Statistics and Probability Letters 109 (2016) 30–38

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Coupling bounds for approximating birth–death processes by truncation

Forrest W. Crawford [a,b,*], Timothy C. Stutz [c], Kenneth Lange [c,d,e]

[a] Department of Biostatistics, Yale School of Public Health, PO Box 208034 New Haven, CT 06510, USA
[b] Department of Ecology and Evolutionary Biology, Yale University, PO Box 208106 New Haven, CT 06520, USA
[c] Department of Biomathematics, University of California, Los Angeles, USA
[d] Department of Human Genetics, University of California, Los Angeles, USA
[e] Department of Statistics, University of California, Los Angeles, USA

ARTICLE INFO

Article history:
Received 5 July 2015
Received in revised form 16 October 2015
Accepted 19 October 2015
Available online 4 November 2015

Keywords:
Chernoff bound
Continuous-time Markov chain
Counting process

ABSTRACT

Birth–death processes are continuous-time Markov counting processes. Approximate moments can be computed by truncating the transition rate matrix. Using a coupling argument, we derive bounds for the total variation distance between the process and its finite approximation.

1. Introduction

A general birth–death process (BDP) is a continuous-time Markov chain $X(t)$, $t > 0$ taking values on the non-negative integers $\mathbb{N}$ (Karlin and McGregor, 1957). A "birth" from state $k \in \mathbb{N}$ to $k+1$ occurs with instantaneous non-negative rate λ_k, and a "death" to $k-1$ occurs with rate μ_k. Typically λ_k and μ_k are functions of k and are time-homogeneous. To conserve probability on $\mathbb{N}$, it is customary to set $\mu_0 = 0$. BDPs are widely used as phenomenological and mechanistic models in ecology, evolution, genetics, and other fields (Novozhilov et al., 2006). For example, the Poisson process results from taking $\lambda_k = \lambda$ and $\mu_k = 0$. Kendall's simple linear birth–death process has $\lambda_k = k\lambda$ and $\mu_k = k\mu$ (Kendall, 1948, 1949). The Yule (pure-birth) process has $\lambda_k = k\lambda$ and $\mu_k = 0$. The Moran process in population genetics (Moran, 1958), logistic population growth (Tan and Piantadosi, 1991), and the susceptible–infective–susceptible infectious disease model (Andersson and Britton, 2000) have more complicated rates. BDPs are also useful for specification of arbitrary probability distributions on $\mathbb{N}$ with features such as power-law tail behavior (Klar et al., 2010), or over- and under-dispersion relative to the Poisson distribution (Faddy, 1997).

Researchers have explored the theoretical properties of BDPs (Karlin and McGregor, 1957; Guillemin and Pinchon, 1998; Flajolet and Guillemin, 2000). Although many properties of BDPs have been characterized analytically, moments still can be difficult or impossible to compute in practical settings (Novozhilov et al., 2006; Renshaw, 2011). Moments are important for both prospective modeling to determine the dynamics of a specified process and statistical inference to estimate parameters using realizations from the process. Applied researchers must either use simpler BDPs that have analytic solutions or resort

* Corresponding author at: Department of Biostatistics, Yale School of Public Health, PO Box 208034 New Haven, CT 06510, USA. Tel.: +1 203 785 6125.
E-mail address: forrest.crawford@yale.edu (F.W. Crawford).

http://dx.doi.org/10.1016/j.spl.2015.10.013
0167-7152/

to numerical simulations to determine the mean dynamics of the process. Methods for calculating moments of general BDPs are available when the birth and death rates λ_j and μ_j are linear functions of j, but for more sophisticated models in which analytic information is unavailable, finite-time solutions remain elusive (Novozhilov et al., 2006), except in some special cases (e.g. Nåsell, 2003). There also exist approximations to the moments using a polynomial representation (e.g. Engblom and Pender, 2014). Recently some progress has been made on computational routines for evaluating probabilities for arbitrary BDPs, when no analytic information is known *a priori* about the BDP beyond its birth and death rates. Following Murphy and O'Donohoe (1975), Crawford and Suchard (2012) and Crawford et al. (2014) provide computational methods to evaluate transition probabilities and perform statistical estimation for BPDs. But there is currently no general method for computing the finite-time conditional moments $m_i^k(t) = \mathrm{E}[X^k(t)|X(0) = i]$ of a general BDP with arbitrarily specified birth and death rates.

Although a BDP is fully characterized by its stochastic rate matrix (infinitesimal generator), this matrix is infinite-dimensional when the state space is $\mathbb{N}$. One straightforward approach to computing moments is to truncate the rate matrix at a sufficiently large index. After truncation, the problem of computing moments becomes tractable using established tools from linear algebra. Choice of the truncation index n determines the accuracy of the approximation. But without knowing its behavior in advance, it is unclear how to choose n for a given BDP: if n is chosen too small, then $X(t) > n$ might occur with high probability, resulting in error; on the other hand, if n is too large, matrix computations should be sufficiently precise, but may be unnecessarily slow. Finding the right truncation index without explicitly computing transition probabilities for the process in advance is a major challenge for modelers of BDPs.

In this paper, we show how to compute the moments of a general BDP by dynamic truncation of the state space. We illustrate upper and lower bounding processes for the conditioned process $X(t)|X(0) = i$ that provide insight into its dynamics. To choose the truncation index n we provide coupling bounds for the total variation distance between lower and upper bounding processes and the true process. We discuss methods for numerical computation of the relevant matrices by singular value decomposition and provide numerical illustrations.

2. Background

Let $X(t)$ be the state of a BDP at time $t > 0$, and suppose $X(0) = i \in \mathbb{N}$. Define the transition probability

$$P_{ij}(t) = \Pr(X(t) = j|X(0) = i),$$

where i and j are non-negative integers. These probabilities obey the backward equations

$$\frac{d}{dt}P_{ij}(t) = \lambda_i P_{i+1,j}(t) + \mu_i P_{i-1,j}(t) - (\lambda_i + \mu_i)P_{ij}(t), \tag{1}$$

with initial conditions $P_{ii}(0) = 1$ and $P_{ij}(0) = 0$ for $j \neq i$. This system can be represented as $\frac{d}{dt}\mathbf{P}(t) = \mathbf{AP}(t)$, where

$$\mathbf{A} = \begin{pmatrix} -\lambda_0 & \lambda_0 & & & & & 0 \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & & & & \\ & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & & & \\ & & \mu_3 & -(\lambda_3+\mu_3) & \lambda_3 & & \\ 0 & & & & \ddots & \ddots & \ddots \end{pmatrix}$$

is the infinite-dimensional stochastic rate matrix, subject to the initial condition $\mathbf{P}(0) = \mathbf{I}$. The solution may be written informally as $\mathbf{P}(t) = \exp[\mathbf{A}t]$. Karlin and McGregor (1957) show that the transition probability for any general BDP can be expressed as

$$P_{ij}(t) = \pi_j \int_0^\infty e^{-xt} Q_i(x) Q_j(x) \psi(dx), \tag{2}$$

where $\pi_j = (\lambda_0 \cdots \lambda_{j-1})/(\mu_1 \cdots \mu_j)$ for $n \geq 1$, $\pi_0 = 1$, the functions $Q_j(x)$ are a family of orthogonal polynomials, and ψ is an orthogonalizing spectral measure. Unfortunately, expression (2) can be difficult or impossible to evaluate. Indeed, the orthogonal polynomials $Q_j(x)$ and corresponding measure ψ are only known for a few special BDPs (Novozhilov et al., 2006; Renshaw, 2011). In what follows, we assume that the process is well behaved and the transition probabilities and moments exist.

2.1. Moments

The ability to rapidly prototype a new BDP by studying its moments would be an important benefit to applied researchers who wish to determine whether a given set of rates $\{\lambda_k\}$, $\{\mu_k\}$ provides the desired dynamics on average. Let

$$m_i^k(t) = \mathrm{E}[X^k(t)|X(0) = i] = \sum_{j=0}^{\infty} j^k P_{ij}(t) \tag{3}$$