Statistics and Probability Letters 88 (2014) 91–98

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Trajectory composition of Poisson time changes and Markov counting systems

Carles Bretó *

Departamento de Estadística and Instituto Flores de Lemus, Universidad Carlos III de Madrid, C/Madrid 126, Getafe, 28903, Madrid, Spain

ARTICLE INFO

Article history:
Received 19 September 2013
Received in revised form 24 January 2014
Accepted 25 January 2014
Available online 7 February 2014

Keywords:
Compartmental models
Subordination
Random time change
Compound process
Infinitesimal dispersion

ABSTRACT

Changing time of simple continuous-time Markov counting processes by independent unit-rate Poisson processes results in Markov counting processes for which we provide closed-form transition rates via composition of trajectories and with which we construct novel, simpler infinitesimally over-dispersed processes.

1. Introduction

The statistical analysis of dynamical systems plays an important role in scientific research. When these systems involve counts, such analysis may be carried out using continuous-time Markov processes, which are often approximations to real systems and hence fail to capture some features of real data. Luckily, some of these features may be better captured after replacing time in those processes by a Markov random time. Such a time randomization approach to improving statistical modeling was recently proposed and studied in detail in Bretó and Ionides (2011), where the resulting Markov time-changed processes are defined via transition rates. Unfortunately, transition rates of such time-changed processes are in general unavailable and can be difficult to obtain in closed form. This lack of closed-form transition rates limits the appeal of this time randomization approach and may even discourage applied researchers from using it at all. To help make this approach more appealing, this paper considers changing by a Poisson process the time of a large family that includes many continuous-time Markov counting processes used in applications (for example from epidemiology, biochemistry or sociology). For the resulting time-changed process, transition rates are provided in the required closed form. These closed-form transition rates constitute the main result of this paper, which we obtain by composing trajectories of the counting process with those of the random time (instead of by integrating out the random time). Our choice of a Poisson time change seems to be unusual in the applied literature and produces time-changed models simpler than those previously considered, as we illustrate by constructing several novel over-dispersed counting processes, which can be used as building blocks to construct multivariate over-dispersed Markov counting systems.

Dynamical systems that involve counts have been studied in many disciplines by considering Markov counting systems without simultaneous events, although compound systems (which allow simultaneity) have also received some attention.

* Tel.: +34 916245855; fax: +34 916249848.
E-mail addresses: carles.breto@uc3m.es, cbreto@umich.edu, cbreto@est-econ.uc3m.es.

http://dx.doi.org/10.1016/j.spl.2014.01.032

Fields where counting systems have been modeled as continuous-time Markov chains include epidemiology and ecology (Kermack and McKendrick, 1927; Shrestha et al., 2011), pharmacokinetics (Matis and Wehrly, 1979; Haseltine and Rawlings, 2002; Srivastava, 2002) and engineering and operations research (Doig, 1957; Jackson, 2002). In these fields, most such processes are in fact Markov counting systems (Bretó and Ionides, 2011), mainly networks of queues (Brémaud, 1999) that are often referred to as compartmental models in the biological sciences (Jacquez, 1996; Matis and Kiffe, 2000) and that rule out the possibility of simultaneous transitions or events. When simultaneous events are possible, these counting systems are called compound (Bretó and Ionides, 2011).

Compound counting systems can capture better the variability in real data thanks to being infinitesimally over-dispersed and have been constructed relying on random time changes and defined via closed-form transition rates, which is what this paper is concerned with. Compound Markov counting systems have been considered as a means to increase compatibility of theoretical models with real data, for example in the context of DNA sequence alignment and genomic data (Thorne et al., 1992) and of environmental stochasticity and epidemiological data (Bretó et al., 2009). They are also infinitesimally over-dispersed (Bretó, 2012a), which is a feature of these systems that is favored by infectious-disease data (Bretó et al., 2009; He et al., 2009; Ionides et al., 2006; Shrestha et al., 2011). In this context, over-dispersion requires that the variance-to-mean ratios of random variables counting events over some time interval be larger than the ratios implied by some reference (e.g., Poisson) counting random variable (McCullagh and Nelder, 1989). However, infinitesimal over-dispersion requires in addition that over-dispersion does not vanish as the length of this time interval tends to zero. Such infinitesimal over-dispersion can be modeled with compound processes, which can be constructed under mild conditions (Bretó, 2012b) via the well-known operation of random change of time. Such time randomization approach was considered in detail in Bretó and Ionides (2011), after being first illustrated in Bretó et al. (2009) where a compound compartmental model was constructed and defined by transition rates expressed in closed form. Investigating such closed-form rates for other models in general is our main concern.

The problem that this paper addresses is the difficulty deriving closed-form transition rates of time-changed processes, which lies in the non-linearity of expected values of transition probabilities and which limits the appeal of time randomization in applications. Transition rates of Markov counting processes can be understood as appropriate limits of transition probabilities (Brémaud, 1999). These probabilities are most likely non-linear in time. After time is randomized, transition rates are instead determined by the expected transition probabilities (with respect to the randomized time), but such expected values need not be readily available in closed form due to the non-linearity. Consider a unit-rate Poisson process whose time index t is changed by random time $R(t)$. Its expected probability of k transitions over time interval $[0, l]$ is the left hand side of (1) below, which is an analytic expression. A corresponding closed-form expression can be obtained, for example, assuming that $\{R(t)\}$ is a gamma process with $E[R(t)] = t$ and $V[R(t)] = t/\tau$. Such expression is (if Γ is the gamma function)

$$E_R\left[R(l)^k e^{-R(l)}\right]/k! = \Gamma\left(l/\tau + k\right) / \left(k!\, \Gamma\left(l/\tau\right) \left(1+\tau\right)^{l/\tau} \left(1+\tau^{-1}\right)^k\right), \tag{1}$$

as derived in Bretó and Ionides (2011) and in Kozubowski and Podgórski (2009). However, this closed-form expression is based on derivations specific to the Poisson gamma process of this example and it need not extend straightforwardly to other random times or counting processes (like the non-linear death processes considered in Section 3). Seemingly technical difficulties like this one can prevent applied researchers from randomizing time.

The discouraging limitations imposed on time randomization by unavailable closed-form expressions stem from the necessity to define time-changed models as implicit hierarchies and the resulting complications on model interpretation, which we seek to alleviate in this paper. Implicit definitions of a model are those given in terms of numerical procedures to generate realizations or sample paths (Bretó et al., 2009), e.g., for our Poisson gamma example above, a realization at time t would come from the following hierarchy of random draws: use a value drawn from a gamma random variable with mean t (and variance t/τ) as the mean of a Poisson random variable from which to draw the desired process realization. Implicit definitions like this one are all that is needed to do inference using "plug-and-play" methods (Bretó et al., 2009), without the need to work out any closed-form expressions like (1). However, implicit models can be harder to interpret. Consider what an applied researcher might ask when deciding what to make of and how to interpret results obtained from an implicit model: Is the time-changed model well-behaved? What aspects of the original model vary after changing time? Should interpretation of the original parameters change? How does the choice of time change affect the answers to these questions? Interpretation issues like these might be tackled considering implicit definitions only but answers may reflect numerical artifacts and may not be as apparent as with closed-form expressions, as illustrated in next section. Helping mitigate such interpretation issues to make time randomization more attractive is the ultimate goal of this paper.

The main contribution of this paper is to provide closed-form transition rates for a large family of Markov counting processes time-changed by Poisson times via composition of trajectories and to illustrate how these closed-form rates facilitate the use of time randomization to improve Markov counting systems used in applications. The sought closed-form expressions are provided in Section 2 under mild requirements satisfied by many well-behaved processes considered in the applied literature. These expressions provide the desired details about the time-changed process to help address interpretation issues, promoting the use of time randomization. They are obtained by focusing on process trajectories to get around non-linear expectations like (1), for which the unusual Poisson time change turns out to be convenient. Not only does our Poisson time choice facilitate the derivation of the expressions, it also avoids increasing the number of parameters,