statistic θ as the sample size n increases; that is, for each $\epsilon > 0$,

$$\lim_{n \to \infty} P\left(|\hat{\theta}_n - \theta| < \epsilon\right) = 1.$$

This condition is written in shorthand as

$$\hat{\theta}_n \xrightarrow{P} \theta.$$

We now show that this condition holds for the matrix estimators $\hat{\boldsymbol{p}} \equiv [\hat{p}_{ij}]$ and $\hat{\boldsymbol{e}} \equiv [\hat{e}_{ij}]$.

Lemma 5.1 *The matrix estimators $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{e}}$ are consistent.*

Proof Let $\{X_n\}$ be a sequence of Bernoulli random variables such that $X_n = 1$ when a transition from i to j occurs at the nth transition, and is 0 otherwise. Accordingly, if $N > 0$ transitions are observed in the time interval $(0, T]$, then the estimated probability of transition from i to j becomes

$$\hat{p}_{ij} = \frac{1}{N} \sum_{n=1}^{N} X_n,$$

with the following equivalences

$$N = \sum_{k \in \mathcal{S}} n_{ik}, \qquad n_{ij} = \sum_{n=1}^{N} X_n.$$

The Markov property at transitions of the embedded DTMC of the SMP implies that the X_n are independent and identically distributed (i.i.d.) random variables. Hence, by the Weak Law of Large Numbers (see Theorem 5.5.2 of [6, p.232]), we have

$$\hat{p}_{ij} \xrightarrow{P} p_{ij},$$

which demonstrates consistency.

Likewise, we see that the x_{ijK_1} are independent of x_{ijK_2} so long as $K_1 \neq K_2$. Thus, the collection $\{x_{ijK}\}_{K=1}^{n_{ij}}$ is i.i.d. By the same reasoning as above, we obtain the convergence in probability

$$\hat{e}_{ij} \xrightarrow{P} e_{ij},$$

Hence, the $\hat{e}_{ij}$ are consistent.

We are now in a position to define the estimators of the rth moments of first passage from state i to state $j \in \mathcal{S}$. By replacing $\boldsymbol{p}$ and $\boldsymbol{e}$ with the matrix estimators $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{e}}$, respectively, in formulas (7) and (8), we obtain the natural estimators

$$\hat{\boldsymbol{\mu}}_j \equiv [\boldsymbol{I} - \hat{\boldsymbol{p}}\boldsymbol{I}(-j)]^{-1} \left(\hat{\boldsymbol{p}} \circ \hat{\boldsymbol{e}}\right)\boldsymbol{J}_j, \tag{10}$$

$$\hat{\boldsymbol{\mu}}_j^{(r)} \equiv [\boldsymbol{I} - \hat{\boldsymbol{p}}\boldsymbol{I}(-j)]^{-1} \times \left[\left(\hat{\boldsymbol{p}} \circ \hat{\boldsymbol{e}}^{(r)}\right) \boldsymbol{J}_j + \sum_{s=1}^{r-1} \binom{r}{s} \left[\left(\hat{\boldsymbol{p}} \circ \hat{\boldsymbol{e}}^{(r-s)}\right) \left((\boldsymbol{J} - \boldsymbol{I})_j \circ \hat{\boldsymbol{\mu}}_j^{(s)} \right) \right] \right], \quad r \geq 2. \tag{11}$$

As expected, estimators (10) and (11) are also consistent.

Lemma 5.2 *For a state $j \in \mathcal{S}$ that is UA with respect to the digraph $\mathcal{G}(\hat{\boldsymbol{p}})$, the estimators $\hat{\boldsymbol{\mu}}_j^{(r)}$, $r \geq 1$, are consistent.*

Proof By Theorem 2.1.4 in [15, pg. 51], a continuous function of consistent estimators is itself a consistent estimator. By Lemma 4.1, and the assumption that j is UA with respect to $\mathcal{G}(\boldsymbol{p})$ (i.e., every state in $\mathcal{S}$ is linked to the state j in the digraph of $\hat{\boldsymbol{p}}$) we may assert the existence of $[\boldsymbol{I} - \hat{\boldsymbol{p}}\boldsymbol{I}(-j)]^{-1}$. The remainder of the terms in (10) and (11) are linear, and hence continuous. We thus conclude that $\hat{\boldsymbol{\mu}}_j^{(r)}$, for moments $r \geq 1$ and for each UA state $j \in \mathcal{S}$, are consistent estimators.

In this section we proposed consistent estimates for first passage moments of SMP. These estimates can be obtained if sufficient data is collected from observing the process.

6 Example

We give an example of an SMP and show how the first passage moments can be estimated. Therefore, given the process depicted in Figure 1 we have 3 transition distributions and a probability p. We will calculate the first passage moments using the direct transition moments, $\boldsymbol{e}$.

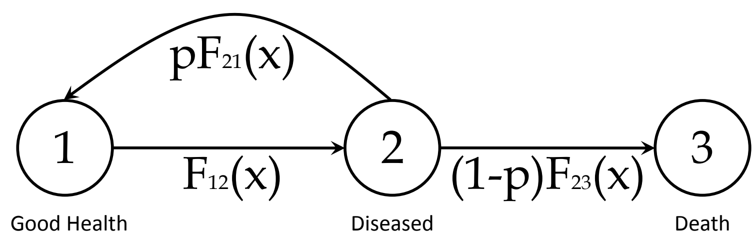

Figure 1: An example SMP of a medical patient.

To begin we have

$$\boldsymbol{p} = \begin{bmatrix} 0 & 1 & 0 \\ p & 0 & 1-p \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \boldsymbol{e} = \begin{bmatrix} 0 & e_{12} & 0 \\ e_{21} & 0 & e_{23} \\ 0 & 0 & e_{33} \end{bmatrix}.$$

Therefore

$$\boldsymbol{\mu}_3 = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 \\ p & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right)^{-1} \left(\begin{bmatrix} 0 & e_{12} & 0 \\ p\, e_{21} & 0 & (1-p)\, e_{23} \\ 0 & 0 & e_{33} \end{bmatrix} \right) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \frac{1}{1-p} \begin{bmatrix} 1 & 1 & 0 \\ p & 1 & 0 \\ 0 & 0 & 1-p \end{bmatrix} \begin{bmatrix} e_{12} \\ p\, e_{21} + (1-p)\, e_{23} \\ e_{33} \end{bmatrix}$$

$$= \frac{1}{1-p} \begin{bmatrix} e_{12} + p\, e_{21} + (1-p)\, e_{23} \\ p\, e_{12} + p\, e_{21} + (1-p)\, e_{23} \\ (1-p)\, e_{33} \end{bmatrix} \qquad (12)$$

Looking closely at the values in Equation 12 we see they are logical. As p gets small we see $\mu_{13} \to e_{12} + e_{23}$ and $\mu_{23} \to e_{23}$. This simple example demonstrates the theory discussed earlier; how even for large systems finding the first passage moments is only constrained by the computational burden of computing the inverse of $\boldsymbol{I} - \mathbf{pI}(-j)$.

If numerical values are substituted then numerical computer programs can handle these types of problems with relative ease. Now suppose we have

$$\boldsymbol{p} = \begin{bmatrix} 0 & 1 & 0 \\ 0.8 & 0 & 0.2 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \boldsymbol{e} = \begin{bmatrix} 0 & 6 & 0 \\ 0.7 & 0 & 1.1 \\ 0 & 0 & 0 \end{bmatrix}.$$

We get the following result

$$\boldsymbol{\mu}_3 = \begin{bmatrix} 33.9 \\ 27.9 \\ 0 \end{bmatrix}.$$

The R-code for this example is included in the appendix. The methods presented in this paper provide a fairly comprehensive method to determine the first passage moments of a SMP.

7 Conclusion

In this paper we devised an exact time-domain approach to derive the moments μ_{ij} of first passage time distributions in irreducible or *reducible* SMPs, given that the terminal state j fulfills the conditions of universal accessibility. Beyond the expanded generality of this method, it also has the advantage of obtaining the solution of first passage moments to only *single* UA states j, rather than to all states, thereby reducing the computational load, particularly for large SMPs. We have also demonstrated the existence of consistent point estimators for the first passage moments of processes that may be modeled as SMPs.

8 Acknowledgments

The authors thank Brian McBee and Dustin Mixon who have provided suggestions and advice in this work.

A R-code

```
First_Passage_Moments <- function(j,p,E) {
# j is the state in the process of interest
# p is the transition probability matrix
# E is a list of r matrices containing 1st, 2nd,...,rth moments

# The output of this function is a list of vectors of first
  passage moments

#checking for valid inputs
if (!is.list(E)) {stop("E must be a list of matrices")}
r <- length(E)
t1 <- prod(c(unlist(lapply(E,is.matrix)),is.matrix(p)))
if(t1 != 1) {stop("p must be a matrix & E a list of matrices")}
n = unique(c(dim(p),unlist(lapply(E,dim))))
 if ((n<2)==1 || length(n)!=1) {
   stop("p & the matrices in E must be nxn w/ n>1")}
W = prod(c(cbind(p),unlist(E))>=0)
 if(is.na(W) || W!=1) {
   stop("p & E must have valid nonnegative entries")}
j <- as.integer(j)
if(!(sum(j == 1:n))) {stop("j is not a valid state")}
if(!(prod(p%*%rep(1,n)==1)))
  {stop("p is not a stochastic matrix")}

#calculating the first passage moments
J=rep(1,n); I=diag(n); Id=I; Id[j,j]=0
result <- vector("list", r+1); result[[1]]=rep(1,n)
```

```
inv_matrix = solve(I-p%*%Id)
for (i in 1:r) {
  result[[i+1]]=0
  for (k in 0:(i-1)) {
    result[[i+1]] = result[[i+1]] +
    (choose(i,k)*p*(E[[i-k]]))%*%(((J-I)[,j])^k*result[[k+1]])
  }
  result[[i+1]] = inv_matrix%*%result[[i+1]]
}
return(result[2:(r+1)])
}

p <- matrix(c(0,1,0,.8,0,.2,0,0,1),nrow=3,byrow=T)
E <- matrix(c(0,6,0,0.7,0,1.1,0,0,0),nrow=3,byrow=T)
L <- list(E)

First_Passage_Moments(3,p,L)
```

References

[1] V. Barbu, M. Boussemart, and N. Limnios, *Discrete-time semi-Markov model for reliability and survival analysis*, Commun Stat-Theor M **33** (2004), no. 11, 28332868.

[2] Vlad Stefan Barbu and Nikolaos Liminios, *Semi-Markov chains and hidden semi-Markov models toward applications: Their use in reliability and DNA analysis*, Springer, New York, 2008.

[3] R.E. Barlow and F. Proschan, *Mathematical theory of reliability*, John Wiley & Sons, London and New York, 1965.

[4] Shelby Brumelle and Darius Walczak, *Dynamic airline revenue management with multiple semi-Markov demand*, Operations Research **51** (2003), no. 1, 137–148.

[5] Ronald W. Butler and Aparna V. Huzurbazar, *Stochastic network models for survival analysis*, J. Amer. Statist. Assoc. **92** (1997), 246–257. MR MR1436113 (98e:62140)

[6] George Casella and Roger L. Berger, *Statistical inference*, second ed., Duxbury, Pacific Grove, CA., 2002.

[7] Miroslav Fiedler, *Special matrices and their applications in numerical mathematics*, 2nd ed., Dover Publications, New York, 2008.

[8] J.J. Hunter, *On the moments of Markov renewal processes*, Advances in Applied Probability (1969), 188–210.

[9] Jacques Janssen and Raimondo Manca, *Semi-Markov risk models for finance, insurance, and reliability*, Springer, New York, 2007.

[10] J.G. Kemény and J.L. Snell, *Finite Markov chains*, University series in undergraduate mathematics, Van Nostrand, 1960.

[11] Jeffrey P. Kharoufeh, Christopher J. Solo, and M.Y. Ulukus, *Semi-Markov models for degradation-based reliability*, IIE Transactions **42** (2010), no. 8, 599–612.

[12] David Kincaid and Ward Cheney, *Numerical analysis: mathematics of scientific computing*, 3rd ed., vol. 2, American Mathematical Society, Providence, 2002.

[13] Leonard Kleinrock, *Queueing systems volume I: Theory*, John Wiley & Sons, New York, 1975.

[14] V.V. Kolpakov, *Matrix seminorms and related inequalities*, Journal of Mathematical Sciences **23** (1983), no. 1, 2094–2106.

[15] E. L. Lehmann, *Elements of large-sample theory*, Springer, New York, 1999.

[16] Steven R Lerman, *The use of disaggregate choice models in semi-Markov process models of trip chaining behavior*, Transportation Science **13** (1979), no. 4, 273–291.

[17] P. Levy, *Processus semi-Markoviens*, Proc. Intern. Congr. Math. **3** (1954), 416–426, Amsterdam, The Netherlands.

[18] ———, *Systems semi-Markoviens ayant au plus une inifinite denombrable d'etats possibles*, Proc. Intern. Congr. Math. **2** (1954), 294–295, Amsterdam, The Netherlands.

[19] N. Limnios and G. Oprişan, *Semi-Markov processes and reliability*, Birkhäuser, Boston, 2001.

[20] Carl D. Meyer, *Matrix analysis and applied linear algebra*, Society for Industrial Mathematics, Philadelphia, 2000.

[21] M.F. Neuts, *Structured stochastic matrices of $M/G/1$ type and their applications*, Probability: Pure and Applied, Marcel Dekker, Inc., New York and Basel, 1989.

[22] Ronald Pyke, *Markov renewal processes: Definitions and preliminary properties*, The Annals of Mathematical Statistics **32** (1961), no. 4, 1231–1242.

[23] ______, *Markov renewal processes with finitely many states*, The Annals of Mathematical Statistics **32** (1961), no. 4, 1243–1259.

[24] Sheldon M. Ross, *Applied probability models with optimization applications*, Dover, New York, 1992.

[25] Jia-Yu Shao, *Products of irreducible matrices*, Linear Algebra and Its Applications **68** (1985), 131–143.

[26] W. L. Smith, *Regenerative stochastic processes*, Proc. Roy. Soc (GB), series A, **232** (1955), 6–31, Amsterdam, The Netherlands.

[27] David D. Yao, *First-passage-time moments of Markov processes*, Journal of Applied Probability **22** (1985), no. 4, 939–945.

[28] H.L.S. Younes and R.G. Simmons, *Solving generalized semi-Markov decision processes using continuous phase-type distributions*, Proceeding of the National Conference on Artificial Intelligence, 2004, pp. 742–747.

[29] Xuan Zhang and Zhenting Hou, *The first-passage times of phase semi-Markov processes*, Statistics & Probability Letters **82** (2012), no. 1, 40 – 48.