



Universidad
Carlos III de Madrid



This is a postprint version of the following published document:

Bretó, C. (2014). Trajectory composition of Poisson time changes and Markov counting systems, *Statistics & Probability Letters*, v. 88, pp. 91–98.
Available in: <http://dx.doi.org/10.1016/j.spl.2014.01.032>

© Elsevier



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

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

A B S T R A C T

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.

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)

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

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,

which results in a simpler interpretation of the parameters of the time-changed models. This is illustrated in Section 3, where time randomization is considered for several processes of interest in the biological and social sciences and applied in detail to the widespread Poisson, linear birth and linear death processes to construct novel, infinitesimally over-dispersed processes and multivariate systems without the interpretation issues of time-changed models defined only implicitly.

2. Closed-form transition rates of simple Markov counting processes time-changed by Poisson processes

Before giving the closed-form time-changed rates in [Theorem 1](#) below, we introduce our notation as we progress through some fundamental aspects of counting processes, random change of time and Poisson processes, which provides a context for the theorem.

Continuous-time Markov counting processes are fully characterized by transition rates, which are intimately related to the process compoundness and intensity, and which will also be used to define our time-changed processes. We denote a time-homogeneous continuous-time Markov counting process by $\{X(t)\}$ and define it via its transition rates or transition semigroup local characteristics ([Brémaud, 1999](#)), which we write as

$$q_X(x, k) \equiv \lim_{h \downarrow 0} \frac{P(X(t+h) = x+k | X(t) = x)}{h} \quad (2)$$

where $h, t \in \mathbb{R}_{\geq 0}$, $x \in \mathbb{N}_{\geq 0}$ and $k \in \mathbb{N}_{\geq 1}$. Transition rates determine whether $\{X(t)\}$ is infinitesimally over-dispersed, which occurs if and only if simultaneous events are possible ([Bretó and Ionides, 2011](#))—i.e., if and only if $\{X(t)\}$ is compound so that there exists at least one x and $k^* > 1$ with $q_X(x, k^*) > 0$. If $\{X(t)\}$ is not compound, then it is *simple* ([Daley and Vere-Jones, 2003](#)) and $q_X(x, 1)$ is the only non-zero transition rate. Transition rates are also related to the process rate function or intensity ([Daley and Vere-Jones, 2003](#)), which we write as

$$\lambda_X(x) \equiv \lim_{h \downarrow 0} \frac{1 - P(X(t+h) = x | X(t) = x)}{h}.$$

If the rate function preserves changes in the transition rates and is finite—i.e., if it satisfies $\lambda_X(x) = \sum_{k \geq 1} q_X(x, k) < \infty$ for all x , we say that $\{X(t)\}$ is conservative and stable ([Brémaud, 1999](#)). Conservation, stability and simpleness are satisfied by many well-behaved processes used in applications (including those in Section 3 below) and are the only conditions that [Theorem 1](#) imposes on the counting process to be time-changed. (Actually, time homogeneity is also imposed but it can be relaxed at the cost of complicating notation and derivations.)

Random change of time only needs to be sketched at this point and details are postponed to [Appendix](#), but we do elaborate here on our reasons for choosing a Poisson time change over other time changes. In a nutshell, random time change involves a base process, say $\{X(t)\}$, whose time index t is replaced by a random time process, say $\{N(t)\}$, giving the time-changed process, say $\{S(t)\} \equiv \{X(N(t))\}$. While the base process considered in this paper is rather general, i.e., simple conservative stable processes, our time change is very specific: a unit-rate Poisson process.

The choice of Poisson time change is unusual in the literature but key if we wish to avoid directly integrating out the random time from the hierarchy $X(N(t))$ by following our alternative approach based on trajectory composition to prove [Theorem 1](#) below. In the time randomization literature, Lévy processes ([Sato, 1999](#)) are routinely considered as time changes. One such process that has recurrently been chosen to obtain closed-form expressions is the gamma process. Examples include time-changed diffusions ([Madan et al., 1998](#)) or, in the context of counting processes, time-changed Poisson processes ([Hougaard et al., 1997](#); [Kozubowski and Podgórski, 2009](#); [Lee and Whitmore, 1993](#)) and time-changed pure-birth and pure-death processes ([Bretó and Ionides, 2011](#)). In spite of processes other than gamma having been considered ([Marion and Renshaw, 2000](#); [Varughese and Fatti, 2008](#); [Kumar et al., 2011](#)), Poisson processes seem to have been dispensed with. This literature emphasizes the random-variable hierarchy perspective, which drags us to the problematic non-linear expected transition probabilities analogous to (1). This hierarchical perspective gives the following analytical expression for the time-changed transition rates using (2) directly

$$\lim_{h \downarrow 0} \frac{E_{N(t+h)} \left[P(B(N(t+h)) = s+k | B(N(t)) = s) \right]}{h}. \quad (3)$$

This expression is not in closed form. To obtain one, the limit could be handled using Taylor expansions but the expectation is likely to be non-linear (as argued in the introduction). Such non-linear expectations can be circumvented by composing the trajectories of the processes (more on this in the [Appendix](#)). Trajectories are easier to compose, as we argue in the [Appendix](#), if the time change is a Poisson process like in [Theorem 1](#).

Theorem 1 (*Transition Rates of Simple Markov Counting Processes Time-changed by Poisson Processes*). *Let $\{X(t)\}$ be a continuous-time Markov counting process that is time-homogeneous, simple, conservative, stable, and defined by rate function $\lambda_X(x) = q_X(x, 1) < \infty$. Let $\{N(t)\}$ be a Poisson process defined by $\lambda_N(n) = 1$. Let also both processes be independent. Then, the transition rates and rate function of the time-changed process $\{S(t)\} \equiv \{X(N(t))\}$ are*

$$q_S(s, k) = P(X(1) = s+k | X(0) = s) \quad (4)$$

$$\lambda_S(s) = 1 - e^{-\lambda_X(s)}. \quad (5)$$

Proof. [Theorem 1](#) is proved in the [Appendix](#) to preserve the flow of the paper. \square

3. Applications: making time change more appealing to model over-dispersed systems

The closed-form expressions in [Theorem 1](#) make time randomization more attractive to applied researchers by providing them with details about their univariate time-changed process $\{S(t)\}$ that refer to the overall process behavior and to the role of specific parameters, as we illustrate below with different time changes and base processes (even when the distribution of $X(1)$ in (4) is difficult to obtain). This distribution only needs to be considered for univariate processes, which can then be used as building blocks to construct infinitesimally over-dispersed multivariate systems of counts.

The overall behavior of a process is a subjective concept, which could include features like whether or not it is well-behaved, compound or whether its transition rates and rate function have any distinctive characteristics, all of which can be complemented by additional information on the role of specific parameters. The Poisson time-changed process $\{S(t)\}$ of [Theorem 1](#) is well-behaved in the sense that its transition rates $q_S(s, k) \in [0, 1] < \infty$ and that it remains conservative and stable, since $\lambda_S(s) = 1 - P(X(1) = s | X(0) = s) = \sum_{k \geq 1} q_S(s, k)$. This assures modelers that their time-changed model is not pathological and that it fits in the framework of most existing Markov process theory. $\{S(t)\}$ is also compound, which implies it is infinitesimally over-dispersed. Such over-dispersion informs applied scientists of the additional variability demanded by the data, which may also be interpreted as the base model $\{X(t)\}$ being misspecified ([Bretó et al., 2009](#); [Bretó and Ionides, 2011](#)). Compoundness also has implications for describing and summarizing $\{S(t)\}$. It suggests complementing the information given by the more conventional infinitesimal mean and variance with that given by the rate function and the mean and variance of the transition rates (or jump sizes), which have the following distinctive features in the case of $\{S(t)\}$. One on hand, moments of the transition rates in (4) are given directly by the distribution of random variable $X(1)$. On the other hand, the rate function in (5) is necessarily restricted to the interval $[0, 1]$. These two features clue in modelers that event size moments will inherit any constraint present in the distribution of $X(1)$ and that the event rate function of $\{S(t)\}$ is bounded from above. Moreover, the closed form for the rate function given in (5) permits quantitatively assessing the impact on $\lambda_S(s)$ of parameters appearing in the original $\lambda_X(x)$.

The role of specific parameters can be analyzed in more detail by using additional information on the distribution of $X(1)$ given $X(0)$ to measure the impact of a parameter on transition rates, on the rate function and on moments, which in addition facilitates connecting and comparing the time-changed process to other processes, as we illustrate in several examples below. The required distribution of $X(1)$ is a ubiquitous result for Poisson base processes $\{X(t)\}$ with rate α (independent of Poisson time change $\{N(t)\}$), for which $X(1) \sim \text{Poisson}(\alpha)$. This distribution is all that is needed to apply [Theorem 1](#) (from which we borrow the notation) to prove the following corollary.

Corollary 2 (*Poisson Poisson Process*). Consider [Theorem 1](#) and let $\lambda_X(x) = \alpha$, i.e., let $\{X(t)\}$ be a Poisson process with rate $\alpha \in \mathbb{R}_{>0}$. The transition rates of $\{S(t)\}$ are the probability mass function of a Poisson random variable parameterized by rate α , i.e., $q_S(s, k) = \alpha^k e^{-\alpha} / k!$ and its rate function is $\lambda_S(s) = 1 - e^{-\alpha}$.

Note 3 (*On Poisson Poisson Processes*). The name ‘‘Poisson Poisson process’’ has been used in the literature to refer to a different process, i.e., to refer to compound Poisson processes with Poisson event size distribution (see for example [Hanson, 2007](#)).

[Corollary 2](#) gives a rate function and transition rates for $\{S(t)\}$ parameterized solely by α . The impact that a change in α has on the transition rates and rate function of $\{S(t)\}$ and on the moments of $S(t)$ is different from the impact that the same change has on the transition rates and rate function of $\{X(t)\}$ and on the moments of $X(t)$. Before the time change, an increase in α linearly increases the rate λ_X at which events occur in $\{X(t)\}$ and events have invariably size one. After the time change, its impact on the event rate λ_S becomes non-linear and, more importantly, it changes the mean and variance of event sizes of $\{S(t)\}$. These event size moments are simply those of $X(1)$, i.e., both moments are equal to α . Although knowing these moments is informative, they differ from the infinitesimal moments of $\{S(t)\}$. These infinitesimal moments and their ratio (the dispersion index D_{dS}) can be derived by noticing that $\{S(t)\}$ is a compound Poisson process, so that

$$\mu_{dS}(s) = \lim_{h \downarrow 0} \frac{E[S(t+h)|S(t)=s]}{h} = \lim_{h \downarrow 0} \frac{\alpha(1 - e^{-\alpha})h}{h} = \alpha(1 - e^{-\alpha}) \quad (6)$$

$$\sigma_{dS}^2(s) = \lim_{h \downarrow 0} \frac{V[S(t+h)|S(t)=s]}{h} = \lim_{h \downarrow 0} \frac{\alpha(1 + \alpha)(1 - e^{-\alpha})h}{h} = \alpha(1 + \alpha)(1 - e^{-\alpha}) \quad (7)$$

$$D_{dS}(s) = \frac{\sigma_{dS}^2(s)}{\mu_{dS}(s)} = 1 + \alpha. \quad (8)$$

The infinitesimal quantities (6)–(8) allow further analysis of the difference in the role of α in $\{S(t)\}$ and in $\{X(t)\}$. In addition, they are also important for applied modelers interested in bounding simulation computational load, which is a key element when applying the plug-and-play inference methods pointed to in the introduction. Computational load can be bound by approximating $\{S(t)\}$ at certain times using deterministic ordinary differential equations (ODEs), as explored for example

in Haseltine and Rawlings (2002). For example, the unit-rate Poisson time $\{N(t)\}$ of Corollary 2 could be approximated by function $n(t) = t$, giving the approximating ODE $dn(t)/dt = 1$. Other approximating ODEs regarding Corollary 2 could be $dx(t)/dt = \alpha$ and $ds(t)/dt = \alpha(1 - e^{-\alpha}) \neq dx(n(t))/dt = \alpha$. The approximating ODE $ds(t)/dt = \alpha$ is correct nevertheless if the time change used is instead the gamma process used to derive (1) (Bretó and Ionides, 2011).

The gamma time change used to derive (1) has disadvantages when compared to our Poisson time change, regardless of the base process. A fundamental disadvantage is that our approach to finding the closed-form expressions in Theorem 1 does not work, basically because the gamma trajectories are more complex (more details in Note 6 in the Appendix). A minor disadvantage is that it introduces an additional parameter, τ . This additional parameter complicates the preceding non-trivial comparison of the role of α in $\{S(t)\}$ with its role in $\{X(t)\}$. Of course, τ makes $\{S(t)\}$ more flexible, e.g., $D_{dS} = 1 + \alpha\tau$ (Bretó and Ionides, 2011), and could always be fixed at some arbitrary value, say $\tau = 1$, which would give the same D_{dS} for both the Poisson and the gamma time change. In this sense, the Poisson choice makes $\{S(t)\}$ simpler and still infinitesimally over-dispersed. The Poisson time change could be replaced by a compound Poisson time change. This extension would allow for additional parameters that could play the role of τ , providing additional flexibility. It would also still permit proving Theorem 1 with some minor modifications (which are beyond the scope of this paper). The above analysis comparing the role of parameters before and after changing time and with different time changes is also possible for base processes other than Poisson.

Base processes for which the distribution of $X(1)$ is well-known also include linear birth and linear death processes (see for example Bharucha-Reid, 1960), for which we next provide closed-form rates and moments before considering applying Theorem 1 to processes for which this distribution is more difficult to obtain.

Corollary 4 (Negative-Binomial Poisson Process). Let $\lambda_X(x) = \beta x \mathbb{I}\{x > 0\}$, i.e., let $\{X(t)\}$ be a linear birth process with individual birth rate $\beta \in \mathbb{R}_{>0}$. The transition rates of $\{S(t)\}$ are the probability mass function of a negative binomial random variable parameterized by success probability $1 - e^{-\beta}$ and number of failures until stopping s , i.e., $q_S(s, k) = \binom{s+k-1}{k} (e^{-\beta})^s (1 - e^{-\beta})^k$ for $s \in \mathbb{N}_{>0}$ and its rate function is $\lambda_S(s) = 1 - e^{-\beta s \mathbb{I}\{s > 0\}}$.

Corollary 5 (Binomial Poisson Process). Let $\lambda_X(x) = \delta(d_0 - x) \mathbb{I}\{x < d_0\}$, i.e., let $\{X(t)\}$ be the counting process associated with a linear death process with individual death rate $\delta \in \mathbb{R}_{>0}$ and initial population size $d_0 \in \mathbb{N}_{>0}$. The transition rates of $\{S(t)\}$ are the probability mass function of a binomial random variable parameterized by success probability $1 - e^{-\delta}$ and number of trials $(d_0 - s)$, i.e., $q_S(s, k) = \binom{d_0-s}{k} (1 - e^{-\delta})^k (e^{-\delta})^{(d_0-s)-k}$ with $k = 1, \dots, d_0 - s$ for each $s = 0, \dots, d_0 - 1$ and its rate function is $\lambda_S(s) = 1 - e^{-\delta(d_0-s) \mathbb{I}\{s < d_0\}}$.

In both Corollaries 4 and 5, the event size moments are the moments of the corresponding negative binomial and binomial distributions. Infinitesimal moments are provided below for the linear death process of Corollary 5 only. (Providing them for Corollary 4 would require a justification to pass the h -limit inside the expectation, a technicality that at this point we prefer leaving for future research.) If $\{S(t)\}$ is the process from Corollary 5, then

$$\begin{aligned} \mu_{dS}(s) &= \sum_{k=1}^{d_0-s} k \lim_{h \downarrow 0} h^{-1} P(S(t+h) = s+k \mid S(t) = s) \\ &= \sum_{k=1}^{d_0-s} k P(X(1) = s+k \mid X(0) = s) = (d_0 - s) (1 - e^{-\delta}) \\ \sigma_{dS}^2(s) &= \sum_{k=1}^{d_0-s} k^2 P(X(1) = s+k \mid X(0) = s) + \lim_{h \downarrow 0} \frac{\mu_{dS}^2(s)h^2 + o(h)}{h} = \mu_{dS}(s) [1 + (d_0 - s - 1) (1 - e^{-\delta})] \\ D_{dS}(s) &= 1 + (d_0 - s - 1) (1 - e^{-\delta}). \end{aligned}$$

These infinitesimal moments are again different from those of a binomial gamma process, but in both cases $\{S(t)\}$ is infinitesimally equi-dispersed for $s = d_0 - 1$ (Bretó and Ionides, 2011). Regarding approximating ODEs corresponding to Corollary 5, they are similar to those of Corollary 2: $dx(t)/dt = \delta(d_0 - x) \mathbb{I}\{x < d_0\}$ and $ds(t)/dt = (d_0 - s) (1 - e^{-\delta}) \mathbb{I}\{s < d_0\}$. All of the preceding analysis of the role of parameters has been greatly facilitated by nice and tractable distributions of $X(1)$.

When the distribution of $X(1)$ is difficult to obtain, the closed-form results provided by Theorem 1 inherit this lack of tractability, which, in the most extreme case, can force the use of approximations in both univariate and multivariate settings. An example of a distribution that is difficult to handle is that of non-linear death processes with $\lambda_X(x) = x(d_0 - x) \mathbb{I}\{x < d_0\}$, which is of current interest in the context of bacterial disinfection (Chou et al., 2005) and for which Theorem 3 in Billard et al. (1979) gives intricate closed-form expressions for $P(X(1) = s+k \mid X(0) = s)$. In this case, results can be obtained without the need of any approximation. Approximation of the distribution of $X(1)$ is an option if this distribution is unavailable in closed form. Such approximation can be done numerically but numerical approximations can be avoided at least for death processes with general death rate, including the following non-linear death processes: logistic and exponential death processes (which are of interest to model invasion of larvae, Faddy and

Fenlon, 1999); Bass death processes (to model innovation diffusion in social groups, e.g., Karmershu, 2007); and power death process (to model chemical formation of activated carbons, Fan and Argoti, 2011). For such processes, it is possible to truncate the infinite sum for $P(X(1) = s + k | X(0) = s)$ provided in Theorem 2 in Billard et al. (1980). These numerical or truncation approximations should in any case be expected to give more reliable results than empirically attempting to approximate the limit and expectation in (3) to obtain approximate time-changed transition rates. The transition rates considered so far refer to univariate counting processes only but can also be used to construct multivariate Markov counting systems.

Over-dispersed Markov counting systems have been constructed combining the transition rates of Poisson gamma and binomial gamma processes as building blocks of a multivariate counting system in the context of epidemiological susceptible-infectious-recovered (SIR) compartmental models (Bretó and Ionides, 2011; Bretó, 2012a). These blocks based on gamma time changes can be replaced by the Poisson-based blocks presented in this paper, resulting in alternative SIR compartmental models. Such alternative SIR models can then be considered as a more parsimonious alternative to the gamma-based models to test for over-dispersion (i.e., for model misspecification) and to improve the fit of the base model to data.

Acknowledgments

This work was supported by Spanish Government Project ECO2012-32401 and Spanish Program *Juan de la Cierva* (JCI-2010-06898). I thank the editor and an anonymous referee for their suggestions, which led to improvements in the exposition of the paper.

Appendix

Before proving Theorem 1, we give a formal definition of time change. Time change can be approached rigorously by defining an underlying probability space and by paying attention to the measure theory involved (as in for example Sato, 1999; Barndorff-Nielsen and Shiryaev, 2010). Such rigor is not needed for our proof but beginning with a probability space (Ω, \mathcal{F}, P) clarifies it. On this probability space, the collection (indexed by $t \in \mathbb{R}_{\geq 0}$) of random variables $X_t(\omega) : \Omega \rightarrow \mathbb{N}_{\geq 0}$ defines the base counting process $\{X(t)\}$ of Theorem 1. If instead of indexing by t , we index by ω (which we stress by switching to lowercase), $x_\omega(t)$ is an $\mathbb{N}_{\geq 0}$ -valued deterministic function called the process trajectory or sample path. Composing trajectory $x_\omega(t)$ with that of Poisson process $\{N(t)\}$ is equivalent to defining the time-changed process $\{S(t)\}$ as the collection of random variables such that, for all t ,

$$S_t(\omega) = X_{N_t(\omega)}(\omega) \quad (\text{A.1})$$

(as in Sato, 1999) with trajectories $s_\omega(t) = x_\omega(n_\omega(t))$.

Proof of Theorem 1. The result is proved by relying on the properties of the composed trajectory to infer both the conditional distribution of inter-event times in $\{S(t)\}$ (which determines the rate function in (5)) and the conditional distribution of event sizes in $\{S(t)\}$ (which determines the transition rates in (4)).

A key property of the trajectories of the processes involved is that they are flat with step or jump discontinuities, which carry information about the random inter-event times. Let these discontinuity points of trajectories $x_\omega(t)$, $n_\omega(t)$ and $s_\omega(t)$ be denoted by $\{\delta_1^s, \delta_2^s, \dots\}$, $\{\delta_1^n, \delta_2^n, \dots\}$ and $\{\delta_1^s, \delta_2^s, \dots\}$ respectively. These three collections are related, as can be seen from trajectory $s_\omega(t)$ implied in (A.2) by definition (A.1):

$$s_\omega(t) \equiv x_\omega(n_\omega(t)) = \begin{cases} x_\omega(0) & \text{for } 0 \leq t < \delta_1^n \\ x_\omega(1) & \text{for } \delta_1^n \leq t < \delta_2^n \\ x_\omega(2) & \text{for } \delta_2^n \leq t < \delta_3^n \\ \vdots & \vdots \end{cases} \quad (\text{A.2})$$

From (A.2), $\{\delta_1^s, \delta_2^s, \dots\}$ will be the subsequence $\{\delta_{g_1}^n, \delta_{g_2}^n, \dots\}$ of $\{\delta_1^n, \delta_2^n, \dots\}$ that skips δ_i^n if and only if $x_\omega(i) = x_\omega(i-1)$, e.g., $g_1 = 1$ (or equivalently $\delta_1^s = \delta_1^n$) if and only if $x_\omega(1) \neq x_\omega(0)$, otherwise δ_1^n will be skipped and δ_1^s will be some later discontinuity point of $n_\omega(t)$. Points $\{\delta_1^s, \delta_2^s, \dots\}$ are in fact a set of realizations of the random event times in $\{S(t)\}$, providing a means to characterizing the distribution of the i th inter-event time of $\{S(t)\}$. This random inter-event time is denoted by T_i^S and its realizations by t_i^S . (In addition, we artificially set the first event time to 0, so that T_1^S coincides with the first non-zero event time.)

The conditional distribution of inter-event times T_i^S is easiest described via an example of trajectories, which later facilitates describing the conditional distribution of event sizes. Consider for example the trajectories in Fig. A.1 below: the realized first inter-event time $t_1^S = t_1^N + t_2^N + t_3^N$. In general, $T_1^S = T_1^N + \dots + T_{G_1}^N$, where G_1 follows a geometric distribution with success probability $\pi(X(0)) = 1 - e^{-\lambda X(0)}$. Since $X(0) = S(0)$ by definition of $N(0)$, conditionally on $S(0) = s$, $T_1^S \sim \text{exponential}(\pi(s))$. Returning to Fig. A.1, $t_2^S = t_4^N + t_5^N$. In general, $T_2^S = T_{G_1}^N + \dots + T_{G_1+G_2}^N$, where G_2 follows again a geometric distribution now with success probability $\pi(X(G_1))$. Since $X(G_1) = S(T_1^S)$, conditionally on

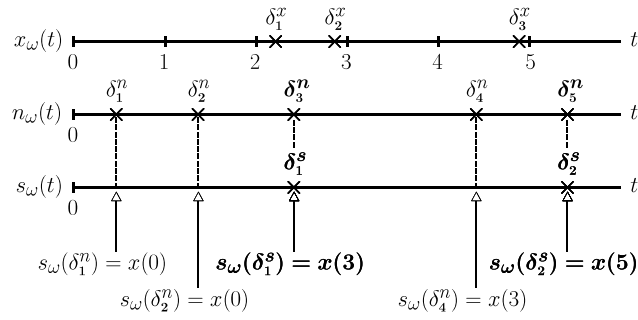


Fig. A.1. Example of composition of trajectories. The three horizontal lines represent time and show (marked with \times) realizations of several event times (i.e., trajectory discontinuities) for each process. Since the first event in $\{X(t)\}$ does not occur until time interval $(2, 3)$, event time realizations δ_1^n and δ_2^n have no effect on $\{S(t)\}$ and δ_3^n triggers the first event time realization of $\{S(t)\}$ (i.e., $\delta_1^s = \delta_3^n$), turning the two events of size one in $\{X(t)\}$ occurred at $\delta_1^x, \delta_2^x \in (2, 3]$ into a simultaneous event of size two in $\{S(t)\}$. After δ_4^n , there is $\delta_2^s = \delta_5^n$, at which the single event (again of size one) occurred at $\delta_3^x \in (4, 5]$ is shifted to δ_2^s and remains of size one.

$S(T_1^S) = s, T_2^S \sim \text{exponential}(\pi(s))$. It follows that in general, conditionally on $S\left(\sum_{i=1}^{i-1} T_i^S\right) = s, T_i^S \sim \text{exponential}(\pi(s))$, which shows (5).

Regarding jump or event sizes, define the first jump size $J_1 \equiv S(T_1^S) - S(0)$. In Fig. A.1, $J_1 = X(3) - X(0) = X(3) - X(2)$. In general, $J_1 = X(G_1) - X(G_1 - 1)$. Since $X(G_1 - 1) = X(0) = S(0)$, conditionally on $S(0) = s$, the probability of a first jump of size $k > 0$ is

$$\begin{aligned} P(J_1 = k \mid S(0) = s) &= \frac{P(X(G_1) = s + k \mid X(G_1 - 1) = s)}{\pi(s)} \\ &= \frac{P(X(1) = s + k \mid X(0) = s)}{\pi(s)} \end{aligned}$$

where the last equality follows by time homogeneity of $\{X(t)\}$. This result generalizes to the i th jump J_i by an argument analogous to the preceding one (used to generalize the result for T_1^S to T_i^S), which shows (4). \square

Note 6 (On Extending the Proof to Other Time Changes). The distributions of T_i^S and J_i^S are not as straightforward to derive using the strategy used in this proof if $\{N(t)\}$ is not a Poisson process. For example, if it is a gamma process, T_i^S is still exponential but the distribution of J_i^S depends on that of the over-shoot of the hitting time of $\{N(t)\}$ to the corresponding δ_i^x . However, this over-shoot distribution seems to have remained elusive and we are only aware of approximations to its moments (see for example Nicolai et al., 2009). If instead of a gamma process, $\{N(t)\}$ is some counting process (of finite activity) other than Poisson, the strategy followed in this proof remains effective, although the proof becomes more involved.

References

- Barndorff-Nielsen, O.E., Shiryaev, A., 2010. Change of Time and Change of Measure. World Scientific Publishing.
- Bharucha-Reid, A.T., 1960. Elements of the Theory of Markov Processes and their Applications. McGraw-Hill.
- Billard, L., Lacayo, H., Langberg, N.A., 1979. A new look at the simple epidemic process. J. Appl. Probab. 16 (1), 198–202.
- Billard, L., Lacayo, H., Langberg, N.A., 1980. Generalizations of the simple epidemic process. J. Appl. Probab. 17 (4), 1072–1078.
- Brémaud, P., 1999. Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues. Springer, New York.
- Bretó, C., 2012a. On infinitesimal dispersion of multivariate Markov counting systems. Statist. Probab. Lett. 82, 720–725.
- Bretó, C., 2012b. Time changes that result in multiple points in continuous-time Markov counting processes. Statist. Probab. Lett. 82, 2229–2234.
- Bretó, C., He, D., Ionides, E., King, A., 2009. Time series analysis via mechanistic models. Ann. Appl. Stat. 3, 319–348.
- Bretó, C., Ionides, E., 2011. Compound Markov counting processes and their applications to modeling infinitesimally over-dispersed systems. Stochastic Process. Appl. 121, 2571–2591.
- Chou, S.T., fan, L.T., Argoti, A., Vidal-Michel, R., More, A., 2005. Stochastic modeling of thermal disinfection of bacteria according to the logistic law. Amer. Inst. Chem. Eng. J. 51 (9), 2615–2618.
- Daley, D., Vere-Jones, D., 2003. An Introduction to the Theory of Point Processes. Vol. I: Elementary Theory and Methods. Springer.
- Doig, A., 1957. A bibliography on the theory of queues. Biometrika 44, 490–514.
- Faddy, M.J., Fenlon, J.S., 1999. Stochastic modelling of the invasion process of nematodes in fly larvae. J. Appl. Stat. 48, 31–37.
- Fan, L.T., Argoti, A., 2011. Stochastic modeling for the formation of activated carbons: nonlinear approach. Ind. Eng. Chem. Res. 50, 8836–8841.
- Hanson, F.B., 2007. Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis, and Computation. In: Advances in Design and Control, Society for Industrial and Applied Mathematics.
- Haseltine, E.L., Rawlings, J.B., 2002. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. J. Chem. Phys. 117, 6959–6969.
- He, D., Ionides, E., King, A., 2009. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. J. R. Soc. Interface 7, 271–283.
- Hougaard, P., Lee, M.-L.T., Whitmore, G.A., 1997. Analysis of overdispersed count data by mixtures of Poisson variables and Poisson processes. Biometrics 53 (4), 1225–1238.
- Ionides, E.L., Bretó, C., King, A.A., 2006. Inference for nonlinear dynamical systems. Proc. Natl. Acad. Sci. USA 103, 18438–18443.
- Jackson, J., 2002. How networks of queues came about. Oper. Res. 50 (1), 112–113.

- Jacquez, J.A., 1996. *Compartmental Analysis in Biology and Medicine*, third ed. BioMedware, Ann Arbor, MI.
- Karmershu, Sharma, P., 2007. Truncating the hierarchy of moment equations based on point distribution—applications to innovation diffusion. *Math. Comput. Modelling* 45, 233–240.
- Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A* 115, 700–721.
- Kozubowski, T.J., Podgórski, K., 2009. Distributional properties of the negative binomial Lévy process. *Probab. Math. Statist.* 29, 43–71.
- Kumar, A., Nane, E., Vellaisamy, P., 2011. Time-changed Poisson processes. *Statist. Probab. Lett.* 81, 1899–1910.
- Lee, M.-L.T., Whitmore, G.A., 1993. Stochastic processes directed by randomized time. *J. Appl. Probab.* 30 (2), 302–314.
- Madan, D.B., Carr, P., Chang, E.C., 1998. The variance gamma process and option pricing. *Eur. Finance Rev.* 2, 79–105.
- Marion, G., Renshaw, E., 2000. Stochastic modelling of environmental variation for biological populations. *Theor. Popul. Biol.* 57, 197–217.
- Matis, J.H., Kiffe, T.R., 2000. *Stochastic Population Models. A Compartmental Perspective*. Springer.
- Matis, J.H., Wehrly, T.E., 1979. Models of compartmental systems. *Biometrics* 35 (1), 199–220.
- McCullagh, P., Nelder, J.A., 1989. *Generalized Linear Models*, second ed. Chapman and Hall, London.
- Nicolai, R., Frenk, J., Dekker, R., 2009. Modelling and optimizing imperfect maintenance of coatings on steel structures. *Struct. Saf.* 31 (3), 234–244.
- Sato, K., 1999. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.
- Shrestha, S., King, A., Rohani, P., 2011. Statistical inference for multi-pathogen systems. *PLoS Comput. Biol.* 7 (8), e1002135.
- Srivastava, R., 2002. Stochastic vs. deterministic modeling of intracellular viral kinetics. *J. Theoret. Biol.* 218, 309–321.
- Thorne, J.L., Kishino, H., Felsenstein, J., 1992. Inching toward reality: an improved likelihood model of sequence evolution. *J. Mol. Evol.* 34, 3–16.
- Varughese, M., Fatti, L., 2008. Incorporating environmental stochasticity within a biological population model. *Theor. Popul. Biol.* 74, 115–129.