Adding noise to Markov cohort models

Rowan Iskandar, PhD, MA
Department of Health Services, Policy, & Practice
Brown University School of Public Health
rowan_iskandar@brown.edu

This study received no financial suppport from any institutions or grants.

Abstract

Following its introduction over thirty years ago, the Markov state-transition cohort model has been used extensively to model population trajectories over time in decision modeling and cost-effectiveness studies. We recently showed that a cohort model represents the average of a continuous-time stochastic process on a multidimensional integer lattice governed by a master equation (ME), which represents the time-evolution of the probability function of a integer-valued random vector. From this theoretical connection, this study introduces an alternative modeling method, stochastic differential equation (SDE), which captures not only the mean behavior but also the variance. We first derive the continuous approximation to the master equation by relaxing integrality constraint of the state space in the form of Fokker Planck equation (FPE), which represents the time-evolution of the probability function of a real-valued random vector. Instead of working with the FPE, the SDE method constitutes time-evolution of the random vector of population counts. We derive the SDE from first principles and describe an algorithm to construct an SDE and solve the SDE via simulation for use in practice. We show the applications of SDE in two case studies. The first example demonstrates that the population trajectories, the mean and the variance, from the SDE and other commonly-used methods match. The second examples shows that users can readily apply the SDE method in their existing works without the need for additional inputs. In addition, in both examples, the SDE is superior to microsimulation in terms of computational speed. In summary, the SDE provides an alternative modeling framework and is less computationally expensive that microsimulation for a typical modeling problem in decision analyses.

1 Introduction

Decision models have been used in various applications from health technology assessment to clinical guideline development. In decision modeling, a Markov state-transition cohort model is often used to simulate the prognosis of a patient or a group of patients following an intervention. Beck and Pauker [1] introduced the modeling method over 30 years ago with the aim to provide a simple tool for prognostic modeling and for practical use in medical decision making. In principle, an multi-state Markov cohort model [1–3] is a recursive algebraic formula that calculates the *average* number of individuals in each state (*state-configuration*) in the next time step as the sum of current state-configuration and the change in the state-configuration during the time step. This calculation produces a deterministic quantity, i.e. the average of the process, and thus does not completely specify a stochastic representation of the modeled system.

We recently explicated the stochastic process underlying a cohort model, defined as a continuous-time discrete-state stochastic process, in the context where population counts are of interest.[3] The stochastic process is governed by a well-known evolution equation of a discrete probability function, known as a master equation. Given this knowledge of evolution equation associated with a cohort model, we can identify the relationship between cohort models and other known methodologies in stochastic process.[4] If we relaxed the integrality constraint on the state space to allow for non-negative real numbers, a cohort model can be immediately extended to represent the change in the *random* state-configuration by adding the variance of the process to the average of the process: *a random vector is expressed as the sum of its average and noise*. Information on the variance of a process is useful for quantifying the inherent variability of the population trajectories[5] and the extent of accuracy in the mean prediction as well as for deciding whether it is worth collecting more information about the decision problem.[6]

This paper aims to introduce an alternative modeling framework by capitulating the relationship between the stochastic process underlying a cohort model and the SDE. For this paper, we adopt a continuous-time perspective, i.e. for a discrete-time process, a corresponding continuous-time process exists (*embeddability assumption*). In addition, we assume that transition rates are time-invariant and the process of interest is Markovian, i.e. knowledge of the current state conveys all information that is relevant to forecasting future states. This paper is organized into three main parts. First, we formulate the stochastic cohort model process under three presentations. We first reintroduce the stochastic process underlying a Markov state-transition cohort model in the form of a master equation. Then, we briefly summarize the Fokker-Planck equation as a continuous approximation of the master equation underlying a cohort model. Lastly, we postulate an equivalent representation of the Fokker Planck equation in the form of an SDE which captures the average and the variance of the process. Secondly, we provide a step-by-step guidance for formulating a stochastic differential equation given a state-space and a set of transition rules and its intensity. Finally, we demonstrate two applications of SDE in decision modeling. The first example shows the equivalence between existing methods[3] and the SDE. In the second example, we develop an SDE for a recently published decision model evaluating the effect of chemotherapy-induced toxicity among breast cancer patients.[7] Throughout this exposition, we try to strike a balance between mathematical rigor and accessibility to practitioners.

2 Stochastic cohort model

A stochastic cohort model where each individual follows a Markov chain on a finite set of mutually exclusive and completely exhaustive health states: $\mathbb{S} = \{S_1, S_2, \ldots, S_s\}$, is defined as a stochastic process, $\{\mathbf{N}(t)\}_{t \geq t_0}$ where s is the number of states and $t \geq t_0$ for some initial time t_0. $\mathbf{N}(t)$ represents a random vector of population counts (*state-configuration*) :

$$\mathbf{N}(t) = [N_1(t)\ N_2(t)\ \ldots\ N_s(t)], \tag{1}$$

in which $\mathbf{N}(t) = \mathbf{n} \in \mathbb{Z}^s$ with the probability of observing a particular state-configuration, $\mathbf{N}(t) = \mathbf{n} \in \mathbb{Z}^s$ at time t or $P(\mathbf{N}(t) = \mathbf{n}) = P(\mathbf{n}, t)$. In this section, we first present the master equation which governs the stochastic process underlying a cohort model: a stochastic process on $\mathbb{Z}^s$. [3] Then, we show a continuous approximation of the master equation to obtain the Fokker Planck Equation, a stochastic process on $\mathbb{R}^s$, and introduce the SDE formulation.

2.1 Master equation

An individual transitioning from S_k to S_l is denoted by the *transition* R_i where $k, l \in \{1 \ldots, s\}$ and $i \in \{1 \ldots, r\}$ (r is the total number of transitions and s^2 is the maximum number of possible transitions). Since the occurrence of each transition alters the state-configuration, each R_i is associated with a vector representing this change, δ_i, and is defined as $\delta_i = e_l - e_k$ where e_j is the j-th column of an identity matrix of size s. We associate each transition with a *propensity function*: $\nu_i(\mathbf{n}, t) = c_i(t)h(n_k, n_l)$, where $c_i(t)$ is the transition rate for transition R_i, and $N_k(t) = n_k$ and $N_l(t) = n_l$. The function $h(n_k, n_l)$ determines the number of ways R_i can occur which is equal to the number of individuals in S_k, $h_{kl}(n_k, n_l) = n_k$, in the case where interaction between individuals in two different states has no effect. The propensity function denotes the rate of a (any) transition occurring in continuous time as a function of the state-configuration and rate constant. The probability of observing *one particular* transition R_i in a small time interval τ is assumed to be linear in time: $c_i(t)\tau + o(\tau)$, where $o(\tau)$ represents other terms that decay to zero faster than τ. Therefore, the probability of a transition from S_k to S_l in time step τ is given by $n_k c_i(t)\tau + o(\tau)$.[3] A transition R_i occurring with a probability $\nu_i(\mathbf{n}, t)\tau$ in a time step τ alters the current state $\mathbf{N}(t) = \mathbf{n}$ to:

$$\mathbf{N}(t + \tau) = \mathbf{n} + \delta_i \tag{2}$$

The master equation is obtained by enumerating all possible events contributing to a change in the probability of observing a particular state-configuration in an infinitesimal time step [3] and is given by:

$$\frac{dP(\mathbf{n}, t)}{dt} = \sum_{i=1}^{r} \left(\nu_i(\mathbf{n} - s_i, t)P(\mathbf{n} - \delta_i, t) - \nu_i(\mathbf{n}, t)P(\mathbf{n}, t)\right) \tag{3}$$

The terms, $\nu_i(\mathbf{n}, t)P(\mathbf{n}, t)$, represent the probability transferred from one to another state-configuration resulting from the transition R_i. In principle, a master equation expresses the change per unit of time of the probability of observing a state-configuration $\mathbf{n}$ as the sum of two opposite effects: the probability flux into and out of $\mathbf{n}$, and exactly represents the dynamics of the stochastic process underlying a cohort model.[1]

2.2 Fokker-Planck equation

In this section, we seek an alternative representation of the stochastic cohort model. Our aim is to approximate the Master equation by assuming that the change in the state-configuration is allowed to be non-integer, i.e. $\mathbf{N}(t)$ has a continuous sample path or $N_i(t+\tau) - N_i(t) \to 0$ as $\tau \to 0$ for $i \in \{1, 2, \ldots, s\}$. A Fokker-Planck equation (FPE) for the continuous-version of stochastic cohort model $\mathbf{N}(t)$ is a linear s-dimensional partial differential equation in the form of

$$\frac{\partial P(\mathbf{n},t)}{\partial t} = -\sum_{i=1}^{s} \frac{\partial \hat{A}_i(\mathbf{n},t)P(\mathbf{n},t)}{\partial n_i} + \frac{1}{2}\sum_{i=1}^{s}\sum_{j=1}^{s} \frac{\partial^2 \hat{B}_{ij}(\mathbf{n},t)P(\mathbf{n},t)}{\partial n_i \partial n_j}, \tag{4}$$

with $\hat{A}_i(\mathbf{n},t)$s are the elements of $1 \times s$ vector $\hat{\mathbf{A}}(\mathbf{n},t)$ and $\hat{B}_{ij}(\mathbf{n},t)$s are the elements of a $s \times s$ symmetric positive definite matrix $\hat{\mathbf{B}}(\mathbf{n},t)$. Furthermore, an FPE can be constructed by simply specifying $\hat{\mathbf{A}}(\mathbf{n},t)$ and $\hat{\mathbf{B}}(\mathbf{n},t)$ which are termed as infinitesimal mean and variance, respectively and defined as the following after taking $\Delta t \to 0$: $\frac{<\Delta n_i>_{\mathbf{n}}}{\Delta t} = A_i(\mathbf{n},t)$ and $\frac{<\Delta n_i \Delta n_j>_{\mathbf{n}}}{\Delta t} = B_{ij}(\mathbf{n},t)$ with $< . >_{\mathbf{n}}$ is an average with respect to the density $P(\mathbf{n},t)$.

The derivation of an FPE is based on the ideas of taking the discrete change in the master equation to the limits of infinitely small jumps.[8] The propensity function ($\nu(\mathbf{n})$) (which specifies a discrete jump of magnitude one in master equation) is redefined to allow smaller jump of magnitude t, by putting $\mathbf{n}' = \mathbf{n} + \mathbf{r}$, i.e. $\hat{\nu}(\mathbf{n},\mathbf{r})$. $\hat{\nu}(\mathbf{n},\mathbf{r})$) is assumed to concentrated at r and is slowly changing with the current configuration $\mathbf{n}$. Furthermore, $P(\mathbf{n},t)$ is assumed to vary slowly with $\mathbf{n}$ so that $P(\mathbf{n} \pm \mathbf{r},t)$ can be approximated by $P(\mathbf{n},t)$. If we put $\nu'(\mathbf{n},\mathbf{r})$ in the CME and expand $P(\mathbf{n},t)$ up to the second order, after further rearrangements Equation 3 then becomes Equation 4 (see [8] for full derivation). In principle, an FPE also represents, albeit approximately (vs. master equation), the dynamics of the stochastic process underlying a cohort model. The main difference lies in the domain of the state space and the form of the differential equation, where the problem of a (potentially) large system of ordinary differential equation is transformed to a relatively more solvable partial differential equation.

2.3 Stochastic differential equation

Most applications of cohort models include time-dependent parameters. In the context of time-varying transition rates, the analytical solution to an FPE is generally difficult to obtain, a problem akin to solving a master equation. In this section, we introduce a different approach for modeling the stochastic population dynamics, which is closely related to the FPE. Instead of working with the evolution equation for the $P(\mathbf{n},t)$, the stochastic differential equation (SDE) method considers the evolution of the random vector, $\mathbf{N}(t)$. We first derive an SDE model heuristically. Then, we outline the steps to develop an SDE for a given decision modeling problem in Section 3.

Given $\mathbb{S}$ and set of transitions R_i where $i \in \{1 \ldots, r\}$ (as in Section 2.1), the time evolution of a random vector, $\mathbf{N}(t)$, can be determined by enumerating all possible changes to $\mathbf{N}(t)$ in a small time step, τ, which is governed by R_i and δ_i, and their associated probabilities, which is governed by $\nu_i(\mathbf{n},t)\tau$. From Equation 2, we write the change in state-configuration in τ step as $\Delta\mathbf{N}(t) = \mathbf{N}(t+\tau) - \mathbf{N}(t)$.

The $1 \times s$ vector of the expected change in the state-configurations at time t is given by:

$$E[\Delta \mathbf{N}(t)] = \sum_{i=1}^{r} \delta_i \nu_i(\mathbf{n}, t)\tau$$

or in matrix form:

$$E[\Delta \mathbf{N}(t)] = \mathbf{v}(\mathbf{n}, t)\mathbf{d}\tau = \mathbf{A}(\mathbf{n}, t)\tau \tag{5}$$

where $\mathbf{v}$ is an $1 \times r$ vector of propensities, $\mathbf{d}$ is a $r \times s$ matrix of δs where the i-th row corresponds to the vector of changes associated with transition i, and $\mathbf{A}(\mathbf{n}, t)$ is a $1 \times s$ vector in which the i-th element corresponds to the rate of change in the number of individuals in state S_i. The $\mathbf{A}(\mathbf{n}, t)$ is essentially the right-hand side of the ordinary differential equation-equivalence to a cohort model [3, 9] and is equal to $\hat{\mathbf{A}}(\mathbf{n}, t)$ (Equation 4).[10] Similarly the covariance matrix of the change in the state-configuration in τ step can be calculated via the definition of the covariance matrix for a random vector:

$$\begin{aligned}
\Sigma(\Delta \mathbf{N}(t)) &= E[\Delta \mathbf{N}(t)\Delta \mathbf{N}(t)']E[\Delta \mathbf{N}(t)]E[\Delta \mathbf{N}(t)]' \\
&= \begin{pmatrix} \Sigma_{i=1}^{r} v_i(\mathbf{n}, t)d_{i1}^2 & \Sigma_{i=1}^{r} v_i(\mathbf{n}, t)d_{i1}d_{i2} & \dots & \Sigma_{i=1}^{r} v_i(\mathbf{n}, t)d_{i1}d_{is} \\ \Sigma_{i=1}^{r} v_i(\mathbf{n}, t)d_{i2}^2 & \Sigma_{i=1}^{r} v_i(\mathbf{n}, t)d_{i2}^2 & \dots & \Sigma_{i=1}^{r} v_i(\mathbf{n}, t)d_{i2}d_{is} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{i=1}^{r} v_i(\mathbf{n}, t)d_{is}^2 & \Sigma_{i=1}^{r} v_i(\mathbf{n}, t)d_{is}d_{i2} & \dots & \Sigma_{i=1}^{r} v_i(\mathbf{n}, t)d_{is}d_{is} \end{pmatrix} \\
&= \mathbf{B}(\mathbf{n}, t)\tau
\end{aligned} \tag{6}$$

if the terms with order higher than τ vanishes. We note that $\mathbf{B}(\mathbf{n}, t)$ is an $s \times s$ positive semi-definite matrix. Therefore, $\Sigma(\Delta \mathbf{N}(t))$ has a unique positive semi-definite square root, i.e. $\mathbf{S} = \sqrt{\mathbf{B}} = \sqrt{\frac{\Sigma(\Delta \mathbf{N}(t))}{\tau}}$, where $S = \mathbf{P}\sqrt{\mathbf{H}}\mathbf{P}$, $\mathbf{P}\mathbf{P}^{-1} = I$, and $\mathbf{H}$ is a diagonal matrix of eigenvalues of $\mathbf{B}$. For a sufficiently large $\mathbf{N}(t)$, $\Delta \mathbf{N}(t)$ is approximately normal with mean $\mathbf{A}(\mathbf{n}, t)\tau$ and variance $\mathbf{B}(\mathbf{n}, t)\tau$. Let $\xi(t)$ be a s-dimensional standard Wiener process, i.e. $\xi(t) \sim \mathcal{N}(0, \mathbf{I})$, where $\mathbf{I}$ is an $s \times s$ identity matrix. Therefore $\Delta \mathbf{N}(t) = \mathbf{A}(\mathbf{n}, t)\tau + \mathbf{S}\sqrt{\tau}\xi(t) \sim \mathcal{N}(\mathbf{A}(\mathbf{n}, t), \mathbf{V}\tau)$. Then, the evolution of a continuous stochastic cohort model can be represented by the following updating formula:

$$\begin{aligned}
\mathbf{N}(t + \tau) &= \mathbf{N}(t) + \Delta \mathbf{N}(t) \\
&= \mathbf{N}(t) + \mathbf{A}(\mathbf{n}, t)\tau + \mathbf{S}\sqrt{\tau}\xi(t)
\end{aligned} \tag{7}$$

If τ approaches 0, then we $\xi_i(t)\sqrt{(}\tau) \to W_i$ and arrive at the *Itô* SDE:

$$d\mathbf{N}(t) = \mathbf{A}(\mathbf{n}, t)dt + \mathbf{B}(\mathbf{n}, t)d\mathbf{W}(t) \tag{8}$$

where $\mathbf{W}(t)$ be a s-dimensional Wiener process. if for all t and t_0, the following equation holds:

$$\mathbf{N}(t) - \mathbf{N}(t_0) = \int_{t_0}^{t} \mathbf{A}(\mathbf{n}, s)ds + \int_{t_0}^{t} \mathbf{B}(\mathbf{n}, s)d\mathbf{W}(s) \tag{9}$$

The state-wise SDE is:

$$dN_i(t) = A_i(\mathbf{N}(t))dt + \sum_{j=1}^{s} B_{ij}(\mathbf{N}(t))dW_j$$

A more theoretical discussion on stochastic differential equations can be found in [11] and [10].

3 Developing an SDE

In this section, we describe an algorithm for developing an SDE for applications in decision modeling, which involves the following steps:

1. determine the state space ($\mathbb{S}$), all possible transitions (R_i), and their corresponding propensity functions, ν_i, for $i \in \{1, \ldots, r\}$;
2. construct the $1 \times r$ vector of propensities, $\mathbf{v}(\mathbf{n}, t)$, and the $r \times s$ matrix of changes, $\mathbf{d}$,
3. calculate the matrices of mean (Equation 5) and covariance (Equation 6) of changes;
4. choose a numerical method for approximately solving the SDE by generating a sample path: $\hat{\mathbf{N}}(t)^{\tau}_{[t_0,T],k}$, where: τ, $[t_0, T]$, and k, refers to the numerical time step, the time period of analysis, and the index of the sample, respectively;
5. and estimate a quantity of interest as a function of the sample path(s), i.e., $q = f(\hat{\mathbf{N}}(t)^{\tau}_{[t_0,T],k})$, where: q is the quantity of interest.

The first two steps are identical to specifying a deterministic cohort model. To calculate the drift and the diffusion matrices, we use Equation 5 and Equation 6, respectively. Standard methods to diagonalize a positive definite matrix can be utilized to calculate the square root of the diffusion matrix.[12]

For solving an SDE model, a numerical method, such as the Euler-Maruyama method [13], is often used since in most applications the analytical solution to an SDE would not be feasible. A typical numerical method is based on the idea of dividing the time interval of interest, $[t_0, T]$, to n discrete time points, which may not necessarily be uniformly-spaced: $t_0 = \tau_0 < \tau_1 < \tau_2 < \ldots < \tau_n - 1 < \tau_n = T$, and estimating the values of $\mathbf{N}(t)$ at these time points sequentially. For example, the Euler-Maruyama approximation, $\hat{\mathbf{N}}(t)$, of the stochastic process, $\mathbf{N}(t)$, is given by the following recursive equation:

$$\hat{\mathbf{N}}(\tau_{i+1}) = \hat{\mathbf{N}}(\tau_i) + A(\hat{\mathbf{N}}(\tau_i))\Delta\tau_i + \hat{\mathbf{B}}(\hat{\mathbf{N}}(\tau_i))d\mathbf{W_i} \tag{10}$$

where $\Delta\tau_i = \tau_{i+1} - \tau_i$ and $d\mathbf{W_i} = \mathbf{W}(\tau_{i+1}) - \mathbf{W}(\tau_i)$ denotes the increment of the Wiener process in the interval $[\tau_i, \tau_{i+1}]$ and is represented by independent Gaussian random variables with mean zeros and variance $\Delta\tau_i$. In practice, the $d\mathbf{W_i}$ is approximated by $\sqrt{\tau_{i+1} - \tau_i}\mathcal{N}(0, \mathbf{I})$. An application the numerical method will yield a sample path: $\hat{\mathbf{N}}(t)^{\tau}_{[t_0,T],k}$.

A quantity of interest, such as survival and prevalence, can be empirically estimated by generating a *sufficiently* large number of sample paths. For example if q represents life expectancy, then its estimator, $\hat{q}$ can be generated by simulating L number of sample paths:

$$\hat{q} = \frac{1}{L}\sum_{k=1}^{L} f(\hat{\mathbf{N}}(t)^{\tau}_{[t_0,T],k})$$

where f is defined as the ratio of the total sum of individuals across non-death states in $[t_0, T]$ to the initial population size. In general, the accuracy of the estimator depends on the choice of τs and L, which are problem-dependent.

4 Applications of SDE

We conduct two numerical exercises to demonstrate: (1) the equivalence of SDE to other methods and (2) how to develop a stochastic cohort model using the SDE method and to compare the computational time relative to microsimulation. In both exercises, the models are coded in R [14] and Julia [15] as open-source programs have emerged as the lingua franca for practitioners in decision modeling. [16, 17] The code for the both examples is available under a GNU GPL license and can be found at https://github.com/rowaniskandar/SDE.

4.1 Case study: 4-state Markov model

For the first example, we emulate the example in [3] to show the connection between an SDE and other modeling methods in the field of decision modeling. We consider a stochastic process with $s = 4$, i.e. $\mathbb{S} = \{S_1, S_2, S_3, S_4\}$ (Figure 1), rate constants: $c_1 = 0.05$, $c_2 = 0.01$, $c_3 = 0.001$, $c_4 = 0.1$, $c_5 = 0.05$, and $c_6 = 2$, and an initial condition in which all individuals start in S_1, i.e. $N_1(0) = 1000$. We simulate the population trajectories, i.e. the counts of individuals in all states, (1) using a cohort model with $\tau = 1$-year, (2) a ordinary differential equation (ODE) cohort model, (3) a microsimulation model, and (4) an SDE solved with Euler-Maruyama method and $\tau = 1$-year. To calculate the mean and the variances of $\mathbf{N}(t)$ based on the SDE model, we generate 10000 sample paths. The benchmark for variance is based on 10000 Monte Carlo samples of *microsimulation* population with $\tau = 1$-year. Each microsimulation population is conceived as a realization of 10000 Markov chains (or individual trajectories) with the generator matrix Q. The calculation of the 1-year transition probability matrix from the Q matrix is based on the formulas from Welton and Ades. [18] The results of the population trajectories are given in Figure (Figure 2). We observe no difference in the trajectories between the empirical estimates from the SDE and the cohort model and its ODE-equivalence.[3] Also, the estimates of the standard deviation from both SDE and microsimulation are approximately equal.

4.2 Case study: Breast cancer toxicity model

For the second example, we develop a cohort model for a published Markov decision model which evaluates the life expectancy associated with receiving chemotherapy among 55 year-old breast cancer patients.[7] In developing the SDE, we follow the steps in Section 3. The published (deterministic) cohort model has seven states: $\{S_H, S_T, S_M, S_{TM}, S_{DT}, S_{DM}, S_{DO}\}$ and 12 possible transitions (Table 1). The 1×7 drift vector is given by:

$$\mathbf{A}(\mathbf{n}, t) = \begin{pmatrix} -(c_T + c_M + c_{DO})N_H(t) \\ c_T N_H(t) - (c_M + c_T + c_{DO})N_T(t) \\ c_M(N_H(t) + N_T(t)) - (c_T + c_{DM} + c_{DO})N_M(t) \\ c_M(N_T(t) + N_M(t)) - (c_{DT} + c_{DM} + c_{DO})N_{TM}(t) \\ c_{DT}(N_T(t) + N_{TM}(t))) \\ c_{DM}(N_M(t) + N_{TM}(t)) \\ c_{DO}(N_H(t) + N_T(t) + N_M(t) + N_{TM}(t)) \end{pmatrix}$$

The 7×7 diffusion matrix can be similarly constructed. For example, the first and the fourth columns, $\mathbf{B}(\mathbf{n}, t)$, are equal to

$$B_{:1}(\mathbf{n}, t) = \begin{pmatrix} (c_T + c_M + c_{DO})N_H(t) \\ -c_T N_H(t) \\ -c_M N_H(t) \\ 0 \\ 0 \\ 0 \\ -C_{DO} N_H(t) \end{pmatrix} \text{ and } B_{:s}(\mathbf{n}, t) = \begin{pmatrix} 0 \\ -c_M N_T(t) \\ -c_T N_M(t) \\ c_M N_T(t) + c_T N_M(t) + (c_{DT} + c_{DM} + c_{DO}) N_{TM}(t) \\ -c_{DT} N_{TM}(t) \\ -c_{DM} N_{TM}(t) \\ -C_{DO} N_{TM}(t) \end{pmatrix},$$

respectively. The description and estimate of each parameter are given in [7]. Given the drift vector and the diffusion matrix, we apply the Euler-Maruyama method to generate sample paths of the SDE using a 1-year time step. We use the R package YUIMA [19] and simulate 10000 trajectories of $\mathbf{N}(t)$. As a comparison, We also develop a microsimulation model with the following generator matrix:

$$Q = \begin{pmatrix} -(c_T + c_M + c_{DO}) & c_T & c_M & 0 & 0 & 0 & c_{DO} \\ 0 & -(c_M + c_{DT} + c_{DO}) & c_M & 0 & c_{DT} & 0 & c_{DO} \\ 0 & 0 & -(c_T + c_{DM} + c_{DO}) & c_T & 0 & c_{DM} & c_{DO} \\ 0 & 0 & 0 & -(c_{DT} + c_{DM} + c_{DO}) & c_{DT} & c_{DM} & c_{DO} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & C0_M & 0 & 0 & 0 & 0 \end{pmatrix},$$

and generate 10000 realizations of $\mathbf{N}(t)$. From these sample paths, we estimate the life expectancies for the chemotherapy arm: 16.9(0.42) and 17.2(0.41) for the SDE and microsimulation methods, respectively. The SDE model requires a considerably less computational time than the microsimulation model to generate the same number of sample paths and to estimate the life expectancy.

5 Discussions

In this study, we introduce the SDE method for approximating the commonly used Markov state-transition cohort model, in the context where population counts are of interest. We first review the exact representation of the stochastic state-transition cohort model on $\mathbb{N}^{|\mathbb{S}|}$ and its evolution equation (master equation).[3] Then, we describe a continuous approximation to the master equation in the form of an FPE. Based on this approximation, we then establish a theoretical linkage between a cohort model (i.e. an evolution of a deterministic vector) and its stochastic counterpart, the SDE (an evolution of a stochastic vector). As a corollary, we place cohort models [1, 20] in a larger framework based on existing literature on stochastic process.[4]

In general, one can model the behavior of a *stochastic* dynamics system by considering either the time-evolution of the random vector (exemplified by SDE) or the probability function of the random vector (exemplified by master equation and FPE). Although the probability function approach potentially provides a way to derive an explicit form of the probability function (i.e. a complete knowledge of the stochastic system) without relying on numerous numerical realizations of the system, the random vector approach is more applicable in the following context. In most applications, an analytical solution to the probability function method is almost inaccessible; thereby rendering numerical methods as the common solution method. In particular, some of the transition rates in most decision rates are time-varying, for example: rate of death due to other causes is a time-dependent parameter. Practitioners often use Monte Carlo simulation (microsimulation) to estimate decision-relevant quantities from a cohort model. The first numerical exercise show that to estimate

the variance of the population counts across states, we need to replicate the microsimulation population a number of times *in addition to* the individual Monte Carlo runs within each population, which is a potentially computationally intensive task and is similar to applying uncertainty analysis to a microsimulation model.[21] In particular, if some of the events in the model are rare and/or the expected differences in the effect sizes among strategies are small, we expect the required number of Monte Carlo runs to be high. Although an SDE model also requires a Monte Carlo solution approach, the sample path of a cohort is generated more efficiently compared to microsimulation: SDE simulates a cohort *in a batch vs. individually* (as in microsimulation).

SDE models have been applied in various fields such as finance[22], infectious disease [23], physics[24], and biology[25]. To date, no study has examined the potential application of SDE, as an alternative to Markov cohort model, in decision modeling. The stochastic process underlying the commonly used cohort model, defined as a continuous-time discrete-state stochastic process, has only been recently demonstrated.[3] Hence, the theoretical connection between a cohort model and an SDE was unknown prior to [3]. In this study, we provide a tutorial to guide users through the process of developing an SDE. Developers of Markov cohort models should be familiar with the first few steps which involves:enumerating the possible health states, assigning the allowable transitions to all pairs of states, and specifying the intensity of each transition. One unique step to SDE requires the user to specify the vector of changes for each transition, i.e. a vector denoting changes in the two populations involved in a transition, the initial state (-1) and the destination states $(+1)$. After providing these *inputs*, the infinitesimal mean and variance can be calculated, which completely specifies an SDE model. The users are then tasked to choose a solver and conduct Monte Carlo simulation to estimate decision-relevant quantities. The implementation of the solver and the Monte Carlo method depends on the user's choice of programming language. In this study, the SDE models are implemented in R and we use an R package with a comprehensive SDE-related functionality to solve the SDEs.[19]

In this study, the derivation of an SDE model follows a heuristic approach. A more rigorous method seeks to prove that the form of the equation (Equation 8) $\Delta\mathbf{N}(t)$ satisfies is a direct consequence of the following assumptions: $\Delta\mathbf{N}(t)$ is Markovian, smooth in its argument, and has finite mean and variance. Readers seeking a more detailed derivation should consult [26]. In addition, we do not explicitly show the relationship between an FPE and SDE. In theory, the probability function of the solution to an SDE satisfies the FPE. This fundamental fact can be shown through the machinery of stochastic calculus which is beyond the scope of this study. For a rigorous proof, we refer readers to [27]. In summary, we opt for a simpler approach so users can readily appreciate the relationship between different representations of a cohort model and apply the methods in their works.

Another potential limitation of the study concerns the focus on continuous-time cohort models endowed with the embeddability property: the discrete-time transition probability matrix is embeddable to a continuous-time model, i.e. the limit of discrete-time transition probability matrix exists and its equal to the generator matrix of the process. Discrete-time Markov models have been the preferred method in many applications as the dynamic processes of interest, e.g. disease progression, evolve continuously; thereby the appropriate modeling framework would be a continuous-time model.[28] In some cases, an individual-level model naturally follows a discrete-time process, for example the transitions among living arrangements for an elderly patient occur in discrete time steps.[29] However, if one considers a sufficiently large cohort of individuals, the probability of one

individual transitioning in a small time interval would very likely to be nonzero. Nevertheless, future studies investigating whether the typical Markov models in decision modeling are embeddable are warranted.

6 Conclusions

Stochastic models can serve as powerful tools for aiding decision-makers in estimating decision-relevant outcomes and quantifying the impact of uncertainty on decision-making. Cohort model provides an easy-to-implement method for modeling recurrent events over time and has been used extensively in many applications ranging from clinical decision making to policy evaluations. The stochastic differential equation method expands the functionality of cohort models in practice by increasing the flexibility of cohort models and decreasing computational time of Monte Carlo methods.

Figure

Figure 1. A 4-states continuous-time stochastic process used in case study 1

Figure 2. Population trajectories across 4 states [CM: cohort model, MSM: microsimulation, ODE: continuous-time cohort model, SDE: stochastic differential equation; sd: standard deviation. Note: plots are on top of each other]

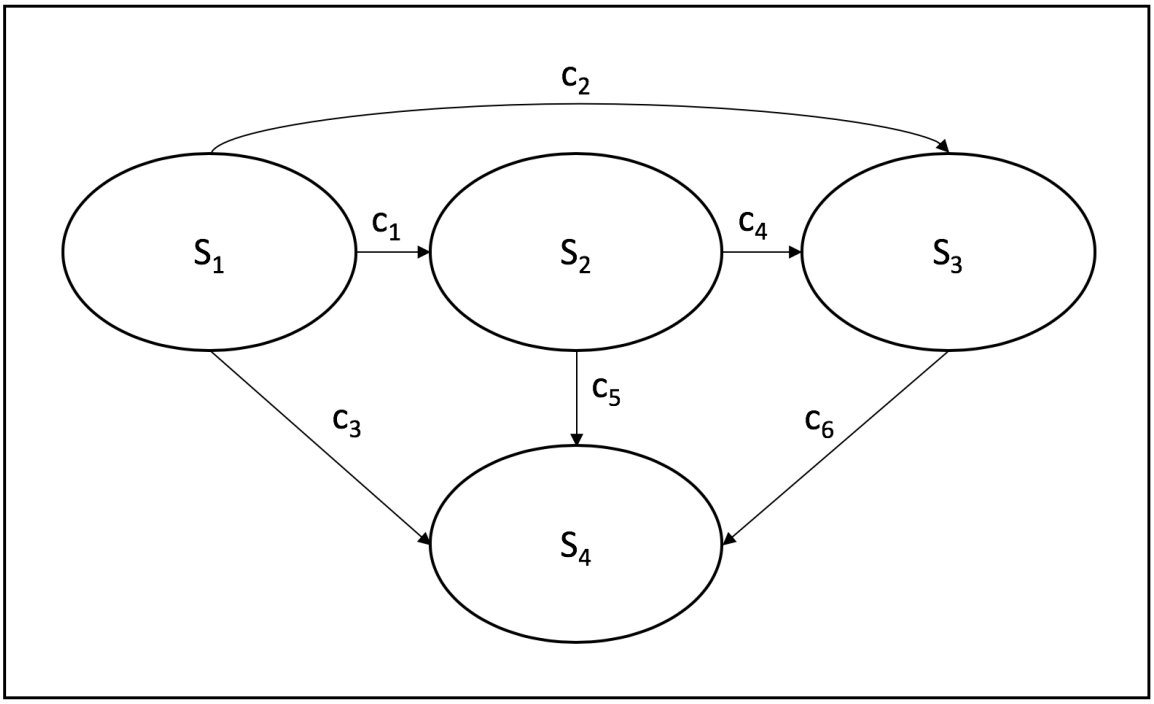

Figure 1:

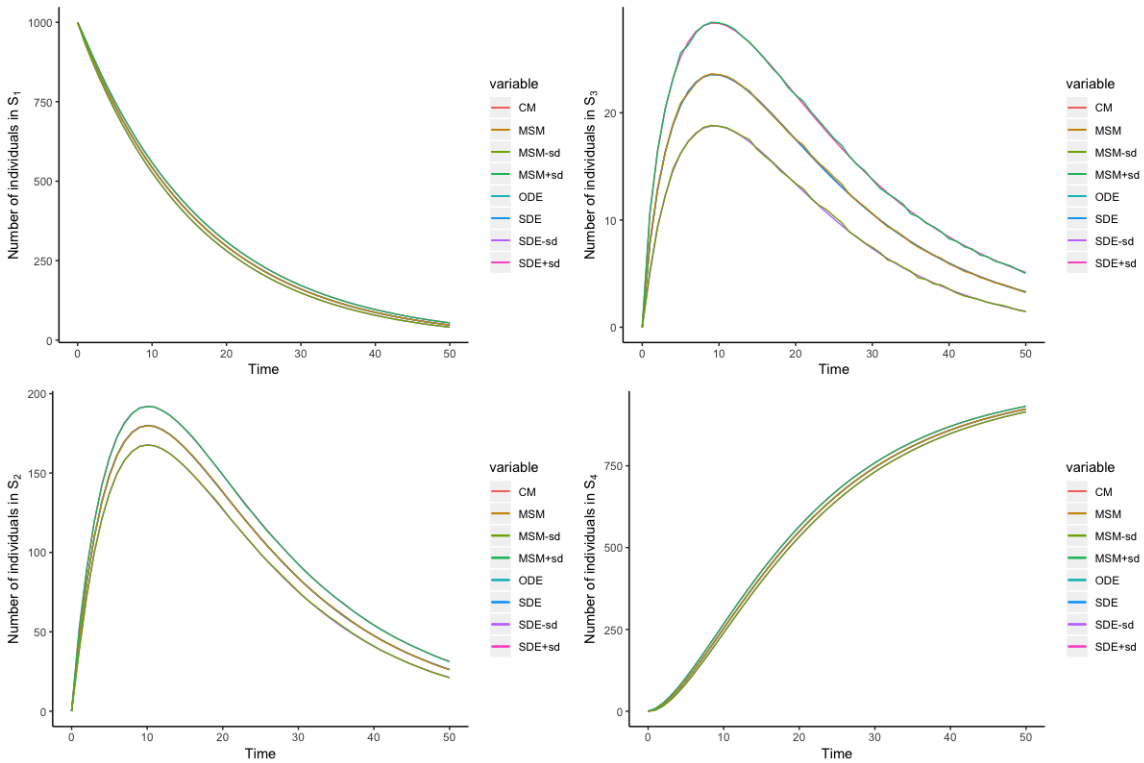
Number of individuals in S_1
Number of individuals in S_2
Number of individuals in S_3
Number of individuals in S_4
Time
variable
CM
MSM
MSM-sd
MSM+sd
ODE
SDE
SDE-sd
SDE+sd

Figure 2:

Table

Table 1. Propensities and vectors of changes for breast cancer toxicity model

Transitions	Propensities ($\mathbf{v}(\mathbf{n}, t)$)	Vectors of changes (δ elements of $\mathbf{d}$)
$S_H \to S_T$	$c_T N_H(t)$	$[-1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$
$S_H \to S_M$	$c_M N_H(t)$	$[-1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0]$
$S_H \to S_{DO}$	$c_{DO} N_H(t)$	$[-1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1]$
$S_T \to S_M$	$c_M N_T(t)$	$[0 \quad -1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0]$
$S_T \to S_{DT}$	$c_{DT} N_T(t)$	$[0 \quad -1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0]$
$S_T \to S_{DO}$	$c_{DO} N_T(t)$	$[0 \quad -1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1]$
$S_M \to S_{TM}$	$c_T N_M(t)$	$[0 \quad 0 \quad -1 \quad 1 \quad 0 \quad 0 \quad 0]$
$S_M \to S_{DM}$	$c_{DM} N_M(t)$	$[0 \quad 0 \quad -1 \quad 0 \quad 0 \quad 1 \quad 0]$
$S_M \to S_{DO}$	$c_{DO} N_M(t)$	$[0 \quad 0 \quad -1 \quad 0 \quad 0 \quad 0 \quad 1]$
$S_{TM} \to S_{DT}$	$c_{DT} N_{TM}(t)$	$[0 \quad 0 \quad 0 \quad -1 \quad 1 \quad 0 \quad 0]$
$S_{TM} \to S_{DM}$	$c_{DM} N_{TM}(t)$	$[0 \quad 0 \quad 0 \quad -1 \quad 0 \quad 1 \quad 0]$
$S_{TM} \to S_{DO}$	$c_{DO} N_{TM}(t)$	$[0 \quad 0 \quad 0 \quad -1 \quad 0 \quad 0 \quad 1]$

Table 1:

References

[1] J Robert Beck and Stephen G Pauker. The markov process in medical prognosis. *Med Decis Making*, 3(4): 419–458, 1983.

[2] MG Myriam Hunink, Milton C Weinstein, Eve Wittenberg, Michael F Drummond, Joseph S Pliskin, John B Wong, and Paul P Glasziou. *Decision making in health and medicine: integrating evidence and values*. Cambridge University Press, 2014.

[3] Rowan Iskandar. A theoretical foundation for state-transition cohort models in health decision analysis. *PloS one*, 13(12):e0205543, 2018.

[4] Mukhtar Ullah and Olaf Wolkenhauer. Family tree of markov models in systems biology. *IET systems biology*, 1(4):247–254, 2007.

[5] Aaron A Stinnett and John Mullahy. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Medical decision making*, 18(2_suppl):S68–S80, 1998.

[6] Karl Claxton. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *Journal of health economics*, 18(3):341–364, 1999.

[7] Fernando Alarid-Escudero, Anne H Blaes, and Karen M Kuntz. Trade-offs between efficacy and cardiac toxicity of adjuvant chemotherapy in early-stage breast cancer patients: Do competing risks matter? *The breast journal*, 23(4):401–409, 2017.

[8] NG Van Kampen. *Stochastic processes in physics and chemistry*. North Holland, 2007.

[9] Gerardus WJ Frederix, Johan GC van Hasselt, Johan L Severens, Anke M Hövels, Alwin DR Huitema, Jan AM Raaijmakers, and Jan HM Schellens. Development of a framework for cohort simulation in cost-effectiveness analyses using a multistep ordinary differential equation solver algorithm in r. *Medical Decision Making*, 33(6): 780–792, 2013.

[10] Edward Allen. *Modeling with Itô stochastic differential equations*, volume 22. Springer Science & Business Media, 2007.

[11] Thomas C Gard. *Introduction to Stochastic Differential Equations. Monographs and Text-books in pure and applied mathematics*. Dekker, Inc, 1988.

[12] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[13] Gisiro Maruyama. Continuous markov processes and stochastic equations. *Rendiconti del Circolo Matematico di Palermo*, 4(1):48–90, 1955.

[14] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL `https://www.R-project.org/`.

[15] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98, 2017. URL `https://doi.org/10.1137/141000671`.

[16] Decision Analysis in R for Technologies in Health (DARTH) workgroup. Decision Analysis in R for Technologies in Health, 2019.

[17] Hawre Jalal, Petros Pechlivanoglou, Eline Krijkamp, Fernando Alarid-Escudero, Eva Enns, and MG Myriam Hunink. An overview of r in health decision sciences. *Medical Decision Making*, 37(7):735–746, 2017.

[18] Nicky J Welton and AE Ades. Estimation of markov chain transition probabilities and rates from fully and partially observed data: uncertainty propagation, evidence synthesis, and model calibration. *Medical Decision Making*, 25(6):633–645, 2005.

[19] Alexandre Brouste, Masaaki Fukasawa, Hideitsu Hino, Stefano M. Iacus, Kengo Kamatani, Yuta Koike, Hiroki Masuda, Ryosuke Nomura, Teppei Ogihara, Yasutaka Shimuzu, Masayuki Uchida, and Nakahiro Yoshida. The yuima project: A computational framework for simulation and inference of stochastic differential equations. *Journal of Statistical Software*, 57(4):1–51, 2014. doi: 10.18637/jss.v057.i04. URL `http://www.jstatsoft.org/v57/i04/`.

[20] Frank A Sonnenberg and J Robert Beck. Markov models in medical decision making: a practical guide. *Medical decision making*, 13(4):322–338, 1993.

[21] Anthony O'hagan, Matt Stevenson, and Jason Madan. Monte carlo probabilistic sensitivity analysis for patient level simulation models: efficient estimation of mean and variance using anova. *Health economics*, 16(10):1009–1023, 2007.

[22] Nicole El Karoui, Shige Peng, and Marie Claire Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.

[23] Alison Gray, David Greenhalgh, Liangjian Hu, Xuerong Mao, and Jiafeng Pan. A stochastic differential equation sis epidemic model. *SIAM Journal on Applied Mathematics*, 71(3):876–902, 2011.

[24] Emil B Iversen, Juan M Morales, Jan K Møller, and Henrik Madsen. Short-term probabilistic forecasting of wind speed using stochastic differential equations. *International Journal of Forecasting*, 32(3):981–990, 2016.

[25] Kuang-Chi Chen, Tse-Yi Wang, Huei-Hun Tseng, Chi-Ying F Huang, and Cheng-Yan Kao. A stochastic differential equation model for quantifying transcriptional regulatory network in saccharomyces cerevisiae. *Bioinformatics*, 21(12):2883–2890, 2005.

[26] Daniel T Gillespie. The multivariate langevin and fokker–planck equations. *American Journal of Physics*, 64 (10):1246–1257, 1996.

[27] Iosif Ilyich Gikhman and Anatoli Vladimirovich Skorokhod. Stochastic differential equations. In *The Theory of Stochastic Processes III*, pages 113–219. Springer, 2007.

[28] Marta O Soares and Luísa Canto e Castro. Continuous time simulation and discretized models for cost-effectiveness analysis. *Pharmacoeconomics*, 30(12):1101–1117, 2012.

[29] Eric Jutkowitz, Robert L Kane, Joseph E Gaugler, Richard F MacLehose, Bryan Dowd, and Karen M Kuntz. Societal and family lifetime cost of dementia: implications for policy. *Journal of the American Geriatrics Society*, 65(10):2169–2175, 2017.