J. Appl. Prob. **52,** 457–472 (2015)
Printed in England

ON BINOMIAL OBSERVATIONS OF CONTINUOUS-TIME MARKOVIAN POPULATION MODELS

N. G. BEAN * AND
R. ELLIOTT,* *The University of Adelaide*
A. ESHRAGH,** *The University of Newcastle*
J. V. ROSS,* *The University of Adelaide*

Abstract

In this paper we consider a class of stochastic processes based on binomial observations of continuous-time, Markovian population models. We derive the conditional probability mass function of the next binomial observation given a set of binomial observations. For this purpose, we first find the conditional probability mass function of the underlying continuous-time Markovian population model, given a set of binomial observations, by exploiting a conditional Bayes' theorem from filtering, and then use the law of total probability to find the former. This result paves the way for further study of the stochastic process introduced by the binomial observations. We utilize our results to show that binomial observations of the simple birth process are non-Markovian.

Keywords: Continuous-time Markovian population model; binomial observation; simple birth process; filtering

2010 Mathematics Subject Classification: Primary 60J27
Secondary 60J80; 62M09; 62M20

1. Introduction

Stochastic processes which are partially observed are a common phenomenon when attempting to recognize patterns (see, e.g. [10]), to analyze digital signals (see, e.g. [26]), or to understand financial processes (see, e.g. [7], [28]), or biological processes (see, e.g. [1], [2]). Given the observations, one wishes to make optimal estimates concerning the underlying, hidden, process. That is, one would like to evaluate the conditional probability mass function (PMF) of the true state of the process conditioned on the sequence of observations to date. The procedure to achieve this goal is called *filtering* and has been extensively studied in the engineering literature; see [8] and [9] for an introduction.

Throughout the study of biology there is a strong interest in observation error, as partial observation is a feature common to most studies in that discipline. This can be seen in ecology (see, e.g. [6], [11], [24], [25]), genetics (see, e.g. [20]), and epidemiology (see, e.g. [22]). A usual form of partial observation is one where the state of the system, or each component of the system, can be observed with a fixed probability p at each observation time. For example, consider the incidence of infection in a population, where each infectious individual seeks medical attention and, hence, is recorded in the data, and then tests positive to infection

Received 30 August 2013; revision received 9 April 2014.
* Postal address: School of Mathematical Sciences, The University of Adelaide, Adelaide, SA 5005, Australia.
** Postal address: School of Mathematical and Physical Sciences, The University of Newcastle, Callaghan, NSW 2308, Australia. Email address: ali.eshragh@newcastle.edu.au

https://doi.org/10.1239/jap/1437658609 Published online by Cambridge University Press

(conditional on being infected), with a fixed probability (see, e.g. [12], [13]). Yet another example is in biological invasions, where the species, or each individual of the species, is only detected with a certain probability upon each survey (see, e.g. [19]). Furthermore, a class of stochastic processes often used in modeling biological phenomena is the model of continuous-time Markovian population (see, e.g. [5], [15], [21]). Hence, the study of *continuous-time Markovian population models* (CTMPMs) under a binomial observation process is of interest and should prove useful in biological applications. We refer below to such processes as a *partially-observable continuous-time Markovian population model* (POCTMPM).

The first contribution of this paper is the explicit evaluation of the conditional PMF of the true state of a partially-observed continuous-time Markovian population model, and of the conditional PMF of the observed state, given a set of past observations. This result for general continuous-time Markovian population models makes use of the reference probability method, an abstract form of Bayes' theorem, as detailed in [9]. It paves the way for a more detailed study and further applications of this class of processes.

One of the simplest continuous-time Markovian population models is the *simple birth process* (SBP); it is a pure-birth process, (that is, the state of the chain only increases, and then only increases by one, at each transition), with transition rate λx_t when the state of the chain is $x_t (> 0)$ for some *birth* rate λ. Despite its simplicity, it has found application in modeling many biological processes, including evolutionary processes (see, e.g. [14], [16], [18], [27]) and epidemiological processes (see, e.g. [4]) and ecological processes (see, e.g. [21]).

The second contribution of this paper is a proof that the partially-observable SBP (that is, the SBP with binomial observations with fixed probability p) is non-Markovian for $0 < p < 1$. In fact, we prove that it is not Markovian of any order. Whilst for many this statement might not be surprising, its proof is not trivial and exploits our earlier results. It establishes the necessity of a new line of analysis when considering such a process, as considered in [3].

2. Partially-observable continuous-time Markovian population models

Suppose that $\{X_t, t \geq 0\}$ is a CTMPM with unknown parameter vector $\boldsymbol{\theta}$. The vector $\boldsymbol{\theta}$ parameterises the q-matrix (generator) $Q(\boldsymbol{\theta})$ of the CTMPM. We restrict our attention to CTMPMs where the range of the random variable X_t is the nonnegative integers and the initial value of this process, x_0, is known. Moreover, we suppose that the process is time-homogeneous, that is the conditional probability $\mathbb{P}_{\{X_{t_2} \mid X_{t_1}\}}\{x_{t_2} \mid x_{t_1}\}$ for any values of $t_2 > t_1 > 0$ depends only on $t_2 - t_1$, (and, of course, x_{t_1} and x_{t_2}). Hence, for simplicity, we denote the latter conditional probability by $\mathcal{P}_{x_{t_1}, x_{t_2}}(t_2 - t_1)$. In order to estimate the unknown parameter vector $\boldsymbol{\theta}$, we take n observations of $\{X_t, t \geq 0\}$ at times $0 < t_1 \leq t_2 \leq \cdots \leq t_n$ (Clearly, having the first observation at $t_1 = 0$ is pointless as we assume that x_0 is known. Thus, we set $t_1 > 0$.). Suppose that at each observation time t_i, we do not observe X_{t_i} directly, but rather only a random sample. This may be due to practical restrictions such as time or budget constraints which limit the ability to survey comprehensively, or might be because of an implicit component of the data collection process.

A common model for the sampling is binomial, where the state of the system, or each component of the system, is observed with a fixed probability p at each observation time. We therefore define the partially-observable continuous-time Markovian population model as follows.

Definition 1. Consider the CTMPM $\{X_t, t \geq 0\}$ with the parameter vector $\boldsymbol{\theta}$. Suppose that the random variables Y_t are defined such that the conditional random variable $\{Y_t \mid X_t = x_t\}$ follows the $\text{bin}(x_t, p)$ distribution, that is

$$\mathbb{P}_{\{Y_t \mid X_t\}}\{y_t \mid x_t\} = \binom{x_t}{y_t} p^{y_t}(1-p)^{x_t - y_t} \quad \text{for } y_t = 0, 1, \ldots, x_t.$$

Then the stochastic process $\{Y_t, t \geq 0\}$ is called a POCTMPM with parameters $(\boldsymbol{\theta}, p)$.

Remark 1. It is readily seen that a POCTMPM with parameter vector $(\boldsymbol{\theta}, 1)$ reduces to a CTMPM with parameter vector $\boldsymbol{\theta}$.

In order to investigate the POCTMPM, we first need to find the conditional PMF of the random variable $\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$, where the random vector $\boldsymbol{Y}_n := (Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n})$ and the vector $\boldsymbol{y}_n := (y_{t_1}, y_{t_2}, \ldots, y_{t_n})$. Knowing the distribution of $\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$, almost all probabilistic properties of the POCTMPM can be extracted. Moreover, it can also be useful for statistical inference. For instance, the likelihood function $\mathcal{L}_{\boldsymbol{Y}_n}\{\boldsymbol{y}_n \mid \boldsymbol{\theta}, p\}$, which is utilized extensively in estimation theory, can be expressed in terms of the conditional PMF of $\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$ as follows:

$$\mathcal{L}_{\boldsymbol{Y}_n}\{\boldsymbol{y}_n \mid \boldsymbol{\theta}, p\} = \mathbb{P}\{\boldsymbol{Y}_n = \boldsymbol{y}_n \mid \boldsymbol{\theta}, p\} = \prod_{i=1}^{n-1} \mathbb{P}_{\{Y_{t_{i+1}} \mid \boldsymbol{Y}_i\}}\{y_{t_{i+1}} \mid \boldsymbol{y}_i; \boldsymbol{\theta}, p\} \mathbb{P}_{\{Y_{t_1}\}}\{y_{t_1} \mid \boldsymbol{\theta}, p\}.$$

3. The distribution of $\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$

This section is mainly devoted to finding the conditional PMF of $\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$. We first need to derive the conditional PMF of $\{X_{t_n} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$. For this purpose, we exploit a technique explained in Elliott *et al.* [9, Chapter 2]. This replaces the original probability distribution with an artificial probability distribution under which calculations are made. Then as stated in Theorem 1, we can convert these results to the results with respect to the original probability distribution. This is helpful if calculations with respect to the artificial probability distribution are easier than under the original one. We quote Theorem 1, namely a conditional Bayes' theorem, from Elliott *et al.* [9, Chapter 2].

Theorem 1. *Consider a stochastic process $\{Z_t, t \geq 0\}$ where the discrete random variables Z_t take values from the nonnegative integers and have the PMF $\mathbb{P}_{Z_t}$. Suppose that $\tilde{\mathbb{P}}_{Z_t}$ is another family of PMFs with the same support as $\mathbb{P}_{Z_t}$ and define a Radon–Nikodym factor $\Delta := \mathbb{P}_{Z_t}/\tilde{\mathbb{P}}_{Z_t}$. If Φ is a continuous function of the random variable Z_t, then*

$$\mathbb{E}_{\mathbb{P}_{Z_t}}\{\Phi \mid \boldsymbol{Z}_n = \boldsymbol{z}_n\} := \frac{\mathbb{E}_{\tilde{\mathbb{P}}_{Z_t}}\{\Delta\Phi \mid \boldsymbol{Z}_n = \boldsymbol{z}_n\}}{\mathbb{E}_{\tilde{\mathbb{P}}_{Z_t}}\{\Delta \mid \boldsymbol{Z}_n = \boldsymbol{z}_n\}}.$$

Here $\boldsymbol{Z}_n := (Z_{t_1}, \ldots, Z_{t_n})$, $\boldsymbol{z}_n := (z_{t_1}, \ldots, z_{t_n})$ for some fixed observation times $t_1, \ldots, t_n$ and, $\mathbb{E}_{\mathbb{P}_{Z_t}}$ and $\mathbb{E}_{\tilde{\mathbb{P}}_{Z_t}}$ represent the expected value with respect to $\mathbb{P}_{Z_t}$ and $\tilde{\mathbb{P}}_{Z_t}$, respectively.

We shall use this result to determine in Theorem 2 the conditional PMF of the values of the underlying process given the sequence of binomial observations from that process, that is $\{X_{t_n} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$. This result is of interest in its own right. The recursive equation provided for the conditional PMF of $\{X_{t_n} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$ can be useful in the study of the POCTMPM. Obviously, in the study of any stochastic process, the main information that we need is an explicit form for the corresponding PMF.

https://doi.org/10.1239/jap/1437658609 Published online by Cambridge University Press

Theorem 2. *Consider the CTMPM $\{X_t, t \geq 0\}$ with parameter vector $\boldsymbol{\theta}$ and known initial population size of x_0, and the corresponding POCTMPM $\{Y_t, t \geq 0\}$ with parameters $(\boldsymbol{\theta}, p)$. Then the conditional PMF of $\{X_{t_n} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$ for $x_{t_n} = y_{t_n}, y_{t_n} + 1, \ldots$ is given, for $n = 1, 2, \ldots$, by $\mathbb{P}_{\{X_{t_n} \mid \boldsymbol{Y}_n\}}\{x_{t_n} \mid \boldsymbol{y}_n\} = \varrho_n^{x_{t_n}} / \sum_{\ell=y_{t_n}}^{\infty} \varrho_n^{\ell}$. Here*

$$\varrho_n^{\ell} := \mathrm{e} y_{t_n}! \binom{\ell}{y_n} p^{y_{t_n}} (1-p)^{\ell - y_{t_n}} \sum_{j=y_{t_{n-1}}}^{\infty} \varrho_{n-1}^{j} \mathcal{P}_{j,\ell}(t_n - t_{n-1})$$

for $\ell = y_{t_n}, y_{t_n} + 1, \ldots$ and $n = 1, 2, \ldots$, with initial conditions $\varrho_0^{x_0} = 1$ and $\varrho_0^{\ell} = 0$ for $\ell \neq x_0$.

Proof. Motivated by Theorem 1, we define a PMF $\tilde{\mathbb{P}}$ such that the random variable Y_t is independent of the random variables Y_s for all $s \neq t$. Moreover, the random variable Y_t is independent of the random variables X_s for all s, t, that is $\tilde{\mathbb{P}}_{\{Y_t \mid X_s\}}\{y_t \mid x_s\} = \tilde{\mathbb{P}}_{Y_t}(y_t)$. We also require that the random variables X_t have the same distribution under both $\mathbb{P}$ and $\tilde{\mathbb{P}}$ for all values of t. More precisely, we require $\tilde{\mathbb{P}}_{X_t} \equiv \mathbb{P}_{X_t}$. In fact, we shall suppose that under $\tilde{\mathbb{P}}$, Y_t has a Poisson distribution with parameter one, that is $\tilde{\mathbb{P}}_{Y_t}(y_t) = \mathrm{e}^{-1}/y_t!$ for $y_t = 0, 1, \ldots$. Note that the choice of Poisson distribution with parameter one is arbitrary, but convenient, and the final result is independent of this choice. We then define the random variable

$$\mathring{\Delta}_i := \frac{\mathbb{P}_{\{Y_{t_i} \mid X_{t_i}\}}\{Y_{t_i} \mid X_{t_i}\}}{\tilde{\mathbb{P}}_{\{Y_{t_i} \mid X_{t_i}\}}\{Y_{t_i} \mid X_{t_i}\}} = \frac{\binom{X_{t_i}}{Y_{t_i}} p^{Y_{t_i}} (1-p)^{X_{t_i} - Y_{t_i}}}{\mathrm{e}^{-1}/Y_{t_i}!} = \mathrm{e} Y_{t_i}! \binom{X_{t_i}}{Y_{t_i}} p^{Y_{t_i}} (1-p)^{X_{t_i} - Y_{t_i}}$$

for $i = 1, 2, \ldots$ and $\mathring{\Delta}_0 := 1$. Write

$$\Delta_n := \prod_{i=1}^{n} \mathring{\Delta}_i. \tag{1}$$

From Theorem 1, we have

$$\mathbb{P}\{X_{t_n} = x_{t_n} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\} = \mathbb{E}_{\mathbb{P}}\{\mathbf{1}_{\{X_{t_n} = x_{t_n}\}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\} = \frac{\mathbb{E}_{\tilde{\mathbb{P}}}\{\Delta_n \mathbf{1}_{\{X_{t_n} = x_{t_n}\}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}}{\mathbb{E}_{\tilde{\mathbb{P}}}\{\Delta_n \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}},$$

where $\mathbf{1}$ is an indicator function. In the last fraction, if we let the numerator be denoted by $\varrho_n^{x_{t_n}} := \mathbb{E}_{\tilde{\mathbb{P}}}\{\Delta_n \mathbf{1}_{\{X_{t_n} = x_{t_n}\}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$, then the denominator can be expressed as $\sum_{\ell=y_{t_n}}^{\infty} \varrho_n^{\ell}$. Therefore, we just need to show that $\varrho_n^{x_{t_n}}$ satisfies the recursive equation given in Theorem 2. By definition, we have

$$\begin{aligned}
\varrho_n^{x_{t_n}} &= \mathbb{E}_{\tilde{\mathbb{P}}}\{\Delta_n \mathbf{1}_{\{X_{t_n} = x_{t_n}\}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\} \\
&= \mathbb{E}_{\tilde{\mathbb{P}}}\{\Delta_{n-1} \mathring{\Delta}_n \mathbf{1}_{\{X_{t_n} = x_{t_n}\}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\} \\
&= \mathbb{E}_{\tilde{\mathbb{P}}}\left\{\Delta_{n-1} \mathrm{e} Y_{t_n}! \binom{X_{t_n}}{Y_{t_n}} p^{Y_{t_n}} (1-p)^{X_{t_n} - Y_{t_n}} \mathbf{1}_{\{X_{t_n} = x_{t_n}\}} \;\middle|\; \boldsymbol{Y}_n = \boldsymbol{y}_n\right\}.
\end{aligned}$$

In the latter, by considering the indicator random variable $\mathbf{1}_{\{X_{t_n} = x_{t_n}\}}$ as well as the given information $\boldsymbol{Y}_n = \boldsymbol{y}_n$, we can simplify it to

$$\varrho_n^{x_{t_n}} = \mathrm{e} y_{t_n}! \binom{x_{t_n}}{y_{t_n}} p^{y_{t_n}} (1-p)^{x_{t_n} - y_{t_n}} \mathbb{E}_{\tilde{\mathbb{P}}}\{\Delta_{n-1} \mathbf{1}_{\{X_{t_n} = x_{t_n}\}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}.$$

Recall that under $\tilde{\mathbb{P}}$, each random variable Y_t is independent of all the other random variables. Therefore, the given information $Y_{t_n} = y_{t_n}$ does not assist in evaluating the last expectation. Hence, we change the given information from $\boldsymbol{Y}_n = \boldsymbol{y}_n$ to $\boldsymbol{Y}_{n-1} = \boldsymbol{y}_{n-1}$. By partitioning on the values of $X_{t_{n-1}}$, we have

$$\begin{aligned}\varrho_n^{x_{t_n}} &= \mathrm{e}y_{t_n}!\binom{x_{t_n}}{y_{t_n}}p^{y_{t_n}}(1-p)^{x_{t_n}-y_{t_n}}\\ &\quad\times\sum_{x_{t_{n-1}}=y_{t_{n-1}}}^{\infty}\mathbb{E}_{\tilde{\mathbb{P}}}\{\Delta_{n-1}\,\mathbf{1}_{\{X_{t_n}=x_{t_n}\}}\,\mathbf{1}_{\{X_{t_{n-1}}=x_{t_{n-1}}\}}\mid \boldsymbol{Y}_{n-1}=\boldsymbol{y}_{n-1}\}\\ &= \mathrm{e}y_{t_n}!\binom{x_{t_n}}{y_{t_n}}p^{y_{t_n}}(1-p)^{x_{t_n}-y_{t_n}}\\ &\quad\times\sum_{x_{t_{n-1}}=y_{t_{n-1}}}^{\infty}\mathbb{E}_{\tilde{\mathbb{P}}}\{\mathbb{E}_{\tilde{\mathbb{P}}}\{\Delta_{n-1}\,\mathbf{1}^2_{\{X_{t_{n-1}}=x_{t_{n-1}}\}}\,\mathbf{1}_{\{X_{t_n}=x_{t_n}\}}\mid \boldsymbol{Y}_{n-1}=\boldsymbol{y}_{n-1},X_{t_{n-1}}\}\}.\end{aligned}$$

Considering the definition (1) it follows that Δ_{n-1} is a function of $\boldsymbol{X}_{n-1}$ and $\boldsymbol{Y}_{n-1}$, and using the Markovian property of the stochastic process $\{X_t, t \geq 0\}$, we have

$$\begin{aligned}\varrho_n^{x_{t_n}} &= \mathrm{e}y_{t_n}!\binom{x_{t_n}}{y_{t_n}}p^{y_{t_n}}(1-p)^{x_{t_n}-y_{t_n}}\\ &\quad\times\sum_{x_{t_{n-1}}=y_{t_{n-1}}}^{\infty}\mathbb{E}_{\tilde{\mathbb{P}}}\{\mathbb{E}_{\tilde{\mathbb{P}}}\{\Delta_{n-1}\,\mathbf{1}_{\{X_{t_{n-1}}=x_{t_{n-1}}\}}\mid \boldsymbol{Y}_{n-1}=\boldsymbol{y}_{n-1},X_{t_{n-1}}\}\}\\ &\qquad\qquad\times\mathbb{E}_{\tilde{\mathbb{P}}}\{\mathbb{E}_{\tilde{\mathbb{P}}}\{\mathbf{1}_{\{X_{t_{n-1}}=x_{t_{n-1}}\}}\,\mathbf{1}_{\{X_{t_n}=x_{t_n}\}}\mid \boldsymbol{Y}_{n-1}=\boldsymbol{y}_{n-1},X_{t_{n-1}}\}\}\\ &= \mathrm{e}y_{t_n}!\binom{x_{t_n}}{y_{t_n}}p^{y_{t_n}}(1-p)^{x_{t_n}-y_{t_n}}\\ &\quad\times\sum_{x_{t_{n-1}}=y_{t_{n-1}}}^{\infty}\mathbb{E}_{\tilde{\mathbb{P}}}\{\mathbb{E}_{\tilde{\mathbb{P}}}\{\Delta_{n-1}\,\mathbf{1}_{\{X_{t_{n-1}}=x_{t_{n-1}}\}}\mid \boldsymbol{Y}_{n-1}=\boldsymbol{y}_{n-1},X_{t_{n-1}}\}\}\\ &\qquad\qquad\times\mathbb{P}\{X_{t_n}=x_{t_n}\mid X_{t_{n-1}}=x_{t_{n-1}}\}\\ &= \mathrm{e}y_{t_n}!\binom{x_{t_n}}{y_{t_n}}p^{y_{t_n}}(1-p)^{x_{t_n}-y_{t_n}}\sum_{x_{t_{n-1}}=y_{t_{n-1}}}^{\infty}\varrho_{n-1}^{x_{t_{n-1}}}\mathcal{P}_{x_{t_{n-1}},x_{t_n}}(t_n-t_{n-1}).\end{aligned}$$

As the initial population size is almost surely equal to x_0, the initial condition for the recursive equation given in Theorem 2 follows immediately.

The conditional PMF of $\{X_{t_n} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$ found in Theorem 2 can be exploited in developing further properties of the POCTMPM. The first calculation that we pursue is to find the conditional PMF of the next observation given the vector of obtained observations. We use the main result of Theorem 2 to derive an analogous recursive equation for the conditional PMF of $\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$.

Theorem 3. *Consider the CTMPM $\{X_t, t \geq 0\}$ with parameter vector $\boldsymbol{\theta}$ and known initial population size of x_0 and the corresponding POCTMPM $\{Y_t, t \geq 0\}$ with parameters $(\boldsymbol{\theta}, p)$.*

https://doi.org/10.1239/jap/1437658609 Published online by Cambridge University Press

Then the conditional PMF of $\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$ *is equal to*

$$\frac{\sum_{x_{t_{n+1}}=y_{t_{n+1}}}^{\infty} \sum_{x_{t_n}=y_{t_n}}^{\infty} \binom{x_{t_{n+1}}}{y_{t_{n+1}}} p^{y_{t_{n+1}}} (1-p)^{x_{t_{n+1}}-y_{t_{n+1}}} \mathcal{P}_{x_{t_n}, x_{t_{n+1}}}(t_{n+1}-t_n) \varrho_n^{x_{t_n}}}{\sum_{\ell=y_{t_n}}^{\infty} \varrho_n^{\ell}}$$

for $y_{t_{n+1}} = 0, 1, 2, \ldots$.

Proof. By considering the conditional independence, utilizing the law of total probability, conditioning on the random variable X_{t_n} as well as the binomial distribution of the conditional random variable $\{Y_{t_{n+1}} \mid X_{t_{n+1}}\}$, we have

$$\begin{aligned}
&\mathbb{P}_{\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n\}}\{y_{t_{n+1}} \mid \boldsymbol{y}_n\} \\
&\quad = \sum_{x_{t_{n+1}}=y_{t_{n+1}}}^{\infty} \mathbb{P}\{Y_{t_{n+1}} = y_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n, X_{t_{n+1}} = x_{t_{n+1}}\} \mathbb{P}\{X_{t_{n+1}} = x_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\} \\
&\quad = \sum_{x_{t_{n+1}}=y_{t_{n+1}}}^{\infty} \mathbb{P}\{Y_{t_{n+1}} = y_{t_{n+1}} \mid X_{t_{n+1}} = x_{t_{n+1}}\} \mathbb{P}\{X_{t_{n+1}} = x_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\} \\
&\quad = \sum_{x_{t_{n+1}}=y_{t_{n+1}}}^{\infty} \binom{x_{t_{n+1}}}{y_{t_{n+1}}} p^{y_{t_{n+1}}} (1-p)^{x_{t_{n+1}}-y_{t_{n+1}}} \mathbb{P}\{X_{t_{n+1}} = x_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}. \qquad (2)
\end{aligned}$$

The last conditional probability can be calculated in a similar way and using Theorem 2, as follows:

$$\mathbb{P}\{X_{t_{n+1}} = x_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\} = \sum_{x_{t_n}=y_{t_n}}^{\infty} \mathcal{P}_{x_{t_n}, x_{t_{n+1}}}(t_{n+1}-t_n) \left(\frac{\varrho_n^{x_{t_n}}}{\sum_{\ell=y_{t_n}}^{\infty} \varrho_n^{\ell}} \right). \qquad (3)$$

Substituting (3) into (2) the desired result is achieved.

In the next section we exploit Theorem 3 to show that the process of binomial observation of a SBP is not Markovian of any order. The simple birth process has many applications in biological modeling (see, e.g. [4], [14], [16], [18], [21], [27]).

4. The simple birth process

Let the stochastic process $\{X_t, t \geq 0\}$ be the time-homogeneous SBP, or Yule process, with the parameter λ (known as the birth rate). This means that if $X_t = x_t$ at time t, then the *transition rate* is equal to λx_t. We suppose that the initial population size x_0 is known. It is shown, see, e.g. [23], that the PMF of the random variable X_t over the values of $x_t = x_0, x_0 + 1, \ldots$ is equal to

$$\mathbb{P}_{X_t}(x_t) = \binom{x_t - 1}{x_0 - 1} {\vartheta_t}^{x_0} (1 - \vartheta_t)^{x_t - x_0},$$

where $\vartheta_t := \mathrm{e}^{-\lambda t}$. Furthermore, for $t_2 > t_1$, the PMF of the conditional random variable $\{X_{t_2} \mid X_{t_1} = x_{t_1}\}$ over the values of $x_{t_2} = x_{t_1}, x_{t_1} + 1, \ldots$ is given by

$$\mathbb{P}_{\{X_{t_2} \mid X_{t_1}\}}(x_{t_2} \mid x_{t_1}) = \binom{x_{t_2} - 1}{x_{t_1} - 1} {\upsilon_{1,2}}^{x_{t_1}} (1 - \upsilon_{1,2})^{x_{t_2} - x_{t_1}}, \qquad (4)$$

where $\upsilon_{1,2} := \mathrm{e}^{-\lambda(t_2 - t_1)}$. Note that we generally define $\upsilon_{j,k} := \vartheta_{t_k - t_j}$ for $j \leq k$.

https://doi.org/10.1239/jap/1437658609 Published online by Cambridge University Press

Let the stochastic process $\{Y_t, t \geq 0\}$ be the corresponding *partially-observable simple birth process* (POSBP) with parameters (λ, p). The partially-observable definition states that, if x_t is the population size at time t and each of these x_t individuals can be observed independently with probability p, then the random variable Y_t counts the total number of observations at time t. Bean *et al.* [3] extensively studied the POSBP. They showed that [3, Theorem 3.2 and Corollary 3.1] the PMF of the random variable Y_t is equal to

$$\mathbb{P}_{Y_t}(y_t) = ((1-p)\beta_t)^{x_0} \quad \text{for } y_t = 0, \tag{5a}$$

$$\begin{aligned}\mathbb{P}_{Y_t}(y_t) &= ((1-p)\beta_t)^{x_0}(1-\beta_t)^{y_t} \\ &\quad \times \sum_{\xi=1}^{\min\{x_0, y_t\}} \binom{y_t - 1}{\xi - 1}\binom{x_0}{\xi}\left(\frac{1-(1-p)\beta_t}{(1-p)(1-\beta_t)}\right)^{\xi} \quad \text{for } y_t = 1, 2, \dots,\end{aligned} \tag{5b}$$

where $\beta_t := \vartheta_t/(p + (1-p)\vartheta_t)$.

Corollary 1. *Consider the SBP* $\{X_t, t \geq 0\}$ *with parameter* λ *and known initial population size of* x_0 *and the corresponding POSBP* $\{Y_t, t \geq 0\}$ *with parameters* (λ, p). *Then*

$$\mathbb{P}_{\{X_{t_n} \mid \boldsymbol{Y}_n\}}(x_{t_n} \mid \boldsymbol{y}_n) = \frac{\varrho_n^{x_{t_n}}}{\sum_{\ell=\underline{x_{t_n}}}^{\infty} \varrho_n^{\ell}} \quad \textit{for } x_{t_n} = \underline{x_{t_n}}, \underline{x_{t_n}} + 1, \dots, \tag{6}$$

where $\underline{x_{t_n}} := \max\{x_0, y_{t_1}, \dots, y_{t_n}\}$ *and*

$$\varrho_n^{\ell} := \mathrm{e} y_n! \binom{\ell}{y_n} p^{y_n}(1-p)^{\ell - y_n} \sum_{j=\underline{x_{t\,n-1}}}^{\ell} \varrho_{n-1}^{j} \binom{\ell-1}{j-1} \upsilon_{n-1,n}{}^{j}(1-\upsilon_{n-1,n})^{\ell-j}$$

for $\ell = \underline{x_{t_n}}, \underline{x_{t_n}} + 1, \dots$ *and* $n = 1, 2, \dots$ *with initial conditions* $\varrho_0^{x_0} = 1$ *and* $\varrho_0^{\ell} = 0$ *for* $\ell > x_0$.

Proof. This result follows directly from Theorem 2 by replacing $\mathcal{P}_{j,\ell}(t_n - t_{n-1})$ with (4). However, as the SBP is a nondecreasing stochastic process, that is for $t_1 < t_2$, almost surely $X_{t_1} \leq X_{t_2}$, the notation $\underline{x_{t_n}}$ is as defined above.

An important question that may arise here is the dependency structure of the random variables Y_t for different values of $t \in (0, \infty)$. Reference to Theorem 4 will explain such behavior. Firstly, we recall the definition of a Markov stochastic process (see, e.g. [17]).

Definition 2. A stochastic process $\{Z_t, t \geq 0\}$ is 'Markovian' of order k (for a fixed value of $k = 1, 2, \dots$), if and only if for any real values of $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, $n = 3, 4, \dots$ and $n > k$,

$$\begin{aligned}&\mathbb{P}\{Z_{t_n} = z_{t_n} \mid Z_{t_1} = z_{t_1}, Z_{t_2} = z_{t_2}, \dots, Z_{t_{n-1}} = z_{t_{n-1}}\} \\ &\quad = \mathbb{P}\{Z_{t_n} = z_{t_n} \mid Z_{t_{n-k}} = z_{t_{n-k}}, \dots, Z_{t_{n-1}} = z_{t_{n-1}}\}.\end{aligned} \tag{7}$$

In the remainder of this section, we utilize Theorem 3 to prove that the POSBP is not Markovian. For this purpose, we first show that the POSBP with $x_0 = 1$ is not Markovian. Then we exploit this result to prove that the POSBP for any value of $x_0 \geq 1$ is not Markovian. At the outset, we calculate two probabilities in Lemmata 2 and 3 that will be used in our proof. However, in order to find the probability in Lemma 2, we first consider the following recursive equation.

Lemma 1. *A general solution of the recursive equation*

$$\psi_n = (1-p)\upsilon_{n-1,n}\psi_{n-1} + 1 - \upsilon_{n-1,n} \tag{8}$$

for $n = 2, 3, \ldots$*, where* $\upsilon_{j,k} = \vartheta_{t_k - t_j}$ *for* $j \leq k$*, and subject to the initial condition* $\psi_1 = 1 - \upsilon_{0,1}$*, is equal to*

$$\varphi_n := 1 - (1-p)^{n-1}\upsilon_{0,n} - p\sum_{i=1}^{n-1}(1-p)^{n-1-i}\upsilon_{i,n} \quad \text{for } n \geq 1. \tag{9}$$

Proof. We prove this by induction. It is readily seen that (9) for φ_1 simplifies to the initial condition ψ_1. For $n = 2$, (8) becomes

$$\varphi_2 = (1-p)\upsilon_{1,2}(1-\upsilon_{0,1}) + 1 - \upsilon_{1,2} = 1 - (1-p)\upsilon_{0,2} - p\upsilon_{1,2},$$

which agrees with (9). Assume that the expression (9) satisfies the recursive equation (8) for all positive integers $n \leq k$. By considering (8) and the induction hypothesis, we have

$$\begin{aligned}
\varphi_{k+1} &= (1-p)\upsilon_{k,k+1}\varphi_k + 1 - \upsilon_{k,k+1} \\
&= (1-p)\upsilon_{k,k+1}\bigg(1 - (1-p)^{k-1}\upsilon_{0,k} - p\sum_{i=1}^{k-1}(1-p)^{k-1-i}\upsilon_{i,k}\bigg) + 1 - \upsilon_{k,k+1} \\
&= 1 - (1-p)^k\upsilon_{0,k+1} - p\sum_{i=1}^{k}(1-p)^{k-i}\upsilon_{i,k+1},
\end{aligned}$$

and the result follows.

Lemma 2. *Consider the SBP* $\{X_t, t \geq 0\}$ *with parameter* λ*, initial population of size* $x_0 = 1$ *and its corresponding POSBP* $\{Y_t, t \geq 0\}$ *with parameters* (λ, p)*. The conditional random variable* $\{X_{t_n} \mid \boldsymbol{Y}_n = \boldsymbol{0}_n\}$ *(here* $\boldsymbol{0}_n = (0, 0, \ldots, 0)$*), follows the geometric distribution with parameter* $(1 - (1-p)\varphi_n)$*, where* φ_n *is given in Lemma 1. That is,* $\mathbb{P}_{\{X_{t_n} \mid \boldsymbol{Y}_n\}}\{x_{t_n} \mid \boldsymbol{0}_n\} = (1-(1-p)\varphi_n)((1-p)\varphi_n)^{x_{t_n}-1}$ *for* $x_{t_n} = 1, 2, \ldots$*.*

Proof. First, we show that for the observation vector $\boldsymbol{y}_n = \boldsymbol{0}_n$, the corresponding $\varrho_n^{x_n}$ is equal to

$$\varrho_n^{x_{t_n}} = \mathrm{e}^n(1-p)^{x_{t_n}+n-1}\upsilon_{0,n}\varphi_n^{x_{t_n}-1}, \tag{10}$$

where $\varphi_n := \psi_n$ as given in Lemma 1. We proceed by induction. For $n = 1$, the recursive equation given in Corollary 1 is equal to

$$\varrho_1^{x_1} = \mathrm{e}0!\binom{x_1}{0}p^0(1-p)^{x_1-0}\sum_{j=1}^{x_1}\varrho_0^j\binom{x_1-1}{j-1}\upsilon_{0,1}^j(1-\upsilon_{0,1})^{x_1-j} = \mathrm{e}(1-p)^{x_1}\upsilon_{0,1}\varphi_1^{x_1-1}.$$

Now, suppose that (10) holds for all nonnegative integers $n \leq k$. We prove that they then hold for $n = k+1$. For $n = k+1$, we have

$$\varrho_{k+1}^{x_{k+1}} = \mathrm{e}(1-p)^{x_{k+1}}\sum_{j=1}^{x_{k+1}}\varrho_k^j\binom{x_{k+1}-1}{j-1}\upsilon_{k,k+1}^j(1-\upsilon_{k,k+1})^{x_{k+1}-j}.$$

Considering the induction hypothesis, we can now replace ϱ_k^j with (10). Thus,

$$\varrho_{k+1}^{x_{k+1}} = \mathrm{e}^{k+1}(1-p)^{x_{k+1}+k}\upsilon_{0,k}\upsilon_{k,k+1}\sum_{j=1}^{x_{k+1}}\binom{x_{k+1}-1}{j-1}((1-p)\upsilon_{k,k+1}\varphi_k)^{j-1} \times (1-\upsilon_{k,k+1})^{x_{k+1}-j}.$$

From the definition of $\upsilon_{n-1,n}$, it is easy to see that $\upsilon_{0,k}\upsilon_{k,k+1} = \upsilon_{0,k+1}$. Changing the index j to $i := j-1$ and considering (8), we have

$$\begin{aligned}\varrho_{k+1}^{x_{k+1}} &= \mathrm{e}^{k+1}(1-p)^{x_{k+1}+k}\upsilon_{0,k+1}((1-p)\upsilon_{k,k+1}\varphi_k + 1 - \upsilon_{k,k+1})^{x_{k+1}-1} \\ &= \mathrm{e}^{k+1}(1-p)^{x_{k+1}+k}\upsilon_{0,k+1}\varphi_{k+1}^{x_{k+1}-1}.\end{aligned}$$

Now, having an explicit expression for $\varrho_n^{x_{t_n}}$, we can exploit Corollary 1 to calculate the desired conditional PMF for $x_{t_n} = 1, 2, \ldots$, as follows:

$$\begin{aligned}\mathbb{P}_{\{X_{t_n}\,|\,\boldsymbol{Y}_n\}}(x_{t_n} \mid \mathbf{0}_n) &= \frac{\varrho_n^{x_{t_n}}}{\sum_{\ell=\underline{x_{t_n}}}^{\infty}\varrho_n^{\ell}} \\ &= \frac{\mathrm{e}^n(1-p)^{x_{t_n}+n-1}\upsilon_{0,n}\varphi_n^{x_{t_n}-1}}{\sum_{\ell=1}^{\infty}\mathrm{e}^n(1-p)^{\ell+n-1}\upsilon_{0,n}\varphi_n^{\ell-1}} \\ &= \frac{((1-p)\varphi_n)^{x_{t_n}-1}}{\sum_{\ell=1}^{\infty}((1-p)\varphi_n)^{\ell-1}} \\ &= (1-(1-p)\varphi_n)((1-p)\varphi_n)^{x_{t_n}-1}.\end{aligned}$$

From the last line, we see that $\{X_{t_n} \mid \boldsymbol{Y}_n = \mathbf{0}_n\}$ follows the geometric distribution-success model with parameter $(1-(1-p)\varphi_n)$.

Lemma 3. *Consider the SBP $\{X_t, t \geq 0\}$ with parameter λ, initial population size $x_0 = 1$ and its corresponding POSBP $\{Y_t, t \geq 0\}$ with parameters (λ, p). Then the conditional PMF of $\mathbb{P}_{\{X_{t_n}\,|\,\boldsymbol{Y}_n\}}\{x_{t_n} \mid \boldsymbol{\mu}_n\}$ is equal to*

$$\frac{(1-p)(1-(1-p)\varphi_n)^2(\varphi_n + (1-p)^{n-1}(\upsilon_{1,n}-\upsilon_{0,n})(x_{t_n}-1))((1-p)\varphi_n)^{x_{t_n}-2}}{(1-(1-p)\varphi_n + (1-p)^n(\upsilon_{1,n}-\upsilon_{0,n}))},$$

where $\boldsymbol{\mu}_n := (1, \mathbf{0}_{n-1})$ for $x_{t_n} = 1, 2, \ldots$, $n = 2, 3, \ldots$, and φ_n is given in (9).

Proof. We first show that under the observation vector $\boldsymbol{y}_n = \boldsymbol{\mu}_n$ for values of $n = 1, 2, \ldots$,

$$\varrho_n^{x_{t_n}} = ((1-p)\mathrm{e})^n p\upsilon_{0,n}((1-p)\varphi_n)^{x_{t_n}-2}(\varphi_n + (1-p)^{n-1}(\upsilon_{1,n}-\upsilon_{0,n})(x_{t_n}-1)), \tag{11}$$

where φ_n is as defined in the recursive equation (8). Similarly to Lemma 2, we show that under the observation vector $\boldsymbol{y}_n = \boldsymbol{\mu}_n$, $\varrho_n^{x_{t_n}}$ satisfies (11) by using induction. From the recursive equation (6) of Corollary 1 and considering the assumption that $x_0 = 1$, we have

$$\varrho_1^{x_1} = \mathrm{e}p\upsilon_{0,1}x_1((1-p)(1-\upsilon_{0,1}))^{x_1-1}.$$

https://doi.org/10.1239/jap/1437658609 Published online by Cambridge University Press

Accordingly,

$$\begin{aligned}
\varrho_2^{x_2} &= \mathrm{e}0!\binom{x_2}{0}p^0(1-p)^{x_2-0}\sum_{j=1}^{x_2}\varrho_1^j\binom{x_2-1}{j-1}\upsilon_{1,2}^j(1-\upsilon_{1,2})^{x_2-j} \\
&= \mathrm{e}^2 p\upsilon_{0,1}\upsilon_{1,2}(1-p)^{x_2}\sum_{j=1}^{x_2} j\binom{x_2-1}{j-1}((1-p)(1-\upsilon_{0,1})\upsilon_{1,2})^{j-1}(1-\upsilon_{1,2})^{x_2-j} \\
&= \mathrm{e}^2 p\upsilon_{0,2}(1-p)^{x_2}(1-\upsilon_{1,2}+(1-p)(\upsilon_{1,2}-\upsilon_{0,2}))^{x_2-2} \\
&\quad\times(1-\upsilon_{1,2}+(1-p)(\upsilon_{1,2}-\upsilon_{0,2})x_2).
\end{aligned} \tag{12}$$

From Lemma 2, we have

$$\varphi_2 = 1-(1-p)\upsilon_{0,2}-p\upsilon_{1,2}. \tag{13}$$

It is easy to show that by setting $n=2$ and substituting (13) into (11), the result coincides with (12). Now, suppose that (11) holds for nonnegative integers $2\le n\le k$. Then for $n=k+1$, we have

$$\begin{aligned}
\varrho_{k+1}^{x_{k+1}} &= \mathrm{e}0!\binom{x_{k+1}}{0}p^0(1-p)^{x_{k+1}-0}\sum_{j=1}^{x_{k+1}}\varrho_k^j\binom{x_{k+1}-1}{j-1}\upsilon_{k,k+1}^j(1-\upsilon_{k,k+1})^{x_{k+1}-j} \\
&= \mathrm{e}(1-p)^{x_{k+1}}\sum_{j=1}^{x_{k+1}}((1-p)\mathrm{e})^k p\upsilon_{0,k}((1-p)\varphi_k)^{j-2} \\
&\qquad\qquad\times(\varphi_k+(1-p)^{k-1}(\upsilon_{1,k}-\upsilon_{0,k})(j-1)) \\
&\qquad\qquad\times\binom{x_{k+1}-1}{j-1}\upsilon_{k,k+1}^j(1-\upsilon_{k,k+1})^{x_{k+1}-j} \\
&= \frac{\mathrm{e}^{k+1}p(1-p)^{x_{k+1}+k-1}\upsilon_{0,k+1}}{\varphi_k} \\
&\quad\times\sum_{j=1}^{x_{k+1}}(\varphi_k+(1-p)^{k-1}(\upsilon_{1,k}-\upsilon_{0,k})(j-1)) \\
&\quad\times\binom{x_{k+1}-1}{j-1}((1-p)\upsilon_{k,k+1}\varphi_k)^{j-1}(1-\upsilon_{k,k+1})^{x_{k+1}-j} \\
&= \frac{\mathrm{e}^{k+1}p(1-p)^{x_{k+1}+k-1}\upsilon_{0,k+1}}{\varphi_k}\varphi_k((1-p)\upsilon_{k,k+1}\varphi_k+1-\upsilon_{k,k+1})^{x_{k+1}-2} \\
&\quad\times((1-p)\upsilon_{k,k+1}\varphi_k+1-\upsilon_{k,k+1}+(1-p)^k(\upsilon_{1,k}-\upsilon_{0,k})\upsilon_{k,k+1}(x_{t_{k+1}}-1)).
\end{aligned}$$

Using the recursive equation (8), we have

$$\begin{aligned}
\varrho_{k+1}^{x_{k+1}} &= ((1-p)\mathrm{e})^{k+1}p\upsilon_{0,k+1}((1-p)\varphi_{k+1})^{x_{k+1}-2} \\
&\quad\times(\varphi_{k+1}+(1-p)^k(\upsilon_{1,k+1}-\upsilon_{0,k+1})(x_{t_{k+1}}-1)),
\end{aligned}$$

which agrees with (11) for $n=k+1$. Thus, similarly to Lemma 2, we can find the conditional

PMF $\mathbb{P}_{\{X_{t_n} \mid \boldsymbol{Y}_n\}}\{x_{t_n} \mid \boldsymbol{\mu}_n\}$ by calculating

$$\begin{aligned}
&\frac{\varrho_n^{x_{t_n}}}{\sum_{\ell=\underline{x_{t_n}}}^{\infty} \varrho_n^{\ell}} \\
&\quad = \frac{((1-p)\mathrm{e})^n p\upsilon_{0,n}((1-p)\varphi_n)^{x_{t_n}-2}(\varphi_n + (1-p)^{n-1}(\upsilon_{1,n}-\upsilon_{0,n})(x_{t_n}-1))}{\sum_{\ell=1}^{\infty}((1-p)\mathrm{e})^n p\upsilon_{0,n}((1-p)\varphi_n)^{\ell-2}(\varphi_n + (1-p)^{n-1}(\upsilon_{1,n}-\upsilon_{0,n})(\ell-1))} \\
&\quad = \frac{(1-p)(1-(1-p)\varphi_n)^2(\varphi_n + (1-p)^{n-1}(\upsilon_{1,n}-\upsilon_{0,n})(x_{t_n}-1))((1-p)\varphi_n)^{x_{t_n}-2}}{(1-(1-p)\varphi_n + (1-p)^n(\upsilon_{1,n}-\upsilon_{0,n}))}
\end{aligned}$$

as required.

Proposition 1. *Consider the SBP $\{X_t, t \geq 0\}$ with parameter λ and initial population size $x_0 = 1$. Let $\{Y_t, t \geq 0\}$ be its corresponding POSBP with parameters (λ, p). The stochastic process $\{Y_t, t \geq 0\}$ is not Markovian of any order for any value of* $0 < p < 1$.

Proof. According to Definition 2, we need only show that for some particular values of the random variables, condition (7) does not hold. For this purpose, we shall show that for any value of $n = 3, 4, \ldots$, $\mathbb{P}\{Y_{t_n} = 0 \mid \boldsymbol{Y}_{n-1} = \boldsymbol{0}_{n-1}\} \neq \mathbb{P}\{Y_{t_n} = 0 \mid \boldsymbol{Y}_{n-1} = \boldsymbol{\mu}_{n-1}\}$. In order to calculate the first conditional probability, we exploit Lemma 2 along with (5a), and the geometric structure of the SBP to obtain

$$\begin{aligned}
&\mathbb{P}\{Y_{t_n} = 0 \mid \boldsymbol{Y}_{n-1} = \boldsymbol{0}_{n-1}\} \\
&\quad = \sum_{x_{t_{n-1}}=\underline{x_{t_{n-1}}}}^{\infty} \mathbb{P}\{Y_{t_n} = 0 \mid X_{t_{n-1}} = x_{t_{n-1}}, \boldsymbol{Y}_{n-1} = \boldsymbol{0}_{n-1}\} \\
&\qquad\qquad \times \mathbb{P}\{X_{t_{n-1}} = x_{t_{n-1}} \mid \boldsymbol{Y}_{n-1} = \boldsymbol{0}_{n-1}\} \\
&\quad = \sum_{x_{t_{n-1}}=1}^{\infty} \mathbb{P}\{Y_{t_n} = 0 \mid X_{t_{n-1}} = x_{t_{n-1}}\}\mathbb{P}\{X_{t_{n-1}} = x_{t_{n-1}} \mid \boldsymbol{Y}_{n-1} = \boldsymbol{0}_{n-1}\} \\
&\quad = \sum_{x_{t_{n-1}}=1}^{\infty} \mathbb{P}\{Y_{t_n - t_{n-1}} = 0 \mid X_0 = x_{t_{n-1}}\}\mathbb{P}\{X_{t_{n-1}} = x_{t_{n-1}} \mid \boldsymbol{Y}_{n-1} = \boldsymbol{0}_{n-1}\} \\
&\quad = \sum_{x_{t_{n-1}}=1}^{\infty} ((1-p)\beta_{t_n - t_{n-1}})^{x_{t_{n-1}}}(1-(1-p)\varphi_{n-1})((1-p)\varphi_{n-1})^{x_{t_{n-1}}-1} \\
&\quad = \frac{(1-p)\beta_{t_n - t_{n-1}}(1-(1-p)\varphi_{n-1})}{1-(1-p)^2\beta_{t_n - t_{n-1}}\varphi_{n-1}}. \qquad (14)
\end{aligned}$$

Analogously, by utilizing (5a) as well as Lemma 3, the second conditional probability

$$\begin{aligned}
&\mathbb{P}\{Y_{t_n} = 0 \mid \boldsymbol{Y}_{n-1} = \boldsymbol{\mu}_{n-1}\} \\
&\quad = \frac{(1-p)\beta_{t_n - t_{n-1}}(1-(1-p)\varphi_{n-1})^2(1+(1-p)^2\beta_{t_n - t_{n-1}}((1-p)^{n-2}(\upsilon_{1,n-1}-\upsilon_{0,n-1})-\varphi_{n-1}))}{(1-(1-p)^2\beta_{t_n - t_{n-1}}\varphi_{n-1})^2(1-(1-p)\varphi_{n-1}+(1-p)^{n-1}(\upsilon_{1,n-1}-\upsilon_{0,n-1}))}.
\end{aligned}$$

It is readily seen that $\mathbb{P}\{Y_{t_n} = 0 \mid \boldsymbol{Y}_{n-1} = \boldsymbol{\mu}_{n-1}\} = \gamma\mathbb{P}\{Y_{t_n} = 0 \mid \boldsymbol{Y}_{n-1} = \boldsymbol{\mu}_{n-1}\}$, where

$$\gamma := \frac{(1-(1-p)\varphi_{n-1})(1+(1-p)^2\beta_{t_n - t_{n-1}}((1-p)^{n-2}(\upsilon_{1,n-1}-\upsilon_{0,n-1})-\varphi_{n-1}))}{(1-(1-p)^2\beta_{t_n - t_{n-1}}\varphi_{n-1})(1-(1-p)\varphi_{n-1}+(1-p)^{n-1}(\upsilon_{1,n-1}-\upsilon_{0,n-1}))}.$$

https://doi.org/10.1239/jap/1437658609 Published online by Cambridge University Press

Hence, if we can show that $\gamma \neq 1$, the proof will be complete. For this purpose, we show that the denominator of γ is strictly greater than its numerator, that is

$$\begin{aligned}
&(1-(1-p)^2\beta_{t_n-t_{n-1}}\varphi_{n-1})(1-(1-p)\varphi_{n-1}+(1-p)^{n-1}(\upsilon_{1,n-1}-\upsilon_{0,n-1}))\\
&\quad -(1-(1-p)\varphi_{n-1})(1+(1-p)^2\beta_{t_n-t_{n-1}}((1-p)^{n-2}(\upsilon_{1,n-1}-\upsilon_{0,n-1})-\varphi_{n-1}))\\
&\quad =(1-p)^{n-1}(\upsilon_{1,n-1}-\upsilon_{0,n-1})(1-(1-p)\beta_{t_n-t_{n-1}})\\
&\quad >0.
\end{aligned}$$

This inequality holds because of the assumption that $0 < p < 1$,

$$\upsilon_{1,n-1}-\upsilon_{0,n-1}=\mathrm{e}^{-\lambda(t_{n-1}-t_1)}-\mathrm{e}^{-\lambda t_{n-1}}=\mathrm{e}^{-\lambda t_{n-1}}(\mathrm{e}^{\lambda t_1}-1)>0,$$

due to $t_1 > 0$, and $\beta_{t_n-t_{n-1}} \leq 1$ as it is a probability.

Having all these results at hand, we are now able to prove Theorem 4.

Theorem 4. *Consider the SBP $\{X_t, t \geq 0\}$ with parameter λ and initial population size $x_0 \geq 1$. Let $\{Y_t, t \geq 0\}$ be its corresponding POSBP with parameters (λ, p). The stochastic process $\{Y_t, t \geq 0\}$ is not Markovian of any order for any value of $0 < p < 1$.*

Proof. To prove this theorem, we use induction on x_0 to show that for any integer value $n \geq 3$,

$$\mathbb{P}\{Y_{t_n}=0 \mid \boldsymbol{Y}_{n-1}=\boldsymbol{0}_{n-1}\} > \mathbb{P}\{Y_{t_n}=0 \mid \boldsymbol{Y}_{n-1}=\boldsymbol{\mu}_{n-1}\}, \tag{15}$$

which implies that

$$\mathbb{P}\{Y_{t_n}=0 \mid \boldsymbol{Y}_{n-1}=\boldsymbol{0}_{n-1}\} \neq \mathbb{P}\{Y_{t_n}=0 \mid \boldsymbol{Y}_{n-1}=\boldsymbol{\mu}_{n-1}\}.$$

From Proposition 1, we see that the inequality (15) holds for $x_0 = 1$, and so the corresponding stochastic process $\{Y_t, t \geq 0\}$ is not Markovian. Suppose that the inequality (15) holds for all values of $x_0 \leq k$. We shall then show that it is also the case for $x_0 = k+1$. Label all $k+1$ individuals in the initial population from $i = 1, 2, \ldots, k+1$ and suppose that $\{Y_t^i, t \geq 0\}$ is the corresponding POSBP of ancestor i at time t. Obviously, all Y_t^is are independent of each other and $Y_t = \sum_{i=1}^{k+1} Y_t^i$. Hence, by considering (14) for the first probability in (15), we have

$$\begin{aligned}
\mathbb{P}\{Y_{t_n}=0 \mid \boldsymbol{Y}_{n-1}=\boldsymbol{0}_{n-1}\} &= \mathbb{P}\Bigg\{\sum_{i=1}^{k+1} Y_{t_n}^i = 0 \;\Bigg|\; \sum_{i=1}^{k+1} Y_{t_1}^i = \cdots = \sum_{i=1}^{k+1} Y_{t_{n-1}}^i = 0\Bigg\}\\
&= \mathbb{P}\{Y_{t_n}^1 = \cdots = Y_{t_n}^{k+1} = 0 \mid Y_{t_1}^1 = Y_{t_1}^2 = \cdots = Y_{t_{n-1}}^{k+1} = 0\}\\
&= \prod_{i=1}^{k+1} \mathbb{P}\{Y_{t_n}^i = 0 \mid Y_{t_1}^i = \cdots = Y_{t_{n-1}}^i = 0\}\\
&= \left(\frac{(1-p)\beta_{t_n-t_{n-1}}(1-(1-p)\varphi_{n-1})}{1-(1-p)^2\beta_{t_n-t_{n-1}}\varphi_{n-1}}\right)^{k+1}.
\end{aligned}$$

In order to find the second probability in (15), let us define the following random events for simplicity: $\mathbb{E}_k := \{Y_{t_n}^1 = \cdots = Y_{t_n}^k = 0\}$ and $F_k := \{Y_{t_2}^1 = Y_{t_2}^2 = \cdots = Y_{t_{n-1}}^k = 0\}$.

https://doi.org/10.1239/jap/1437658609 Published online by Cambridge University Press

The second probability in (15) is equal to

$$
\begin{aligned}
&\mathbb{P}\{Y_{t_n} = 0 \mid \boldsymbol{Y}_{n-1} = \boldsymbol{\mu}_{n-1}\} \\
&\quad = \mathbb{P}\biggl\{\sum_{i=1}^{k+1} Y_{t_n}^i = 0 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, \sum_{i=1}^{k+1} Y_{t_2}^i = \cdots = \sum_{i=1}^{k+1} Y_{t_{n-1}}^i = 0\biggr\} \\
&\quad = \mathbb{P}\biggl\{\mathbb{E}_{k+1} \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\} \\
&\quad = \mathbb{P}\biggl\{\mathbb{E}_{k+1} \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, Y_{t_1}^{k+1} = 0, F_{k+1}\biggr\}\mathbb{P}\biggl\{Y_{t_1}^{k+1} = 0 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\} \\
&\qquad + \mathbb{P}\biggl\{\mathbb{E}_{k+1} \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, Y_{t_1}^{k+1} = 1, F_{k+1}\biggr\}\mathbb{P}\biggl\{Y_{t_1}^{k+1} = 1 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\} \\
&\quad = \mathbb{P}\biggl\{\mathbb{E}_{k+1} \biggm| \sum_{i=1}^{k} Y_{t_1}^i = 1, Y_{t_1}^{k+1} = 0, F_{k+1}\biggr\}\mathbb{P}\biggl\{Y_{t_1}^{k+1} = 0 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\} \\
&\qquad + \mathbb{P}\{\mathbb{E}_{k+1} \mid Y_{t_1}^1 = \cdots = Y_{t_1}^k = 0, Y_{t_1}^{k+1} = 1, F_{k+1}\} \\
&\qquad\quad \times \mathbb{P}\biggl\{Y_{t_1}^{k+1} = 1 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\}.
\end{aligned}
$$

As the observations from the ancestors of different initial individuals are independent from each other, we have

$$
\begin{aligned}
&\mathbb{P}\{Y_{t_n} = 0 \mid \boldsymbol{Y}_{n-1} = \boldsymbol{\mu}_{n-1}\} \\
&\quad = \mathbb{P}\{Y_{t_n}^{k+1} = 0 \mid Y_{t_1}^{k+1} = \cdots = Y_{t_{n-1}}^{k+1} = 0\}\mathbb{P}\biggl\{\mathbb{E}_k \biggm| \sum_{i=1}^{k} Y_{t_1}^i = 1, F_k\biggr\} \\
&\qquad \times \mathbb{P}\biggl\{Y_{t_1}^{k+1} = 0 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\} \\
&\qquad\quad + \mathbb{P}\{Y_{t_n}^{k+1} = 0 \mid Y_{t_1}^{k+1} = 1, Y_{t_2}^{k+1} = \cdots = Y_{t_{n-1}}^{k+1} = 0\} \\
&\qquad\qquad \times \mathbb{P}\{\mathbb{E}_k \mid Y_{t_1}^1 = \cdots = Y_{t_1}^k = 0, F_k\} \\
&\qquad\qquad\quad \times \mathbb{P}\biggl\{Y_{t_1}^{k+1} = 1 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\}.
\end{aligned}
$$

Considering the results of Proposition 1 and the hypothesis of induction, we have

$$
\begin{aligned}
&\mathbb{P}\{Y_{t_n} = 0 \mid \boldsymbol{Y}_{n-1} = \boldsymbol{\mu}_{n-1}\} \\
&\quad < \mathbb{P}\{Y_{t_n}^{k+1} = 0 \mid Y_{t_1}^{k+1} = \cdots = Y_{t_{n-1}}^{k+1} = 0\} \\
&\qquad \times \mathbb{P}\{\mathbb{E}_k \mid Y_{t_1}^1 = \cdots = Y_{t_1}^k = 0, F_k\} \\
&\qquad\quad \times \mathbb{P}\biggl\{Y_{t_1}^{k+1} = 0 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\}
\end{aligned}
$$

https://doi.org/10.1239/jap/1437658609 Published online by Cambridge University Press

$$
\begin{aligned}
&\quad + \mathbb{P}\{Y_{t_n}^{k+1} = 0 \mid Y_{t_1}^{k+1} = 1, Y_{t_2}^{k+1} = \cdots = Y_{t_{n-1}}^{k+1} = 0\} \\
&\qquad \times \mathbb{P}\{\mathbb{E}_k \mid Y_{t_1}^1 = \cdots = Y_{t_1}^k = 0, F_k\} \\
&\qquad \times \mathbb{P}\biggl\{Y_{t_1}^{k+1} = 1 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\} \\
&< \mathbb{P}\{Y_{t_n}^{k+1} = 0 \mid Y_{t_1}^{k+1} = \cdots = Y_{t_{n-1}}^{k+1} = 0\} \\
&\quad \times (\mathbb{P}\{Y_{t_n}^1 = 0 \mid Y_{t_1}^1 = Y_{t_2}^1 = \cdots = Y_{t_{n-1}}^1 = 0\})^k \\
&\quad \times \mathbb{P}\biggl\{Y_{t_1}^{k+1} = 0 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\} \\
&\quad + \mathbb{P}\{Y_{t_n}^{k+1} = 0 \mid Y_{t_1}^{k+1} = \cdots = Y_{t_{n-1}}^{k+1} = 0\} \\
&\qquad \times (\mathbb{P}\{Y_{t_n}^1 = 0 \mid Y_{t_1}^1 = Y_{t_2}^1 = \cdots = Y_{t_{n-1}}^1 = 0\})^k \\
&\qquad \times \mathbb{P}\biggl\{Y_{t_1}^{k+1} = 1 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\} \\
&= \frac{(1-p)\beta_{t_n - t_{n-1}}(1-(1-p)\varphi_{n-1})}{1-(1-p)^2\beta_{t_n - t_{n-1}}\varphi_{n-1}} \\
&\quad \times \biggl(\frac{(1-p)\beta_{t_n - t_{n-1}}(1-(1-p)\varphi_{n-1})}{1-(1-p)^2\beta_{t_n - t_{n-1}}\varphi_{n-1}}\biggr)^k \\
&\quad \times \mathbb{P}\biggl\{Y_{t_1}^{k+1} = 0 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\} \\
&\quad + \frac{(1-p)\beta_{t_n - t_{n-1}}(1-(1-p)\varphi_{n-1})}{1-(1-p)^2\beta_{t_n - t_{n-1}}\varphi_{n-1}} \\
&\qquad \times \biggl(\frac{(1-p)\beta_{t_n - t_{n-1}}(1-(1-p)\varphi_{n-1})}{1-(1-p)^2\beta_{t_n - t_{n-1}}\varphi_{n-1}}\biggr)^k \\
&\qquad \times \mathbb{P}\biggl\{Y_{t_1}^{k+1} = 1 \biggm| \sum_{i=1}^{k+1} Y_{t_1}^i = 1, F_{k+1}\biggr\} \\
&< \biggl(\frac{(1-p)\beta_{t_n - t_{n-1}}(1-(1-p)\varphi_{n-1})}{1-(1-p)^2\beta_{t_n - t_{n-1}}\varphi_{n-1}}\biggr)^{k+1} \\
&= \mathbb{P}\{Y_{t_n} = 0 \mid \boldsymbol{Y}_{n-1} = \boldsymbol{0}_{n-1}\}.
\end{aligned}
$$

The last inequality is true since $Y_{t_1}^{k+1} \in \{0, 1, 2, \ldots\}$.

5. Conclusion

In this paper we studied a new class of stochastic processes based on binomial observations of a continuous-time Markovian population model. In order to investigate these processes, the conditional PMF of the next binomial observation, given a set of binomial observations, $\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$, is determined. For this, we first derive the conditional PMF of the true value of the stochastic process given the binomial observations, that is $\{X_{t_n} \mid \boldsymbol{Y}_n = \boldsymbol{y}_n\}$, using a filtering approach. Then the former result was derived by applying the law of total probability.

This result should help further research on the class of POCTMPMs.

Note that, although we assume the initial value of the continuous-time Markovian population model is almost surely known, this does not restrict our calculations. All results stated in Theorems 2 and 3 can be easily converted to the case with an unknown initial population size. More precisely, in these two theorems we derived the conditional PMFs of the random variables $\{X_{t_n} \mid \boldsymbol{Y}_n\}$ and $\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n\}$, when in fact they are $\{X_{t_n} \mid \boldsymbol{Y}_n, X_0\}$ and $\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n, X_0\}$. Then knowing the distribution of the random variable X_0 and exploiting the law of total probability, one can derive the conditional PMFs of $\{X_{t_n} \mid \boldsymbol{Y}_n\}$ and $\{Y_{t_{n+1}} \mid \boldsymbol{Y}_n\}$.

Next, the SBP is discussed. This has many applications in modeling biological and ecological systems. We employed the above result to show that the POSBP is not Markovian of any order. We believe that this result could be extended. One important direction for this line of research would be to show that the following conjecture is true.

Conjecture 1. *Let us consider a CTMPM* $\{X_t, t \geq 0\}$ *with vector parameter* $\boldsymbol{\theta}$ *and a related POCTMPM* $\{Y_t, t \geq 0\}$ *with vector parameter* $(\boldsymbol{\theta}, p)$*. The stochastic process* $\{Y_t, t \geq 0\}$ *is not Markovian of any order for* $0 < p < 1$.

Theorems 2 and 3 should lead to a greater understanding of the dynamics of many biological systems in ecology, genetics, and epidemiology and possible applications of the control of the spread of infectious diseases and to a reduction of the risk of extinction of endangered species.

Acknowledgements

We are grateful to the Editor and the referee for their valuable comments that helped improve and condense the previous version. This research was supported under the Australian Research Council's Discovery Projects funding scheme (Project numbers DP110101929 and DP08677639).

References

[1] Aggoun, L. and Elliott, R. J. (1998). Recursive estimation in capture recapture models. *Sultan Qaboos Univ. Oman Sci. Tech.* **3,** 67–75.

[2] Baldi, P. and Brunak, S. (2001). *Bioinformatics*, 2nd edn. MIT Press, Cambridge, MA.

[3] Bean, N. G., Eshragh, A. and Ross, J. V. (2015). Fisher information for a partially-observable simple birth process. To appear in *Commun. Statist. Theory Meth.*

[4] Becker, N. (1972). Vaccination programmes for rare infectious diseases. *Biometrika* **59,** 443–453.

[5] Black, A. J. and McKane, A. J. (2012). Stochastic formulation of ecological models and their applications. *Trends Ecology Evolution* **27,** 337–345.

[6] Clark, J. S. and Bjørnstad, O. N. (2004). Population time series: process variability, observation errors, missing values, lags, and hidden states. *Ecology* **85,** 3140–3150.

[7] Cvitaníc, J., Liptser, R. and Rozovskii, B. (2006). A filtering approach to tracking volatility from prices observed at random times. *Ann. Appl. Prob.* **16,** 1633–1652.

[8] Elliott, R. J. (1982). *Stochastic Calculus and Applications*. Springer, New York.

[9] Elliott, R. J., Aggoun, L. and Moore, J. B. (1995). *Hidden Markov Models: Estimation and Control*. Springer, New York.

[10] Fink, G. A. (2008). *Markov Models for Pattern Recognition: From Theory to Applications*. Springer, London.

[11] Fuller, W. A. (1987). *Measurement Error Models*. John Wiley, New York.

[12] Gerberry, D. J. (2009). Trade-off between BCG vaccination and the ability to detect and treat latent tuberculosis. *J. Theoret. Biol.* **261,** 548–560.

[13] Kao, R. R. (2003). The impact of local heterogeneity on alternative control strategies for foot-and-mouth disease. *Proc. R. Soc. London B* **270,** 2557–2564.

[14] Kaplan, S., Itzkovitz, S. and Shapiro, E. (2007). A universal mechanism ties genotype to phenotype in trinucleotide diseases. *PLoS Comput. Biol.* **3,** 2291–2298.

[15] Keeling, M. J. and Ross, J. V. (2008). On methods for studying stochastic disease dynamics. *J. R. Soc. Interface* **5,** 171–181.

[16] MORAN, P. A. P. (1951). Estimation methods for evolutive processes. *J. R. Statist. Soc. B* **13,** 141–146.
[17] NORRIS, J. R. (1997). *Markov Chains*. Cambridge University Press.
[18] NOVOZHILOV, A. S., KAREV, G. P. AND KOONIN, E. V. (2006). Biological applications of the theory of birth-and-death processes. *Briefings Bioinformatics* **7,** 70–85.
[19] OLSON, C. A., BEARD, K. H., KOONS, D. N. AND PITT, W. C. (2012). Detection probabilities of two introduced frogs in Hawaii: implications for assessing non-native species distributions. *Biol. Invasions* **14,** 889–900.
[20] PONCIANO, J. M., DE GELDER, L., TOP, E. M. AND JOYCE, P. (2007). The population biology of bacterial plasmids: a hidden Markov model approach. *Genetics* **176,** 957–968.
[21] RENSHAW, E. (1991). *Modelling Biological Populations in Space and Time*. Cambridge University Press.
[22] ROSS, J. V. (2012). On parameter estimation in population models III: time-inhomogeneous processes and observation error. *Theoret. Pop. Biol.* **82,** 1–17.
[23] ROSS, S. M. (1996). *Stochastic Processes*, 2nd edn. John Wiley, New York.
[24] SEBER, G. A. F. (1982). *The Estimation of Animal Abundance and Related Parameters*, 2nd edn. Macmillan, New York.
[25] SOLOW, A. R. (1998). On fitting a population model in the presence of observation error. *Ecology* **79,** 1463–1466.
[26] VASEGHI, S. V. (2009). *Advanced Digital Signal Processing and Noise Reduction*, 4th edn. John Wiley, Chichester.
[27] YULE, G. U. (1925). A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Philos. Trans. R. Soc. London B* **213,** 21–87.
[28] ZENG, Y. (2003). A partially observed model for micromovement of asset prices with Bayes estimation via filtering. *Math. Finance* **13,** 411–444.