

Throwing stones and collecting bones: Looking for Poisson-like random measures

Caleb Deen Bastian¹  | Grzegorz A. Rempala² 

¹Program in Applied and Computational Mathematics, Princeton University, Princeton, New Jersey

²Division of Biostatistics and Department of Mathematics, The Ohio State University, Columbus, Ohio

Correspondence

Caleb Deen Bastian, Program in Applied and Computational Mathematics, Washington Road, Fine Hall, Princeton University, Princeton, NJ. 08544.
Email: cbastian@princeton.edu

Communicated by: S. Wise

Funding information

National Science Foundation, Grant/Award Number: DMS 1853587

We show that in a broad class of random counting measures, one may identify only three that are rescaled versions of themselves when restricted to a subspace. These are Poisson, binomial, and negative binomial random measures. We provide some simple examples of possible applications of such measures.

KEYWORDS

Laplace functional, Poisson-type (PT) distributions, random counting measure, stone throwing construction, strong invariance, thinning

MSC CLASSIFICATION

60G57; 60G18; 60G55

1 | INTRODUCTION

Random counting measures, also known as point processes, are the central objects of this note. For general introduction, see, for instance, monographs by Olav^{1,2} or Erhan.³ Random counting measures have numerous uses in statistics and applied probability, including representation and construction of stochastic processes, Monte Carlo schemes, etc. For example, the *Poisson* random measure is a fundamental random counting measure that is related to the structure of Lévy processes, Markov jump processes, or the excursions of Brownian motion, and is prototypical to the class of completely random (additive) random measures.³ In particular, it is also well known that the Poisson random measure is self-similar in the sense of being invariant under restriction to a subspace (invariant under thinning). The *binomial* random measure is another fundamental random counting measure that underlies the theory of autoregressive integer-valued processes.^{4,5}

In this note, we explore a broad class of random counting measures to identify those that share the Poisson self-similarity property and discuss their possible applications. The paper is organized as follows. In Section 2, we provide necessary background and lay out the main mathematical results, whereas in Section 3, we give examples of possible applications in different areas of modern sciences, from epidemiology to consumer research to traffic flows.

The main result of the note is Theorem 3, which identifies in a broad class of random counting measures those that are closed under restriction to subspaces, ie, invariant under thinning. They are the Poisson, negative binomial, and binomial random measures. We show that the corresponding counting distributions are the *only* distributions in the power series family that are invariant under thinning. We also give simple examples to highlight calculus of PT random measures and their possible applications.

2 | THROWING STONES AND LOOKING FOR BONES

Consider measurable space $(E, \mathcal{E})$ with some collection $\mathbf{X} = \{X_i\}$ of iid random variables (stones) with law ν and some non-negative integer-valued random variable K ($K \in \mathbb{N}_{\geq 0} = \mathbb{N}_{>0} \cup \{0\}$) with law κ that is independent of $\mathbf{X}$ and has finite mean c . Whenever it exists, the variance of K is denoted by $\delta^2 > 0$. Let $\mathcal{E}_+$ be the set of positive $\mathcal{E}$ -measurable functions.

It is well known³ that the random counting measure N on $(E, \mathcal{E})$ is uniquely determined by the pair of deterministic probability measures (κ, ν) through the so-called *stone throwing construction* (STC) as follows. For every outcome $\omega \in \Omega$,

$$N_\omega(A) = N(\omega, A) = \sum_{i=1}^{K(\omega)} \mathbb{I}_A(X_i(\omega)) \quad \text{for } A \in \mathcal{E}, \quad (1)$$

where K has law κ , the iid $X_1, X_2, \dots$ have law ν , and $\mathbb{I}_A(\cdot)$ denotes the indicator function for set A . Below, we write $N = (\kappa, \nu)$ to denote the random measure N determined by (κ, ν) through STC. We note that N may be also regarded as a mixed binomial process.² In particular, when κ is the Dirac measure, then N is a binomial process.² Note that on any test function $f \in \mathcal{E}_+$,

$$N_\omega f = \sum_i^{K(\omega)} f \circ X_i(\omega) = \sum_i^{K(\omega)} f(X_i(\omega)).$$

Below, for brevity, we write Nf , so that eg, $N(A) = N\mathbb{I}_A$. It follows from the above and the independence of K and $\mathbf{X}$ that

$$\mathbb{E}Nf = c\nu f \quad (2)$$

$$\text{Var}Nf = c\nu f^2 + (\delta^2 - c)(\nu f)^2, \quad (3)$$

and that the Laplace functional for N is

$$\mathbb{E}e^{-Nf} = \mathbb{E}(\mathbb{E}e^{-f(X)})^K = \mathbb{E}(\nu e^{-f})^K = \psi(\nu e^{-f}),$$

where $\psi(t) = \mathbb{E}t^K$ is the probability generating function (pgf) of K . In what follows, we will also sometimes consider the alternate pgf (apgf) defined as $\tilde{\psi}(t) = \mathbb{E}(1-t)^K$. Note also that for any measurable partition of E , say $\{A, \dots, B\}$, the joint distribution of the collection $N(A), \dots, N(B)$ is for $i, \dots, j \in \mathbb{N}$ and $i + \dots + j = k$

$$\begin{aligned} \mathbb{P}(N(A) = i, \dots, N(B) = j) &= \mathbb{P}(N(A) = i, \dots, N(B) = j | K = k) \mathbb{P}(K = k) \\ &= \frac{k!}{i! \dots j!} \nu(A)^i \dots \nu(B)^j \mathbb{P}(K = k). \end{aligned} \quad (4)$$

The following result extends construction of a random measure $N = (K, \nu)$ to the case when the collection $\mathbf{X}$ is expanded to $(\mathbf{X}, \mathbf{Y}) = \{(X_i, Y_i)\}$, where Y_i is a random transformation of X_i . Heuristically, Y_i represents some properties (marks) of X_i . We assume that the conditional law of Y follows some transition kernel according to $\mathbb{P}(Y \in B | X = x) = Q(x, B)$.

Theorem 1 (Marked STC). *Consider random measure $N = (K, \nu)$ and the transition probability kernel Q from $(E, \mathcal{E})$ into $(F, \mathcal{F})$. Assume that given the collection $\mathbf{X}$ the variables $\mathbf{Y} = \{Y_i\}$ are conditionally independent with $Y_i \sim Q(X_i, \cdot)$. Then, $M = (K, \nu \times Q)$ is a random measure on $(E \times F, \mathcal{E} \otimes \mathcal{F})$. Here, $\mu = \nu \times Q$ is understood as $\mu(dx, dy) = \nu(dx)Q(x, dy)$. Moreover, for any $f \in (\mathcal{E} \otimes \mathcal{F})_+$*

$$\mathbb{E}e^{-Mf} = \psi(\nu e^{-g}),$$

where $\psi(\cdot)$ is pgf of K , and $g \in \mathcal{E}_+$ satisfies $e^{-g(x)} = \int_F Q(x, dy) e^{-f(x, y)}$.

The proof of this result is standard but for convenience, we provide it in the appendix. For any $A \subset E$ with $\nu(A) > 0$, define the conditional law ν_A by $\nu_A(B) = \nu(A \cap B)/\nu(A)$. The following is a simple consequence of Theorem 1 upon taking the transition kernel $Q(x, B) = \mathbb{I}_A(x) \nu_A(B)$.

Corollary 1. $N_A = (N\mathbb{I}_A, \nu_A)$ is a well-defined random measure on the measurable subspace $(E \cap A, \mathcal{E}_A)$ where $\mathcal{E}_A = \{A \cap B : B \in \mathcal{E}\}$. Moreover, for any $f \in \mathcal{E}_+$

$$\mathbb{E}e^{-N_A f} = \psi(\nu e^{-f} \mathbb{I}_A + b),$$

where $b = 1 - \nu(A)$.

In many practical situations, one is interested in analyzing random measures of the form $N = (K, \nu \times Q)$ while having some information about the restricted measure $N_A = (N\mathbb{I}_A, \nu_A \times Q)$. Note that the counting variable for N_A is $K_A = N\mathbb{I}_A$, the original counting variable K restricted to (thinned by) the subset $A \subset E$. The purpose of this note is to identify the families of counting distributions K for which the family of random measures $\{N_A : A \subset E\}$ belongs to the same family of distributions. We refer to such families of counting distributions as “bones” and give their formal definition below. The term reflects the prototypical or foundational nature of these families within the class of random measures considered here. One obvious example is the Poisson family of distributions, but it turns out that there are also others. The definite result on the existence and uniqueness of random measures based on such “bones” in a broad class is given in Theorem 3 of Section 2.2.

2.1 | Subset invariant families (bones)

Let $N = (\kappa_\theta, \nu)$ be the random measure on $(E, \mathcal{E})$, where κ_θ is the distribution of K parametrized by $\theta > 0$, that is, $\mathbb{P}(K = k)_{k \geq 0} = (p_k(\theta))_{k \geq 0}$, where we assume $p_0(\theta) > 0$. For brevity, we write below $K \sim \kappa_\theta$.

Consider the family of random variables $\{N\mathbb{I}_A : A \subset E\}$ and let $\psi_A(t)$ be the pgf of $N\mathbb{I}_A$ with $\psi_\theta(t) = \psi_E(t)$ being the pgf of K (since $N\mathbb{I}_E = K$). Let $a = \nu(A)$, $b = 1 - a$ and note that

$$\psi_A(t) = \mathbb{E}(\mathbb{E} t^{\mathbb{I}_A})^K = \mathbb{E}(at + b)^K = \psi_\theta(at + b),$$

or equivalently, in terms of apgf, $\tilde{\psi}_A(t) = \tilde{\psi}_\theta(at)$.

Definition 1 (Bones). We say that the family $\{\kappa_\theta : \theta \in \Theta\}$ of counting probability measures is *strongly invariant* with respect to the family $\{N\mathbb{I}_A : A \subset E\}$ (is a “bone”) if for any $0 < a \leq 1$ there exists a mapping $h_a : \Theta \rightarrow \Theta$ such that

$$\psi_\theta(at + 1 - a) = \psi_{h_a(\theta)}(t). \quad (5)$$

Note that in terms of apgf, the above condition becomes simply $\tilde{\psi}_\theta(at) = \tilde{\psi}_{h_a(\theta)}(t)$.

In Table 1, we give some examples of such invariant (“bone”) families.

2.2 | Finding bones in power series family

Consider the family $\{\kappa_\theta : \theta \in \Theta\}$ to be in the form of the non-negative power series (NNPS) where

$$p_k(\theta) = a_k \theta^k / g(\theta). \quad (6)$$

and $p_0 > 0$. We call NNPS canonical if $a_0 = 1$. Setting $b = 1 - a$, we see that for canonical NNPS, the bone condition in Definition 1 becomes

$$g((at + b)\theta) = g(b\theta)g(h_a(\theta)t). \quad (7)$$

The following is a fundamental result on the existence of “bones” in the NNPS family.

Theorem 2 (Bones in NNPS). Let ν be diffuse (ie, non-atomic). For canonical NNPS κ_θ satisfying additionally $a_1 > 0$, the relation (7) holds iff $\log g(\theta) = \theta$ or $\log g(\theta) = \pm c \log(1 \pm \theta)$, where $c > 0$.

Name	Parameter θ	$\psi_\theta(t)$	$h_a(\theta)$
Poisson	λ	$\exp[\theta(t - 1)]$	$a\theta$
Bernoulli	$p/(1 - p)$	$(1 + \theta t)/(1 + \theta)$	$a\theta/(1 + (1 - a)\theta)$
Geometric	p	$(1 - \theta)/(1 - t\theta)$	$a\theta/(1 - (1 - a)\theta)$

TABLE 1 Some examples of “bone” distributions with corresponding pgfs and mappings of their canonical parameters

Proof. The proof follows from Lemma 1 in the appendix and the assumptions on NNPS family. $\square$

Remark 1 (Enumerating bones in NNPS). There are only three “bones” in canonical NNPS such that $a_1 > 0$, namely, κ_θ is either Poisson, negative binomial, or binomial. Note that the entries in Table 1 are all special cases.

The “bone” families of distributions $\{\kappa_\theta : \theta \in \Theta\}$ are sometimes referred to as *Poisson-type* or *PT*.⁶ We also refer to the random measures $N = (\kappa_\theta, \nu)$, where κ_θ is a “bone” family as Poisson-type or PT random measures. The following is the main result of this note.

Theorem 3 (Existence and uniqueness of PT random measures). *Assume that $K \sim \kappa_\theta$, where pgf ψ_θ belongs to the canonical NNPS family of distributions and $\{0, 1\} \subset \text{supp}(K)$. Consider the random measure $N = (\kappa_\theta, \nu)$ on the space $(E, \mathcal{E})$ and assume that ν is diffuse. Then, for any $A \subset E$ with $\nu(A) = a > 0$, there exists a mapping $h_a : \Theta \rightarrow \Theta$ such that the restricted random measure is $N_A = (\kappa_{h_a(\theta)}, \nu_A)$, that is,*

$$\mathbb{E}e^{-N_A f} = \psi_{h_a(\theta)}(\nu_A e^{-f}) \quad \text{for } f \in \mathcal{E}_+ \quad (8)$$

iff K is Poisson, negative binomial, or binomial.

Proof. The sufficiency part follows by direct verification of (8) for K Poisson, binomial, and negative binomial. The appropriate mappings are given in the last column of Table 1. The necessity part follows upon taking in (8) constant f of the form $f(x) \equiv -\log t$ for some $t \in (0, 1]$ and applying Corollary 1 and Theorem 2. $\square$

Remark 2. It follows from Theorem 1 that in Theorem 3, we may replace the laws ν and ν_A with $\nu \times Q$ and $\nu_A \times Q$, respectively.

Sometimes, it may be more convenient to parametrize PT distributions by their mean and variance (instead of θ) and write $PT(c, \delta^2)$. The following is useful in computations related to PT random measures.

Remark 3 (PT random measures can be thinned on average). Note that if $N = (\kappa_\theta, \nu)$ is a PT random measure and $K \sim \kappa_\theta = PT(c, \delta^2)$, then for any random variable $K_A = N \mathbb{I}_A$ where $A \subset E$ such that $\nu(A) = a > 0$, it follows from (5) that

$$\begin{aligned} \mathbb{E}K_A &= a\mathbb{E}K = ac \\ \mathbb{E}K_A(K_A - 1) &= a^2\mathbb{E}K(K - 1) = a^2(\delta^2 + c^2 - c). \end{aligned}$$

Remark 4 (Atomic measure and a nondifferentiable mapping). The iff result of Theorem 3 holds for diffuse measures ν . When ν is atomic, the sufficiency part holds but the necessity part (uniqueness) fails if we also relax the differentiability condition for the mapping h . To see this, consider the following simple example where we may construct a bone mapping for K that is not PT. This example was generously pointed out to us by one of the reviewers. Let $E = \{\blacklozenge, \blacksquare\}$ with $\nu\{\blacklozenge\} = 1/2$. There are four subsets of $A \subseteq E$ with functions

$$\psi_\emptyset(t) = 1, \quad \psi_{\blacklozenge}(t) = \mathbb{E}((t+1)/2)^K = \psi_{\blacksquare}(t), \quad \psi_E(t) = \mathbb{E}t^K$$

For $A = \{\blacklozenge\}$ with $a = \nu\{\blacklozenge\} = 1/2$, the restriction is $\psi_\theta(at + 1 - a) = \mathbb{E}((t+1)/2)^{K_\theta}$ with

$$K_\theta = \begin{cases} \tilde{K} = \sum_{i=1}^K C_i & \theta = 1 \\ K & \theta = 2 \end{cases}$$

where $\tilde{K}$ is the restricted or thinned version of K by independent coin tosses $\{C_i\}$ (Bernoulli random variables) and $\Theta = \{1, 2\}$. Then, the bone condition

$$\mathbb{E}((t+1)/2)^{K_\theta} = \mathbb{E}t^{K_{h_a(\theta)}}$$

is satisfied with the mapping

$$h_a(\theta) = \begin{cases} 1 & a = 1/2 \\ 2 & a = 1. \end{cases}$$

3 | EXAMPLES

Below, we discuss some simple examples of applications of PT random measures. The first one is an extension of the well-known construction for compound Poisson random measures. The second one is (to our knowledge) an original idea for application of binomial random measure to monitoring epidemics. Finally, the third one is an extension of a Poisson point process to a PT process in a particle system having birth and death dynamics, applied to traffic flows of spacecraft.

3.1 | Compound PT processes

Assume that the number of customers and their arrivals times over n days form a PT random measure (K, ν) with $K \sim PT(c, \delta^2)$ either Poisson or negative binomial. Consider the associated mark random measure $N = (K, \nu \times Q \times Q_2)$, where $T \sim \nu$ gives customer arrival times, and the transition kernels $Q(t, x) = P(X = x | T = t)$ and $Q_2(x, y) = \nu(y | X = x)$ describe, respectively, customer's "state" $x = 1, \dots, s$ and his/her amount Y spent at the store, so that each customer may be represented by the triple (T, X, Y) . We further assume that customers are independent with the conditional variable $(X | T = t) \sim \text{Multinom}(1, p_1^t, \dots, p_s^t)$ and the conditional variable $(Y | X = x)$, with mean α_x and variance β_x^2 . Assume that we only have information about customers on a specific subset A of n days. We would like to decompose the average total amount $\mathbb{E}Z$ spent by customers over the entire n days period into two components, corresponding to the observed and unobserved subsets (A and A^c). Let therefore

$$\mathbb{E}Z = \mathbb{E}Z_A + \mathbb{E}Z_{A^c}, \quad (9)$$

where Z_B is the total amount spent in time set $B \in \{A, A^c\}$. Recall PT random measure $N = (K, \tilde{\nu})$, where $\tilde{\nu} = \nu \times Q \times Q_2$, and consider two restricted measures $N_B = (K_B, \tilde{\nu}_B)$ where $\tilde{\nu}_B = \nu_B \times Q \times Q_2$ for $B \in \{A, A^c\}$. Then,

$$Z = Nf \quad \text{and} \quad Z_B = N_B f, \quad B \in \{A, A^c\}$$

where $f(t, x, y) = y$. By Theorem 3, Z_A and Z_{A^c} are also PT random measures with the corresponding $h_a(\theta)$ transformation as presented in the last row of Table 1. Setting $b = \nu(B)$ and recalling Remark 3, it follows from (2) that for $B \in \{A, A^c\}$

$$\begin{aligned} \mathbb{E}Z_B &= cb \tilde{\nu}_B f = c \tilde{\nu} f \mathbb{I}_B \\ &= c \int \mathbb{I}_B(t) \nu(dt) Q(t, dx) Q_2(x, dy) y \\ &= c \int_B \nu(dt) \sum_{x=1}^s p_x^t \alpha_x \end{aligned}$$

and

$$\begin{aligned} \text{Var} Z_B &= cb \tilde{\nu}_B f^2 + b^2 (\tilde{\nu}_B f)^2 (\delta^2 - c) = c \tilde{\nu} f \mathbb{I}_B + (\delta^2 - c) (\tilde{\nu} f \mathbb{I}_B)^2 \\ &= c \int_B \nu(dt) \sum_{x=1}^s p_x^t (\alpha_x^2 + \beta_x^2) + (\delta^2 - c) \left(\int_B \nu(dt) \sum_{x=1}^s p_x^t \alpha_x \right)^2. \end{aligned}$$

Similarly, we find

$$\text{Cov}(Z_A, Z_{A^c}) = (\delta^2 - c) \left(\int_A \nu(dt) \sum_{x=1}^s p_x^t \alpha_x \right) \left(\int_{A^c} \nu(dt) \sum_{x=1}^s p_x^t \alpha_x \right).$$

Consequently, from (9)

$$\mathbb{E}Z = c \tilde{\nu} f \mathbb{I}_A + c \tilde{\nu} f \mathbb{I}_{A^c} = c \tilde{\nu} f = c \int \nu(dt) \sum_{x=1}^s p_x^t \alpha_x,$$

as well as

$$\begin{aligned} \text{Var} Z &= \text{Var} Z_A + \text{Var} Z_{A^c} + 2 \text{Cov}(Z_A, Z_{A^c}) \\ &= c \int \nu(dt) \sum_{x=1}^s p_x^t (\alpha_x^2 + \beta_x^2) + (\delta^2 - c) \left(\int \nu(dt) \sum_{x=1}^s p_x^t \alpha_x \right)^2. \end{aligned}$$

Note that the last expression is equivalent to $c \tilde{\nu} f^2 + (\delta^2 - c) (\tilde{\nu} f)^2$ as obtained from (2). Note also that the term $\delta^2 - c$ is zero for K Poisson (since then $N_A f$ and $N_{A^c} f$ are independent) but is strictly positive for K negative binomial. Intuitively,

this implies that in this case, the observed variable $N_A f$ carries some information about the unobserved $N_{A^c} f$. This idea appears to be closely related to negative binomial thinning.⁴ Observe that Theorem 3 states that this type of thinning operation cannot be extended to other NNPS distributions.

3.2 | SIR epidemic model

Assume that the independency of individuals (U_i) surveyed for symptoms of infectious (or sexually transmitted) disease forms a random measure $N = (K, \nu \times Q)$ on the space $(E, \mathcal{E})$, where $E = \{(x, y) : 0 < x < y\}$. Each individual $U_i = (X_i, Y_i)$ is described by a pair of infection and recovery times and $K \sim \text{Binom}(n, p)$, where $n \geq 1$ and $p > 0$ (to be specified later). Assume that at time $t > 0$, the collection of labels $L_t(U_i) \in \{S, I, R\}$ for $i = 1, \dots, n$ is observed.

To describe the relevant mean law $\nu \times Q$, consider a standard SIR model describing the evolution of proportions of susceptible (S), infectious (I), and removed (R) units according to the ODE system

$$\begin{aligned}\dot{S}_t &= -\beta I_t S_t \\ \dot{I}_t &= \beta I_t S_t - \gamma I_t \\ \dot{R}_t &= \gamma I_t,\end{aligned}\tag{10}$$

with the initial conditions $S_0 = 1, I_0 = \rho > 0, R_0 = 0$. Define $\mathcal{R}_0 = \beta/\gamma > 1$ and note that

$$S_t = e^{-\mathcal{R}_0 R_t}\tag{11}$$

$$I_t - \rho e^{-\gamma t} = - \int_0^t \dot{S}_u e^{-\gamma(t-u)} du.\tag{12}$$

Interpreting (0) as the *mass transfer model* (see previous study⁷) with initial mass $S_0 = 1$, the function S_t is the probability of an initially susceptible unit remaining uninfected at time $t > 0$. Since $S_t + I_t + R_t = 1 + \rho$ and $I_\infty = 0$, then $S_\infty = 1 - \tau$, where $\tau \in (0, 1)$ is the solution of

$$1 - \tau = e^{-\mathcal{R}_0(\tau + \rho)}.\tag{13}$$

By the law of total probability $S_t = \tau \tilde{S}_t + 1 - \tau$ where $\tilde{S}_t$ is a proper survival function conditioned on the fact that the unit will eventually get infected, an event with probability $\tau < 1$ given by (13). Note that the Lebesgue density function of the proper conditional distribution function $1 - \tilde{S}_t$ is simply

$$\nu(x) = -\dot{\tilde{S}}_x/\tau.\tag{14}$$

Define now $\tau \tilde{I}_t := I_t - \rho e^{-\gamma t}$ and note that from (12) and the last equation in (0), we may interpret $\gamma \tilde{I}_t$ as the Lebesgue density of the (conditional) recovery time t given by the Lebesgue density of the sum of two independent random variables, one of them being exponential with rate γ . Hence, we may define the mean law $\nu \times Q$ by taking (14) along with the transition kernel $Q(x, \cdot)$ in the form of the shifted exponential Lebesgue density

$$Q(x, y) = H_x(y) \sim \text{Exp}(\gamma) \mathbb{I}_{\{x < y\}}(y).$$

To complete the definition of N , take $K \sim \text{Binom}(n, p)$ with $p = \tau$ defined in (13) so that $\mathbb{E}K = n\tau$.

For fixed $t > 0$, let the sets $E_S^t = \{(x, y) : x > t\}$, $E_I^t = \{(x, y) : x \leq t < y\}$ and $E_R^t = \{(x, y) : x < y \leq t\}$ define the t -induced partition of the space E . Define the label on the i th individual observed at time t as

$$L_t(U_i) = \begin{cases} S & \text{if } U_i \in E_S, \\ I & \text{if } U_i \in E_I, \\ R & \text{if } U_i \in E_R. \end{cases}$$

Setting $k = k_S + k_I + k_R$ from (0), we obtain that

$$\begin{aligned}\mathbb{P}(N(E_S^t) = k_S, N(E_I^t) = k_I, N(E_R^t) = k_R) \\ &= \mathbb{P}(N(E_S^t) = k_S, N(E_I^t) = k_I, N(E_R^t) = k_R | K = k) \mathbb{P}(K = k) \\ &= \frac{n!}{k_S! k_I! k_R! (n-k)!} (\tau \tilde{S}_t)^{k_S} (\tau \tilde{I}_t)^{k_I} \tau^{k_R} (1 - \tilde{S}_t - \tilde{I}_t)^{k_R} (1 - \tau)^{n-k}.\end{aligned}\tag{15}$$

Since the overall count of susceptible labels is $k_S + n - k$, marginalizing over the unobserved counts k_S and k gives the final distribution of I, R labels among n individuals at time t

$$\mathbb{P}(N(E_I^t) = k_I, N(E_R^t) = k_R) = \frac{n!}{k_I!k_R!(n - k_I - k_R)!} (\tau \tilde{I}_t)^{k_I} (1 - S_t - \tau \tilde{I}_t)^{k_R} S_t^{n - k_I - k_R}.$$

Hence, it follows in particular that for the i th individual, its label probabilities at t are $\mathbb{P}(L_t(U_i) = S) = S_t$, $\mathbb{P}(L_t(U_i) = I) = \tau \tilde{I}_t$ and $\mathbb{P}(L_t(U_i) = R) = 1 - S_t - \tau \tilde{I}_t$.

Let $A = (0, t]$ and define the conditional infection Lebesgue density by rescaling (14)

$$\nu_A(x) = \nu(x) \mathbb{I}_{\{x < t\}}(x) / \nu(A) = -\dot{S}_x \mathbb{I}_{\{x < t\}}(x) / (1 - S_t).$$

Then, by Theorem 3 and Remark 2, the restricted random measure $N_A = (K_A, \nu_A \times Q)$ is a binomial random measure and according to Remark 3,

$$\mathbb{E}K_A = n(1 - \tilde{S}_t)\tau = n(1 - S_t)$$

$$\text{Var}K_A = n\tau(1 - \tau)(1 - \tilde{S}_t)^2 + n\tau\tilde{S}_t(1 - \tilde{S}_t) = nS_t(1 - S_t),$$

so we see that $K_A \sim \text{Binom}(n, 1 - S_t)$.

3.3 | Spacecraft traffic flows

Consider a particle system of vehicles moving about in $E \subset \mathbb{R}^3$. We are interested in the locations of the vehicles in space and time. We assume the vehicles form an independency, ie, are mutually independent, implied by weak gravitational interaction, and their configuration forms a random counting measure N with number of vehicles $K \sim \kappa_\theta$. Particle system ideas have been applied in air traffic control, for example, in an “interacting” particle system of aircraft for estimating collision probabilities.⁸ We consider the scenario of space traffic control, now in its infancy, by taking E as the Solar System and vehicles as spacecraft (such as satellites, rockets, space planes, space stations, probes, etc), although these ideas may be readily applied to air traffic control, which is in a mature state.

A key issue for space traffic control is modeling the counts of the particle system in various subspaces $\{N \mathbb{I}_A : A \subset E\}$, such as in regions of interest,⁹ eg, space traffic control thinning (restriction) of the particle system into orbital regimes has been considered a topical issue in a recent Presidential Memorandum.¹⁰ Traffic flows can be subject to complex dynamics, with varying degrees of “interactions” among spacecraft (in the sense of correlated counts in time and space).

An obvious extension of the Poisson point process model used in Maria et al⁸ is to use random counting measures closed under thinning with general covariance, ie, PT random measures. We discuss the role of PT random measures in describing the dynamics of the arrivals of spacecraft into subspaces of time and space.

To describe the atomic structure of the particle system, first we label the spacecraft with integers i in $\mathbb{N}_{>0}$. Let X_i be the initial location of spacecraft i in $(E, \mathcal{E})$ and $Y_i = (Y_i(t))_{t \in \mathbb{R}_+}$ be its motion in $(F, \mathcal{F})$. Each Y_i is a stochastic process with state-space $(E, \mathcal{E})$, a path in space and time called a *world line* (also known as a *trajectory* or *orbit*) and regarded as a random element of the function space $(F, \mathcal{F}) = (E, \mathcal{E})^{\mathbb{R}_+}$. The quantity $Y_i(t)$ is the location of spacecraft i in $(E, \mathcal{E})$ at time t , where $Y_i(0) = X_i$ is the initial location. Therefore, each spacecraft i is described by a pair (X_i, Y_i) . Assume X has law ν and the conditional variable $(Y|X = x)$ has transition probability kernel $Q(x, B) = \mathbb{P}(Y \in B|X = x)$ for $B \in \mathcal{F}$. We construct random measures from independencies using STC. Let $K \sim \kappa_\theta$ where κ_θ is PT. The independency $\mathbf{Y} = \{Y_i\}$ forms a PT random measure $M = (\kappa_\theta, \mu)$ on $(F, \mathcal{F})$ through STC as

$$M(A) = \sum_{i=1}^K \mathbb{I}_A(Y_i) \quad \text{for } A \in \mathcal{F},$$

with mean measure $\mu = \nu Q$ defined by

$$\mu(A) = \int_E \nu(dx) Q(x, A) \quad \text{for } A \in \mathcal{F}.$$

Consider the mapping $h : F \mapsto E$ as $h(w) = w(t)$ for $w \in F$ and t fixed. The PT image random measure $N_t = M \circ h^{-1} = (\kappa_\theta, \mu_t)$ on $(E, \mathcal{E})$ is formed by $\mathbf{Y}(t) = \{Y_i(t)\}$ through STC as

$$N_t(A) = \sum_{i=1}^K \mathbb{I}_A(Y_i(t)) \quad \text{for } A \in \mathcal{E}, \quad (16)$$

with mean $\mu_t = \mu \circ h^{-1} = \nu P_t$ defined by

$$\mu_t(A) = \int_E \nu(dx) \int_F Q(x, dw) \mathbb{I}_A(w(t)) = \int_E \nu(dx) P_t(x, A) \quad \text{for } A \in \mathcal{E} \quad (17)$$

and the $\{Y_i(t)\}$ having conditional distributions $\{P_t(X_i, \cdot)\}$ defined by transition kernel

$$P_t(x, A) = \int_F Q(x, dw) \mathbb{I}_A(w(t)) = \mathbb{P}(w \in F : w(t) \in A, w(0) = x) \quad \text{for } A \in \mathcal{E}.$$

The (16) and (17) $N_t = (\kappa_\theta, \mu_t)$ defines an immortal particle system on $(E, \mathcal{E})$. The family of transition kernels $(P_t)_{t \in \mathbb{R}_+}$ is the transition semigroup in the theory of Markov processes.³ Queries about the particle system $f \in \mathcal{E}_+$ form random variables $N_t f$ with mean (2) and variance (3).

The concept of thinning is well established for particle systems, such as in the Bienaymé-Galton-Watson branching process literature¹¹ as well as in the analysis of count time-series using PT thinning operators.¹² Space traffic control thinning (restriction) of the particle system into disjoint subspaces is a key operation. Using PT random measures, let $A \subset E$ with $\mu_t(A) = a > 0$ be a subspace and $N_A = (\kappa_{h_a(\theta)}, \mu_A)$ be the restricted random measure of N_t with $\mu_A(\cdot) = \mu_t(A \cap \cdot) / \mu_t(A)$. Theorem 3 says all such thinnings $\{N_A : A \subset E\}$ are PT. Hence, Theorem 3 is archetypical for space traffic control. Moreover, the PT family members identified in Theorem 3 convey distinct dynamic meanings for the counting process of the particle system, reflected in their covariances. For Poisson, the counts of spacecraft arrivals in disjoint subspaces are independent and Markov and correspond to low-density flows of freely passing spacecraft.¹³ For binomial, the counts in disjoint subspaces are negatively correlated and are identified to following behaviors, platoons, or congestion.¹⁴ For negative binomial, counts in disjoint subspaces are positively correlated and are identified to flows having cycles, control intersections, or contagion.¹⁴ These ideas carry over to the random variables $\{N_t f : f \in \mathcal{E}_+\}$.

Additional frills for the particle system include a notion of birth and death, manufacture and destruction respectively. Death is achieved through a single point extension of the state-space to contain a point ∂ outside of E called a cemetery with measure space $(\bar{E}, \bar{\mathcal{E}})$, where $\bar{E} = E \cup \{\partial\}$ and $\bar{\mathcal{E}} = \mathcal{E} \cup \{A \cup \{\partial\} : A \in \mathcal{E}\}$. The world line space becomes $(F, \mathcal{F}) = (\bar{E}, \bar{\mathcal{E}})^{\mathbb{R}_+}$. Manufacturing is the notion of an arrival time for each spacecraft T_i on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, independent of spacecraft location or motion. $Y_i(t)$ is the location of spacecraft i in $(\bar{E}, \bar{\mathcal{E}})$ at time $T_i + t$, and $Y_i(0) = X_i$ is the (manufacturing) location at time T_i . For Earth or Moon manufacturing, the motion $(Y(t))_{t \geq 0}$ involves moving the manufactured spacecraft to a spaceport, launching, and bringing into orbit. Some spacecraft undergo repeated orbital maneuvers, such as landing at a spaceport, launching, and bringing into orbit, repeating many times.¹⁵ Note that under this setup, the measure P_t is defective on $(E, \mathcal{E})$ as some spacecraft that are manufactured are destroyed with probability $1 - P_t(E)$.

The independency $(\mathbf{T}, \mathbf{X}, \mathbf{Y}) = \{(T_i, X_i, Y_i)\}$ forms the random measure $N = (K, \eta \times \nu \times Q)$ on $\mathbb{R} \times E \times F$ through STC as

$$N(A) = \sum_{i=1}^K \mathbb{I}_A(T_i, X_i, Y_i) \quad \text{for } A \in \mathcal{B}_{\mathbb{R}} \otimes \mathcal{E} \otimes \mathcal{F}.$$

To describe spacecraft manufactured and not yet destroyed, let

$$h(s, x, w) = \begin{cases} (s, x, w(t-s)) & \text{for } s \leq t \\ (s, x, \partial) & \text{for } s > t \end{cases}$$

and put $N \circ h^{-1}$ as the image of N under h . Then spacecraft manufactured and not yet destroyed at time t are represented by the trace of $N \circ h^{-1}$ on E . This is formed by $(\mathbf{T}, \mathbf{X}, \mathbf{Y}(t - \mathbf{T})) = \{(T_i, X_i, Y_i(t - T_i))\}$ through STC as

$$N_t(A) = \sum_{i=1}^K \mathbb{I}_{(-\infty, t] \times E \times A}(T_i, X_i, Y_i(t - T_i)) \quad \text{for } A \in \mathcal{E} \quad (18)$$

with mean μ_t defined by

$$\mu_t(A) = \int_{(-\infty, t]} \eta(ds) \int_E v(dx) \int_F Q(x, dw) \mathbb{I}_A(w(t-s)) \quad \text{for } A \in \mathcal{E}. \quad (19)$$

Other elaborations of the model $N_t = (\kappa_\theta, \mu_t)$ include expanding the state-space of the particle system to provide additional mark spaces, such as radiation detection and crew and passenger health monitoring systems for each spacecraft.

4 | DISCUSSION AND CONCLUSIONS

It is well known that the PT distributions are invariant under thinning.^{12,16} The “if” part of our Theorem 3 gives a different proof of this result in terms of a certain functional equation called the “bone” condition. To the best of our knowledge, the “only if” part of the theorem is novel. Therefore, the main result is the definite one on the existence and uniqueness of PT random measures as random counting measures invariant under thinning.

We characterize PT distributions as those discrete distributions whose generating functions satisfy the “bone” condition. Hence, we can refer to the PT distributions as the “bone” class of distributions. It turns out that there are other characterizations for PT distributions aside from the “bone” condition. For example, the PT distributions arise when considering discrete distributions whose mass functions obey a certain recursive relation and are called the *Panjer* or $(a, b, 0)$ classes of distributions.¹⁷ Yet another (similar) recursive relation involving mass functions recapitulates the PT distributions as the *Katz* family of distributions.¹⁸ Another route to attaining the PT distributions is starting with and generalizing the Poisson distribution to the Conway-Maxwell-Poisson distribution, each PT member being a special or limiting case.¹⁹ These highlights show PT distributions possess rich structure and are independently retrievable using multiple distinct hypotheses.

Given the ubiquity of random count data, PT random measures have wide utility in the sciences. We illustrate this with several examples. First, we give an extension to the compound model applied to modeling the amount of money spent by customers in a store, using compound Poisson and negative binomial random measures. We also give an application to monitoring epidemics, showing that the popular SIR model has the structure of a binomial random measure. Finally, we give an application to closed particle systems, highlighting how the distinct covariances of the PT random measures confer multiple dynamical meanings to the particle system.

ACKNOWLEDGEMENTS

This work was funded by the US National Science Foundation grant DMS 1853587. The authors are indebted to the two reviewers for many valuable comments and suggestions that helped improve the original manuscript. This work does not have any conflicts of interest.

ORCID

Caleb Deen Bastian  <https://orcid.org/0000-0002-2027-7491>

Grzegorz A. Rempala  <https://orcid.org/0000-0002-6307-4555>

REFERENCES

1. Olav K. *Foundations of Modern Probability*. New York: Springer-Verlag; 2002.
2. Olav K. *Random Measures, Theory and Applications*. Cham: Springer; 2017.
3. Erhan C. *Probability and Stochastics*. New York: Springer-Verlag; 2011.
4. Steutel FW, Van Harn K. Discrete analogues of self-decomposability and stability. *Ann Probab*. 1979;7:893-899.
5. Al-Osh MA, Alzaid AA. First-order integer-valued autoregressive (INAR(1)) process. *J Time Ser Anal*. 1987;8(3):261-275.
6. Jacobsen KA, Burch MG, Tien JH, Rempala GA. The large graph limit of a stochastic epidemic model on a dynamic multilayer network. *J Biol Dyn*. 2018;12(1):746-788.
7. KhudaBukhsh WR, Boseung C, Eben K, Rempala GA. Survival Dynamical Systems for the Population-level Analysis of Epidemics. arXiv preprint arXiv:1901.00405; 2019.
8. Maria P, Blom Henk AP, Bakker GJ. (Bert). Air Traffic Complexity and the Interacting Particle System Method: An Integrated Approach for Collision Risk Estimation. In: *Proceedings of the 2011 American Control Conference*; 2011; San Francisco, CA, USA:2154-2159.

9. Stuart E. *Space traffic control*: American Institute of Aeronautics and Astronautics: Reston (Virginia); 2017.
10. Trump DJ. Space policy directive-3 national space traffic management policy. Presidential Memoranda; 2018.
11. Feller W. *An Introduction to Probability Theory and its Applications*: John Wiley & Sons: New York; 1968.
12. Scotto Manuel G, Weiß CH, Sónia G. Thinning-based models in the analysis of integer-valued time series: a review. *Stat Model*. 2015;15(6):590-618.
13. Leo B. The poisson tendency in traffic distribution. *Ann Math Statist*. 1963;34(1):308-311.
14. Daniel G, Flow HM. Traffic theory. Transportation Research Board, National Research Council: Washington DC; 1975.
15. Elon M. Making humans a multi-planetary species. *New Space*. 2017;5(2):46-61.
16. Harremoës P, Johnson O, Kontoyiannis I. Thinning, Entropy, and the Law of Thin Numbers. *IEEE Trans Info Theory*. 2010;56(9):4228-4244.
17. Panjer Harry H. Recursive evaluation of a family of compound distributions. *ASTIN Bulletin: J IAA*. 1981;12(1):22-26.
18. Classical. Distributionsch, Contagious Discrete. *Unified treatment of a broad class of discrete probability distributions*. Oxford: Pergamon Press; 1965:175-182.
19. Conway RW, Maxwell WL. A queuing model with state dependent service rates. *J Indust Eng*. 1962:12.

How to cite this article: Bastian CD, Rempala GA. Throwing stones and collecting bones: Looking for Poisson-like random measures. *Math Meth Appl Sci*. 2020;43:4658-4668. <https://doi.org/10.1002/mma.6224>

APPENDIX A: PROOFS

Proof of Theorem 1. It suffices to verify the claimed identity for the Laplace functional of $M = (K, \nu \times Q)$ with arbitrary $f \in (\mathcal{E} \otimes \mathcal{F})_+$ as it will in particular imply the existence of M . To this end consider

$$\mathbb{E}e^{-Mf} = \mathbb{E}(\mathbb{E}e^{-f(X,Y)})^K = \psi((\nu \times Q)e^{-f}),$$

where $\psi(\cdot)$ is pgf of K . Since

$$(\nu \times Q)e^{-f} = \mathbb{E} \int_F Q(X, dy) e^{-f} \circ (X, y) = \nu e^{-g},$$

where $g \in \mathcal{E}_+$ is defined by

$$e^{-g(x)} = \int_F Q(x, dy) e^{-f(x,y)},$$

therefore,

$$\mathbb{E}e^{-Mf} = \psi((\nu \times Q)e^{-f}) = \psi(\nu e^{-g}) = \mathbb{E}e^{-Ng} < \infty. \quad \square$$

Proof of Theorem 2. The result follows from the following lemma. $\square$

Lemma 1 (Modified Cauchy equation). Assume that $f(t)$ is twice continuously differentiable in some neighborhood of the origin, satisfies $f(0) = 0$ and $f'(0) > 0$ as well as

$$f(s+t) - f(s) = f(h(s)t), \quad (\text{A1})$$

where $h(s)$ is t free. Then f is of the form $f(t) = At$ or $f(t) = B \log(1+At)$ for some $A, B \neq 0$. Moreover, $h(s) = f'(s)/f'(0)$.

Proof. Differentiating (A1) with respect to t we obtain

$$f'(s+t) = h(s)f'(h(s)t). \quad (\text{A2})$$

Taking the above at $t = 0$ and denoting $C_1 = f'(0) > 0$ gives

$$h(s) = f'(s)/C_1. \quad (\text{A3})$$

Differentiating (A1) with respect to s yields likewise (note that h is differentiable in view of (A3))

$$f'(s+t) = f'(s) + t h'(s) f'(h(s) t).$$

Equating the two right hand side expressions and using (A3) we have

$$C_1 h(s) + t h'(s) f'(h(s) t) = h(s) f'(h(s) t)$$

$$f'(h(s) t) = \frac{C_1}{1 - t \frac{h'(s)}{h(s)}}.$$

In the last expression, we take now $s = 0$, denote $C_2 = h'(0)$ and consider two cases according to $C_2 = 0$ and $C_2 \neq 0$. Since by (A3) $h(0) = 1$, for the case $C_2 = 0$

$$f(t) = At \tag{A4}$$

where ($A = C_1$) and we have one solution. Consider now $C_2 \neq 0$, then

$$f'(t) = \frac{C_1}{1 - C_2 t}$$

and hence the general form of f when it is not linear is

$$f(t) = B \log(1 + At)$$

where $B = -C_1/C_2$ and $A = -C_2$. This as well as (A4) and (A3) give the hypothesis of the theorem. □