Statistics and Probability Letters 135 (2018) 32–37

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Integrability conditions for compound random measures

Alan Riva Palacio, Fabrizio Leisen *

School of Mathematics, Statistics and Actuarial Sciences, University of Kent, UK

ARTICLE INFO

Article history:
Received 19 July 2017
Received in revised form 25 October 2017
Accepted 10 November 2017
Available online 5 December 2017

Keywords:
Bayesian non-parametrics
Multivariate Lévy measure
Partial exchangeability
Exchangeable Partition Probability Function

ABSTRACT

Compound random measures (CoRM's) are a flexible and tractable framework for vectors of completely random measure. In this paper, we provide conditions to guarantee the existence of a CoRM. Furthermore, we prove some interesting properties of CoRM's when exponential scores and regularly varying Lévy intensities are considered.

1. Introduction

Recently, a growing literature in Bayesian non-parametrics (BNP) proposed new priors which can take into account different features of the data, such as *partial exchangeability*, see De Finetti (1938). In this case, one would like to consider different densities for different groups instead of a single common density for all the data. After the seminal paper of MacEachern (1999), the problem of modeling a finite number of dependent densities has become an active area of research in Bayesian non-parametrics. A common approach is to construct BNP priors based on functions of *Completely Random Measures* (CRM's), see Lijoi and Prünster (2010). For example, special attention has been given to the normalization of CRM's starting with the work of Regazzini et al. (2003). Roughly speaking, a CRM is a generalization of a subordinator, that is a process with independent increments and almost surely increasing paths; for a full account of CRM's see Kingman (1993). This property is very helpful to derive the Laplace functional transform which is the basis to derive some analytical quantities of interest such as posterior and predictive distributions or the Exchangeable Partition Probability Function (EPPF), see James et al. (2009). To build more flexible priors in possibly higher dimensional spaces, vectors of dependent CRM's are constructed for example in Leisen and Lijoi (2011), Leisen et al. (2013) and Zhu and Leisen (2015) where respectively they build vectors of Poisson–Dirichlet and Dirichlet processes. These papers deal with the Lévy-Copula approach introduced in Cont and Tankov (2004) to induce dependence among the components of the vector. In a similar fashion, Lijoi et al. (2014) introduce a vector of random probability measures where the dependence arises by virtue of a suitable construction of the Poisson random measures underlying the CRM's; furthermore, in the framework of survival analysis, Lijoi and Nipoti (2014) introduce a new class of vectors of random hazard rate functions that are expressed as kernel mixtures of dependent CRM's. Camerlenghi et al. (2017) focus on partial exchangeable models which arise from hierarchical specifications of CRM's.

Compound random measures (CoRM's), introduced by Griffin and Leisen (2017), are a flexible and tractable framework for many dependent random measures including many of the superposition and Lévy copula approaches. They have recently been applied to modeling graphs for overlapping communities by Todeschini et al. (2017). Griffin and Leisen (2017, 2018) described posterior sampling methods for a particular class of normalized compound random measure mixtures which

* Correspondence to: School of Mathematics, Statistics and Actuarial Sciences Cornwallis Building, University of Kent, Canterbury CT2 7NF, UK.
E-mail addresses: ar515@kent.ac.uk (A. Riva Palacio), fabrizio.leisen@gmail.com (F. Leisen).

https://doi.org/10.1016/j.spl.2017.11.005
0167-7152/

exploits a representation of the Laplace transform of the CoRM through a univariate integral of a moment generating function.

In this paper we aim to provide explicit existence conditions for CoRM's in order to guarantee the existence of the marginal Lévy intensities. On the other hand, we prove that the resulting CoRM is well posed in the sense that it satisfies the usual integrability condition for multivariate Lévy processes. Furthermore, this paper provides an interesting result for CoRM's when regularly varying Lévy intensities are considered. The paper closes highlighting a representation on the multivariate Lévy intensity of a CoRM when the score distribution is the result of marginal independent and identically distributed exponential scores.

The outline of the paper is as follows. Section 2 will set the scene by introducing the basic definitions which are required in the CoRM setting. Section 3 is devoted to prove our main results. Section 4 deals with CoRM's built with regularly varying Lévy intensities and exponential scores. Section 5 concludes.

2. Preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathbb{X}$ a Polish space with corresponding Borel σ-algebra $\mathcal{X}$. We denote by $\mathbb{M}_{\mathbb{X}}$ the space of boundedly finite measures on the measurable space $(\mathbb{X}, \mathcal{X})$ and by $\mathcal{M}_{\mathbb{X}}$ the associated Borel σ-algebra, see Appendix 2 in Daley and Vere-Jones (2003) for technical details.

Definition 1. A random measure μ on $\mathbb{X}$ is called a completely random measure (CRM) if for any $n > 1$ and disjoint sets $A_1, \ldots, A_n \in \mathcal{X}$ the random variables $\mu(A_1), \ldots, \mu(A_n)$ are mutually independent.

In the following we consider only CRM's without fixed jumps, namely CRM's of the form $\mu = \sum_{i=1}^{\infty} w_i \delta_{u_i}$ for collections of random variables $\{w_i\}_{i=1}^{\infty}$ in $\mathbb{R}^+$ and $\{u_i\}_{i=1}^{\infty}$ in $\mathbb{X}$. Such CRM's can be characterized by their Laplace transform

$$\mathbb{E}\left[e^{-\mu(f)}\right] = e^{-\int_{\mathbb{R}^+\times\mathbb{X}}(1-e^{-sf(x)})\tilde{\nu}(\mathrm{d}s,\mathrm{d}x)}$$

where $\mu(f) = \int_{\mathbb{X}} f(x)\mu(\mathrm{d}x)$, $f : \mathbb{X} \to \mathbb{R}^+$ is such that $\mu(f) < \infty$ and $\tilde{\nu}(\mathrm{d}s, \mathrm{d}x)$ is a measure in $(\mathbb{R}^+ \times \mathbb{X}, \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{X})$ such that

$$\int_{\mathbb{R}^+\times X} \min\{1, s\}\tilde{\nu}(\mathrm{d}s, \mathrm{d}x) < \infty \tag{1}$$

for any bounded set $X \in \mathcal{X}$. A measure $\tilde{\nu}$ satisfying the condition displayed in Eq. (1) is called the *Lévy intensity* of μ. We say that $\tilde{\nu}$ is homogeneous when $\tilde{\nu}(\mathrm{d}s, \mathrm{d}x) = \rho(\mathrm{d}s)\alpha(\mathrm{d}x)$ with ρ a measure in $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ and α a measure in $(\mathbb{X}, \mathcal{X})$. The notion of a completely random measure can be generalized to higher dimensions in a similar fashion to Definition 1, see for instance Griffin and Leisen (2017). As a result, we have a representation in terms of a Laplace functional transform. Precisely, for a vector of completely random measures $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)$ on $\mathbb{X}$ we have that

$$\mathbb{E}\left[e^{-\mu_1(f_1)-\cdots-\mu_d(f_d)}\right] = e^{-\int_{(\mathbb{R}^+)^d\times\mathbb{X}}\left(1-e^{-s_1f_1(x)-\cdots-s_df_d(x)}\right)\tilde{\nu}_d(\mathrm{d}\boldsymbol{s},\mathrm{d}x)}$$

with $f_j : \mathbb{X} \to \mathbb{R}^+, j \in \{1, \ldots, d\}$ such that $\mu_j(f_j) < \infty$, where for $g : \mathbb{X} \to \mathbb{R}^+$ we have $\mu_j(g) = \int_{\mathbb{X}} g(x)\mu_j(\mathrm{d}x)$. The measure $\tilde{\nu}_d$ in $\left((\mathbb{R}^+)^d \times \mathbb{X}, \mathcal{B}(\mathbb{R}^+)^d \otimes \mathcal{X}\right)$ must be such that

$$\int_{(\mathbb{R}^+)^d\times\mathbb{X}} \min\{1, \|\boldsymbol{s}\|\}\tilde{\nu}_d(\mathrm{d}\boldsymbol{s}, \mathrm{d}x) < \infty \tag{2}$$

for any bounded set $X \in \mathcal{X}$; we call $\tilde{\nu}_d$ a multivariate Lévy intensity. We set the notation

$$\nu_j(A, X) = \int_{(\mathbb{R}^+)^{d-1}} \tilde{\nu}_d(\mathrm{d}s_1, \ldots, \mathrm{d}s_{j-1}, A, \mathrm{d}s_{j+1}, \ldots, \mathrm{d}s_d, X)$$

with $j \in \{1, \ldots, d\}$ and $A \in \mathcal{B}\left(\mathbb{R}^+\right)$. We call ν_j the jth marginal of the d-variate Lévy intensity $\tilde{\nu}_d$; it follows that for each $j \in \{1, \ldots, \mathrm{d}\}$, μ_j has Lévy intensity ν_j. In this framework we can define the concept of *Compound Random Measure* (CoRM). The following definition differs from the one in Griffin and Leisen (2017) since it takes into account the *inhomogeneous* case, where the locations and associated weights in the CRM are not independent.

Definition 2. A *Compound Random Measure* (CoRM) is a vector of CRM's whose Lévy intensity is given by

$$\tilde{\nu}_d(\mathrm{d}\boldsymbol{s}, \mathrm{d}x) = \int_{\mathbb{R}^+} z^{-d} h\left(\frac{s_1}{z}, \ldots, \frac{s_d}{z}\right) \mathrm{d}\boldsymbol{s}\,\nu^\star(\mathrm{d}z, \mathrm{d}x) \tag{3}$$

where h, *the score distribution*, is a d-variate probability density function and, $\nu^\star$, *the directing Lévy measure*, is a Lévy intensity.

By performing a simple change of variable we note that

$$\int_{(\mathbb{R}^+)^d} z^{-d} h\left(\frac{s_1}{z}, \ldots, \frac{s_d}{z}\right) \mathrm{d}\boldsymbol{s} = 1.$$