Appendix J

Completely Random Measures

Completely random measures are measures whose values on disjoint sets are independent random variables. They arise as priors, or building blocks for priors, in many Bayesian nonparametric applications. In this appendix we first discuss the special case of measures that assume only (nonnegative) integer values and are Poisson distributed. Next we generalize to general completely random measures. Finally we specialize to the case that the sample space is the positive half line, when the measures can be described as processes with independent increments through their cumulative distribution functions.

Throughout the chapter $(\Omega, \mathscr{A}, P)$ is a probability space, rich enough to support countably many independent nontrivial random variables, and $(\mathfrak{X}, \mathscr{X})$ is a Polish space with its Borel σ-field. Furthermore $\mathfrak{M}_\infty = \mathfrak{M}_\infty(\mathfrak{X}, \mathscr{X})$ is the set of all measures on $(\mathfrak{X}, \mathscr{X})$, and $\mathscr{M}_\infty$ is the σ-field on $\mathfrak{M}_\infty$ generated by the evaluation maps $\mu \mapsto \mu(A)$, for $A \in \mathscr{X}$.

J.1 Poisson Random Measures

A "Poisson random subset" is a model for randomly spreading out points in space. It can be viewed as a random subset, a stochastic process, or a random measure with values in the integers, depending on the aspect one focuses on.

Definition J.1 (Poisson random subset) A *Poisson random subset* (*PRS*) of $\mathfrak{X}$ is a map Π from Ω into the collection of subsets of $\mathfrak{X}$ of at most countably many elements such that $N(A) := \text{card}(\Pi \cap A)$ is a random variable for every $A \in \mathscr{X}$ and $N(A_i) \overset{\text{ind}}{\sim} \text{Poi}(\mu(A_i))$, for every finite collection of disjoint sets $A_1, \ldots, A_k \in \mathscr{X}$[1] and a measure μ on $(\mathfrak{X}, \mathscr{X})$, called the *intensity measure*. The stochastic process $N = (N(A): A \in \mathscr{X})$ is called a *Poisson process* on $\mathfrak{X}$ with intensity measure μ. The corresponding counting measure on the points Π is called a *Poisson random measure*.

It is included in the definition that the total number of points $N(\mathfrak{X})$ is also a Poisson variable and hence it is either infinite or finite almost surely. That μ is a measure is actually implied by the other requirements (see Problem J.2). It is the mean measure of N, and also called the *compensator* of the Poisson process. Because $N(\{x\}) \leq 1$ by construction and a nontrivial Poisson variable has support $\mathbb{N}$, it follows that $\mu(\{x\}) = 0$ for every x, i.e. μ is atomless. In the case that $\mathfrak{X} = \mathbb{R}^d$ and μ is absolutely continuous with respect to the

[1] By convention, Poi(0) and Poi(∞) are the measures that are degenerate at 0 and ∞, respectively.

Lebesgue measure, the density function of μ is called the *intensity function* (and is often denoted by λ).

By the additivity of N, the joint distribution of $(N(A_1), \ldots, N(A_k))$ for sets $A_1, \ldots, A_k$ that are not necessarily disjoint can be computed from the distribution of the vector of counts $(N(B_1), \ldots, N(B_n))$, in the collection $\{B_1, \ldots, B_n\}$ of all intersections of sets from $\{A_1, \ldots, A_k\}$. In particular, the distribution of a PRS is completely determined by its intensity measure. In view of the bilinearity of the covariance it also follows that $\operatorname{cov}(N(A_1), N(A_2)) = \mu(A_1 \cap A_2)$, for every pair of measurable sets.

The following properties of a PRS will be important. Their proofs, which are not too difficult, can be found in Kingman (1993).

(i) *Disjointness:* If Π_1, Π_2 are independent PRS on $\mathfrak{X}$ with intensity measures μ_1 and μ_2 respectively, then $\Pi_1 \cap \Pi_2 \cap A = \varnothing$ with probability one, for any $A \in \mathscr{A}$ with $\mu_1(A), \mu_2(A) < \infty$.

(ii) *Superposition:* If $\Pi_1, \Pi_2, \ldots$ are independent PRS on $\mathfrak{X}$ with finite intensity measures $\mu_1, \mu_2, \ldots$ respectively, then $\Pi = \cup_{i=1}^{\infty} \Pi_i$ is a PRS on $\mathfrak{X}$ with intensity measure $\mu = \sum_{i=1}^{\infty} \mu_i$.

(iii) *Restriction:* If $\mathfrak{X}' \in \mathscr{X}$ and Π is a PRS on $\mathfrak{X}$, then $\Pi' = \Pi \cap \mathfrak{X}'$ is a PRS on $\mathfrak{X}'$ with intensity measure μ' given by e $\mu'(A) = \mu(A \cap \mathfrak{X}')$.

(iv) *Transformation:* If Π is a PRS on $\mathfrak{X}$ and $\psi: \mathfrak{X} \to \mathfrak{Y}$ is a measurable map into another measurable space $(\mathfrak{Y}, \mathscr{Y})$ such that the induced measure $\mu\psi^{-1}$ is atomless, then $\psi(\Pi)$ is a PRS on $\mathfrak{Y}$ with intensity measure $\mu\psi^{-1}$.

(v) *Conditioning:* If Π is a PRS on $\mathfrak{X}$ and $\mathfrak{X}' \in \mathscr{X}$ satisfies $0 < \mu(\mathfrak{X}') < \infty$, then

$$(N(A_1 \cap \mathfrak{X}'), \ldots, N(A_k \cap \mathfrak{X}') | N(\mathfrak{X}') = n) \sim \mathrm{MN}_k(n; \bar{\mu}(A_1), \ldots, \bar{\mu}(A_k)),$$

for any measurable partition $A_1, \ldots, A_k$ of $\mathfrak{X}$, where $\bar{\mu}(A) = \mu(A \cap \mathfrak{X}')/\mu(\mathfrak{X}')$.

(vi) *Thinning:* If Π is a PRS on $\mathfrak{X}$ with σ-finite intensity measure μ and $f: \mathfrak{X} \to [0, 1]$ is a measurable function, then the random set of points obtained by deleting every point x of Π independently with probability $f(x)$, is a PRS with intensity measures $A \mapsto \int_A f \, d\mu$.

The process $\Pi' = \Pi \cap \mathfrak{X}'$ with counting measure $N'(A) = \operatorname{card}(\Pi \cap \mathfrak{X}' \cap A)$, as in (v), conditioned on the total number of points in $\mathfrak{X}'$ being n is a suitable model for spreading out a fixed finite number of points randomly in a space. It will be called a *multinomial random set* with parameters n and $\bar{\mu}$. The special case corresponding to $n = 1$ is equivalent to a single random variable on $\mathfrak{X}$ with distribution $\bar{\mu}$.

The existence of a PRS can be proved constructively by using these properties in a converse direction. Let μ be a measure on $(\mathfrak{X}, \mathscr{X})$ such that there exist finite measures μ_n, $n \in \mathbb{N}$, with $\mu = \sum_{n=1}^{\infty} \mu_n$. (This certainly includes every σ-finite measure μ.) Let $\bar{\mu}_n = \mu_n/\mu_n(\mathfrak{X})$ be the probability measures obtained by normalizing μ_n, and on $(\Omega, \mathscr{A}, P)$ construct mutually independent random variables $N_n \sim \mathrm{Poi}(\mu_n(\mathfrak{X}))$ and $X_{n,r} \overset{\mathrm{iid}}{\sim} \bar{\mu}_n$, for $r, n \in \mathbb{N}$. Then define $\Pi_n = \{X_{n,1}, \ldots, X_{n,N_n}\}$, for $n \in \mathbb{N}$, and

$$\Pi = \bigcup_{n=1}^{\infty} \Pi_n = \bigcup_{n=1}^{\infty} \{X_{n,1}, \ldots, X_{n,N_n}\}.$$

Clearly, $\Pi_n | \, N_n = m$ is a multinomial random set with parameters m and $\bar{\mu}_n$, whence Π is a PRS on $\mathfrak{X}$ with intensity measure μ, by the properties (ii) and the converse of (v).

Interpreting N as a (random) counting measure that puts weight 1 at every point in Π, we can write $\int f \, dN = \sum_{x \in \Pi} f(x)$, for a given measurable function $f: \mathfrak{X} \to \mathbb{R}$. The distribution of this random variable can be easily characterized by its Laplace transform. As a function of f (and with $\theta = 1$) the left side of the following lemma is called the *Laplace functional* of N.

Lemma J.2 *If N is a Poisson random measure with intensity measure μ and $f: \mathfrak{X} \to \mathbb{R}$ is nonnegative or $\int f \, dN$ is finite almost surely, then, for $\theta > 0$,*

$$\mathrm{E}\Big[e^{-\theta \int f \, dN}\Big] = \exp\Big[-\int (1 - e^{-\theta f(x)}) \, d\mu(x)\Big].$$

Proof For an indicator function $f(x) = \mathbb{1}_A(x)$ the formula follows from the formula for the Laplace transform of the Poisson variable $N(A)$. It extends to linear combinations of indicators of disjoint sets, by the independence of the corresponding Poisson variables. A nonnegative function f can be approximated from below by such simple functions; the formula is preserved under such approximation by the monotone convergence theorem. If $\int f \, dN < \infty$ almost surely, then it is the difference of the variables $\int f^+ \, dN$ and $\int f^- \, dN$, which are independent, because f^+ and f^- have disjoint supports. Both left and right sides factorize over this decomposition. □

J.1.1 Palm Theory

The *Campbell measure* of a Poisson process N is the measure on $(\mathfrak{X} \times \mathfrak{M}_\infty, \mathscr{X} \otimes \mathscr{M}_\infty)$ determined by $C(A \times B) = \mathrm{E}\big[N(A) \mathbb{1}\{N \in B\}\big]$, for $A \in \mathscr{X}$ and $B \in \mathscr{M}_\infty$. Here the expectation is relative to N, which we view as a random element in $\mathfrak{M}_\infty$ by identifying it with the counting measure on its point set. By monotone approximation, for every nonnegative (or suitably integrable) measurable function $g: \mathfrak{X} \times \mathfrak{M}_\infty \to \mathbb{R}$,

$$\int_{\mathfrak{X} \times \mathfrak{M}_\infty} g(x, \xi) \, C(dx, d\xi) = \mathrm{E} \int g(x, N) \, N(dx) = \int_{\mathfrak{M}_\infty} \int_{\mathfrak{X}} g(x, N) \, N(dx) \, \mathcal{P}(dN),$$

where $\mathcal{P}$ is the law of N on $(\mathfrak{M}_\infty, \mathscr{M}_\infty)$ (and we abuse notation by employing the symbol N also as a dummy variable for integration). This shows that the measure $N(dx) \, \mathcal{P}(dN)$ is a disintegration of C. Since $C(A \times \mathfrak{M}_\infty) = \mathrm{E}\big[N(A)\big]$, the marginal distribution of C on its first coordinate is the intensity measure μ of N. Thus the disintegration of C with its two coordinates swapped takes the form

$$\mathrm{E} \int g(x, N) \, N(dx) = \int_{\mathfrak{X}} \int_{\mathfrak{M}_\infty} g(x, N) \, \mathcal{P}_x(dN) \, \mu(dx).$$

The measure $\mathcal{P}_x$ in the right side is known as the *Palm measure* of N at x. By definition of *disintegration* every $\mathcal{P}_x$ is a measure on $(\mathfrak{M}_\infty, \mathscr{M}_\infty)$ and the map $x \mapsto \mathcal{P}_x(B)$ is measurable for every $B \in \mathscr{M}_\infty$.

For a more detailed treatment of these concepts and the (not so difficult) proof of the following proposition, see Chapter 13 of Daley and Vere-Jones (2008).

Proposition J.3 (Palm measure) *The Palm measure at x of the Poisson random measure N is the counting measure on the points $N \cup \{x\}$.*

The proposition is equivalent to the assertion that, for every nonnegative measurable function $g: \mathfrak{X} \times \mathfrak{M}_\infty \to \mathbb{R}$,

$$\mathrm{E} \int g(x, N)\, N(dx) = \int \mathrm{E} g(x, N \cup \{x\})\, \mu(dx). \tag{J.1}$$

(Both E-signs mean expectation relative to N.) The formula is often read as saying that "a Poisson process conditioned to have a point at x (referring to $\mu(dx)$ at the far right) is distributed as before but with the point at x added." In the simpler case that the function g is free of N the formula reduces to *Campbell's theorem*. Various extensions (to other random measures or processes, higher-order versions) carry additional names, including Mecke, Hardy and Palm.

One might interpret the definition of the Campbell measure C as a recipe for generating a pair (x, N) in $\mathfrak{X} \times \mathfrak{M}_\infty$ by first generating N and next given N generating x from N viewed as a measure on $\mathfrak{X}$. Because C will typically be infinite, the notion of "generating" does not carry immediate meaning. Ignoring this difficulty, we might interpret the Palm measures as a recipe for generating the same point (x, N) by first generating x from μ and next N from $\mathcal{P}_x$. It is possible to introduce and interpret Palm measures for general random measures in this way. Proposition J.3 shows that for a Poisson random measure the second step comes down to generating a realization of the original process N and adding the point x to it.

In a similar spirit, but thinking of the law of N as a prior on $\mathfrak{M}_\infty$ and x as an observation from N, we can also think of the Palm measure as the posterior distribution given the data. Consideration of multiple data points could be implemented through an extension to multivariate Campbell measures of the form $C_n(A_1 \times \cdots \times A_n \times B) = \mathrm{E}\big[\prod_{i=1}^n N(A_i) \mathbb{1}\{N \in B\}\big]$.

One application of Proposition J.3 is to the distribution of a Poisson process when a point is removed. Let $w: \mathfrak{X} \to \mathbb{R}^+$ be a given weight function such that $\int w(x)\, N(dx) < \infty$ almost surely, and suppose that given N a point Y_1 from N is chosen with probability proportional to $w(Y_1)$, i.e. for every measurable set A,

$$\mathrm{P}(Y_1 \in A \,|\, N) = \frac{\int_A w(x)\, N(dx)}{\int w(x)\, dN(x)}.$$

Let $N_1 = N - \{Y_1\}$ be the process N with the chosen point removed, and let $T = \int w(x)\, N(dx)$ and $T_1 = \int w(x)\, N_1(dx)$ be the total weights of the points in N and N_1, respectively (so that $T = T_1 + w(Y_1)$).

Lemma J.4 *If N is a Poisson process with intensity measure μ, then for $\mathcal{P}$ the distribution of N on $(\mathfrak{M}_\infty, \mathscr{M}_\infty)$,*

(i) *(Y_1, N_1) has density $(y, \xi) \mapsto w(y)/(\int w(x)\, \xi(dx) + w(y))$ relative to the measure $\mu \times \mathcal{P}$;*
(ii) *T_1 has density $t \mapsto \int w(x)/(t + w(x))\, d\mu(x)$ relative to the distribution of T;*
(iii) *Y_1 and N_1 are conditionally independent given T_1;*
(iv) *$N_1 \,|\, T_1 = t \sim N \,|\, T = t$, for almost every t, $[P^T]$.*

Proof (i). The definition of Y_1 implies that $\mathrm{E}\big[f(Y_1)|\,N\big] = \int f(x)w(x)\,dN(dx)/T$, for every bounded, measurable function $f\colon \mathfrak{X} \to \mathbb{R}$. We apply this with the function $y \mapsto f(y)g(N-\{y\})$, for given N and given bounded, measurable functions $f\colon \mathfrak{X} \to \mathbb{R}$ and $g\colon \mathfrak{M}_\infty \to \mathbb{R}$, to find that

$$\mathrm{E}\big[f(Y_1)g(N_1)\big] = \mathrm{E}\,\mathrm{E}\big[f(Y_1)g(N-\{Y_1\})|\,N\big] = \mathrm{E}\Big[\frac{\int f(x)g(N-\{x\})w(x)\,N(dx)}{\int w(y)\,dN(y)}\Big],$$

where the expectation on the right is with respect to N. By the definition of the Palm measures, the right side can also be written as

$$\iint \frac{f(x)g(\xi-\{x\})w(x)}{\int w(y)\,\xi(dy)}\,\mathcal{P}_x(d\xi)\,\mu(dx) = \iint \frac{f(x)g(N)w(x)}{\int w(y)\,N(dy)+w(x)}\,\mathcal{P}(dN)\,\mu(dx).$$

The last step follows because the Palm measure $\mathcal{P}_x$ is the distribution of $N\cup\{x\}$, by Proposition J.3. Assertion (i) follows, since it has been established that the left side of the preceding display is equal to the right side of the last display for every bounded, measurable functions f and g.

(ii), (iii) and (iv). All three assertions are consequences of the fact that the density of (Y_1, N_1) in (i) relative to the product dominating measure $\mathcal{P}\times\mu$ depends on (Y_1, T_1) only. For a precise proof, let (Y', N') be distributed according to $\mu\times\mathcal{P}$ and let $T' = \int w(x)\,N'(dx)$. Then, by (i), for every bounded, measurable functions g and h,

$$\begin{aligned}\mathrm{E}\big[g(N_1)h(Y_1,T_1)\big] &= \mathrm{E}\Big[g(N')h(Y',T')\frac{w(Y')}{T'+w(Y')}\Big]\\ &= \mathrm{E}\Big[\mathrm{E}(g(N')|\,Y',T')h(Y',T')\frac{w(Y')}{T'+w(Y')}\Big].\end{aligned}$$

Taking g equal to 1, we can conclude from the equation that (Y_1, T_1) has density $(y,t)\mapsto w(y)/(t+w(y))$ relative to the distribution of (Y', T'). Since the law of (Y', T') is a product distribution, the marginal distribution of T_1 has density relative to the distribution of T' equal to the integral of the joint density of (Y_1, T_1) relative to the density of Y'. Since $T' =_d T \sim \mu$, this gives (ii). If we set $k(Y', T') = \mathrm{E}(g(N')|\,Y', T')$, then the far right side of the display is equal to the expectation of $k(Y', T')h(Y', T')$ times the density of (Y_1, T_1) relative to (Y', T'), whence it is equal to $\mathrm{E}[k(Y_1, T_1)h(Y_1, T_1)]$. Since this is true for every bounded, measurable h, it follows that $k(Y_1, T_1) = \mathrm{E}[g(N_1)|\,Y_1, T_1]$. By the independence of Y' and N', the function k actually depends only on its second argument, and hence $\mathrm{E}[g(N_1)|\,Y_1, T_1]$ is free of Y_1, proving (iii). Moreover, taken at a value $T_1 = t$ the conditional expectation is equal to $k(Y_1, t) = \mathrm{E}[g(N')|\,T' = t]$, which is identical to $\mathrm{E}[g(N)|\,T = t]$, since $N =_d N'$. This proves (iv). □

J.2 Completely Random Measures

A Poisson process N takes its values in the integers $\{0, 1, \ldots, \infty\}$, a restriction that is too rigid for a general prior. This constraint is released in the following definition of a general completely random measure.

Definition J.5 (Completely random measure) A measurable map $\Phi\colon \Omega \to (\mathfrak{M}_\infty, \mathscr{M}_\infty)$ is a *completely random measure* (*CRM*) on $\mathfrak{X}$ if the random variables $\Phi(A_1), \ldots, \Phi(A_k)$ are mutually independent, for any disjoint sets $A_1, \ldots, A_k \in \mathscr{X}$.

As in the case of a Poisson process the joint distribution of $(\Phi(A_1), \ldots, \Phi(A_k))$ for arbitrary measurable subsets $A_1, \ldots, A_k$ of $\mathfrak{X}$ can be obtained through additivity of Φ from its distribution over the partition generated by these sets. By the assumed independence this is fixed by the distributions of the univariate random variables $\Phi(A)$. Thus the distribution of a completely random measure Φ on $\mathfrak{X}$ is completely determined by the marginal distributions of the variables $\Phi(A)$, for $A \in \mathscr{X}$.

The Laplace transform $\theta \mapsto \mathrm{E}e^{-\theta\Phi(A)}$ of such a variable is defined and finite (at least) for $\theta > 0$. The *cumulant measures* of Φ are defined as minus the logarithms of these Laplace transforms:

$$\lambda_\theta(A) = -\log \mathrm{E}(e^{-\theta\Phi(A)}).$$

From the countable additivity of Φ and independence over disjoint sets, it can be checked that the cumulant measures are indeed measures on $(\mathfrak{X}, \mathscr{X})$. By uniqueness of Laplace transforms (Feller 1971, Theorem XIII.1.1) they collectively, for $\theta > 0$, determine the distribution of any $\Phi(A)$, and hence the full distribution of Φ. It can also be seen that $\lambda_\theta(A) = 0$ if and only if $\Phi(A) = 0$ a.s.; and $\lambda_\theta(A) = \infty$ if and only if $\Phi(A) = \infty$ a.s. In particular $\lambda_\theta(A)$ is zero or nonzero finite or infinite for some $\theta > 0$ if and only if the same is true for *every* $\theta > 0$.

We shall assume that λ_θ is σ-finite for some (and hence by the preceding observation all) $\theta > 0$. Then it has at most countably many atoms and the sample space can be partitioned as $\mathfrak{X} = \{a_1, a_2, \ldots\} \cup \cup_j \mathfrak{X}_j$, where $a_1, a_2, \ldots$ are the atoms and the $\mathfrak{X}_j$ are measurable sets with $0 < \lambda_\theta(\mathfrak{X}_j) < \infty$. The set of atoms is the same for every θ, and the sets $\mathfrak{X}_j$ can be chosen independent of θ as well. By the definition of a CRM the random measures in the corresponding decomposition

$$\Phi(\cdot) = \sum_j \Phi(\{a_j\})\delta_{a_j}(\cdot) + \sum_j \Phi(\cdot \cap \mathfrak{X}_j) \tag{J.2}$$

are jointly independent. Because $\Phi(\{a\})$ is a non-degenerate variable if and only if a is an atom of λ_θ (and otherwise is identically zero), $a_1, a_2, \ldots$ are called *fixed atoms* of Φ. Every of the processes $\Phi(\cdot \cap \mathfrak{X}_j)$ is a CRM without fixed atoms that is neither identically zero nor identically infinite; its λ_θ-measure is the restriction of λ_θ to $\mathfrak{X}_j$. Thus to construct a CRM we may construct countably many independent CRMs with disjoint, finite and atomless λ_θ-measures and countably many independent variables $\Phi(\{a_j\})$, and add them together as in (J.2).

It turns out that apart from a deterministic component an arbitrary CRM can be described as a purely atomic measure with atoms of random weights at random locations in $\mathfrak{X}$. Furthermore, the "non-fixed atoms" are placed at the points of a PRS, and also the pairs (x, s) of a location x and a weight s are given by a PRS, on the product space $\mathfrak{X} \times (0, \infty]$. That the pairs (x, s) are spread according to a PRS means that they are "spatially independent" in the product space, but location x and weight s are not independent in general.

The representation is in "in law" (or "weak") in that the proposition shows how a CRM with a given distribution can be constructed on a suitable probability space.

Proposition J.6 *Any CRM Φ with σ-finite cumulant measures λ_θ can be represented uniquely as $\Phi = \sum_j \Phi(\{a_j\})\delta_{a_j} + \beta + \Psi$, for fixed points $a_1, a_2, \ldots$ in $\mathfrak{X}$, independent nonnegative random variables $\Phi(\{a_1\}), \Phi(\{a_2\}), \ldots$, a deterministic σ-finite Borel measure β on $\mathfrak{X}$, and a CRM Ψ of the form*

$$\Psi(A) = \sum_{(x,s)\in\Pi^c, x\in A} s \quad = \iint \mathbb{1}_A(x) s \, N^c(dx, ds), \tag{J.3}$$

for Π^c and N^c a PRS on $\mathfrak{X} \times (0, \infty]$, independent of $\Phi(\{a_1\}), \Phi(\{a_2\}), \ldots$, with intensity measure ν^c such that $\nu^c(\{x\} \times (0, \infty]) = 0$ for every $x \in \mathfrak{X}$, and $\int_{\mathfrak{X}_j} \int (s \wedge 1)\, \nu^c(dx, ds) < \infty$, for every set $\mathfrak{X}_j$ in a countable measurable partition of $\mathfrak{X}$. The cumulant measures of Ψ take the form

$$-\log \mathrm{E}(e^{-\theta\Psi(A)}) = \int_A \int_{(0,\infty]} (1 - e^{-\theta s})\, \nu^c(dx, ds). \tag{J.4}$$

Conversely, every $a_1, a_2, \ldots$, $\Phi(\{a_1\}), \Phi(\{a_2\}), \ldots$, β and N^c and ν^c with the given properties define a CRM with σ-finite cumulant measures λ_θ.

Proof As shown in (J.2) the fixed atoms of the CRM can be separated off and the CRM can be decomposed over subsets $\mathfrak{X}_\mathrm{j}$, of finite λ_θ measure. Thus it is not a loss of generality to assume that the λ_θ are nontrivial, finite, atomless measures. In this case the distribution of $\Phi(A)$ (on $[0, \infty]$) is necessarily *infinitely divisible*: for every $n \in \mathbb{N}$ it can be written as the convolution of the distributions of n independent variables that are asymptotically negligible as $n \to \infty$. Indeed, $\Phi(A) = \sum_{j=1}^n \Phi(A_{n,j})$ for any measurable partition $A = \cup_{j=1}^n A_{n,j}$; this can be chosen so that $\lambda_1(A_{n,j}) = \lambda_1(A)/n$ if λ_1 is atomless, in which case, for any $\epsilon > 0$ as $n \to \infty$,

$$\max_{1\le j\le n} \mathrm{P}(\Phi(A_{n,j}) \ge \epsilon) = \max_{1\le j\le n} \mathrm{P}(1 - e^{-\Phi(A_{n,j})} \ge 1 - e^{-\epsilon}) \le \frac{1 - e^{-\lambda_1(A)/n}}{1 - e^{-\epsilon}} \to 0.$$

Since the variable $\Phi(A)$ is infinitely divisible, its negative log-Laplace transform admits the *Lévy-Khinchine representation* (Kallenberg 1986, Corollary 15.8, or Feller 1971, Theorem XIII.7.2[2]), for all $\theta > 0$,

$$\lambda_\theta(A) = -\log \mathrm{E}(e^{-\theta\Phi(A)}) = \theta\beta(A) + \int_{(0,\infty]} (1 - e^{-\theta s})\, \gamma(A, ds). \tag{J.5}$$

Here $\beta(A)$ is a nonnegative constant, the *deterministic component* of $\Phi(A)$, and $\gamma(A, \cdot)$ is a measure on $(0, \infty]$ with the property $\int_{(0,\infty]} (s \wedge 1)\, \gamma(A, ds) < \infty$, the *Lévy measure* of $\Phi(A)$. The characteristics $\beta(A)$ and $\gamma(A, \cdot)$ are uniquely determined by the distribution of $\Phi(A)$. From the σ-additivity of λ_θ it follows that $\sum_i \beta(A_i)$ and $\sum_i \gamma(A_i, \cdot)$ can be

[2] Unlike in these references, the present $\Phi(A)$ can be infinite with positive probability. This is accommodated by allowing $\gamma(A, \cdot)$ to have an atom at ∞: $\mathrm{P}\big(\Phi(A) < \infty\big) = e^{-\gamma(A,\{\infty\})}$, as follows by letting $\theta \downarrow 0$ in the formula. The references incorporate the term $\theta\beta(A)$ in the integral by allowing an atom at 0 in the (redefined) Lévy measure.

used as the number and measure in the right side of (J.5) to represent $\lambda_\theta(\cup_i A_i)$, while the number 0 and the zero measure represent $\lambda_\theta(\varnothing)$ in its left side. From the uniqueness of the representation we conclude that the maps $A \mapsto \beta(A)$ and $A \mapsto \gamma(A, B)$, for a fixed measurable $B \subset (0, \infty]$, are atomless measures.

First assume that the measure μ defined by $\mu(A) = \gamma(A, (0, \infty])$ is σ-finite. Then the measures μ_s defined by $\mu_s(A) = \gamma(A, (0, s])$ are right-continuous and increasing in $s > 0$ and $\mu_s \ll \mu$. By the Radon-Nikodym theorem there exists for every s a measurable function $F(\cdot, s)\colon \mathfrak{X} \to [0, \infty)$ such that $\mu_s(A) = \int_A F(x, s)\, d\mu(x)$ for every A. By the monotonicity in s these functions automatically satisfy $F(x, s) \le F(x, t)$ for μ-almost every x whenever $s \le t$. This and right continuity can be used to show that the functions $\inf_{s\in\mathbb{Q}, s>t} F(x, s)$ also represent the measures μ_t. As the latter functions are right-continuous and monotone in t for *every* x, it is not a loss of generality to assume that the original function F possesses this property. Then $F(x, \cdot)$ is a Lebesgue-Stieltjes distribution function, and we can define a measure ν^c by $d\nu^c(x, s) = F(x, ds)\, d\mu(x)$. By construction this satisfies $\nu^c(A \times B) = \gamma(A, B)$.

If μ is not σ-finite, define measures $\bar\mu_k(A) = \gamma(A, (1/(k+1), 1/k])$, for $k \in \mathbb{N}$. Since $\gamma(A, (\epsilon, \infty]) < \infty$ for all A with $\lambda_1(A) < \infty$ and $\epsilon > 0$, by the condition $\int_{(0,\infty]}(s \wedge 1)\, \gamma(A, ds) < \infty$, and λ_1 is finite by assumption, it follows that $\bar\mu_k$ is σ-finite for every k. Now define $\nu^c = \sum_k \nu_k$ for the measures ν_k defined by $d\nu_k(x, s) = F_k(x, ds)\, d\mu_k(x)$, where F_k is constructed from $\bar\mu_k$ as previously F from μ. Again $\nu(A \times B) = \gamma(A, B)$, and hence $\int_A \int (s \wedge 1)\, \nu^c(dx, ds) = \int (s \wedge 1)\, \gamma(A, ds)$, which is finite, also for $A = \mathfrak{X}$.

For disjoint sets $A_1, \ldots, A_k$ the variables $\Psi(A_1), \ldots, \Psi(A_k)$ obtained from the process (J.3) are mutually independent, as they are functions of the independent Poisson processes $\{N^*(A_i \times (0, s])\colon s > 0\}$, for $i = 1, \ldots, k$. Adding the deterministic component β does not change this independence. Therefore $\Phi = \beta + \Psi$ is a CRM with characteristics β and γ as soon as (J.5) holds. To verify this we first validate formula (J.4) by applying Lemma J.2 to the Poisson process N^c on $\mathfrak{X} \times [0, \infty)$ and the function $f(x, s) = \mathbb{1}_A(x)s$. Next we note that $\int_A \int_0^\infty f(s)\, \nu^c(dx, ds) = \int_0^\infty f(s)\, \gamma(A, ds)$ for the functions $f(s) = 1 - e^{-\theta s}$, as $\nu^c(A \times B) = \gamma(A, B)$ by construction.

Taking $A = \{x\}$ in (J.4) gives that $\int_{\{x\}} \int (1 - e^{-\theta s})\, \nu^c(dx, ds) = \lambda_\theta(\{x\}) = 0$ and hence $\nu^c(\{x\} \times (0, \infty]) = 0$. □

The measure ν^c in the proposition, the intensity measure of the Poisson process N^c associated with the CRM Φ, is by an abuse of terminology also referred to as the *intensity measure* of the CRM Φ.

Combination of the Poisson representation and Lemma J.2 gives for the integral $\int f\, d\Psi$ of a nonnegative measurable function $f\colon \mathfrak{X} \to \mathbb{R}$ with respect to the CRM Ψ in (J.3) (or if f is integrable with respect to Ψ a.s.) as

$$\log \mathrm{E}(e^{-\theta \int f\, d\Psi}) = -\iint (1 - e^{-\theta s f(x)})\, \nu^c(dx, ds). \tag{J.6}$$

The exponent of the left side viewed as a function of f (and with $\theta = 1$) is called the *Laplace functional* of Ψ. Expanding the exponential in the right side and integrating term by term gives the cumulant generating function of $\int f\, d\Psi$, for Ψ as in (J.3), as

$$\log \mathrm{E}(e^{-\theta \int f\, d\Psi}) = \sum_{j=1}^{\infty} \frac{(-\theta)^j}{j!} \iint_{\mathfrak{X}\times\mathbb{R}^+} s^j f^j(x)\, \nu^c(dx, ds). \tag{J.7}$$

The jth cumulant of $\int f\, d\Psi$ is the coefficient of $(-\theta)^j/j!$ in the right side (provided the integrals exist). In particular,

$$\mathrm{E}\Big(\int f\, d\Psi\Big) = \iint_{\mathfrak{X}\times\mathbb{R}^+} sf(x)\, \nu^c(dx, ds), \tag{J.8}$$

$$\mathrm{var}\Big(\int f\, d\Psi\Big) = \iint_{\mathfrak{X}\times\mathbb{R}^+} s^2 f^2(x)\, \nu^c(dx, ds). \tag{J.9}$$

These formulas are valid for CRMs without fixed atoms and without deterministic component. Adding a fixed component β shifts the mean by $\int f\, d\beta$, but does not affect the higher cumulants. Adding a fixed atom at a simply adds the cumulant of the variable $f(a)\Phi(\{a\})$.

The preceding construction exhibits the CRM without fixed atoms as a purely atomic random measure, with weights of size s at points x, for (x, s) the points in the associated PRS. Since finite-dimensional distributions do not determine path properties, it is not immediate that *any* version of the CRM is purely atomic, but this is nevertheless true. For a gamma process on a Polish space, this was established by Blackwell (1973), and for general CRMs by Kingman (1975).

The counting process $A \mapsto N^c(A \times (0, \infty])$ giving the "marginal distribution of the locations" will often be infinite, but is a PRS with intensity measure $A \mapsto \nu^c(A \times (0, \infty])$. The equations $\nu^c(\{x\} \times (0, \infty]) = 0$ show that the latter is atomless, so that with probability one the Poisson process N^c has at most one point (x, s) on every half line $\{x\} \times (0, \infty]$.[3] This is natural, as per given location x only one weight s need be specified.

The fixed atoms are special in two ways: their locations $a_1, a_2, \ldots$ are deterministic; and their weights need not be "Poisson" distributed. It is possible to represent the fixed atoms through the random measure N^d on $\mathfrak{X} \times (0, \infty]$ given by

$$N^d = \sum_j \delta_{a_j, \Phi(\{a_j\})}.$$

Then $\int_A \int s\, N^d(dx, ds)$ is identical to the first sum on the right side of (J.2) evaluated at A, and the representation of a CRM Φ as in Proposition J.6 can be written in the form

$$\Phi(A) = \beta(A) + \iint \mathbb{1}_A(x) s\, N(dx, ds), \qquad N = N^c + N^d.$$

The sum N is a CRM on $\mathfrak{X} \times (0, \infty]$, but it is not a Poisson random measure. For instance, the variable $N(\{a_j\} \times A)$, which is almost surely equal to $\mathbb{1}\{\Phi(\{a_j\}) \in A\}$, is not Poisson, but Bernoulli distributed. Jacod and Shiryaev (2003) (in the case that $\mathfrak{X} = \mathbb{R}^+$ and $(0, \infty]$ is replaced by a general space E) call N an *extended Poisson random measure*. The mean measure of N is given by

$$\nu(A) := \mathrm{E}N(A) = \nu^c(A) + \nu^d(A), \qquad \nu^d(A \times B) := \sum_j \mathbb{1}_A(a_j)\mathrm{P}(\Phi(\{a_j\}) \in B).$$

[3] For a finite measure ν^c this interpretation is clear from representing Π^c as $N_n \sim \mathrm{Poi}(|\nu^c|)$ points (x, s) drawn i.i.d. from $\bar{\nu}^c$; the general case follows by decomposition.

As $\nu^c(\{x\} \times (0,\infty]) = 0$ for every x, the measures ν^c and ν^d are concentrated off and on the set $\cup_j \{a_j\} \times (0,\infty]$, respectively, and hence are orthogonal. Their sum $\nu = \nu^c + \nu^d$ is called the *intensity measure* of the CRM Φ, and together with the deterministic part β uniquely identifies its distribution. Because the variable $\Phi(\{a_j\})$ may be zero with positive probability, the measure $\nu^d(\{a_j\} \times \cdot) = \mathrm{P}(\Phi(\{a_j\}) \in \cdot)$ may be a subprobability measure on $(0,\infty]$; the "missing mass" satisfies $1 - \nu^d(\{a_j\} \times (0,\infty]) = \mathrm{P}(\Phi(\{a_j\}) = 0)$. We can compute[4]

$$\mathrm{E}e^{-\theta\Phi(\{a_j\})} = 1 - \nu^d(\{a_j\}, (0,\infty]]) + \int_{(0,\infty]} e^{-\theta s}\, \nu^d(\{a_j\}, ds). \tag{J.10}$$

This may be combined with (J.4) to obtain an expression for the cumulant measure of a general CRM in terms of the intensity measure ν and deterministic component β, but because of the conceptually different roles played by the fixed atoms and continuous part, the resulting formula may be obscure rather than helpful. In particular note that replacing ν^c by ν in the right sight of (J.7) does not give a valid formula for the cumulant generating function of $\int f\, d\Psi$ for a general CRM.

By expressing the mean and variance of the variables $f(a_j)\Phi(\{a_j\})$ in ν^d (for instance by differentiating the preceding display) and some subsequent manipulation, we can extend formulas (J.8) and (J.9) for the mean and variance of $\int f\, d\Phi$ for a CRM without fixed atoms to a general CRM Φ, with intensity measure[5] $\nu = \nu^c + \nu^d$

$$\mathrm{E}\Big(\int f\, d\Psi\Big) = \iint_{\mathfrak{X}\times\mathbb{R}^+} sf(x)\, \nu(dx, ds), \tag{J.11}$$

$$\mathrm{var}\Big(\int f\, d\Psi\Big) = \iint_{\mathfrak{X}\times\mathbb{R}^+} s^2 f^2(x)\, \nu(dx, ds) - \sum_x \Big(\int sf(x)\, \nu(\{x\}, ds)\Big)^2. \tag{J.12}$$

Example J.7 (Gamma process) Perhaps the most important example of a CRM is the *gamma process*. In this case $\Phi(A)$ is gamma distributed with parameters $\alpha(A)$ and 1, for a given atomless σ-finite measure α on $\mathfrak{X}$. The corresponding characteristics are $\beta = 0$ and $\nu(dx, ds) = s^{-1}e^{-s}\, ds\, d\alpha(x)$. Indeed, with these choices equation (J.7) reduces to

$$\begin{aligned}\log \mathrm{E}\left(e^{-\theta\int f\, d\Phi}\right) &= \sum_{j=1}^{\infty} \frac{(-\theta)^j}{j!} \int_0^\infty s^{j-1}e^{-s}\, ds \int f^j(x)\, d\alpha(x) \\ &= -\int \log(1 + \theta f(x))\, d\alpha(x).\end{aligned} \tag{J.13}$$

In the special case $f = \mathbb{1}_A$ this reduces to $-\alpha(A)\log(1+\theta)$, which is indeed the cumulant generating function of a gamma distribution with parameters $\alpha(A)$ and 1.

It is often convenient to disintegrate the intensity measure as $\nu(dx, ds) = \rho_x(ds)\, \alpha(dx)$, where α is a measure on $\mathfrak{X}$ and ρ_x a transition kernel on $\mathfrak{X} \times \mathbb{R}^+$. The more interesting processes have $\rho_x(\mathbb{R}^+) = \infty$, for all x. The measure α refers to the distribution of the locations of the masses, and ρ_x to the size of a mass at x. If ρ_x does not depend on x, then the

[4] The notation $\int f(s)\, \nu(\{a\}, ds)$ means the integral of f relative to the measure $B \mapsto \nu(\{a\} \times B)$.
[5] Note that $\nu\big(\{x\}, ds\big) = \nu^d\big(\{x\}, ds\big)$, but ν in the other terms cannot be replaced by ν^c.

locations and sizes are stochastically independent. The CRM is then called *homogeneous*. (This same word is also employed to refer to uniformity in space or time, for instance (only) if ρ_x is the Lebesgue measure.)

Example J.8 (Product intensities) If $\nu(dx, ds) = \rho(ds)\,\alpha(dx)$ for a finite measure α, then it is no loss of generality to assume that α is a probability measure. Also assume that α is atomless. Below we show that the CRM can then be generated as $\Psi(A) = \sum_{i\in I} S_i \mathbb{1}\{X_i \in A\}$, for a Poisson random set $(S_i : i \in I)$ with intensity measure ρ on $(0, \infty)$, and an independent random sample $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \alpha$.

If $\rho(0, \infty) < \infty$, then the sum (over I) will have finitely many terms (namely a Poisson number), but typically the sum will be infinite. The requirement $\int (s \wedge 1)\,\rho(ds) < \infty$, resulting from Proposition J.6, shows that $\rho([\epsilon, \infty)) < \infty$ for every $\epsilon > 0$, so that the number of jumps bigger than ϵ is finite almost surely, and hence the jumps will cluster at zero only. The requirement also shows that $\int_0^1 s\,\rho(ds) < \infty$ so that the sum of the "small" jumps has finite expectation and hence the sum of all jumps is finite almost surely.

To prove the claim that Ψ has the given representation it suffices to verify that the point process defined by $N(A) := \#(i : (X_i, S_i) \in A)$ is a Poisson process with intensity ν. For product sets $A = C \times D$, the counts $N(A)$ arise as the thinning of the process $\#(i : S_i \in D)$ by removing points i with $X_i \notin C$, and hence $N(C \times D)$ is a Poisson process in D for every fixed C. Together with independence across disjoint C, this identifies N as a Poisson process, since the maps $M \mapsto M(C \times D)$ generate the same σ-field as the maps $A \mapsto M(A)$ (cf. Proposition A.5 together with a truncation argument to accommodate infinite random measures).

Example J.9 (Simulation) The preceding example extends to more general intensities $\nu(dx, ds) = \rho_x(ds)\,\alpha(dx)$ for a probability measure α, not necessarily of product form. We generate a sequence $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \alpha$ of locations as before, but define the corresponding weights as $S_i = L_{X_i}^{-1}(V_i)$, for a standard, homogeneous Poisson process $V_1, V_2, \ldots$ on $(0, \infty)$ and $L_x(s) = \rho_x((s, \infty))$. The CRM is obtained as $\Psi(A) = \sum_i S_i \mathbb{1}\{X_i \in A\}$, as before.

To see that this works it suffices again to show that the point process of the points (X_i, S_i) is a Poisson process with intensity measure ν. Arguing as in the preceding example, we see that the points (X_i, V_i) form a Poisson process with intensity measure $ds\,\alpha(dx)$. The transformed points (X_i, S_i) are obtained by applying the deterministic transformation $(x, v) \mapsto (x, L_x^{-1}(v))$ to the points of this process and hence still form a Poisson process. By the definition of L_x, for every measurable set A and $0 \le u < v$,

$$\begin{aligned}\mathrm{E}\#(X_i \in A, u < S_i \le v) &= \sum_{i=1}^{\infty} \int_A \mathrm{E}\mathbb{1}\{L_x(u) < V_i \le L_x(v)\}\, d\alpha(x) \\ &= \int_A (L_x(v) - L_x(u))\, d\alpha(x) = \nu(A \times (u, v]).\end{aligned}$$

Thus the transformed points possess intensity measure equal to ν.

J.3 Completely Random Measures on the Positive Half Line

CRMs on $\mathfrak{X} = \mathbb{R}^+$ are especially important in survival analysis. If they are finite on finite intervals, then they can be described by their Lebesgue-Stieltjes distribution function, for which we shall use the (unusual) notation X, so that $X(t) = \Phi((0,t])$. The sample paths $(X(t): t \in \mathbb{R}^+)$ of this stochastic process are nondecreasing right-continuous functions with independent increments over disjoint intervals. Thus X is an *independent increment process* (or *IIP*) with nonnegative increments. If the increments are stationary, then X is also a "Lévy process" and a "subordinator." In general, a *Lévy process* has cadlag sample paths and stationary independent increments, and a *subordinator* is a Lévy process with nonnegative increments. Several examples are given below. A Brownian motion is also an example of a Lévy process, but its increments are not nonnegative. A general Lévy process is the sum of an affine linear function, a multiple of Brownian motion, and a pure jump process. For CRMs the Brownian motion is absent and the pure jump processes have nonnegative jumps, and hence sample paths of bounded variation.

A CRM on $\mathfrak{X} = \mathbb{R}^+$ corresponds to a jump measure N on $\mathbb{R}^+ \times (0,\infty)$. Its distribution function can be represented in terms of $N = N^d + N^c$ as

$$\begin{aligned} X(t) &= \beta(t) + \int_{(0,t]} \int_{(0,\infty)} s\, N(dx, ds) \\ &= \beta(t) + \sum_{j: a_j \le t} \Delta X(a_j) + \int_{(0,t]} \int_{(0,\infty)} s\, N^c(dx, ds), \end{aligned}$$

where the variables $\Delta X(a_j)$ give the fixed atoms of the CRM, but are referred to as *fixed jumps* in this context. The fixed jumps are independent nonnegative random variables independent of the PRS N^c, the *fixed jump times* a_j are arbitrary positive numbers, and β is a cumulative distribution function on $\mathbb{R}^+$. If it is understood that $\Delta X(a)$ is the zero variable if $a \notin \{a_1, a_2, \ldots\}$, then the sum in the far right side of the display can also be extended to range over all $a \le t$. The mean measure ν of N can be decomposed as $\nu = \nu^d + \nu^c$ for ν^c the intensity measure of the PRS N^c, and ν^d describing the distribution of the fixed jumps: it concentrates on $\cup_j \{a_j\} \times (0,\infty)$ and $\nu^d(\{a_j\} \times A) = \mathrm{P}(\Delta X(a_j) \in A)$, for $A \subset (0,\infty)$ and every j. The random measure N^c is a PRS, while N is an *extended Poisson random measure*, in the terminology of Jacod and Shiryaev (2003). By formulas (J.11) and (J.12) the mean and variance of $X(t)$ are given in terms of ν by

$$\begin{aligned} \mathrm{E}\big[X(t)\big] &= \beta(t) + \int_{(0,t]} \int s\, \nu(dx, ds), \\ \mathrm{var}\big[X(t)\big] &= \int_{(0,t]} \int s^2\, \nu(dx, ds) - \sum_{x: x \le t} \Big(\int s\, \nu(\{x\}, ds) \Big)^2. \end{aligned}$$

Here $\nu(\{x\}, \cdot) = \nu^d(\{x\}, \cdot)$ is nonzero only at fixed jump times x, and hence the second term in the variance formula disappears if the process is without fixed jumps.

Because its sample paths are nondecreasing, the process X is trivially a *submartingale*, provided it is integrable. By the independence of increments the process $X(t) - \mathrm{E}\big[X(t)\big]$ is a martingale, whence $t \mapsto \mathrm{E}\big[X(t)\big]$ is the *compensator* of X.

Example J.10 (Time change) For a given process X with independent, nonnegative increments and a deterministic cadlag, nondecreasing function A, the process defined by $Y(t) = X(A(t))$ also possesses independent, nonnegative increments. This is called a *time change*. In the representation of X through a point process N^c on $\mathbb{R}^+ \times (0, \infty]$, the transformation simply moves the points, giving $(0, t] \times D \mapsto N^c((0, A(t)] \times D)$ as the point process representing Y, with intensity measure $(0, t] \times D \mapsto \nu^c((0, A(t)] \times D)$, for ν^c the intensity measure of X. Fixed atoms of X are moved similarly to fixed atoms of Y.

Example J.11 (Integration) Another operation that preserves independent, nonnegative increments is the formation of the process $Y(t) = \int_{(0,t]} b\, dX$, for a deterministic, nonnegative function b. Because an increment $Y(t_2) - Y(t_1)$ of the latter process depends only on the variables $\{X(u) - X(t_1) : t_1 < u \leq t_2\}$, the increments of Y are independent if those of X are independent. If X does not have fixed jumps, then the Laplace transform of the process Y can be computed using (J.6): if X and Y also denote the corresponding CRMs, then $\int f\, dY = \int fb\, dX$ and hence the formula gives, for ν the intensity measure of X,

$$-\log \mathrm{E}\left(e^{-\theta \int f\, dY}\right) = \iint (1 - e^{-\theta s f(x) b(x)})\, \nu(dx, ds).$$

In order to identify the intensity measure ν^* of Y the right side must be written in the form $\iint (1 - e^{-\theta s f(x)})\, \nu^*(dx, ds)$. This gives

$$\nu^*(A) = \iint \mathbb{1}_A(x, sb(x))\, \nu(dx, ds). \tag{J.14}$$

Example J.15 gives a concrete illustration.

Example J.12 (Compound Poisson process) A *standard Poisson process* on $\mathbb{R}^+$ is a subordinator. Its intensity measure is given by $\nu(dt, ds) = dt\, \delta_1(ds)$, for δ_1 the Dirac measure at 1. The Dirac measure arises as all jumps have size 1.

In a *compound Poisson process* the jumps occur at the times of a standard Poisson process, but their sizes are independently generated from a fixed probability distribution ρ, on $(0, \infty)$ to obtain nonnegative increments. The corresponding intensity measure is $\nu(dt, ds) = dt\, \rho(ds)$.

These processes are said to have "finite activity," as only finitely many jumps occur in bounded time intervals.

Example J.13 (σ-stable process) The *σ-stable subordinator*, for $0 < \sigma < 1$, corresponds to the intensity measure

$$\nu(dt, ds) = \frac{\sigma}{\Gamma(1 - \sigma)} s^{-1-\sigma}\, dt\, ds.$$

In this case the distribution of $X(t)$ is positive σ-stable: the sum of n independent copies of $X(t)$ is distributed like $n^{1/\sigma} X(t)$, for any n.

Example J.14 (Gamma process) The marginal distribution of the increments $X(t) - X(s)$ of the process resulting from the intensity measure $\nu(dt, ds) = as^{-1}e^{-bs}\, dt\, ds$ are Ga($at -$

$as, b)$, as noted in Example J.7 when $b = 1$. The process X is called a *(time-homogeneous) gamma process* with parameters (a, b); if $b = 1$ it is called a *standard gamma process*.

The more flexible process obtained by allowing a to depend on t, with intensity measure $\nu(dt, ds) = a(t)s^{-1}e^{-bs}\,dt\,ds$, is called the *time-inhomogeneous gamma process*. This process can also be defined for a general increasing, cadlag process A, not necessarily absolutely continuous, by the intensity measure

$$\nu(dt, ds) = s^{-1}e^{-bs}\,dA(t)\,ds.$$

The process can be constructed by a *time change* from a standard gamma process Y as $X(t) = Y(A(t))$. This shows that the marginal distribution of $X(t)$ is $\mathrm{Ga}(A(t), b)$.

The behavior of $X(t)$ as $t \to \infty$ can be controlled by the corresponding behavior of $A(t)$. If A increases to infinity (i.e. $A(\infty) = \infty$), then $X(t) \to \infty$ a.s., whereas in the other case $X(\infty) := \lim_{t\to\infty} X(t)$ is a real-valued random variable distributed as $\mathrm{Ga}(A(\infty), 1)$.

Example J.15 (Extended gamma process) If X is a gamma process with intensity measure $s^{-1}e^{-s}\,dA(t)\,ds$ and b is a given positive, measurable function, then the process $Y(t) = \int_{(0,t]} b\,dX$ is an independent increment process with intensity measure

$$\nu^c(dt, ds) = s^{-1}e^{-s/b(t)}\,dA(t)\,ds. \tag{J.15}$$

This follows from the general formula (J.14), and the change of variables $(t, sb(t)) \to (t, u)$ in the integral. Because $Y(t) = \int \mathbb{1}_{(0,t]} b\,dX$, the cumulant generating function of $Y(t)$ is equal to $-\int_{(0,t]} \log(1 + \theta b)\,dA$, by formula (J.13). It follows that $\mathrm{E}[Y(t)] = \int_{(0,t]} b\,dA$ and $\mathrm{var}[Y(t)] = \int_{(0,t]} b^2\,dA$.

The process Y is known as an *extended gamma process*. When b is constant, the process reduces to an ordinary gamma process. This is the only case in which the extended gamma process is homogeneous.

Example J.16 (Beta process) The *standard beta process* with parameters $a, b > 0$ has intensity measure

$$\nu(dt, ds) = B(a, b)^{-1}s^{a-2}(1 - s)^{b-1}\mathbb{1}_{(0,1)}(s)\,dt\,ds.$$

If $a > 1$, then the function $s \mapsto s^{a-2}(1 - s)^{b-1}$ is integrable, and one could construct the process by generating jump sizes from the beta distribution to be placed at jump locations forming a Poisson process, as in Example J.8 but with the roles of locations and sizes swapped. If $a \leq 1$ the sizes follow a "beta distribution with infinite weight near zero."

Example J.17 (Generalized gamma process) The *generalized gamma process* is the independent increment process with intensity measure, for $0 < \sigma < 1$ and $\tau \geq 0$,

$$\nu(dt, ds) = \frac{\sigma}{\Gamma(1 - \sigma)}s^{-(1+\sigma)}e^{-\tau s}dA(t)\,ds.$$

For $\tau = 0$ this reduces to the σ-stable process, whereas letting $\sigma \to 0$ takes it to the homogeneous gamma process. The *inverse-Gaussian process* is obtained for $\sigma = 1/2$.

Two CRMs with a similar intensity measure are close. The following proposition gives a simple criterion for the convergence of random cumulative distribution functions in terms of their intensity measures. The convergence is in the *Skorohod space* $\mathfrak{D}[0,\infty)$ relative to the Skorohod topology on compacta (see Billingsley 1968 or Pollard 1984, Chapter VI). The use of the latter topology, which is weaker than the topology of uniform convergence on compacta, is natural as the jumps of the processes occur at random times, and are not restricted to a given countable set. For simplicity we restrict to the situation that the limiting CRM possesses a continuous intensity measure.[6]

Proposition J.18 (Convergence of CRMs) *If X_n are processes with nonnegative, independent increments with intensity measures ν_n such that $(s\wedge 1)\,\nu_n(dx,ds) \rightsquigarrow (s\wedge 1)\,\nu(dx,ds)$ on $[0,t]\times\mathbb{R}^+$, for every $t>0$,[7] for an intensity measure ν with zero fixed jump component ν^d, then $X_n \rightsquigarrow X$ in $\mathfrak{D}[0,\infty)$ relative to the Skorohod topology on compacta, for X the process with nonnegative, independent increments with intensity measure ν.*

Proof By combining (J.6) and (J.10) we see that, for a given continuous function $f\colon\mathbb{R}^+\to[0,1]$ with compact support, and with $R(x)=\log(1-x)+x$,

$$\begin{aligned}-\log \mathrm{E}[e^{-\int f\,dX_n}] &= \int\int (1-e^{-f(x)s})\,\nu_n^c(dx,ds) - \sum_x \log\Bigl(1-\int (1-e^{-f(x)s})\,\nu_n^d(\{x\},ds)\Bigr)\\ &= \int\int (1-e^{-f(x)s})\,\nu_n(dx,ds) - \sum_x R\Bigl(\int (1-e^{-f(x)s})\,\nu_n^d(\{x\},ds)\Bigr).\end{aligned}$$

Since the function $(x,s)\mapsto 1-e^{-f(x)s}$ is continuous, nonnegative and bounded above by $s\wedge 1$, the first term on the far right tends to the same expression, but with ν_n replaced by ν. If it can be shown that the last term on the right tends to zero, then it follows that the Laplace transforms of the variables $\int f\,dX_n$ tend to the Laplace transform of $\int f\,dX$, which implies their convergence in distribution (e.g. Chung 2001, Theorem 6.6.3) and next marginal convergence of X_n.

In view of the assumed continuity of the first marginal of ν, we also have that the maps $t\mapsto L_n(t):=\int_{(0,t]}\int(1-e^{-f(x)s})\,\nu_n(dx,ds)$ converge pointwise to the map L defined in the same way from ν, and then by Pólya's theorem also uniformly on compacta. The integrals $\int(1-e^{-f(x)s})\,\nu_n^d(\{x\},ds)$ are the jumps of the maps L_n, and hence converge uniformly to the jumps of L, which are zero. This allows to complete the argument of the preceding paragraph.

To verify that the sequence X_n is uniformly tight in $\mathfrak{D}[0,\infty)$, assume first that the intensity measures ν_n concentrate on $\mathbb{R}^+\times[0,K]$, for some K. Then the function $(x,s)\mapsto s$ is bounded above by a multiple of $s\wedge 1$, whence, for $t>0$,

[6] For the general case, and also for stronger versions of the proposition, see Theorems VII.3.13 and VII.3.4 in Jacod and Shiryaev (2003). In the case of a CRM with $\mathrm{E}[X(t)]<\infty$, the characteristic triple of X is $(B,C,\nu):=(\mathrm{E}\tilde{X}(t),0,\nu)$, where $\tilde{X}_t$ is X_t with the big jumps removed, and the modified second characteristic $\tilde{C}$ is $\int_0^t\int h^2(s)\,\nu(dx,ds)-\sum_{x:x\le t}(\int h(s)\,\nu(\{x\},ds))^2$, where h is the truncation function.

[7] This means $\int_0^t\int g(x,s)(s\wedge 1)\,\nu_n(dx,ds)\to\int_0^t\int g(x,s)(s\wedge 1)\,\nu(dx,ds)$ for every continuous bounded function $g\colon[0,t]\times(0,\infty)\to\mathbb{R}$.

$$\mathrm{E}[X_n(t)] = \int_{(0,t]} \int s\, \nu_n(dx, ds) \to \int_{(0,t]} \int s\, \nu(dx, ds) = \mathrm{E}[X(t)].$$

Because the limit function is continuous by assumption, this convergence is automatically uniform. It follows that $\mathrm{E}|X_n(v) - X_n(u)|$ converges uniformly in (u, v) belonging to compacta to $\mathrm{E}|X(v) - X(u)|$, which can be made uniformly (arbitrarily) small by restricting u, v to $|u - v| < \delta$, for a small δ. This implies that *Aldous condition* for uniform tightness is satisfied (see Pollard 1984, Theorem VI.16 and Example VI.18).

Finally the weak convergence of the measures $(s \wedge 1)\, \nu_n(dx, ds)$ implies their uniform tightness, so that $\sup_n \nu_n([0, t] \times (K, \infty)) \to 0$ as $K \to \infty$. The number of non-fixed jumps of size bigger than K in X_n is a Poisson variable with mean $\nu_n^c([0, t] \times (K, \infty))$ and hence is zero with probability tending to one, as $K \to \infty$. Similarly, the probability that at least one fixed jump height exceeds K is bounded above by $\nu_n^d([0, t] \times (K, \infty))$ and tends to zero. This means that $\inf_n \mathrm{P}(X_n = X_n^K) \to 1$, for X_n^K the process defined by removing from X_n the jumps of heights bigger than K. By the preceding paragraph $X_n^K \rightsquigarrow X^K$, as $n \to \infty$, for every fixed K, and $X^K \rightsquigarrow X$ as $K \to \infty$. We can choose a suitable $K_n \to \infty$ such that both $\mathrm{P}(X_n = X_n^{K_n}) \to 1$ and $X_n^{K_n} \rightsquigarrow X$ to complete the proof. □

The sample paths of most of the preceding examples of nondecreasing, independent increment processes are delicate, as they increase by countably many jumps in any interval. This complicates a description of their rates of growth. In the following result this is given for the standard gamma process (for a proof see Fristedt and Pruitt 1971)). Similar results can be obtained for other subordinators (see Section III.4 of Bertoin 1996).

Theorem J.19 *Let X be the standard gamma process on $[0, \infty)$ and $h\colon (0, \infty) \to (0, \infty)$.*

(i) *If h is strictly increasing, then either* $\mathrm{P}(\liminf_{t\to 0} X(t)/h(t) = 0) = 1$ *or* $\mathrm{P}(\liminf_{t\to 0} X(t)/h(t) = \infty) = 1$.

(ii) *Let h be strictly increasing and convex. Then*

$$\mathrm{P}(\limsup_{t\to 0} X(t)/h(t) = 0) = \begin{cases} 1, & \text{if } \int_0^1 \int_{h(t)}^\infty e^{-x} x^{-1} dx\, dt < \infty, \\ 0, & \text{if } \int_0^1 \int_{h(t)}^\infty e^{-x} x^{-1} dx\, dt = \infty. \end{cases}$$

(iii) *The variable* $\liminf_{t\to 0} X(t) \exp[r \log|\log t|/t]$ *is equal to* 0 *a.s. if* $r < 1$ *and equal to* ∞ *a.s. if* $r > 1$.

J.4 Historical notes

Kingman (1967, 1975) developed the theory of completely random measures on general sample spaces. For general point processes (treated within the context of random measures) and their historical development, also see Daley and Vere-Jones (2003, 2008). For the special case of Lévy processes, which were developed earlier, see Bertoin (1996).

Lemma J.4 is based on Section 4 of Perman et al. (1992).

Problems

J.1 Using the time change technique of Lévy processes, state and prove a version of Theorem J.19 for a non-homogeneous gamma processes.

J.2 (Kingman 1993) Show that the mean μ of a Poisson process is a measure. Further show that μ is atomless. [Hint: Use $N(\{x\}) \leq 1$ for all $x \in \mathfrak{X}$.]

J.3 (Kingman 1993) Show that $\operatorname{cov}(N(A_1), N(A_2)) = \mu(A_1 \cap A_2)$ for a Poisson process N.

J.4 (Kingman 1993) Let $Z = Z(\Pi)$ be a random variable depending on a PRS Π on $\mathfrak{X}$ with finite mean measure μ. Derive the following compounding formula:

$$\mathrm{E}(Z) = \sum_{n=0}^{\infty} \mathrm{E}_n(Z) e^{-\mu(\mathfrak{X})} \mu(\mathfrak{X})^n / n!,$$

where $\mathrm{E}_n(Z) = \mathrm{E}[Z(\Pi_n)]$, $\Pi_0 = \varnothing$, and Π_n is a multinomial random set with parameters n and $P := \mu/\mu(\mathfrak{X})$.

J.5 If $(X(t): t \geq 0)$ is an independent increment process with cadlag sample paths, then it defines a random distribution function and hence a random measure Φ. The independence of the increments imply that the variables $\Phi(A_1), \ldots, \Phi(A_k)$ are independent whenever $A_1, \ldots, A_k$ are disjoint intervals. Show that these variables are independent for any disjoint Borel sets $A_1, \ldots, A_k$, and $k \in \mathbb{N}$, so that Φ is also a CRM according to the definition. [Hint: use the monotone class theorem and the "good sets principle" in succession on $A_1, \ldots, A_k$.]