Stationary Infinitely-Divisible Markov Processes with Non-negative Integer Values

RLW* & LDB†

Draft 51: April 24, 2011

Abstract

We characterize all stationary time-reversible Markov processes whose finite-dimensional marginal distributions (of all orders) are infinitely divisible. Aside from two trivial cases (iid and constant), every such process with full support in both discrete and continuous time is a branching process with Poisson or Negative Binomial marginal distributions and a specific bivariate distribution at pairs of times.

Key words: Decomposable; Markov branching process; Negative binomial; Negative trinomial; Time reversible. AMS 2010 subject classifications: Primary 60G05; secondary 60J10.

1 Introduction

Many applications feature autocorrelated count data X_t at discrete times t. A number of authors have constructed and studied stationary stochastic processes X_t whose one-dimensional marginal distributions come from an arbitrary infinitely-divisible distribution family $\{\mu^\theta\}$, such as the Poisson $\mathsf{Po}(\theta)$ or negative binomial $\mathsf{NB}(\theta, p)$, and that are "$\mathsf{AR}(1)$-like" in the sense that their autocorrelation function is $\mathsf{Corr}[X_s, X_t] = \rho^{|s-t|}$ for some $\rho \in (0,1)$ (Lewis, 1983; Lewis, McKenzie and Hugus, 1989; McKenzie, 1988; Al-Osh and Alzaid, 1987; Joe, 1996). The most common approach is to build a time-reversible Markov process using *thinning*, in which the process at any two consecutive times may be written in the form

$$X_{t-1} = \xi_t + \eta_t \qquad X_t = \xi_t + \zeta_t$$

with ξ_t, η_t, and ζ_t all independent and from the same infinitely-divisible family (see Sec. (1.1) below for details). A second construction of a stationary time-reversible process with the same one-dimensional marginal distributions and autocorrelation function, with the feature that its finite-dimensional marginal distributions of all orders are infinitely-divisible, is to set $X_t := \mathcal{N}(G_t)$ for a *random measure* $\mathcal{N}$ on some measure space $(E, \mathcal{E}, m)$ that assigns independent infinitely-divisible random variables $\mathcal{N}(A_i) \sim \mu^{\theta_i}$ to disjoint sets $A_i \in \mathcal{E}$ of measure $\theta_i = m(A_i)$, and a family of sets $\{G_t\} \subset \mathcal{E}$ whose intersections have measure $m\big(G_s \cap G_t\big) = \theta\rho^{|s-t|}$ (see Sec. (1.2)).

*Robert L. Wolpert (`wolpert@stat.duke.edu`) is Professor in the Department of Statistical Science, Duke University, Durham NC, USA.

†Lawrence D. Brown (`lbrown@wharton.upenn.edu`) is Miers Busch Professor and Professor of Statistics in the Department of Statistics, Wharton School, University of Pennsylvania, Philadelphia PA, USA.

For the normal distribution $X_t \sim \mathsf{No}(\mu, \sigma^2)$, these two constructions both yield the usual Gaussian $\mathsf{AR}(1)$ process. The two constructions also yield identical processes for the Poisson $X_t \sim \mathsf{Po}(\theta)$ distribution, but they differ for all other nonnegative integer-valued infinitely-divisible distributions. For each nonnegative integer-valued infinitely-divisible marginal distribution except the Poisson, the process constructed by thinning does not have infinitely-divisible marginal distributions of all orders (Theorem 2, Sec. (3.5)), and the process constructed using random measures does not have the Markov property (Theorem 3, Sec. (3.5)). Thus none of these is completely satisfactory for modeling autocorrelated count data with heavier tails than the Poisson distribution.

In the present manuscript we construct and characterize every process that is Markov, infinitely-divisible, stationary, and time-reversible with non-negative integer values. The formal characterization is contained in the statement of Theorem 1 in Sec. (3.5), which follows necessary definitions and the investigation of special cases needed to establish the general result.

1.1 Thinning Process

Any univariate infinitely-divisible (ID) distribution $\mu(dx)$ on $\mathbb{R}^1$ is μ^1 the for a convolution semigroup $\{\mu^\theta : \theta \geq 0\}$ and, for $0 < \theta < \infty$ and $0 < \rho < 1$, determines uniquely a "thinning distribution" $\mu^\theta_\rho(dy \mid x)$ of Y conditional on the sum $X = Y + Z$ of independent $Y \sim \mu^{\rho\theta}$ and $Z \sim \mu^{(1-\rho)\theta}$. This thinning distribution determines a unique stationary time-reversible Markov process with one-step transition probability distribution given by the convolution

$$\mathsf{P}[X_{t+1} \in A \mid \mathcal{F}_t] = \int_{\xi+\zeta\in A} \mu^{(1-\rho)\theta}(d\zeta)\ \mu^\theta_\rho(d\xi \mid X_t)$$

for Borel sets $A \subset \mathbb{R}$, where $\mathcal{F}_t = \sigma\{X_s : s \leq t\}$ is the minimal filtration. By induction the auto-correlation is $\mathsf{Corr}(X_s, X_t) = \rho^{|s-t|}$ for square-integrable X_t. The process can be constructed beginning at any $t_0 \in \mathbb{Z}$ by setting

$$X_{t_0} \sim \mu^\theta(dx) \tag{1a}$$

$$\xi_t \sim \mu^\theta_\rho(d\xi \mid x) \text{ with } x = \begin{cases} X_{t-1} & \text{if } t > t_0 \\ X_{t+1} & \text{if } t < t_0 \end{cases} \tag{1b}$$

$$X_t := \xi_t + \zeta_t \qquad \text{for } \zeta_t \sim \mu^{\theta(1-\rho)}(d\zeta). \tag{1c}$$

Time-reversibility and hence the lack of dependence of this definition on the choice of t_0 can be verified as in the proof of Theorem 2 in Sec. (3.5) below.

1.1.1 Thinning Example 1: Poisson

For $X_t \sim \mu^\theta = \mathsf{Po}(\theta)$, for example, the Poisson distribution with mean θ, the thinning recursion step for $t > t_0$ can be written

$$\begin{aligned} X_t &= \xi_t + \zeta_t \quad \text{for independent:} \\ \xi_t &\sim \mathsf{Bi}\big(X_{t-1}, \rho\big), \qquad \zeta_t \sim \mathsf{Po}\big(\theta(1-\rho)\big) \end{aligned}$$

and hence the joint generating function at two consecutive times is

$$\phi(s, z) = \mathsf{E}\Big[s^{X_{t-1}} z^{X_t}\Big] = \exp\Big\{(s + z - 2)\theta(1-\rho) + (s\,z - 1)\theta\rho\Big\}.$$

This was called the "Poisson AR(1) Process" by McKenzie (1985) and has been studied by many other authors.

1.1.2 Thinning Example 2: Negative Binomial

In the thinning process applied to the Negative Binomial $X_t \sim \mu^\theta = \mathsf{NB}(\theta, p)$ distribution with mean $\theta(1-p)/p$, recursion for $t > t_0$ takes the form

$$\begin{aligned} X_t &= \xi_t + \zeta_t \quad \text{for independent:} \\ \xi_t &\sim \mathsf{BB}\big(X_{t-1};\ \theta\rho,\ \theta(1-\rho)\big), \qquad \zeta_t \sim \mathsf{NB}\big(\theta(1-\rho),\ p\big) \end{aligned}$$

for beta-binomial distributed $\xi_t \sim \mathsf{BB}(n; \alpha, \beta)$ (see Johnson *et al.*, 2005, §2.2) with $n = X_{t-1}$, $\alpha = \theta\rho$, and $\beta = \theta(1-\rho)$, and negative binomial $\zeta_t \sim \mathsf{NB}\big(\theta(1-\rho), p\big)$. Thus the joint generating function is

$$\begin{aligned} \phi(s, z) &= \mathsf{E}\Big[s^{X_{t-1}} z^{X_t}\Big] \\ &= p^{\theta(2-\rho)} (1 - q\,s)^{-\theta(1-\rho)}\, (1 - q\,z)^{-\theta(1-\rho)}\, (1 - q\,s\,z)^{-\theta\rho} \end{aligned} \tag{2}$$

from which one can compute the conditional generating function

$$\phi(z \mid x) = \mathsf{E}\left[z^{X_t} \mid X_{t-1} = x\right] = \left(\frac{p}{1 - qz}\right)^{\theta(1-\rho)} {}_2F_1(\theta\rho, -x; \theta; 1 - z)$$

where ${}_2F_1(a, b; c; z)$ is Gauss' hypergeometric function (Abramowitz and Stegun, 1964, §15) and, from this (for comparison below),

$$\begin{aligned} \mathsf{P}[X_{t-1} = 0, X_{t+1} = 0 \mid X_t = 2] &= [p^{\theta(1-\rho)}\ {}_2F_1(\theta\rho, -x; \theta; 1)]^2 \\ &= \Big[p^{\theta(1-\rho)}(1-\rho)\Big]^2 \left[\frac{1 + \theta(1-\rho)}{1 + \theta}\right]^2. \end{aligned} \tag{3}$$

This process, as we will see below in Theorem 2, is Markov, stationary, and time-reversible, with infinitely-divisible one-dimensional marginal distributions $X_t \sim \mathsf{NB}(\theta, p)$, but the joint marginal distributions at three or more consecutive times are not ID. It appears to have been introduced by Joe (1996, *p.* 665).

1.2 Random Measure Process

Another approach to the construction of processes with specified stationary distribution $\mu^\theta(dx)$ is to set $X_t := \mathcal{N}(G_t)$ for a *random measure* $\mathcal{N}$ and a class of sets $\{G_t\}$, as in (Wolpert and Taqqu, 2005, §3.3, 4.4). We begin with a countably additive random measure $\mathcal{N}(dx\,dy)$ that assigns independent random variables $\mathcal{N}(A_i) \sim \mu^{|A_i|}$ to disjoint Borel sets $A_i \in \mathcal{B}(\mathbb{R}^2)$ of finite area $|A_i|$ (this is possible by the Kolmogorov consistency conditions), and a collection of sets

$$G_t := \Big\{(x, y) :\ x \in \mathbb{R},\ 0 \le y < \theta\lambda\ e^{-2\lambda|t-x|}\Big\}$$

(shown in Fig. (1)) whose intersections satisfy $|G_s \cap G_t| = \theta e^{-\lambda|s-t|}$. For $t \in \mathbb{Z}$, set

$$X_t := \mathcal{N}(G_t). \tag{4}$$

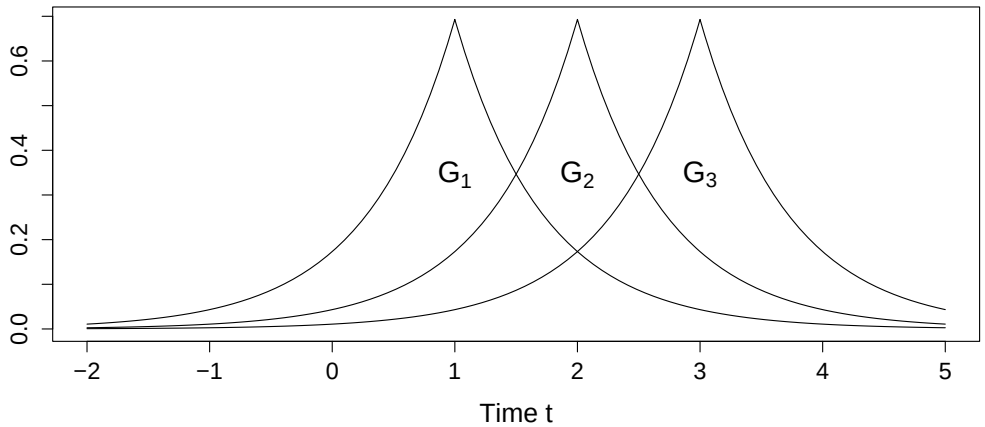

Figure 1: Random measure construction of process $X_t = \mathcal{N}(G_t)$

For any n times $t_1 < t_2 < \cdots < t_n$ the sets $\{G_{t_i}\}$ partition $\mathbb{R}^2$ into $n(n+1)/2$ sets of finite area (and one with infinite area, $(\cup G_{t_i})^c$), so each X_{t_i} can be written as the sum of some subset of $n(n+1)/2$ independent random variables. In particular, any $n = 2$ variables X_s and X_t can be written as

$$X_s = \mathcal{N}(G_s \backslash G_t) + \mathcal{N}(G_s \cap G_t), \qquad X_t = \mathcal{N}(G_t \backslash G_s) + \mathcal{N}(G_s \cap G_t)$$

just as in the thinning approach, so both 1-dimensional and 2-dimensional marginal distributions for the random measure process coincide with those for the thinning process of Sec. (1.1).

Evidently the process X_t constructed from this random measure is stationary, time-reversible and infinitely divisible in the strong sense that all finite-dimensional marginal distributions are ID. Although the 1- and 2-dimensional marginal distributions of this process coincide with those of the thinning process, the k-dimensional marginals may differ for $k \geq 3$, so this process cannot be Markov. We will see in Theorem 3 below that the only nonnegative integer-valued distribution for which it is Markov is the Poisson.

1.2.1 Random Measure Example 1: Poisson

The conditional distribution of $X_{t_n} = \mathcal{N}(G_{t_n})$ given $\{X_{t_j} : j < n\}$ can be written as the sum of n independent terms, $(n-1)$ of them with binomial distributions (all with the same probability parameter $p = \rho^{|t_n - t_{n-1}|}$, and with size parameters that sum to $X_{t_{n-1}}$) and one with a Poisson distribution (with mean $\theta(1 - \rho^{|t_n - t_{n-1}|})$. It follows by induction that the random-measure Poisson process is identical in distribution to the thinning Poisson process of Sec. (1.1.1).

1.2.2 Random Measure Example 2: Negative Binomial

The random variables X_1, X_2, X_3 for the random measure process built on the Negative Binomial distribution $X_t \sim \mathsf{NB}(\theta, p)$ with autocorrelation $\rho \in (0, 1)$ can be written as sums

$$X_1 = \zeta_1 + \zeta_{12} + \zeta_{123} \qquad X_2 = \zeta_2 + \zeta_{12} + \zeta_{23} + \zeta_{123} \qquad X_3 = \zeta_3 + \zeta_{23} + \zeta_{123}$$

of six independent negative binomial random variables $\zeta_s \sim \mathsf{NB}(\theta_s, p)$ with shape parameters

$$\theta_1 = \theta_3 = \theta(1-\rho), \qquad \theta_2 = \theta(1-\rho)^2, \qquad \theta_{12} = \theta_{23} = \theta\rho(1-\rho), \qquad \theta_{123} = \theta\rho^2$$

(each $\zeta_s = \mathcal{N}\big(\cap_{t\in s} G_t\big)$ and $\theta_s = |\cap_{t\in s} G_t|$ in Fig. (1)). It follows that the conditional probability

$$\begin{aligned}\mathsf{P}[X_1 = 0,\ X_3 = 0 \mid X_2 = 2] &= \mathsf{P}[\zeta_1 = \zeta_{12} = \zeta_{123} = \zeta_{23} = \zeta_3 = 0 \mid \zeta_2 + \zeta_{12} + \zeta_{23} + \zeta_{123} = 2] \\ &= \frac{\mathsf{P}[\zeta_2 = 2,\ \text{all other } \zeta_s = 0]}{\mathsf{P}[X_2 = 2]} \\ &= \left[p^{\theta(1-\rho)}(1-\rho)\right]^2 \frac{1+\theta(1-\rho)^2}{1+\theta} \qquad (5)\end{aligned}$$

differs from that of the thinning negative binomial process in Eqn. (3) for all $\theta > 0$ and $\rho > 0$. Thus this process is stationary, time-reversible, and has infinitely-divisible marginal distributions of all orders, but it cannot be Markov since its 2-dimensional marginal distributions coincide with those of the Markov thinning process but its 3-dimensional marginal distributions do not.

In this paper we characterize every process that is Markov, Infinitely-divisible, Stationary, and Time-reversible with non-negative Integer values ("MISTI" for short).

2 MISTI Processes

A real-valued stochastic process X_t indexed by $t \in \mathbb{Z}$ is *stationary* if each finite-dimensional marginal distribution

$$\mu_T(B) := \mathsf{P}\left[X_T \in B\right]$$

satisfies

$$\mu_T(B) = \mu_{s+T}(B) \qquad (6a)$$

for each set $T \subset \mathbb{Z}$ of cardinality $|T| < \infty$, Borel set $B \in \mathcal{B}(\mathbb{R}^{|T|})$, and $s \in \mathbb{Z}$, where as usual "$s+T$" denotes $\{(s+t) :\ t \in T\}$. A stationary process is *time-reversible* if also

$$\mu_T(B) = \mu_{-T}(B) \qquad (6b)$$

(where "$-T$" is $\{-t :\ t \in T\}$) and *Markov* if for every $t \in \mathbb{Z}$ and finite $T \subset \{s \in \mathbb{Z} :\ s \geq t\}$,

$$\mathsf{P}[X_T \in B \mid \mathcal{F}_t] = \mathsf{P}[X_T \in B \mid X_t] \qquad (6c)$$

for all $B \in \mathcal{B}(\mathbb{R}^{|T|})$, where $\mathcal{F}_t := \sigma\{X_s :\ s \leq t\}$. The process X_t is Infinitely Divisible (ID) or, more specifically, *multivariate* infinitely divisible (MVID) if each μ_T is the n-fold convolution of some other distribution $\mu_T^{(1/n)}$ for each $n \in \mathbb{N}$. This is more restrictive than requiring only that the one-dimensional marginal distributions be ID and, for integer-valued processes that satisfy

$$\mu_T(\mathbb{Z}^{|T|}) = 1, \qquad (6d)$$

it is equivalent (by the Lévy-Khinchine formula; see, for example, Rogers and Williams, 2000, *p.* 74) to the condition that each μ_T have characteristic function of the form

$$\int_{\mathbb{R}^{|T|}} e^{i\omega' x}\, \mu_T(dx) = \exp\left\{\int_{\mathbb{Z}^{|T|}} \left(e^{i\omega' u} - 1\right) \nu_T(du)\right\}, \qquad \omega \in \mathbb{R}^{|T|} \qquad (6e)$$

for some finite measure ν_T on $\mathcal{B}(\mathbb{Z}^{|T|})$. Call a process X_t or its distributions $\mu_T(du)$ *MISTI* if it is Markov, nonnegative Integer-valued, Stationary, Time-reversible, and Infinitely divisible, *i.e.*, satisfies Eqns. (6a–6e). We now turn to the problem of characterizing all MISTI distributions.

2.1 Three-dimensional Marginals

By stationarity and the Markov property all MISTI finite-dimensional distributions $\mu_T(du)$ are determined completely by the marginal distribution for X_t at two consecutive times; to exploit the MVID property we will study the three-dimensional marginal distribution for X_t at any set T of $|T| = 3$ consecutive times— say, $T = \{1, 2, 3\}$. By Eqn. (6e) we can represent $X_{\{1,2,3\}}$ in the form

$$X_1 = \sum iN_{i++} \qquad X_2 = \sum jN_{+j+} \qquad X_3 = \sum kN_{++k}$$

for independent Poisson-distributed random variables

$$N_{ijk} \overset{\text{ind}}{\sim} \mathsf{Po}(\lambda_{ijk})$$

with means $\lambda_{ijk} := \nu(\{(i, j, k)\})$; here and hereafter, a subscript "+" indicates summation over the entire range of that index— $\mathbb{N}_0 = \{0, 2 \dots\}$ for $\{N_{ijk}\}$ and $\{\lambda_{ijk}\}$, $\mathbb{N} = \{1, 2, \dots\}$ for $\{\theta_j\}$. The sums $\theta_j := \lambda_{+j+}$ for $j \geq 1$ characterize the univariate marginal distribution of each X_t— for example, through the probability generating function (pgf)

$$\varphi(z) := \mathsf{E}[z^{X_t}] = \exp\left[\sum\nolimits_{j\geq 1} \left(z^j - 1\right)\theta_j\right].$$

To avoid trivial technicalities we will assume that $0 < \mathsf{P}[X_t = 1] = \varphi'(0) = \theta_1 e^{-\theta_+}$, *i.e.*, $\theta_1 > 0$. Now set $r_i := \lambda_{i1+}/\theta_1$, and for later use define functions:

$$\psi_j(s, t) := \sum_{i,k\geq 0} s^i t^k \lambda_{ijk} \qquad p(s) := \psi_1(s, 1)/\theta_1 = \sum_{i\geq 0} s^i\, r_i \qquad P(z) := \sum_{j\geq 1} z^j\, \theta_j. \tag{7}$$

Since r_i and θ_j are nonnegative and summable (by Eqns. (6d, 6e)), $p(s)$ and $P(z)$ are analytic on the open unit ball $\mathbb{U} \subset \mathbb{C}$ and continuous on its closure. Similarly, since λ_{ijk} is summable, each $\psi_j(s, t)$ is analytic on $\mathbb{U}^2$ and continuous on its closure. Note $\psi_j(1, 1) = \theta_j$, $p(0) = r_0$ and $p(1) = 1$, while $P(0) = 0$ and $P(1) = \theta_+$; also $\varphi(z) = \exp\{P(z) - \theta_+\}$. Each $\psi_j(s, t) = \psi_j(t, s)$ is symmetric by Eqn. (6b), as are the conditional probability generating functions:

$$\varphi(s, t \mid z) := \mathsf{E}\big[s^{X_1} t^{X_3} \mid X_2 = z\big].$$

2.1.1 Conditioning on $X_2 = 0$

By the Markov property Eqn. (6c), X_1 and X_3 must be conditionally independent given X_2, so the conditional probability generating function must factor:

$$\begin{aligned}\varphi(s, t \mid 0) &:= \mathsf{E}\big[s^{X_1} t^{X_3} \mid X_2 = 0\big] = \mathsf{E}\big[s^{\sum_{i\geq 0} iN_{i0+}}\ t^{\sum_{k\geq 0} kN_{+0k}}\big] \\ &= \exp\Big\{\sum_{i,k\geq 0} \left(s^i t^k - 1\right)\lambda_{i0k}\Big\} \\ &\equiv \varphi(s, 1 \mid 0)\ \varphi(1, t \mid 0).\end{aligned} \tag{8}$$

Taking logarithms,

$$\sum \left(s^i t^k - 1\right)\lambda_{i0k} \equiv \sum \left(s^i - 1\right)\lambda_{i0k} + \sum \left(t^k - 1\right)\lambda_{i0k}$$

or, for all s and t in the unit ball in $\mathbb{C}$,

$$0 \equiv \sum (s^i - 1)(t^k - 1)\lambda_{i0k}. \tag{9}$$

Thus $\lambda_{i0k} = 0$ whenever both $i > 0$ and $k > 0$ and, by symmetry,

$$\varphi(1, z \mid 0) = \varphi(z, 1 \mid 0) = \exp\left\{\sum_{i\geq 0}(z^i - 1)\lambda_{i00}\right\}.$$

2.1.2 Conditioning on $X_2 = 1$

Similarly

$$\begin{aligned}
\varphi(s,t \mid 1) := \mathsf{E}\big[s^{X_1}t^{X_3} \mid X_2 = 1\big] &= \mathsf{E}\big[s^{\sum_{i\geq 0} i(N_{i0+}+N_{i1+})} \quad t^{\sum_{k\geq 0} k(N_{+0k}+N_{+1k})} \mid N_{+1+} = 1\big] \\
&= \varphi(s,t \mid 0)\mathsf{E}\big[s^{\sum_{i\geq 0} iN_{i1+}} \quad t^{\sum_{k\geq 0} kN_{+1k}} \mid N_{+1+} = 1\big] \\
&= \varphi(s,t \mid 0)\left\{\sum_{i,k\geq 0} s^i t^k \left[\lambda_{i1k}/\lambda_{+1+}\right]\right\}
\end{aligned}$$

since $\{N_{i1k}\}$ is conditionally multinomial given N_{+1+} and independent of $\{N_{i0k}\}$. By the Markov property this too must factor, as $\varphi(s,t \mid 1) = \varphi(s,1 \mid 1)\,\varphi(1,t \mid 1)$, so by Eqn. (8)

$$\theta_1\left\{\sum_{i,k\geq 0} s^i t^k \lambda_{i1k}\right\} = \left\{\sum_{i\geq 0} s^i \lambda_{i1+}\right\}\left\{\sum_{k\geq 0} t^k \lambda_{+1k}\right\}$$

or, since $\lambda_{i1k} = \lambda_{k1i}$ by Eqns. (6b, 7),

$$\begin{aligned}
\psi_1(s,t) &:= \sum_{i,k\geq 0} s^i t^k \lambda_{i1k} = \theta_1 p(s)\,p(t), \\
\varphi(s,t \mid 1) &= \varphi(s,t \mid 0)\,p(s)\,p(t).
\end{aligned}$$

2.1.3 Conditioning on $X_2 = 2$

The event $\{X_2 = 2\}$ for $X_2 := \sum_{j\geq 1} jN_{+j+}$ can happen in two ways: either $N_{+1+} = 2$ and each $N_{+j+} = 0$ for $j \geq 2$, or $N_{+2+} = 1$ and $N_{+j+} = 0$ for $j = 1$ and $j \geq 3$, with N_{+0+} unrestricted in each case. These two events have probabilities $({\theta_1}^2/2)e^{-\theta_+}$ and $(\theta_2)e^{-\theta_+}$, respectively, so the joint generating function for $\{X_1, X_3\}$ given $X_2 = 2$ is

$$\begin{aligned}
\varphi(s,t \mid 2) &:= \mathsf{E}\big[s^{X_1}t^{X_3} \mid X_2 = 2\big] \\
&= \mathsf{E}\big[s^{\sum_{i\geq 0} i(N_{i0+}+N_{i1+}+N_{i2+})} \quad t^{\sum_{k\geq 0} k(N_{+0k}+N_{+1k}+N_{+2k})} \mid N_{+1+} + 2N_{+2+} = 2\big] \\
&= \varphi(s,t \mid 0)\left\{\frac{{\theta_1}^2/2}{{\theta_1}^2/2 + \theta_2}\left[\sum_{i,k\geq 0} s^i t^k \lambda_{i1k}/\lambda_{+1+}\right]^2 + \frac{\theta_2}{{\theta_1}^2/2 + \theta_2}\left[\sum_{i,k\geq 0} s^i t^k \lambda_{i2k}/\lambda_{+2+}\right]\right\} \\
&= \frac{\varphi(s,t \mid 0)}{{\theta_1}^2/2 + \theta_2}\left\{\frac{{\theta_1}^2}{2}\left[\sum_{i,k\geq 0} s^i t^k \lambda_{i1k}/\theta_1\right]^2 + \theta_2\left[\sum_{i,k\geq 0} s^i t^k \lambda_{i2k}/\theta_2\right]\right\} \\
&= \frac{\varphi(s,t \mid 0)}{{\theta_1}^2/2 + \theta_2}\left\{\frac{{\theta_1}^2}{2}p(s)^2 p(t)^2 + \psi_2(s,t)\right\}.
\end{aligned} \tag{10}$$

In view of Eqn. (8), this will factor in the form $\varphi(s,t \mid 2) = \varphi(s,1 \mid 2)\,\varphi(1,t \mid 2)$ as required by Markov property Eqn. (6c) if and only if for all s,t in the unit ball:

$$\left[\frac{{\theta_1}^2}{2} + \theta_2\right]\left[\frac{{\theta_1}^2}{2}p(s)^2p(t)^2 + \psi_2(s,t)\right] = \left[\frac{{\theta_1}^2}{2}p(s)^2 + \psi_2(s,1)\right]\left[\frac{{\theta_1}^2}{2}p(t)^2 + \psi_2(1,t)\right]$$

or

$$\frac{{\theta_1}^2}{2}\Big[\theta_2 p(s)^2p(t)^2 - p(s)^2\psi_2(1,t) - \psi_2(s,1)p(t)^2 + \psi_2(s,t)\Big] \\ = \Big[\psi_2(s,1)\,\psi_2(1,t) - \theta_2\psi_2(s,t)\Big].$$

To satisfy the ID requirement of Eqn. (6e), this must hold with each θ_j replaced by θ_j/n for each integer $n \in \mathbb{N}$. Since the left and right sides are homogeneous in θ of degrees 3 and 2 respectively, this will only happen if each square-bracketed term vanishes identically, *i.e.*, if

$$\theta_2\psi_2(s,t) \equiv \psi_2(s,1)\psi_2(1,t)$$

and

$$\begin{aligned} 0 &= \theta_2\Big[\theta_2 p(s)^2p(t)^2 - p(s)^2\psi_2(1,t) - \psi_2(s,1)p(t)^2\Big] + \psi_2(s,1)\psi_2(1,t) \\ &= \left[\theta_2 p(s)^2 - \psi_2(s,1)\right]\left[\theta_2 p(t)^2 - \psi_2(1,t)\right], \end{aligned}$$

so

$$\begin{aligned} \psi_2(s,t) &:= \sum_{i,k\geq 0} s^i t^k \lambda_{i2k} = \theta_2 p(s)^2\, p(t)^2, \\ \varphi(s,t \mid 2) &= \varphi(s,t \mid 0)\, p(s)^2 p(t)^2. \end{aligned}$$

2.1.4 Conditioning on $X_2 = j$

The same argument applied recursively, using the Markov property for each $j \geq 1$ in succession, leads to:

$$\left[\frac{{\theta_1}^j}{j!} + \cdots + \theta_1\theta_{j-1}\right]\Big[\theta_j p(s)^j p(t)^j - p(s)^j\psi_j(1,t) - \psi_j(s,1)p(t)^j + \psi_j(s,t)\Big] \\ = \Big[\psi_j(s,1)\psi_j(1,t) - \theta_j\psi_j(s,t)\Big]$$

so

$$\psi_j(s,t) := \sum_{i,k\geq 0} s^i t^k \lambda_{ijk} = \theta_j\, p(s)^j p(t)^j, \qquad j \geq 1 \tag{11}$$

and consequently

$$\varphi(s,t \mid j) = \mathsf{E}\big[s^{X_1}t^{X_3} \mid X_2 = j\big] = \left[\varphi(s,1 \mid 0)\, p(s)^j\right]\ \left[\varphi(1,t \mid 0)\, p(t)^j\right].$$

Conditionally on $\{X_2 = j\}$, X_1 and X_3 are distributed independently, each as the sum of j independent random variables with generating function $p(s)$, plus one with generating function $\varphi(s,1 \mid 0)$—

so X_t is a branching process (Harris, 1963) whose unconditional three-dimensional marginal distributions have generating function:

$$\begin{aligned}\varphi(s,z,t) &:= \mathsf{E}\big[s^{X_1} z^{X_2} t^{X_3}\big] \\ &= \varphi(s,t \mid 0) \sum_{j\geq 0} z^j p(s)^j p(t)^j \mathsf{P}[X_2 = j] \\ &= \varphi(s,t \mid 0) \mathsf{E}\left[zp(s)p(t)\right]^{X_2} \\ &= \varphi(s,t \mid 0) \varphi\big(zp(s)p(t)\big) \\ &= \varphi(s,t \mid 0) \exp\big[P\big(zp(s)p(t)\big) - \theta_+\big]. \end{aligned} \tag{12}$$

See Secs. 4.3 and 5 for further development of this branching process representation.

2.2 Stationarity

Without loss of generality we may take $\lambda_{000} = 0$. By Eqn. (11) with $s = 0$ and $t = 1$ we have $\lambda_{0j+} = \theta_j {r_0}^j$; by Eqn. (9) we have $\lambda_{i00} = \lambda_{i0+}$. By time-reversibility we conclude that $\lambda_{i00} = 0$ for $i = 0$ and, for $i \geq 1$,

$$\lambda_{i00} = \theta_i {r_0}^i. \tag{13}$$

Now we can evaluate

$$\varphi(s,t \mid 0) = \exp\left\{P(s\, r_0) + P(t\, r_0) - 2P(r_0)\right\}$$

and, from this and Eqn. (12), evaluate the joint generating function for $X_{\{1,2,3\}}$ as:

$$\varphi(s,z,t) = \exp\left\{P\big(z\, p(s)p(t)\big) - \theta_+ + P(s\, r_0) + P(t\, r_0) - 2P(r_0)\right\}, \qquad j \geq 1 \tag{14}$$

and so that for $X_{\{1,2\}}$ as:

$$\varphi(s,z,1) = \exp\left\{P\big(z\, p(s)\big) - \theta_+ + P(s\, r_0) - P(r_0)\right\}. \tag{15}$$

Now consider Eqn. (11) with $t = 1$,

$$\sum_{i\geq 0} s^i \lambda_{ij+} = \theta_j\, p(s)^j. \tag{16}$$

It follows first for $j = 1$ and then for $i = 1$ that

$$\begin{aligned} \lambda_{i1+} &= \theta_1 r_i && i \geq 1 \\ \lambda_{1j+} &= \theta_j [j {r_0}^{j-1} r_1] && j \geq 1 \end{aligned}$$

so again by time reversibility with $i = j$, since $\theta_1 > 0$, we have

$$r_j = \theta_j [j\, {r_0}^{j-1} r_1]/\theta_1 \qquad j \geq 1. \tag{17}$$

Thus r_0, r_1, and $\{\theta_j\}$ determine all the $\{r_j\}$ and so all the $\{\lambda_{ijk}\}$ by Eqns. (11, 13) and hence the joint distribution of $\{X_t\}$.

Now consider Eqn. (16) first for $j = 2$ and then $i = 2$:

$$\sum_{i\geq 0} s^i \lambda_{ij+} = \theta_j \left[\sum\nolimits_{i\geq 0} s^i r_i\right]^j$$

$$\lambda_{i2+} = \theta_2 \sum_{k=0}^{i} r_k r_{i-k} \qquad i \geq 2$$

$$\lambda_{2j+} = \theta_j \left[j {r_0}^{j-1} r_2 + \binom{j}{2} {r_0}^{j-2} {r_1}^2\right] \qquad j \geq 2$$

Equating these for $i = j \geq 2$ (by time-reversibility) and applying Eqn. (17) for $0 < k < i$ (the cases $k = 0$ and $k = i$ need to be handled separately),

$$ {r_0}^{i-2} {r_1}^2 \left[\theta_2 \sum_{0<k<i} \theta_k \theta_{i-k} k(i-k) - \theta_i \frac{i(i-1)}{2} {\theta_1}^2\right] = 0. \tag{18}$$

3 The Solutions

Eqn. (18) holds for all $i \geq 2$ if $r_0 = 0$ or $r_1 = 0$, leaving $r_j = 0$ by Eqn. (17) for all $j \geq 2$, hence $r_0 + r_1 = 1$ and $\{\theta_j\}$ is restricted only by the conditions $\theta_1 > 0$ and $\theta_+ < \infty$.

3.1 The Constant Case

The case $r_0 = 0$ leads to $r_1 = 1$ and $r_j = 0$ for all $j \neq 1$, so $p(z) \equiv z$. By Eqn. (14) the joint pgf is

$$\varphi(s, z, t) = \exp\left\{P(s\,z\,t) - \theta_+\right\},$$

so $X_1 = X_2 = X_3$ and all $\{X_t\}$ are identical, with an arbitrary ID distribution.

3.2 The IID Case

The case $r_1 = 0$ leads to $r_0 = 1$ and $r_j = 0$ for all $j \neq 0$ so $p(z) \equiv 1$ and

$$\varphi(s, z, t) = \exp\left\{P(s) + P(z) + P(t) - 3\theta_+\right\}$$

by Eqn. (14), making all $\{X_t\}$ independent, with identical but arbitrary ID distributions.

3.3 The Poisson Case

Aside from these two degenerate cases, we may assume $r_0 > 0$ and $r_1 > 0$, and (by Eqn. (17)) rewrite Eqn. (18) in the form:

$$r_i = \frac{r_2}{{r_1}^2 (i-1)} \sum_{k=1}^{i-1} r_k r_{i-k}, \quad i \geq 2,$$

whose unique solution for all integers $i \geq 1$ (by induction) is

$$r_i = r_1 (r_2/r_1)^{i-1}. \tag{19}$$

If $r_2 = 0$, then again $r_i = 0$ for all $i \geq 2$ but, by Eqn. (17), $\theta_j = 0$ for all $j \geq 2$; thus $P(z) = \theta_1 z$ so each $X_t \sim \mathsf{Po}(\theta_1)$ has a Poisson marginal distribution with mean $\theta_1 = \theta_+$. In this case $r_0 + r_1 = 1$, $p(z) = r_0 + r_1 z$, and the two-dimensional marginals (by Eqn. (15)) of X_1, X_2 have joint pgf

$$\begin{aligned}\varphi(s,z) &= \exp\left\{P\big(z\,p(s)\big) - \theta_+ + P(s\,r_0) - P(r_0)\right\} \\ &= \exp\left\{\theta_1 r_0(s+z-2) + \theta_1 r_1(sz-1)\right\},\end{aligned} \tag{20}$$

the bivariate Poisson distribution (Johnson, Kotz and Balakrishnan, 1997, § 37.2), so X_t is the familiar "Poisson AR(1) Process" of McKenzie (1985, 1988) (with autocorrelation $\rho = r_1$) considered in Sec. (1.1.1). Its connection with Markov branching processes was recognized earlier by Steutel, Vervaat and Wolfe (1983). By Eqn. (20) the conditional distribution of X_{t+1}, given $\mathcal{F}_t := \sigma\{X_s : \ s \leq t\}$, is that of the sum of X_t independent Bernoulli random variables with pgf $p(s)$ and a Poisson innovation term with pgf $\exp\{P(r_0 s) - P(r_0)\}$, so the Markov process X_t may be written recursively starting at any t_0 as

$$\begin{aligned}&X_{t_0} \sim \mathsf{Po}(\theta_+) \\ &X_t = \xi_t + \zeta_t, \quad \text{where } \xi_t \sim \mathsf{Bi}(X_{t-1}, r_1) \text{ and } \zeta_t \sim \mathsf{Po}(\theta_t r_0)\end{aligned}$$

(all independent) for $t > t_0$, the thinning construction of Sec. (1)

3.4 The Negative Binomial case

Finally if $r_0 > 0$, $r_1 > 0$, and $r_2 > 0$, then (by Eqn. (19)) $r_i = r_1(qr_0)^{i-1}$ for $i \geq 1$ and hence (by Eqn. (17)) $\theta_j = \alpha q^j/j$ for $j \geq 1$ with $q := (1 - r_0 - r_1)/r_0(1 - r_0)$ and $\alpha := \theta_1/q$. The condition $\theta_+ < \infty$ entails $q < 1$ and $\theta_+ = -\alpha \log(1-q)$. The 1-marginal distribution is $X_t \sim \mathsf{NB}(\alpha, p)$ with $p := (1-q)$, and the functions $P(\cdot)$ and $p(\cdot)$ are $P(z) = -\alpha \log(1 - qz)$, $p(s) = r_0 + r_1 s/(1 - qr_0 s)$, so the joint pgf for the 2-marginal distribution of X_1, X_2 is

$$\begin{aligned}\varphi(s,z) &= \exp\left\{P\big(z\,p(s)\big) - \theta_+ + P(s\,r_0) - P(r_0)\right\} \\ &= p^{2\alpha}[(1 - q\rho) - q(1-\rho)(s+z) + q(q-\rho)sz]^{-\alpha}\end{aligned} \tag{21}$$

with one-step autocorrelation $\rho := (1-r_0)^2/r_1$. This bivariate distribution was introduced by Edwards and Gurland (1961) as the "compound correlated bivariate Poisson", but we prefer to call it the Branching Negative Binomial distribution. In the branching formulation X_t may be viewed as the sum of X_{t-1} iid random variables with pgf $p(s) = r_0 + r_1 s/(1 - qr_0 s)$ and one with pgf $\exp\{P(sr_0) - P(r_0)\} = (1 - qr_0)^\alpha (1 - qr_0\, s)^{-\alpha}$. The first of these may be viewed as Y_t plus a random variable with the $\mathsf{NB}(Y_t, 1-qr_0)$ distribution, for $Y_t \sim \mathsf{Bi}(X_{t-1}, 1 - r_0)$, and the second has the $\mathsf{NB}(\alpha, 1-qr_0)$ distribution, so a recursive updating scheme beginning with $X_{t_0} \sim \mathsf{NB}(\alpha, p)$ is:

$$X_t = Y_t + \zeta_t, \quad \text{where } Y_t \sim \mathsf{Bi}(X_{t-1},\ 1-r_0) \text{ and } \zeta_t \sim \mathsf{NB}(\alpha + Y_t,\ 1-qr_0).$$

In the special case of $\rho = q$ the joint pgf simplifies to $\varphi(s,z) = p^\alpha[1 + q(1 - s - z)]^{-\alpha}$ and the joint distribution of X_1, X_2 reduces to the negative trinomial distribution (Johnson *et al.*, 1997, Ch. 36) with pmf

$$\mathsf{P}[X_1 = i, X_2 = j] = \frac{\Gamma(\alpha + i + j)}{\Gamma(\alpha)\ i!\ j!}\left(\frac{1-q}{1+q}\right)^\alpha \left(\frac{q}{1+q}\right)^{i+j}$$

and simple recursion $X_t \mid X_{t-1} \sim \mathsf{NB}\big(\alpha + X_{t-1},\ \frac{1}{1+q}\big)$.

3.5 Results

We have just proved:

Theorem 1. *Let $\{X_t\}$ be a Markov process indexed by $t \in \mathbb{Z}$ taking values in the non-negative integers $\mathbb{N}_0$ that is stationary, time-reversible, has infinitely-divisible marginal distributions of all finite orders, and satisfies $\mathsf{P}[X_t = 1] > 0$. Then $\{X_t\}$ is one of four processes:*

1. *$X_t \equiv X_0 \sim \mu_0(dx)$ for an arbitrary ID distribution μ_0 on $\mathbb{N}_0$ with $\mu_0(\{1\}) > 0$;*
2. *$X_t \overset{\text{iid}}{\sim} \mu_0(dx)$ for an arbitrary ID distribution μ_0 on $\mathbb{N}_0$ with $\mu_0(\{1\}) > 0$;*
3. *For some $\theta > 0$ and $0 < \rho < 1$, $X_t \sim \mathsf{Po}(\theta)$ with bivariate joint generating function*

$$\mathsf{E}\left[s^{X_1}\ z^{X_2}\right] = \exp\left\{\theta(1-\rho)(s-1) + \theta(1-\rho)(z-1) + \theta\rho(sz-1)\right\}$$

and hence correlation $\mathsf{Corr}(X_s, X_t) = \rho^{|s-t|}$ and recursive update

$$X_t = \xi_t + \zeta_t, \quad \text{where } \xi_t \sim \mathsf{Bi}(X_{t-1},\ \rho) \text{ and } \zeta_t \sim \mathsf{Po}\big(\theta(1-\rho)\big);$$

4. *For some $\alpha > 0$, $0 < p < 1$, and $0 < \rho < 1$, $X_t \sim \mathsf{NB}(\alpha, p)$, with bivariate joint generating function*

$$\mathsf{E}\left[s^{X_1}\ z^{X_2}\right] = p^{2\alpha}[(1-q\rho) - q(1-\rho)(s+z) + q(q-\rho)sz]^{-\alpha}$$

where $q = 1-p$, and hence correlation $\mathsf{Corr}(X_s, X_t) = \rho^{|s-t|}$ and recursive update

$$X_t = Y_t + \zeta_t, \quad \text{where } Y_t \sim \mathsf{Bi}\big(X_{t-1},\ \rho\, p/(1-\rho q)\big) \text{ and } \zeta_t \sim \mathsf{NB}\big(\alpha + Y_t,\ p/(1-\rho q)\big).$$

Note the limiting cases of autocorrelation $\rho = 1$ and $\rho = 0$ in cases 3., 4. are subsumed by the degenerate cases 1. and 2., respectively. From this theorem follows:

Theorem 2. *Let $\left\{\mu^\theta :\ \theta \geq 0\right\}$ be an ID semigroup of probability distributions on the nonnegative integers $\mathbb{N}_0$ with $\mu^\theta(\{1\}) > 0$. Fix $\theta > 0$ and $0 < \rho < 1$ and let $\{X_t\}$ be the "thinning process" of Eqn. (1) in Sec. (1.1) with the representation*

$$X_{t-1} = \xi_t + \eta_t \qquad X_t = \xi_t + \zeta_t \tag{22}$$

for each $t \in \mathbb{Z}$ with independent

$$\xi_t \sim \mu^{\rho\theta}(d\xi) \qquad \eta_t \sim \mu^{(1-\rho)\theta}(d\eta) \qquad \zeta_t \sim \mu^{(1-\rho)\theta}(d\zeta).$$

Then X_t is Markov, stationary, time-reversible, and nonnegative integer valued, but it does not have infinitely-divisible marginal distributions of all orders unless $\{\mu^\theta\}$ is the Poisson family.

Proof. By construction X_t is obviously Markov and stationary. The joint distribution of the process at consecutive times is symmetric (see Eqn. (22)) since the marginal and conditional pmfs

$$p(x) := \mu^\theta(\{x\}), \qquad q(y \mid x) := \frac{\sum_z \mu^{\rho\theta}(\{z\})\ \mu^{(1-\rho)\theta}(\{x-z\})\ \mu^{(1-\rho)\theta}(\{y-z\})}{\mu^\theta(\{x\})}$$

of X_t and $X_t \mid X_{t-1}$ satisfy the symmetric relation

$$p(x)\ q(y \mid x) = q(x \mid y)\ p(y).$$

Applying this inductively, for any $s < t$ and any $\{x_s, \cdots, x_t\} \subset \mathbb{N}_0$ we find

$$\begin{aligned}
\mathsf{P}[X_s = x_s, \cdots, X_t = x_t] &= p(x_s) \quad q(x_{s+1} \mid x_s)\ q(x_{s+2} \mid x_{s+1}) \cdots q(x_t \mid x_{t-1}) \\
&= q(x_s \mid x_{s+1}) p(x_{s+1})\ q(x_{s+2} \mid x_{s+1}) \cdots q(x_t \mid x_{t-1}) \\
&= \cdots \\
&= q(x_s \mid x_{s+1}) q(x_{s+1} \mid x_{s+2}) \cdots q(x_{t-1} \mid x_t)\ p(x_t),
\end{aligned}$$

and so the distribution of X_t is time-reversible. Now suppose that it is also ID. Then by Theorem 1 it must be one of the four specified processes: constant, iid, branching Poisson, or branching negative binomial.

Since $\rho < 1$ it cannot be the constant $\{X_t \equiv X_0\}$ process; since $\rho > 0$ it cannot be the independent $\left\{X_t \overset{\text{iid}}{\sim} \mu^\theta(dx)\right\}$ process. The joint generating function $\phi(s, z)$ at two consecutive times for the negative binomial thinning process, given in Eqn. (2), differs from that for the negative binomial branching process, given in Eqn. (21). The only remaining option is the Poisson branching process of Sec. (1.1.1). □

Theorem 3. *Let $\left\{\mu^\theta :\ \theta \geq 0\right\}$ be an ID semigroup of probability distributions on the nonnegative integers $\mathbb{N}_0$ with $\mu^\theta(\{1\}) > 0$. Fix $\theta > 0$ and $0 < \rho < 1$ and let $\{X_t\}$ be the "random measure process" of Eqn. (4) in Sec. (1.2). Then X_t is ID, stationary, time-reversible, and nonnegative integer valued, but it is not a Markov process unless $\{\mu^\theta\}$ is the Poisson family.*

Proof. By construction X_t is ID, stationary, and time-reversible; suppose that it is also Markov. Then by Theorem 1 it must be one of the four specified processes: constant, iid, branching Poisson, or branching negative binomial.

Since $\rho < 1$ it cannot be the constant $\{X_t \equiv X_0\}$ process; since $\rho > 0$ it cannot be the independent $\left\{X_t \overset{\text{iid}}{\sim} \mu^\theta(dx)\right\}$ process. The joint generating function $\phi(s, z)$ at two consecutive times for the negative binomial random measure process coincides with that for the negative binomial thinning process, given in Eqn. (2), and differs from that for the negative binomial branching process, given in Eqn. (21). The only remaining option is the Poisson branching process of Sec. (1.1.1). □

4 Continuous Time

Now consider $\mathbb{N}_0$-valued time-reversible stationary Markov processes indexed by continuous time $t \in \mathbb{R}$. The restriction of any such process to $t \in \mathbb{Z}$ will still be Markov, hence MISTI, so there can be at most two non-trivial ones— one with univariate Poisson marginal distributions, and one with univariate Negative Binomial distributions. Both do in fact exist.

4.1 Continuous-Time Poisson Branching Process

Fix $\theta > 0$ and $\lambda > 0$ and construct a nonnegative integer-valued Markov process with generator

$$\begin{aligned}
\mathfrak{A} f(x) &= \frac{\partial}{\partial s} \mathsf{E}[f(X_s) - f(X_t) \mid X_t = x]\Big|_{s=t} \\
&= \lambda\theta\big[f(x+1) - f(x)\big] + \lambda x\big[f(x-1) - f(x)\big] \qquad (23a)
\end{aligned}$$

or, less precisely but more intuitively, for all $i, j \in \mathbb{N}_0$ and $\epsilon > 0$,

$$\mathsf{P}\big[X_{t+\epsilon} = i \mid X_t = j\big] = o(\epsilon) + \begin{cases} \epsilon\lambda\theta & i = j+1 \\ 1 - \epsilon\lambda(\theta + j) & i = j \\ \epsilon\lambda j & i = j-1 \end{cases} \tag{23b}$$

X_t could be described as a linear death process with immigration. In Sec. (4.4) we verify that its univariate marginal distribution and autocorrelation are

$$X_t \sim \mathsf{Po}(\theta)$$
$$\mathsf{Corr}(X_s, X_t) = e^{-\lambda|s-t|},$$

and its restriction to integer times $t \in \mathbb{Z}$ is precisely the process described in Sec. (3) item 3, with one-step autocorrelation $\rho = e^{-\lambda}$.

4.2 Continuous-Time Negative Binomial Branching Process

Now fix $\theta > 0$, $\lambda > 0$, and $0 < p < 1$ and construct a nonnegative integer-valued Markov process with generator

$$\begin{aligned} \mathfrak{A}f(x) &= \frac{\partial}{\partial s}\mathsf{E}[f(X_s) - f(X_t) \mid X_t = x]\Big|_{s=t} \\ &= \frac{\lambda(\alpha + x)(1-p)}{p}\big[f(x+1) - f(x)\big] + \frac{\lambda x}{p}\big[f(x-1) - f(x)\big] \end{aligned} \tag{24a}$$

or, for all $i, j \in \mathbb{N}_0$ and $\epsilon > 0$,

$$\mathsf{P}\big[X_{t+\epsilon} = i \mid X_t = j\big] = o(\epsilon) + \begin{cases} \epsilon\lambda(\alpha + j)(1-p)/p & i = j+1 \\ 1 - \epsilon\lambda[(\alpha + j)(1-p) + j]/p & i = j \\ \epsilon\lambda j/p & i = j-1, \end{cases} \tag{24b}$$

so X_t is a linear birth-death process with immigration. The univariate marginal distribution and autocorrelation (see Sec. (4.4)) are now

$$X_t \sim \mathsf{NB}(\alpha, p)$$
$$\mathsf{Corr}(X_s, X_t) = e^{-\lambda|s-t|},$$

and its restriction to integer times $t \in \mathbb{Z}$ is precisely the process described in Sec. (3) item 4, with autocorrelation $\rho = e^{-\lambda}$.

4.3 Markov Branching (Linear Birth/Death) Processes

The process X_t of Sec. (4.1) can also be described as the size of a population at time t if individuals arrive in a Poisson stream with rate $\lambda\theta$ and die or depart independently after exponential holding times with rate λ; as such, it is a continuous-time Markov branching process.

Similarly, that of Sec. (4.2) can be described as the size of a population at time t if individuals arrive in a Poisson stream with rate $\lambda\alpha(1-p)/p$, give birth (introducing one new individual) independently at rate $\lambda(1-p)/p$, and die or depart at rate λ/p. In the limit as $p \to 1$ and $\alpha \to \infty$ with $\alpha(1-p) \to \theta$ this will converge in distribution to the Poisson example of Sec. (4.1).

4.4 Marginal Distributions

Here we verify that the Poisson and Negative Binomial distributions are the stationary distributions for the Markov chains with generators $\mathfrak{A}$ given in Eqn. (23) and Eqn. (24), respectively.

Denote by $\pi_i^0 = \mathsf{P}[X_t = i]$ the pmf for X_t and by $\pi_i^\epsilon = \mathsf{P}[X_{t+\epsilon} = i]$ that for $X_{t+\epsilon}$, and by $\varphi_0(s) = \mathsf{E}[s^{X_t}]$ and $\varphi_\epsilon(s) = \mathsf{E}[s^{X_{t+\epsilon}}]$ their generating functions. The stationarity requirement that $\varphi_0(s) \equiv \varphi_\epsilon(s)$ will determine $\varphi(s)$ and hence $\{\pi_i\}$ uniquely.

4.4.1 Poisson

From Eqn. (23b) for $\epsilon > 0$ we have

$$\pi_i^\epsilon = \epsilon\lambda\theta\pi_{i-1}^0 \qquad + [1 - \epsilon\lambda(\theta + i)]\pi_i^0 \qquad + \epsilon\lambda(i+1)\pi_{i+1}^0 \qquad + o(\epsilon).$$

Multiplying by s^i and summing, we get:

$$\begin{aligned}\varphi_\epsilon(s) &= \epsilon\lambda\theta s \sum_{i\geq 1} s^{i-1}\pi_{i-1}^0 + [1 - \epsilon\lambda\theta]\varphi_0(s) - \epsilon\lambda s \sum_{i\geq 0} i s^{i-1}\pi_i^0 + \epsilon\lambda \sum_{i\geq 0}(i+1)s^i\pi_{i+1}^0 + o(\epsilon)\\ &= \epsilon\lambda\theta s\varphi_0(s) \qquad + [1 - \epsilon\lambda\theta]\varphi_0(s) - \epsilon\lambda s\varphi_0'(s) \qquad + \epsilon\lambda\varphi_0'(s) \qquad + o(\epsilon)\end{aligned}$$

so

$$\varphi_\epsilon(s) - \varphi_0(s) = \epsilon\lambda(s-1)\left[\theta\varphi_0(s) - \varphi_0'(s)\right] + o(\epsilon)$$

and stationarity ($\varphi_0(s) \equiv \varphi_\epsilon(s)$) entails $\lambda = 0$ or $\varphi_0'(s)/\varphi_0(s) \equiv \theta$, so $\log\varphi_0(s) \equiv (s-1)\theta$ and:

$$\varphi_0(s) = \exp\{(s-1)\theta\}$$

so $X_t \sim \mathsf{Po}(\theta)$ is the unique stationary distribution.

4.4.2 Negative Binomial

From Eqn. (24b) for $\epsilon > 0$ we have

$$\begin{aligned}\pi_i^\epsilon &= (\epsilon\lambda(1-p)/p)(\alpha + i - 1)\ \pi_{i-1}^0 + \{1 - (\epsilon\lambda/p)[(\alpha+i)(1-p) + i]\}\ \pi_i^0 + (\epsilon\lambda/p)(i+1)\ \pi_{i+1}^0 + o(\epsilon)\\ \varphi_\epsilon(s) &= (\epsilon\lambda(1-p)/p)\alpha\ s\varphi_0(s) + (\epsilon\lambda(1-p)/p)\ s^2\varphi_0'(s)\\ &\quad + \varphi_0(s) - (\epsilon\lambda(1-p)/p)\alpha\varphi_0(s) - (\epsilon\lambda/p)((1-p)+1)\ s\varphi_0'(s)\\ &\quad + (\epsilon\lambda/p)\ \varphi_0'(s) + o(\epsilon)\\ \varphi_\epsilon(s) - \varphi_0(s) &= (\epsilon\lambda/p)\ \left\{\varphi_0(s)\ \alpha(1-p)(s-1) + \varphi_0'(s)\ [(1-p)s^2 - ((1-p)+1)s + 1]\right\} + o(\epsilon)\\ &= (\epsilon\lambda/p)(s-1)\ \left\{\varphi_0(s)\ \alpha(1-p) + \varphi_0'(s)\ ((1-p)s - 1)\right\} + o(\epsilon)\end{aligned}$$

so either $\lambda = 0$ (the trivial case where $X_t \equiv X$) or $\lambda > 0$ and:

$$\begin{aligned}\varphi_0'(s)/\varphi_0(s) &= \alpha(1-p)(1-(1-p)s)^{-1}\\ \log\varphi_0(s) &= -\alpha\log(1-(1-p)s) + \alpha\log(p)\\ \varphi_0(s) &= p^\alpha(1-(1-p)s)^{-\alpha}\end{aligned}$$

and $X_t \sim \mathsf{NB}(\alpha, p)$ is the unique stationary distribution.

4.4.3 Alternate Proof

A detailed-balance argument (Hoel, Port and Stone, 1972, *p.* 105) shows that the stationary distribution $\pi_i := \mathsf{P}[X_t = i]$ for linear birth/death chains is proportional to

$$\pi_i \propto \prod_{0 \le j < i} \frac{\beta_j}{\delta_{j+1}}$$

where β_j and δ_j are the birth and death rates when $X_t = j$, respectively. For the Poisson case, from Eqn. (23b) this is

$$\pi_i \propto \prod_{0 \le j < i} \frac{\lambda\theta}{\lambda(j+1)} = \theta^i / i!,$$

so $X_t \sim \mathsf{Po}(\theta)$, while for the Negative Binomial case from Eqn. (24b) we have

$$\pi_i \propto \prod_{0 \le j < i} \frac{\lambda(\alpha+j)(1-p)/p}{\lambda(j+1)/p} = \frac{\Gamma(\alpha+i)}{\Gamma(\alpha)\, i!}\, (1-p)^i,$$

so $X_t \sim \mathsf{NB}(\alpha, p)$. In each case the proportionality constant is $\pi_0 = P[X_t = 0]$: $\pi_0 = e^{-\theta}$ for the Poisson case, and $\pi_0 = p^\alpha$ for the negative binomial.

4.4.4 Autocorrelation

Aside from the two trivial (iid and constant) cases, MISTI processes have finite pth moments for all $p < \infty$ and, in particular, have finite variance and well-defined autocorrelation. By the Markov property and induction that autocorrelation must be of the form

$$\mathsf{Corr}[X_s, X_t] = \rho^{-|t-s|}$$

for some $\rho \in [-1, 1]$. In both the Poisson and negative binomial cases the one-step autocorrelation ρ is nonnegative; without loss of generality we may take $0 < \rho < 1$.

5 Discussion

The condition $\mu^\theta(\{1\}) > 0$ introduced in Sec. (2.1) to avoid trivial technicalities is equivalent to a requirement that the support $\mathrm{spt}(\mu^\theta) = \mathbb{N}_0$ be all of the nonnegative integers. Without this condition, for any MISTI process X_t and any integer $k \in \mathbb{N}$ the process $Y_t = k\, X_t$ would also be MISTI, leading to a wide range of essentially equivalent processes.

The branching approach of Sec. (4.3) could be used to generate a wider class of continuous-time stationary Markov processes with ID marginal distributions (Vervaat, 1979; Steutel *et al.*, 1983). If families of size $k \ge 1$ immigrate independently in Poisson streams at rate λ_k, with $\sum_{k \ge 1} \lambda_k \log k < \infty$, and if individuals (after independent exponential waiting times) either die (at rate $\delta > 0$) or give birth to some number $j \ge 1$ of progeny (at rate $\beta_j \ge 0$), respectively, with $\delta > \sum_{j \ge 1} j\, \beta_j$, then the population size X_t at time t will be a Markov, infinitely-divisible, stationary processes with nonnegative integer values. Unlike the MISTI processes, these may have

infinite pth moments if $\sum_{k\geq 1} \lambda_k k^p = \infty$ for some $p > 0$ and, in particular, may not have finite means, variances, or autocorrelations.

Unless $\lambda_k = 0$ and $\beta_j = 0$ for all $k, j > 1$, however, these will not be time-reversible, and hence not MISTI. Decreases in population size are always of unit size (necessary for the Markov property to hold), while increases might be of size $k > 1$ (if immigrating family sizes exceed one) or $j > 1$ (if multiple births occur).

Acknowledgments

The authors would like to thank Xuefeng Li, Avi Mandelbaum, Yosef Rinott, Larry Shepp, and Henry Wynn for helpful conversations. This work was supported in part by National Science Foundation grants DMS–0635449, DMS–0757549, PHY–0941373 and NASA grant NNX09AK60G.

References

Abramowitz, M. and Stegun, I. A., eds. (1964). *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, *Applied Mathematics Series*, vol. 55. National Bureau of Standards, Washington, D.C. Reprinted in paperback by Dover (1974); on-line at http://www.math.sfu.ca/~cbm/aands/.

Al-Osh, M. A. and Alzaid, A. A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.* **8** 261–275.

Edwards, C. B. and Gurland, J. (1961). A class of distributions applicable to accidents. *J. Am. Stat. Assoc.* **56** 503–517.

Harris, T. E. (1963). *The Theory of Branching Processes*, *Die Grundlehren der Mathematischen Wissenschaften*, vol. 119. Springer-Verlag, Berlin, DE.

Hoel, P. G., Port, S. C. and Stone, C. J. (1972). *Introduction to Stochastic Processes.* Houghton Mifflin, Boston, MA.

Joe, H. (1996). Time series models with univariate margins in the convolution-closed infinitely divisible class. *J. Appl. Probab.* **33** 664–677.

Johnson, N. L., Kemp, A. W. and Kotz, S. (2005). *Univariate Discrete Distributions.* John Wiley & Sons, New York, NY, third edn.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997). *Discrete Multivariate Distributions.* John Wiley & Sons, New York, NY, second edn.

Lewis, P. A. W. (1983). Generating negatively correlated gamma variates using the beta-gamma transformation. In *Proceedings of the 1983 Winter Simulation Conference* (S. D. Roberts, J. Banks and B. W. Schmeiser, eds.), 175–176.

Lewis, P. A. W., McKenzie, E. and Hugus, D. K. (1989). Gamma processes. *Commun. Stat. Stoch. Models* **5** 1–30.

McKenzie, E. (1985). Some simple models for discrete variate time series. *Water Resources Bulletin* **21** 645–650.

McKenzie, E. (1988). Some ARMA models for dependent sequences of Poisson counts. *Adv. Appl. Probab.* **20** 822–835.

Rogers, L. C. G. and Williams, D. (2000). *Diffusions, Markov Processes, and Martingales*, vol. 1. Cambridge Univ. Press, Cambridge, UK, second edn.

Steutel, F. W., Vervaat, W. and Wolfe, S. J. (1983). Integer-valued branching processes with immigration. *Adv. Appl. Probab.* **15** 713–725.

Vervaat, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Probab.* **11** 750–783.

Wolpert, R. L. and Taqqu, M. S. (2005). Fractional Ornstein-Uhlenbeck Lévy processes and the Telecom process: Upstairs and downstairs. *Signal Processing* **85** 1523–1545.