

Markov cohort state-transition model: A multinomial distribution representation

Rowan Iskandar and Cassandra Berns

Center of Excellence in Decision-Analytic Modeling and Health Economics
Research, sitem-insel, Bern, Switzerland (RI, CB); Brown University,
Providence, RI, USA (RI, CB); Institute of Social and
Preventive Medicine, University of Bern, Switzerland (RI).

Corresponding Author

Rowan Iskandar, Center of Excellence in Decision-Analytic Modeling
and Health Economics Research, sitem-insel, Freiburgstrasse 3, Bern,
3010, Switzerland; (rowan.iskandar@gmail.com)

Abstract

Markov cohort state-transition models have been the standard approach for simulating the prognosis of patients or, more generally, the life trajectories of individuals over a time period. Current approaches for estimating the variance of a Markov model using a Monte Carlo sampling or a master equation representation are computationally expensive and analytically difficult to express and solve. We introduce an alternative representation of a Markov model in the form of a multinomial distribution. We derive this representation from principles and then verify its veracity in a simulation exercise. This representation provides an exact and fast approach to computing the variance and a way of estimating transition probabilities in a Bayesian setting.

Highlights

- Markov model simulates the average experience of a cohort of patients.
- Monte Carlo simulation, the standard approach for estimating the variance, is computationally expensive.
- A multinomial distribution provides an exact representation of a Markov model.
- Using the known formulas of a multinomial distribution, the mean and variance of a Markov model can be readily calculated.

1 INTRODUCTION

Markov cohort state-transition models have been used in cost-effectiveness analyses and decision-analytic modeling studies to simulate the life trajectories of a patient or a group of patients following the receipt of a health intervention.¹ Briefly, a Markov cohort state-transition model is a recursive matrix formula that calculates the *average* number of individuals in each state using a transition probability matrix and an initial distribution of individuals across the states. This formulation introduced by Beck and Pauker,¹ and its subsequent tutorial,² however, does not fully capture the *stochastic* nature of the Markov model since it only represents the average behavior of the stochastic process.³ A more useful quantity is the probability distribution of all individuals at all times. To obtain this quantity, Iskandar³ explicated the stochastic process underlying a Markov model in the form of a well-known time-evolution equation of a probability function, i.e., a master equation, whose solution is the probability distribution of interest.³ Although the concept of a master equation and its solution provide a complete characterization of a Markov model, including higher-order moments, such a formulation requires some understanding of stochastic processes. This brief report introduces an alternative representation of a Markov model in the form of a multinomial distribution that is intuitive and hence more accessible to practitioners. We start with a description of a Markov cohort state-transition model. Then, we formally derive a multinomial distribution representation. Lastly, we conduct a numerical exercise to verify its equivalence to a canonical approach for simulating cohort models.

2 MARKOV COHORT STATE-TRANSITION MODEL

A cohort of n_0 individuals where each individual (indexed by $j = 1, \dots, n_0$) follows a Markov chain on a finite set of mutually exclusive and completely exhaustive s health states: $\mathbb{S} = \{S_1, S_2, \dots, S_s\}$, is defined as a stochastic process, $\{\mathbf{N}(t)\}_{t \geq t_0}$, where t and t_0 are time and the initial time, respectively. $\mathbf{N}(t)$ represents a random vector of individuals across s states (*state-configuration*):

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

in which $\mathbf{N}(t) = \mathbf{n} \in \mathbb{N}^s$ (\mathbb{N} is the set of non-negative integers). The probability of observing a particular state-configuration $\mathbf{N}(t) = \mathbf{n} \in \mathbb{N}^s$ at time t or $Pr[\mathbf{N}(t) = \mathbf{n}] := G(\mathbf{n}, t)$ (the notation $Pr[\cdot]$ denotes the probability of observing the random event $\{\cdot\}$ at time t) provides the complete information about the cohort model. We will show that $Pr(\mathbf{N}(t) = \mathbf{n})$ is a multinomial distribution.

3 MULTINOMIAL DISTRIBUTION REPRESENTATION

To derive the multinomial representation, we start by deriving the stochastic process for each individual and aggregate the individual processes. We then uncover the multinomial distribution by inspecting the functional form of the aggregated process.

Individual process. For each j -th individual, we define a $1 \times s$ random unit vector of health state occupancy, $\mathbf{Y}_j(t)$, where the value of the i -th element of $\mathbf{Y}_j(t)$, i.e., $Y_{j,i}(t)$, represents whether the individual is in state S_i : 1 (occupied) or 0 (not occupied), i.e., $\mathbf{Y}_j(t) \in \{0, 1\}^s$ and $\sum_{i=1}^s Y_{j,i}(t) = 1$ (an individual must be in one and only one state at any time). The dynamic of $\mathbf{Y}_j(t)$ is governed by the allowed transitions at discrete time steps ($t_z; z = 1, \dots$) between all pairs of states in \mathbb{S} and the intensities of these transitions. We define an $s \times s$ stochastic matrix of transition probabilities governing the intensities of the transitions among states in \mathbb{S} , i.e., $\mathbf{P}_\tau(t_z) = [p_{kl}(t_z)]_{1 \leq k, l \leq s}$, where τ denotes the time step (Markov cycle). Each $p_{kl}(t)$ has the usual interpretation of the probability of an individual transitioning from states S_k to S_l in one time step τ (Markov cycle) at time t_z . We are interested in deriving the probability that the individual j occupies a particular state at any time t_z given an initial condition. Such a probability is also a transition probability and can be derived by iteratively propagating the initial probability distribution of state occupancy (denoted by the $1 \times s$ vector $\mathbf{p}(t_z)$) at the initial time, i.e., $\mathbf{p}_j(t_0)$, forward in time using the transition probability matrix:

$$\mathbf{p}_j(t_z) = \mathbf{p}_j(t_0) \prod_{u=1}^{z-1} \mathbf{P}_\tau^j(t_u) \quad (2)$$

To derive the probability distribution of \mathbf{Y}_j , we note that there are s possible realizations of \mathbf{Y}_j . The probability of a realization, i.e., the probability of an individual in some state k , is equal to $[\mathbf{p}_j(t_z)]_k$, where $[\cdot]_k$ denotes the k -th element of a vector (e.g., $[\mathbf{y}_j(t_z)]_k = y_{j,k}(t_z)$). We write the probability distribution of $\mathbf{Y}_j(t_z)$ as follows:

$$Pr(\mathbf{Y}_j(t_z) = \mathbf{y}_j(t_z)) = \prod_{k=1}^s ([\mathbf{p}_j(t_z)]_k)^{[\mathbf{y}_j(t_z)]_k} \quad (3)$$

Equation 3 has the form of a multinomial distribution, where the number of individuals is equal to one ($n_0 = 1$). Each term in the product corresponds to the probability of individual j being in one of the s states at time t_z .

The expected value of $\mathbf{Y}_j(t_z)$ is equal to $Pr[\mathbf{Y}_j(t_z)]$ since the expected value of each element of $\mathbf{Y}_j(t_z)$ is equal to the expected value of an indicator function of whether the j -th individual occupies the corre-

sponding state, i.e., $E[Y_{j,k}(t_z)] = [\mathbf{p}_j(t_z)]_k = p_{j,k}(t_z)$. The covariance matrix of $\mathbf{Y}_j(t_z)$, denoted by $\Sigma_{\mathbf{Y}_j}(t_z)$, is calculated using the definition of a covariance matrix, $[\Sigma_{\mathbf{Y}_j}(t_z)]_{kl} = E[Y_{j,k}Y_{j,l}] - E[Y_{j,k}]E[Y_{j,l}]$, and is equal to:

$$[\Sigma_{\mathbf{Y}_j}(t_z)]_{kl} = \begin{cases} p_{j,k}(t_z)(1 - p_{j,l}(t_z)) & \text{if } k = l \\ -p_{j,k}(t_z)p_{j,l}(t_z) & \text{if } k \neq l \end{cases} \quad (4)$$

From individuals to a cohort. We relate $\mathbf{N}(t)$ (Equation 1) with $\mathbf{Y}_j(t)$ by using the following relationship:

$$\mathbf{N}(t) = \sum_{j=1}^{n_0} \mathbf{Y}_j(t) \quad (5)$$

where we write a cohort as the sum of its individual members. To derive the probability distribution of $\mathbf{N}(t)$, we need to associate each element in $\mathbf{N}(t)$ with the corresponding element in the sum. Each element in the sum consists of a vector of 1s and 0s, i.e., $\mathbf{Y}_j(t)$. The sum is constrained to be n_0 (a closed cohort). The number of possible ways that $\mathbf{N}(t)$ can be realized, given that ordering of individuals within group does not matter, is equal to the number of possible realizations of the vector $[\mathbf{Y}_1(t)\mathbf{Y}_2(t)\dots\mathbf{Y}_{n_0}(t)]$. Since all individuals are independent and follow the same transition rules and intensities ($\mathbf{P}_\tau^j(t) = \mathbf{P}_\tau(t)$, $p_{j,k}(t_z) = p_k(t_z) =$ for all ks , and $\mathbf{p}_j(t_0) = \mathbf{p}(t_0)$), each realization of $[\mathbf{Y}_1(t_z)\mathbf{Y}_2(t_z)\dots\mathbf{Y}_{n_0}(t_z)]$ occurs with a probability of

$$\begin{aligned} Pr([\mathbf{Y}_1(t_z)\mathbf{Y}_2(t_z)\dots\mathbf{Y}_{n_0}(t_z)]) &= \prod_{j=1}^{n_0} \prod_{k=1}^s p_k(t_z)^{y_{j,k}(t)} \\ &= \prod_{k=1}^s p_k(t_z)^{\sum_{j=1}^{n_0} y_{j,k}(t)} \\ &= \prod_{k=1}^s p_k(t_z)^{n_k(t)} \end{aligned} \quad (6)$$

Multinomial distribution. The number of ways that $\mathbf{N}(t)$ can be realized is identical to the problem of assigning n_0 of people into s groups, which is given by the multinomial coefficient: $\binom{n_0}{n_1 n_2 \dots n_s}$. Putting all components together, the probability of $\mathbf{N}(t_z)$ at time t_z is given by the following multinomial distribution:

$$Pr(\mathbf{N}(t_z)) = \binom{n_0}{n_1 n_2 \dots n_s} \prod_{k=1}^s p_k(t_z)^{n_k(t_z)} \quad (7)$$

To derive the first two moments, we use the first two moments of the individual process ($\mathbf{Y}_j(t)$) and the relationship between each individual and a cohort of individuals (Equation 5). By linearity of expectation,

the expected value of $\mathbf{N}(t_z)$ is equal to:

$$E[\mathbf{N}(t_z)] = n_0 \mathbf{p}(t_z) \quad (8)$$

By independence of individuals, the variance of $\mathbf{N}(t_z)$ is the sum of n_0 identical variance of $\mathbf{Y}_j(t_z)$ and is given by:

$$\Sigma_{\mathbf{N}}(t_z) = n_0 \Sigma_{\mathbf{Y}_j}(t_z) \quad (9)$$

The covariance matrix of $\mathbf{N}(t_z)$, denoted by $\Sigma_{\mathbf{N}}(t_z)$, is calculated using the definition of a covariance matrix, $[\Sigma_{\mathbf{N}}(t_z)]_{uv} = E[N_u N_v] - E[N_u]E[N_v]$, and is equal to:

$$[\Sigma_{\mathbf{N}}(t_z)]_{uv} = \begin{cases} n_0 p_u(t_z)(1 - p_u(t_z)) & \text{if } v = u \\ -n_0 p_u(t_z)p_v(t_z) & \text{if } v \neq u \end{cases} \quad (10)$$

4 NUMERICAL VERIFICATION

We conduct a simulation study to verify whether the multinomial distribution (Equation 7) represents the solution to a Markov cohort state-transition model. We consider a 4-state model with the allowed transitions and their probabilities as follows: $Pr[S_1 \rightarrow S_2] = 0.1$, $Pr[S_1 \rightarrow S_3] = 0.05$, $Pr[S_1 \rightarrow S_4] = 0.14$, $Pr[S_2 \rightarrow S_3] = 0.07$, $Pr[S_2 \rightarrow S_4] = 0.17$, and $Pr[S_3 \rightarrow S_4] = 0.11$. The simulation focuses on comparing the mean (Equation 8) and variance (Equation 10) of the multinomial distribution with those of a microsimulation. We simulate a cohort of 10000 individuals ($n_0 = 10000$) with $\tau = 1$ -year. Each individual's life trajectory is a realization of a Markov chain based on the given transition probabilities. We replicate the simulation 1000 times and calculate the means and variances of the number of individuals across states at all times. The results of the population trajectories across the four states are given in Figure 1. We observe no difference between the empirical estimates of the mean and variance from the microsimulation and the multinomial distribution. The code for the numerical exercise is available under a GNU GPL license and can be found at https://github.com/rowaniskandar/CM_multinomial.

5 CONCLUDING REMARKS

This study explicates an alternative representation of a Markov cohort state-transition model in the form of a multinomial distribution. We derive the equivalent representation by using elementary arguments. First, the derivation starts with specifying a Markov state-transition model for simulating an individual following Beck and Pauker.¹ We then extend the model to a cohort of individuals by imposing independence among

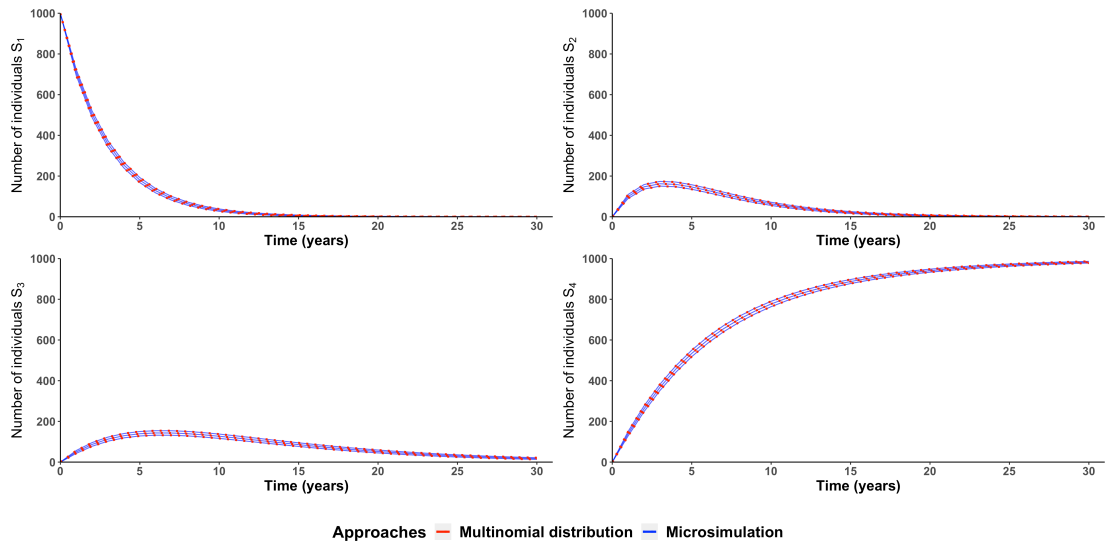


Figure 1. The time trajectories of individual counts in each state using two approaches, i.e., microsimulation (solid blue lines) and multinomial distribution (dotted red lines). For each approach, there are three lines, i.e., mean, mean+standard deviation, mean-standard deviation. The lines are on the top of each other, indicating no difference between the two approaches.

individuals and uncover the multinomial distribution by inspection. We verify the first two moments of the derived multinomial distribution of the number of individuals across health states against using a microsimulation. The formula for the first moment (Equation 8) is indeed the formula for simulating Markov models as introduced by Beck and Pauker.¹ To estimate the variance, practitioners often use microsimulations, which may be computationally expensive since we need to replicate the cohort simulation a number of times in addition to the individual Monte Carlo runs within each simulated cohort. In contrast, the multinomial representation provides a more direct and non-computationally demanding approach, particularly for estimating the variance. This approach relies only on the formula for the mean, which in turn depends only on the initial distribution and the time-dependent or time-invariant transition probability matrix. In addition to the computational advantage, the multinomial representation provides a convenient way to conduct Bayesian inference on the transition probabilities.⁴ In a Bayesian setting, the Markov model would be treated as the likelihood and naturally takes the form of a multinomial distribution using the result of this study. We can then derive the posterior distribution of the transition probabilities by utilizing the conjugacy between the Dirichlet and multinomial distributions. In sum, this study introduces another formulation of a well-established methodology and reinforces the utility of Markov models further.

REFERENCES

- ¹ J Robert Beck and Stephen G Pauker. The markov process in medical prognosis. *Med Decis Making*, 3(4):419–458, 1983.
- ² Frank A Sonnenberg and J Robert Beck. Markov models in medical decision making: a practical guide. *Medical decision making*, 13(4):322–338, 1993.
- ³ Rowan Iskandar. A theoretical foundation for state-transition cohort models in health decision analysis. *PloS one*, 13(12):e0205543, 2018.
- ⁴ Peter Walley. Inferences from multinomial data: learning about a bag of marbles. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):3–34, 1996.