INTERNATIONAL CENTRE FOR ECONOMIC RESEARCH
ICER

WORKING PAPER SERIES

Antonio Lijoi, Igor Prünster and Stephen G. Walker

POSTERIOR ANALYSIS FOR SOME CLASSES OF NONPARAMETRIC MODELS

Working Paper no. 5/2008

APPLIED MATHEMATICS AND QUANTITATIVE METHODS
WORKING PAPER SERIES

Posterior analysis for some classes of nonparametric models

Antonio Lijoi[1], Igor Prünster[2] and Stephen G. Walker[3]

[1] Dipartimento Economia Politica e Metodi Quantitatavi, Università degli Studi di Pavia, via San Felice 5, 27100 Pavia, and CNR-IMATI, Milano, Italy *e-mail*: lijoi@unipv.it.

[2] Dipartimento di Statistica e Matematica Applicata, Collegio Carlo Alberto and ICER, Università degli Studi di Torino, Piazza Arbarello 8, 10122 Torino, Italy. *e-mail*: igor@econ.unito.it

[3] Institute of Mathematics, Statistics and Actuarial Science, University of Kent, Kent CT2 7NZ, UK. *e-mail*: S.G.Walker@kent.ac.uk

Abstract. Recently, James [15, 16] has derived important results for various models in Bayesian nonparametric inference. In particular, he defined a spatial version of neutral to the right processes and derived their posterior distribution. Moreover, he obtained the posterior distribution for an intensity or hazard rate modeled as a mixture under a general multiplicative intensity model. His proofs rely on the so–called Bayesian Poisson partition calculus. Here we provide new proofs based on an alternative technique.

Keywords: Bayesian Nonparametrics; Completely random measure; Hazard rate; Neutral to the right prior; Multiplicative intensity model;

1 Introduction

Recently James [12] introduced a technique, called Bayesian Poisson partition calculus, which allows the derivation of posterior distributions for a large variety of Bayesian nonparametric models. His technique, whose roots lie in [23, 24], consists in a Laplace functional change of measure combined with a Poisson Palm/Fubini calculus on random partitions of the positive integers. See [26] and references therein for an exhaustive account on exchangeable random partitions and applications in areas not directly related to Bayesian Nonparametrics.

In [15] the multiplicative intensity model of Aalen [1] is considered. It is well–known that the multiplicative intensity model covers a large variety of important applied models such as the simple life testing model, the Cox proportional hazards regression model, the multiple decrement model and Poisson process spatial regression models, among others. A typical Bayesian nonparametric approach, in this area, relies on designing the intensity or hazard function as a mixture with respect to gamma (or allied) random measures. See, e.g., [6, 24, 31, 13, 25, 11]. James [15] derived the posterior distribution for general multiplicative intensity models in which the intensity or hazard rate is a mixture driven by any completely random measure thus generalizing all previous posterior representations. Basing upon this result and suitable simulation algorithms, practitioners have now the possibility not only of selecting an appropriate kernel but also to decide which random measure to adopt.

In James [16] a spatial version of the popular neutral to the right (NTR) processes, termed spatial neutral to the right (SPNTR) processes is introduced. Indeed, NTR priors, due to Doksum [5], have been successfully exploited in the context of survival analysis leading to Bayesian nonparametric analogs of the Kaplan–Meier estimator. See, among others, [9, 30, 29, 7]. One of the main drawbacks of NTR priors is represented by the fact that they can be defined only on the real line and not on multidimensional or general abstract spaces. The notion of SPNTR prior obviates to this. Moreover, it relates in a nice way the literature on NTR processes with the other line of research, initiated by Hjort [10], which consists in modeling the cumulative hazard by means of a suitable completely random measure, in particular a so–called beta process. Within this framework we recall, e.g., [20, 21, 4]. James [15], applying Poisson partition calculus, derived the posterior distribution for a general SPNTR prior, again opening up the possibility of exploiting concretely many different alternatives in several important applications.

In this paper we provide alternative proofs of two results of [15, 16] regarding the posterior distribution of SPNTR priors and mixture priors for multiplicative intensity models by means of a different technique. Given the importance of the posterior characterizations in [15, 16], it is useful to have different derivations of them and this could then set a basis for obtaining posterior distributions also in other models involving completely random measures. Our approach consists in reading a suitably transformed version of the data "likelihood" as a derivative of the Laplace functional of the completely random measure upon which the model is built, setting up a recursion and obtaining the posterior Laplace functional in the limit. Such a device has been, at least to authors' knowledge, first employed in [27, 22]. Recently, the posterior distribution of normalized random measures with independent increments, a class of priors introduced in [28], has been derived in [17]. In that paper, additionally to a proof based on the Poisson process calculus technique, one relying on the approach of the present paper is provided as well. Finally, it is worth mentioning that the techniques we adopt are connected to some results obtained in [2] where the authors exploit Faa di Bruno's formula and deduce generalizations of Dobinski's formula.

In Section 2 we recall the definition of completely random measure and some useful notation. In Sections 3 and 4 we provide alternative proofs of the posterior characterizations of SPNTR models and of multiplicative hazard models, respectively.

2 Preliminaries and notation

At the heart of most nonparametric models there is the concept of completely random measure introduced by Kingman [18], which we briefly recall here. It is worth noting that an increasing additive process (or increasing Lévy process with not necessarily stationary increments) can always be seen as the càdlàg distribution function induced by a completely random measure on $\mathbb{R}$.

Consider a measure space $(\mathbb{X}, \mathscr{X})$, where $\mathbb{X}$ is a complete and separable metric space and $\mathscr{X}$ is the usual Borel σ–field. For notational simplicity, set $\mathbb{S} = \mathbb{R}^+ \times \mathbb{X}$ and let $\mathscr{S}$ denote the product σ–algebra $\mathscr{B}(\mathbb{R}^+) \otimes \mathscr{X}$ on $\mathbb{S}$ where, as usual, $\mathscr{B}(\mathbb{R}^+)$ stands for the class of Borel subsets of $\mathbb{R}^+$. Introduce, now, a *Poisson random measure* N, defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and taking values in the set of non–negative counting measures on $(\mathbb{S}, \mathscr{S})$, with *intensity measure* ν, i.e. $\mathbb{E}\,[N(\mathrm{d}v, \mathrm{d}x)] = \nu(\mathrm{d}v, \mathrm{d}x)$. Hence, for any $A \in \mathscr{S}$ such that $\nu(A) < \infty$, $N(A)$ is a Poisson random variable of parameter $\nu(A)$ and, given any finite collection of pairwise disjoint sets, $A_1, \ldots, A_k$, in $\mathscr{S}$, the random variables $N(A_1), \ldots, N(A_k)$ are mutually independent. Throughout the paper, $\mathbb{E}[\,\cdot\,]$ will denote expectation with respect to $\mathbb{P}$. Moreover, the intensity measure ν must satisfy $\int_{\mathbb{R}^+}(v \wedge 1)\nu(\mathrm{d}v, \mathbb{X}) < \infty$, where $a \wedge b = \min\{a, b\}$. See [3] for an exhaustive account on Poisson random measures.

Let now $(\mathbb{M}, \mathscr{B}(\mathbb{M}))$ be the space of boundedly finite measures on $(\mathbb{X}, \mathscr{B}(\mathbb{X}))$. We suppose that $\mathbb{M}$ is equipped with the topology of weak convergence and $\mathscr{B}(\mathbb{M})$ is the corresponding Borel σ–algebra.

Let μ be a random element defined on $(\Omega, \mathscr{F}, \mathbb{P})$ and with values in $(\mathbb{M}, \mathscr{B}(\mathbb{M}))$. It is further assumed that μ can be represented as a linear functional of the Poisson random measure N as follows

$$\mu(B) = \int_{\mathbb{R}^+ \times B} g(v)\, N(\mathrm{d}v, \mathrm{d}x) \qquad \text{for any } B \in \mathscr{X},$$

where $g : \mathbb{X} \to \mathbb{R}^+$ is some measurable function. It can be easily seen from the properties of N that μ is, in the terminology of [18], a *completely random measure* (CRM) on $\mathbb{X}$, i.e. for any collection of disjoint sets $B_1, B_2, \ldots$ in $\mathscr{X}$, the random variables $\mu(B_1), \mu(B_2), \ldots$ are mutually independent.

Let, now, $\mathscr{H}_{\nu,g}$ be the space of functions $h : \mathbb{X} \to \mathbb{R}^+$ such that $\int_{\mathbb{S}}[1 - \mathrm{e}^{-g(v)\,h(x)}]\nu(\mathrm{d}v, \mathrm{d}x)$ is finite. Then, μ is uniquely characterized by its *Laplace functional* which, for any h in $\mathscr{H}_{\nu,g}$, is given by

$$\mathbb{E}\left[\mathrm{e}^{-\int_{\mathbb{X}} h(x)\,\mu(\mathrm{d}x)}\right] = \mathrm{e}^{-\int_{\mathbb{S}}[1-\mathrm{e}^{-g(v)\,h(x)}]\nu(\mathrm{d}v,\mathrm{d}x)} =: \mathrm{e}^{-\psi_{\nu,g}(h)} \tag{1}$$

When $g(v) \equiv v$, we write $\mathscr{H}_{\nu,g} = \mathscr{H}_\nu$ and $\psi_{\nu,g} = \psi_\nu$. See [19] for details and further references on CRMs.

3 Posterior analysis of spatial neutral to the right models

Let us start by recalling the precise definition of a Spatial neutral to the right (SPNTR) random probability measure given in [16]. To this end we set $\mathbb{W} = [0,1] \times \mathbb{S}$, with $\mathscr{W} = \mathscr{B}([0,1]) \otimes \mathscr{S}$, and introduce a Poisson random measure N on $(\mathbb{W}, \mathscr{W})$ with intensity measure of the form

$$\nu(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z) = \rho(\mathrm{d}v|s)\Lambda_0(\mathrm{d}s, \mathrm{d}z). \tag{2}$$

In (2), ρ is a Lévy density and Λ_0 is a hazard measure on $(\mathbb{S}, \mathscr{S})$. Based on a Poisson random measure with intensity of the form (2), define the following CRMs

$$\Lambda(C) = \int_{[0,1]\times C} v\ N(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z) \qquad \text{for any } C \in \mathscr{S} \tag{3}$$

$$Z(C) = \int_{[0,1]\times C} -\log(1-v)\ N(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z) \qquad \text{for any } C \in \mathscr{S} \tag{4}$$

We are now in a position to recall the James' definition of SPNTR random probability measure.

Definition.[16] *Given Λ and Z defined as in* (3) *and* (4)*, respectively, a spatial neutral to the right process (SPNTR) is a random probability measure on $(\mathbb{S}, \mathscr{S})$ defined by means of the relation*

$$F(\mathrm{d}t, \mathrm{d}x) \stackrel{d}{=} \mathrm{e}^{-Z(A_t)}\Lambda(\mathrm{d}t, \mathrm{d}x).$$

where $A_t = (0, t) \times \mathbb{X}$.

Note that by integrating over $\mathbb{X}$ one obtains the usual neutral to the right process [5] on $\mathbb{R}^+$ i.e. $F(\mathrm{d}t) = \mathrm{e}^{-Z(A_t)}\Lambda(\mathrm{d}t, \mathbb{X})$.

We now move to considering the posterior distribution of a SPNTR model. We first remark that a SPNTR measure is almost surely discrete, which can be seen e.g. from [14]: this implies that samples from a SPNTR measure will contain ties with positive probability. Consequently, in deriving the posterior one has to consider the case of samples of size n for which only $k \le n$ observations are distinct. Let $(\boldsymbol{T}, \boldsymbol{X}) = \{(T_1, X_1), \ldots, (T_n, X_n)\}$ be a sample of size n and denote by $(\boldsymbol{T}^*, \boldsymbol{X}^*) =$

$\{(T_1^*, X_1^*), \dots, (T_k^*, X_k^*)\}$ the $k \leq n$ different observations with frequencies $n_1, \dots, n_k$, respectively, and agree that the pairs in $(\boldsymbol{T^*}, \boldsymbol{X^*})$ are set in an increasing order with respect to the first coordinate i.e. (T_i^*, X_i^*) and (T_j^*, X_j^*) are such that $T_i^* < T_j^*$ for any $i < j$. Moreover, define $\bar{n}_j = \sum_{i=j}^k n_i$ and $\bar{N}(s) = \sum_{i=1}^k \bar{n}_i \mathbb{I}_{[T_{i-1}^*, T_i^*)}(s)$, where we agree on $T_0^* = 0$.

Proposition 1. [16, Prop. 4.1] *Let $F(\mathrm{d}s, \mathrm{d}z) = \mathrm{e}^{-Z(A_s)}\Lambda(\mathrm{d}s, \mathrm{d}z)$ be a SPNTR process. The posterior distribution of F, given $(\boldsymbol{T}, \boldsymbol{X})$, is a SPNTR process defined by*

$$F^{(n)}(\mathrm{d}s, \mathrm{d}z) \stackrel{d}{=} \mathrm{e}^{-Z^{(n)}(A_s)}\Lambda^{(n)}(\mathrm{d}s, \mathrm{d}z) \tag{5}$$

with the following specifications:

(i) $\Lambda^{(n)}(\mathrm{d}s, \mathrm{d}z) \stackrel{d}{=} \Lambda^*(\mathrm{d}s, \mathrm{d}z) + \sum_{i=1}^k J_i \delta_{(T_i^*, X_i^*)}(\mathrm{d}s, \mathrm{d}z)$, *where* $\Lambda^*(\mathrm{d}s, \mathrm{d}z) = \int_{[0,1]} v N^*(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z)$ *is a CRM with Poisson intensity*

$$\nu^*(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z) = (1-v)^{\bar{N}(s)} \rho(\mathrm{d}v|s)\Lambda_0(\mathrm{d}s, \mathrm{d}z) \tag{6}$$

and, for $i = 1, \dots, k$, (T_i^*, X_i^*) is a fixed point of discontinuity with corresponding jump J_i distributed as

$$f_{J_i}(dv) = \frac{v^{n_i}(1-v)^{\bar{n}_{i+1}}\rho(\mathrm{d}v|T_i^*)}{\int_{[0,1]} v^{n_i}(1-v)^{\bar{n}_{i+1}}\rho(\mathrm{d}v|T_i^*)}. \tag{7}$$

Moreover, the J_i's are conditionally independent of Λ^*.

(ii) $Z^{(n)}(\mathrm{d}s, \mathrm{d}z) \stackrel{d}{=} Z^*(\mathrm{d}s, \mathrm{d}z) + \sum_{i=1}^k K_i \delta_{(T_i^*, X_i^*)}(\mathrm{d}s, \mathrm{d}z)$, *where* $Z^*(\mathrm{d}s, \mathrm{d}z) = \int_{[0,1]} -\log(1-v) N^*(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z)$ *is a CRM with intensity* (6) *and, for $i = 1, \dots, k$, (T_i^*, X_i^*) is a fixed point of discontinuity with corresponding jump $K_i \stackrel{d}{=} -\log(1 - J_i)$ where J_i is distributed as in* (7).

Finally, (5) *can be rewritten as*

$$F^{(n)}(\mathrm{d}s, \mathrm{d}z) \stackrel{d}{=} \mathrm{e}^{-Z^*(A_s)} \prod_{i=1}^k (1 - \mathbb{I}_{(T_i^*, \infty)}(s) J_i) \Lambda^*(\mathrm{d}s, \mathrm{d}z) + \sum_{i=1}^k P_i^* \delta_{(T_i^*, X_i^*)}(\mathrm{d}s, \mathrm{d}z) \tag{8}$$

where all the quantities are as above and, for $i = 1, \dots, k$, P_i^* is equal in distribution to $J_i \mathrm{e}^{-Z^*(A_{T_i^*})} \prod_{j=1}^{i-1}(1 - J_j)$.

Proof. The proof–strategy relies on the derivation of the posterior Laplace functional of Λ defined in (3) which then uniquely characterizes its posterior distribution. Given the posterior distribution of Λ the other parts of the result follow by simple arguments.

By (1), the Laplace functional of a CRM with intensity (2) is of the form

$$\mathrm{e}^{-\psi_{\nu,g}(h)} := \mathrm{e}^{-\int_{\mathbb{W}}(1-\mathrm{e}^{-h(s,z)g(v)})\rho(\mathrm{d}v|s)\Lambda_0(\mathrm{d}s,\mathrm{d}z)}$$

for any $h \in \mathscr{H}_{\nu,g}$. Consider now a set $A^\epsilon(k) \subset \mathscr{S}$ defined as the product set $\times_{i=1}^k A_{i,\epsilon}^{n_i}$ with $A_{i,\epsilon} = (t_i^* - \epsilon, t_i^* + \epsilon) \times B_\epsilon(x_i^*)$, where $B_\epsilon(x_i^*)$ denotes a ball of size ϵ around the point x_i^*. Note that $(\boldsymbol{T}, \boldsymbol{X}) \in A^\epsilon(k)$ corresponds to having n_i observations in $A_{i,\epsilon}$ for $i = 1, \dots, k$; by letting $\epsilon \downarrow 0$,

$(\boldsymbol{T},\boldsymbol{X}) \in A^\epsilon(k)$ reduces to a sample $(\boldsymbol{T},\boldsymbol{X})$ featuring k distinct values (t_i^*, x_i^*) with frequency n_i, for $i = 1,\dots,k$. Our aim is to derive the posterior Laplace functional of Λ i.e.

$$\lim_{\epsilon\downarrow 0} \mathbb{E}\left[\mathrm{e}^{-\int_{\mathbb{S}} h(s,z)\Lambda(\mathrm{d}s,\mathrm{d}z)} \big| (\boldsymbol{T},\boldsymbol{X}) \in A^\epsilon(k)\right] \tag{9}$$

The conditional expectation in (9), before evaluating the limit, can expressed as

$$\frac{\mathbb{E}\left[\mathrm{e}^{-\int_{\mathbb{S}} h(s,z)\,\Lambda(\mathrm{d}s,\mathrm{d}z)}\ \prod_{i=1}^k \Phi_{i,\epsilon}^{n_i}\right]}{\mathbb{E}\left[\prod_{i=1}^k \Phi_{i,\epsilon}^{n_i}\right]} \tag{10}$$

with $\Phi_{i,\epsilon} := \int_{A_{i,\epsilon}} \mathrm{e}^{Z(A_t)} \Lambda(\mathrm{d}t,\mathrm{d}x)$ denoting the (random) probability that an observation falls in $A_{i,\epsilon}$. Notice that

$$\mathrm{e}^{-Z(A_{t_i^*+\epsilon})} \int_{A_{i,\epsilon}} \Lambda(\mathrm{d}t,\mathrm{d}x) \le \Phi_{i,\epsilon} \le \mathrm{e}^{-Z(A_{t_i^*-\epsilon})} \int_{A_{i,\epsilon}} \Lambda(\mathrm{d}t,\mathrm{d}x) \tag{11}$$

Hence a lower bound for the numerator becomes

$$\mathbb{E}\left[\mathrm{e}^{-\int_{\mathbb{W}} (h(s,z)\,v\,-\bar{N}_\epsilon(s)\log(1-v))\,N(\mathrm{d}v,\mathrm{d}s,\mathrm{d}z)}\ \prod_{i=1}^k \left(\int_{[0,1]\times A_{i,\epsilon}} v\,N(\mathrm{d}v,\mathrm{d}t,\mathrm{d}x)\right)^{n_i}\right]$$

having set $\bar{N}_\epsilon(s) = \sum_{i=1}^k \bar{n}_i \mathbb{I}_{[t_{i-1}^*+\epsilon,t_i^*+\epsilon)}(s)$, with $t_0^* = 0$. Let now

$$g_{\epsilon,\lambda}^C(v,s,z) := \mathbb{I}_C(s,z)[h(s,z)\,v - \bar{N}_\epsilon(s)\,\log(1-v) + \lambda\,v],$$

where λ is a constant and C some set in $\mathscr{S}$ and also set $g_{\epsilon,0}^{\mathbb{S}} := g_\epsilon$. If $C_\epsilon = \cap_{i=1}^k A_{i,\epsilon}^c$, one can exploit the independence of the increments of N in order to decompose the expected value as

$$\mathrm{e}^{-\psi_\nu(g_{\epsilon,0}^{C_\epsilon})} \prod_{i=1}^k (-1)^{n_i} \frac{\mathrm{d}^{n_i}}{\mathrm{d}\lambda^{n_i}}\ \mathrm{e}^{-\psi_\nu(g_{\epsilon,\lambda}^{A_{i,\epsilon}})}\Big|_{\lambda=0} = \mathrm{e}^{-\psi_\nu(g_\epsilon)} \prod_{i=1}^k V_{i,\epsilon}^{(n_i)}$$

where

$$V_{i,\epsilon}^{(n_i)} = \mathrm{e}^{\psi_\nu(g_{\epsilon,0}^{A_{i,\epsilon}})}\,(-1)^{n_i} \frac{\mathrm{d}^{n_i}}{\mathrm{d}\lambda^{n_i}}\ \mathrm{e}^{-\psi_\nu(g_{\epsilon,\lambda}^{A_{i,\epsilon}})}\Big|_{\lambda=0}.$$

In order to evaluate $V_{i,\epsilon}^{(n_i)}$ one can exploit the following recursive relation

$$V_{i,\epsilon}^{(n_i)} = \Lambda_0(A_{i,\epsilon}) \sum_{j=0}^{n_i-1} \binom{n_i-1}{j} \xi_{n_i-j}^{(i,\epsilon)}\,V_{i,\epsilon}^{(j)} = \Lambda_0(A_{i,\epsilon})\,\Delta_{i,\epsilon}^{(n_i)}$$

for any $n_i \ge 1$, where $V_{i,\epsilon}^{(0)} \equiv 1$ and, for any $m \ge 1$ and $i \in \{1,\dots,k\}$,

$$\xi_m^{(i,\epsilon)} = \int_{[0,1]\times A_{i,\epsilon}} v^m\,\mathrm{e}^{-g_\epsilon(v,t,x)}\,\rho(\mathrm{d}v|t)\frac{\Lambda_0(\mathrm{d}t,\mathrm{d}x)}{\Lambda_0(A_{i,\epsilon})}.$$

One observes that $\Delta_{i,\epsilon}^{(n_i)} = \xi_{n_i}^{(i,\epsilon)} + K_{i,\epsilon}$, where $K_{i,\epsilon}$ is such that $\lim_{\epsilon\downarrow 0}(K_{i,\epsilon}/\Lambda_0(A_{i,\epsilon})) = K_i < \infty$. Moreover,

$$\lim_{\epsilon\downarrow 0} \xi_{i,\epsilon}^{(n_i)} = \int_{[0,1]} \mathrm{e}^{-h(t_i^*,x_i^*)\,v}\,v^{n_i}\,(1-v)^{\bar{n}_{i+1}}\,\rho(\mathrm{d}v|t_i^*) =: \xi_i^{(n_i)}.$$

Hence, the numerator of the lower bound of (10) can be rewritten as follows

$$\mathrm{e}^{-\psi_\nu(g_\epsilon)} \prod_{i=1}^{k} \left\{ \Lambda_0(A_{i,\epsilon})\, \xi_{i,\epsilon}^{(n_i)} + o(\Lambda(A_{i,\epsilon})) \right\}, \tag{12}$$

where, as usual, $g(x) = o(f(x))$ as $x \to 0$ means that $\lim_{x\to 0}(g(x)/f(x)) = 0$. On the other hand, by (11), an upper bound for the denominator is

$$\mathbb{E}\left[\mathrm{e}^{-\int_{\mathbb{W}} -\bar{N}_{-\epsilon}(s)\log(1-v) N(\mathrm{d}v,\mathrm{d}s,\mathrm{d}x)} \prod_{i=1}^{k} \left(\int_{[0,1]\times A_{i,\epsilon}} v\, N(\mathrm{d}v,\mathrm{d}s,\mathrm{d}x) \right)^{n_i} \right]. \tag{13}$$

where $\bar{N}_{-\epsilon}(s) := \sum_{i=1}^{k} \bar{n}_i\, \mathbb{I}_{[t^*_{i-1}-\epsilon, t^*_i-\epsilon)}(s)$. One can then resort to arguments analogous to those leading to the lower bound for the numerator and re–express (13) as

$$\mathrm{e}^{-\psi_\nu(l_\epsilon)} \prod_{i=1}^{k} \left\{ \Lambda_0(A_{i,\epsilon}) \int_{[0,1]\times A_{i,\epsilon}} v^{n_i}\, (1-v)^{\bar{N}_{-\epsilon}(s)}\, \rho(\mathrm{d}v|s)\, \frac{\Lambda_0(\mathrm{d}s,\mathrm{d}x)}{\Lambda_0(A_{i,\epsilon})} + o(\Lambda(A_{i,\epsilon})) \right\} \tag{14}$$

having set $l_\epsilon(v,s,z) := \bar{N}_\epsilon(s)\log(1-v)$. Consider now the ratio of the lower bound for the numerator given in (12) and the upper bound for the denominator given in (14) and take the limit as $\epsilon \downarrow 0$. This yields a lower bound for the posterior Laplace functional of Λ (9) coinciding with

$$\mathrm{e}^{-\int_{\mathbb{W}}(1-\mathrm{e}^{-h(s,x)\,v})\,(1-v)^{\bar{N}(s)}\rho(\mathrm{d}v|s)\,\Lambda_0(\mathrm{d}s,\mathrm{d}x)} \prod_{i=1}^{k} \int_{[0,1]} \mathrm{e}^{-h(t^*_i,x^*_i)\,v} \frac{v^{n_i}\,(1-v)^{\bar{n}_{i+1}}\,\rho(\mathrm{d}v|t^*_i)}{\int_{[0,1]} v^{n_i}\,(1-v)^{\bar{n}_{i+1}}\,\rho(\mathrm{d}v|t^*_i)}$$

which agrees with with the posterior representation of Λ given in Point (i). The same result can be obtained by deriving an upper bound for (9) by means of (11) and letting $\epsilon \downarrow 0$. This completes the proof of Point (i).

As for Point (ii), from the posterior distribution of Λ one can easily deduce the posterior distribution of Z and, hence, of the SPNTR measure. Note that the posterior distribution of a linear functional of a random measure μ, conditional on a vector of observations $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$, is given by the linear functional of the posterior distribution of μ. In other terms, $(\int f(y)\mu(\mathrm{d}y)\,|\,\boldsymbol{Y}) \stackrel{d}{=} \int f(y)\,\mu^{(n)}(\mathrm{d}y)$, where $\mu^{(n)}$ is a random measure whose distribution coincides with the conditional distribution of μ, given $\boldsymbol{Y}$. Hence, the distribution of the background driving Poisson measure N in (3), given $(\boldsymbol{T^*}, \boldsymbol{X^*})$, is of the form $N^* + \sum_{i=1}^{k} \delta_{(J_i, T^*_i, X^*_i)}$ where N^* is a Poisson random measure with intensity (6). Consequently, the posterior distribution of Z, defined in terms of N via (4), corresponds to $Z^{(n)}(\mathrm{d}s,\mathrm{d}z) = \int_{[0,1]} -\log(1-v)\ [N^*(\mathrm{d}v,\mathrm{d}s,\mathrm{d}z) + \sum_{i=1}^{k} \delta_{(J_i,T^*_i,X^*_i)}(\mathrm{d}v,\mathrm{d}s,\mathrm{d}z)]$ which leads to the statements in Point (ii). Given $\Lambda^{(n)}$ and $Z^{(n)}$, (5) and (8) now easily follow. □

4 Posterior analysis of multiplicative intensity models

Here we consider a general model for multiplicative intensities as described, e.g., in [15]. Let $\mathbb{X}$ and $\mathbb{Y}$ be Polish spaces endowed with their Borel σ–algebra $\mathscr{X}$ and $\mathscr{Y}$, respectively. Consider a kernel K on $\mathbb{X} \times \mathbb{Y}$ taking values in $\mathbb{R}^+$ such that $y \mapsto K(C|y)$ is $\mathscr{Y}$–measurable for any $C \in \mathscr{X}$ and, for some σ–finite measure τ on $\mathbb{X}$, $\int_\cdot K(x,y)\tau(\mathrm{d}x)$ is a σ–finite measure on $\mathscr{B}(\mathbb{R}^+)$ for any y in $\mathbb{Y}$. Consider, now, the random intensity

$$\lambda(x) = \int_{\mathbb{Y}} K(x,y)\,\mu(\mathrm{d}y).$$

where μ is a general CRM on $(\mathbb{Y},\mathscr{Y})$ as defined in Section 2.1. Write the intensity of μ as $\nu(\mathrm{d}v,\mathrm{d}y) = \rho(\mathrm{d}v|y)\,\alpha(\mathrm{d}y)$ and assume α is non–atomic. Theorem 4.1 in [15] provides a full description of the posterior distribution of μ given the observations $\boldsymbol{X} = (X_1,\dots,X_n)$ generated by a multiplicative intensity model. Here we provide and alternative proof of this result based on techniques analogous to those exploited for the derivation of the posterior distribution of a SPNTR. As in [15], the likelihood function is given by

$$\mathscr{L}(\mu;\boldsymbol{x}) = \mathrm{e}^{-\int_{\mathbb{Y}} g_m(y)\,\mu(\mathrm{d}y)} \prod_{i=1}^{n} \int_{\mathbb{Y}} K(x_i,y)\,\mu(\mathrm{d}y), \tag{15}$$

where $m \ge n$ denotes the number of observations of which $\boldsymbol{x} = (x_1,\dots,x_n)$ are observable and the remaining $m-n$ are censored. The function $g_m(y) := \sum_{i=1}^{m} \int_{\mathbb{X}} U_i(x)\,K(x,y)\,\tau(\mathrm{d}x)$ is defined in terms of the predictable and observable processes $\{U_i(x) : x \in \mathbb{X}\}$, $i = 1,\dots,m$. Moreover, the kernel $K(\,\cdot\,,\,\cdot\,)$ is chosen in such a way that g_m is a non–negative and measurable function with bounded support on $\mathbb{Y}$. See [15] for a discussion on the generality of this model.

If we condition on the latent variables $\boldsymbol{Y} = (Y_1,\dots,Y_n)$, the likelihood above reduces to

$$\mathscr{L}(\mu;\boldsymbol{x},\boldsymbol{y}) = \mathrm{e}^{-\int_{\mathbb{Y}} g_m(y)\,\mu(\mathrm{d}y)} \prod_{i=1}^{n} K(x_i;y_i)\,\mu(\mathrm{d}y_i).$$

The above can be usefully rewritten so to take into account the fact that the Y_i's may feature some ties. In other terms, $k \le n$ latent variables are distinct and we denote their values by $Y_1^*,\dots,Y_k^*$. Correspondingly, one has

$$\mathscr{L}(\mu;\boldsymbol{x},\boldsymbol{y}) = \mathrm{e}^{-\int_{\mathbb{Y}} g_m(y)\,\mu(\mathrm{d}y)} \prod_{i=1}^{k} [\mu(\mathrm{d}y_i^*)]^{n_i} \prod_{j\in C_i} K(x_j;y_i^*)$$

where $C_i = \{r : y_r = y_i^*\}$. Before presenting our alternative proof of the posterior characterization, we introduce some useful notation. Let $\boldsymbol{\Pi}_n = (K_n, \boldsymbol{N}_{K_n})$ be a random vector with the number of classes and the frequencies of each class generated by a random partition of the set of integers $[n] = \{1,\dots,n\}$. In other terms, the realization $\boldsymbol{\Pi_n} = (k,n_1,\dots,n_k)$ corresponds to a partition of $[n]$ into k sets with respective frequencies $n_1,\dots,n_k$. Clearly, $k \in \{1,\dots,n\}$ and $\boldsymbol{n} = (n_1,\dots,n_k) \in A_{k,n}$ where $A_{k,n} := \{(n_1,\dots,n_k) : n_j \ge 1, \sum_{j=1}^{k} n_j = n\}$. Finally, set

$$\omega_{n_j}(y) = \int_{\mathbb{R}^+} v^{n_j}\,\mathrm{e}^{-v g_m(y)}\,\rho(\mathrm{d}v|y).$$

Proposition 2. [15, Th. 4.1] *Let $\lambda(x) = \int_{\mathbb{Y}} K(x,y)\,\mu(\mathrm{d}y)$ be a random intensity on $\mathbb{X}$ with $\nu(\mathrm{d}v,\mathrm{d}y) = \rho(\mathrm{d}v|y)\,\alpha(\mathrm{d}y)$ under the model* (15). *Then, given $\boldsymbol{X}$, the posterior distribution of λ can be characterized as follows:*

(i) *Given $\boldsymbol{Y}$ and $\boldsymbol{X}$, the conditional distribution of μ coincides with the distribution of the random measure*

$$\mu^{(n)} \stackrel{d}{=} \mu^* + \sum_{i=1}^{k} J_i\,\delta_{Y_j^*}$$

where μ^* is a CRM with intensity $\nu^*(\mathrm{d}v, \mathrm{d}y) = \mathrm{e}^{-v g_m(y)}\, \rho(\mathrm{d}v|y)\, \alpha(\mathrm{d}y)$ and, for $i = 1, \ldots, k$, Y_i^* is a fixed point of discontinuity with corresponding jump J_i distributed as

$$f_{J_i}(\mathrm{d}v) = \frac{v^{n_i}\, \mathrm{e}^{-v g_m(y_i^*)}\, \rho(\mathrm{d}v|y_i^*)}{\int_{\mathbb{R}^+} v^{n_i}\, \mathrm{e}^{-v g_m(y_i^*)}\, \rho(\mathrm{d}v|y_i^*)}.$$

Moreover, the J_i's are, conditionally on the Y_j^*'s, independent from μ^*;

(ii) *Conditionally on $\boldsymbol{X}$ and on $\boldsymbol{\Pi}_n = (k, \boldsymbol{n})$, the Y_j^* are independent and Y_j^* has distribution given by*

$$f_{Y_j^*}(\mathrm{d}y) = \frac{\omega_{n_j}(y)\, \prod_{i \in C_j} K(x_i, y)\alpha(\mathrm{d}y)}{\int_{\mathbb{Y}} \omega_{n_j}(y)\, \prod_{i \in C_j} K(x_i, y)\, \alpha(\mathrm{d}y)} \tag{16}$$

and, given $\boldsymbol{X}$, the conditional probability that $\boldsymbol{\Pi}_n = (k, \boldsymbol{n})$ is

$$\frac{\prod_{j=1}^k \int_{\mathbb{Y}} \omega_{n_j}(y)\, \prod_{i \in C_j} K(x_i, y)\, \alpha(\mathrm{d}y)}{\sum_{i=1}^n \sum_{\boldsymbol{n} \in A_{i,n}} \prod_{j=1}^i \int_{\mathbb{Y}} \omega_{n_j}(y)\, \prod_{i \in C_j} K(x_i, y)\, \alpha(\mathrm{d}y)} \tag{17}$$

Proof. Following arguments analogous to those employed for the evaluation of the posterior distribution for a SPNTR process prior, one can proceed to the determination of the posterior Laplace functional of μ, i.e. $\mathbb{E}[\mathrm{e}^{-\mu(f)}|\boldsymbol{X}]$, where $\mu(f) = \int_{\mathbb{Y}} f(y)\, \mu(\mathrm{d}y)$ for any measurable function $f : \mathbb{Y} \to \mathbb{R}^+$ such that $\int_{\mathbb{R}^+ \times \mathbb{Y}} (1 - \mathrm{e}^{-v\, f(y)})\rho(\mathrm{d}v|y)\alpha(\mathrm{d}y) < \infty$. To this end, suppose the vector $\boldsymbol{y}$ consists of k distinct observations $y_1^*, \ldots, y_k^*$ and let $y_i = y_{j(i)}^*$ where $j(i) \in \{1, \ldots, k\}$. If $B_\epsilon(x)$ stands for the ball of radius ϵ around point x, denote by $A_i^\epsilon(k)$ the rectangle $B_\epsilon(x_i) \times B_\epsilon(y_{j(i)}^*)$. Then

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{-\mu(f)} \,\middle|\, \boldsymbol{X}, \boldsymbol{Y}\right] &= \lim_{\epsilon \downarrow 0} \mathbb{E}\left[\mathrm{e}^{-\mu(f)} \,\middle|\, (X_i, Y_i) \in A_i^\epsilon(k),\ i = 1, \ldots, k\right] \\
&= \lim_{\epsilon \downarrow 0} \frac{\mathbb{E}\left[\mathrm{e}^{-\mu(f+g_m)} \prod_{j=1}^k \prod_{i \in C_j} \int_{B_\epsilon(x_i)} \int_{B_\epsilon(y_j^*)} K(x, y)\, \mu(\mathrm{d}y)\, \tau(\mathrm{d}x)\right]}{\mathbb{E}\left[\mathrm{e}^{-\mu(g_m)} \prod_{j=1}^k \prod_{i \in C_j} \int_{B_\epsilon(x_i)} \int_{B_\epsilon(y_j^*)} K(x, y)\, \mu(\mathrm{d}y)\, \tau(\mathrm{d}x)\right]}
\end{aligned}$$

As far as the numerator is concerned, one notes that it can be rewritten as

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{e}^{-\int_{\mathbb{Y}^*} (f(y)+g_m(y))\, \mu(\mathrm{d}y)}\right] \times \\
&\quad \times \prod_{j=1}^k \mathbb{E}\left[\mathrm{e}^{-\int_{B_\epsilon(y_j^*)} [f(y)+g_m(y)]\, \mu(\mathrm{d}y)} \prod_{i \in C_j} \int_{B_\epsilon(y_j^*)} \int_{B_\epsilon(x_i)} K(x, y)\, \tau(\mathrm{d}x)\, \mu(\mathrm{d}y)\right]
\end{aligned} \tag{18}$$

and $\mathbb{Y}^* = (\cup_{j=1}^k B_\epsilon(y_j^*))^c$. For the moment it is useful to set $\gamma_{i,\epsilon}(y) := \int_{B_\epsilon(x_i)} K(x, y)\, \tau(\mathrm{d}x)$ for any $y \in \mathbb{Y}$ and $i \in C_j$ and introduce the function $q_{j,\epsilon}^{\boldsymbol{\lambda}} = \left\{f + g_m + \sum_{i=1}^{n_j} \lambda_i\, \gamma_{i,\epsilon}\right\} \mathbb{I}_{B_\epsilon(y_j^*)}$. Then, recall that n_j is the cardinality of the set of indices and, for simplicity, assume that $C_j = \{1, \ldots, n_j\}$. One notes that (18) also coincides with

$$\mathrm{e}^{-\psi_\nu([f+g_m]\, \mathbb{I}_{\mathbb{Y}^*})} \prod_{j=1}^k (-1)^{n_j} \frac{\partial^{n_j}}{\partial \lambda_1 \cdots \partial \lambda_{n_j}}\, \mathrm{e}^{-\psi_\nu(q_{j,\epsilon}^{\boldsymbol{\lambda}})} \bigg|_{\boldsymbol{\lambda}=\boldsymbol{0}} = \mathrm{e}^{-\psi_\nu(f+g_m)} \prod_{j=1}^k V_{j,\epsilon}^{(n_j)}$$

where $\mathbf{0} = (0, \dots, 0)$ is an n_j–dimensional vector and

$$V_{j,\epsilon}^{(n_j)} = \mathrm{e}^{\psi_\nu(q_{j,\epsilon}^{\mathbf{0}})}\,(-1)^{n_j}\,\frac{\partial^{n_j}}{\partial\lambda_1\,\cdots\,\partial\lambda_{n_j}}\,\mathrm{e}^{-\psi_\nu(q_{j,\epsilon}^{\boldsymbol{\lambda}})}\,\Big|_{\boldsymbol{\lambda}=\mathbf{0}}.$$

The recursive formula recalled in the proof of Proposition 1 applies in this case as well yielding

$$V_{j,\epsilon}^{(n_j)} = \alpha(B_\epsilon(y_j^*))\,\sum_{i=0}^{n_j-1}\binom{n_j-1}{i}\zeta_{n_j-i}^{(j,\epsilon)}\,V_{j,\epsilon}^{(i)} = \alpha(B_\epsilon(y_j^*))\,\Delta_{j,\epsilon}^{(n_j)}$$

where $\Delta_{j,\epsilon}^{(n_j)} = \sum_{i=0}^{n_j-1}\binom{n_j-1}{i}\zeta_{n_j-i}^{(j,\epsilon)}\,V_{j,\epsilon}^{(i)}$ and, for any $r \in \{1, \dots, n_j\}$,

$$\zeta_r^{(j,\epsilon)} = \int_{B_\epsilon(y_j^*)\times\mathbb{R}^+}\left(\prod_{i=1}^{r}\gamma_i(y)\right)\,v^r\,\mathrm{e}^{-v\,q_{j,\epsilon}^{\mathbf{0}}(y)}\,\rho(\mathrm{d}v|y)\,\frac{\alpha(\mathrm{d}y)}{\alpha(B_\epsilon(y_j^*))}.$$

Hence one can again conclude that

$$V_{j,\epsilon}^{(n_j)} = \alpha(B_\epsilon(y_j^*))\,\zeta_j^{(n_j)} + o(\alpha(B_\epsilon(y_j^*)))$$

as $\epsilon \downarrow 0$, where $\zeta_j^{(n_j)} = \prod_{i\in C_j}\gamma_i(y_j^*)\,\int_{\mathbb{R}^+} v^{n_j}\,\mathrm{e}^{-v[f(y_j^*)+g_m(y_j^*)]}\,\rho(\mathrm{d}v|y_j^*)$ and, for any i in C_j, $\gamma_i(y_j^*) = K(x_i, y_j^*)\,\tau(\mathrm{d}x_i)$. This implies that

$$\mathbb{E}\left[\mathrm{e}^{-\mu(f)}\,\middle|\,\boldsymbol{X},\boldsymbol{Y}\right] = \frac{\mathbb{E}\left[\mathrm{e}^{-\mu(f+g_m)}\right]\prod_{j=1}^{k}\int_{\mathbb{R}^+}\mathrm{e}^{-v\,f(y_j^*)}\,v^{n^j}\,\mathrm{e}^{-vg_m(y_j^*)}\rho(\mathrm{d}v|y_j^*)}{\mathbb{E}\left[\mathrm{e}^{-\mu(g_m)}\right]\,\prod_{j=1}^{k}\int_{\mathbb{R}^+} v^{n^j}\,\mathrm{e}^{-vg_m(y_j^*)}\rho(\mathrm{d}v|y_j^*)}$$

from which the statement in (i) follows.

One can analogously determine the conditional distribution of $\boldsymbol{Y}$, given $\boldsymbol{X}$ and $\boldsymbol{\Pi}_n$. Indeed, if $\boldsymbol{n} = (n_1, \dots, n_k)$, one has

$$\begin{aligned}(19)\qquad & P\left[\bigcap_{i=1}^{n}\left\{Y_i \in B_\epsilon(y_{j(i)}^*)\right\}\,\cap\,\bigcap_{i=1}^{n}\{X_i \in B_\epsilon(x_i)\}\,\cap\,\{\boldsymbol{\Pi}_n = (k,\boldsymbol{n})\}\right]\\ &= \mathbb{E}\left[\mathrm{e}^{-\mu(g_m)}\,\prod_{j=1}^{k}\prod_{i\in C_j}\int_{B_\epsilon(y_j^*)\times B_\epsilon(x_i)}K(x,y)\,\tau(\mathrm{d}x)\,\mu(\mathrm{d}y)\right]\\ &= \mathrm{e}^{-\psi_n u(g_m\,\mathbb{I}_{\mathbb{Y}_\epsilon^*})}\,\prod_{j=1}^{k}\mathbb{E}\left[\mathrm{e}^{-\mu(g_m\,\mathbb{I}_{B_\epsilon(y_j^*)})}\,\prod_{i\in C_j}\int_{B_\epsilon(y_j^*)\times B_\epsilon(x_i)}K(x,y)\,\tau(\mathrm{d}x)\,\mu(\mathrm{d}y)\right].\end{aligned}$$

Suppose again $C_j = \{1, \dots, n_j\}$ and, after setting $h_{j,\epsilon}^{\boldsymbol{\lambda}} = [g_m + \sum_{i=1}^{n_j}\lambda_i\,\gamma_{i,\epsilon}]\,\mathbb{I}_{B_\epsilon(y_j^*)}$, one can proceed in a similar fashion as before to obtain the following representation for the probability in (19)

$$\mathrm{e}^{-\psi_\nu(g_m\,\mathbb{I}_{\mathbb{Y}_\epsilon^*})}\,\prod_{j=1}^{k}(-1)^{n_j}\,\frac{\partial^{n_j}}{\partial\lambda_1\,\dots\,\partial\lambda_{n_j}}\,\mathbb{E}\left[\mathrm{e}^{-\mu(h_{j,\epsilon}^{\boldsymbol{\lambda}})}\right]\,\Big|_{\boldsymbol{\lambda}=\mathbf{0}} = \mathrm{e}^{-\psi_\nu(g_m)}\,\prod_{j=1}^{k}W_{j,\epsilon}^{(n_j)}$$

where $\mathbb{Y}_\epsilon^* = (\cup_{j=1}^{k}B_\epsilon(y_j^*))^c$,

$$\begin{aligned}W_{j,\epsilon}^{(n_j)} &= \mathrm{e}^{\psi_\nu(h_{j,\epsilon}^{\mathbf{0}})}\,(-1)^{n_j}\,\frac{\partial^{n_j}}{\partial\lambda_1\,\dots\,\partial\lambda_{n_j}}\,\mathrm{e}^{-\psi_\nu(h_{j,\epsilon}^{\boldsymbol{\lambda}})}\,\Big|_{\boldsymbol{\lambda}=\mathbf{0}}\\ &= \alpha(B_\epsilon(y_j^*))\,\sum_{i=0}^{n_j-1}\binom{n_j-1}{i}\,\zeta_{n_j-i}^{(j,\epsilon)}\,W_{j,\epsilon}^{(i)}\end{aligned}$$

and

$$\zeta_r^{(j,\epsilon)} = \int_{B_\epsilon(y_j^*)} \left(\prod_{i=1}^r \gamma_{i,\epsilon}(y) \right) \int_{\mathbb{R}^+} v^r \, \mathrm{e}^{-v g_m(y)} \, \rho(\mathrm{d}v|y) \, \frac{\alpha(\mathrm{d}y)}{\alpha(B_\epsilon(y_j^*))}.$$

Let δ denote the counting measure on $\{1, \ldots, n\} \times (\cup_{k=1}^n A_{k,n})$, so that the density of the vector $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{\Pi}_n)$ is absolutely continuous with respect to $\alpha^k \times \tau^n \times \delta$. The corresponding density evaluated at $(\boldsymbol{y}^*, \boldsymbol{x}, \boldsymbol{\pi}_n)$ coincides with

$$f(\boldsymbol{y}^*, \boldsymbol{x}, \boldsymbol{\pi}_n) = \mathrm{e}^{-\psi_\nu(g_m)} \prod_{j=1}^k \left\{ \prod_{i \in C_j} K(x_i, y_j^*) \right\} \int_{\mathbb{R}^+} v^{n_j} \, \mathrm{e}^{-v g_m(y_j^*)} \, \rho(\mathrm{d}v|y_j^*)$$

where $\boldsymbol{\pi}_n = (k, n_1, \ldots, n_k)$. At this point, one can easily determine the distribution of $\boldsymbol{Y}$ given $(\boldsymbol{X}, \boldsymbol{\Pi}_n)$, which admits density on $\mathbb{Y}^k$ (with respect to α^k) coinciding with

$$f(\boldsymbol{y}^* \,|\, \boldsymbol{x}, \boldsymbol{\pi}_n) \propto \prod_{j=1}^k \omega_{n_j}(y_j^*) \prod_{i \in C_j} K(x_i, y_j^*)$$

from which (16) and (17) follow. $\square$

Acknowledgements

A. Lijoi and I. Prünster are partially supported by MiUR, grants 2006/ 134525 and 2006/133449, respectively.

References

[1] Aalen, O. (1978). Nonparametric Inference for a Family of Counting Processes. *Ann. Statist.* **6**, 701—726.

[2] Constantine, G.M. and Savits, T.H. (1994). A stochastic process interpretation of partition identities. *SIAM J. Discrete Math.* **7**, 194–202.

[3] Daley, D. J. and Vere-Jones, D. (1988). *An introduction to the theory of point processes.* Springer, New York.

[4] De Blasi, P. and Hjort, N.L. (2007). Proportional hazard regression models. Bayesian analysis using Beta processes. *Scand. J. Statist.* **34**, 229–257.

[5] Doksum, K. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probab.*, **2**, 183–201.

[6] Dykstra, R.L. and Laud, P.W. (1981). A Bayesian nonparametric approach to reliability. *Ann. Statist.* **9**, 356–367.

[7] Epifani, I., Lijoi, A. and Prünster, I. (2003). Exponential functionals and means of neutral-to-the-right priors. *Biometrika*, **90**, 791–808.

[8] Ferguson, T.S. and Phadia, E.G. (1979). Bayesian nonparametric estimation based on censored data. *Ann. Statist.*, **7**, 163–186.

[9] Ferguson, T.S. and Phadia, E.G. (1979). Bayesian nonparametric estimation based on censored data. *Ann. Statist.* **7**, 163-186.

[10] Hjort, N.L. (1990). Nonparametric Bayes estimators based on beta processes in models for life history data. *Ann. Statist.*, **18**, 1259–1294.

[11] Ishwaran, H. and James, L. F. (2004). Computational methods for multiplicative intensity models using weighted gamma processes: Proportional hazards, marked point processes, and panel count data. *J. Amer. Stat. Assoc.* **99**, 175–190.

[12] James, L.F. (2002). Poisson Process Partition Calculus with applications to exchangeable models and Bayesian Nonparametrics. Manuscript. *MatharXiv* arXiv:math/0205093v1.

[13] James L.F. (2003a). Bayesian calculus for gamma processes with applications to semiparametric intensity models. *Sankhya* **65** (2003), 179–206.

[14] James, L.F. (2003b). A simple proof of the almost sure discreteness of a class of random measures. *Statist. Probab. Lett.*, **65**, 363–368.

[15] James L.F. (2005). Bayesian Poisson process partition calculus with an application to Bayesian Lévy moving averages. *Ann. Statist.* **33**, 1771–1799.

[16] James, L.F. (2006). Poisson calculus for spatial neutral to the right processes. *Ann. Statist.* **34**, 416–440.

[17] James, L.F., Lijoi, A. and Prünster, I. (2008). Posterior analysis for normalized random measures with independent increments. *Scand. J. Statist.*, in press.

[18] Kingman, J.F.C. (1967). Completely random measures. *Pacific J. Math.*, **21**, 59–78.

[19] Kingman, J.F.C. (1993). *Poisson processes.* Oxford University Press, Oxford.

[20] Kim, Y. (1999). Nonparametric Bayesian estimators for counting processes. *Ann. Statist.* **27**, 562–588.

[21] Kim, Y. and Lee, J. (2003). Bayesian analysis of proportional hazard models. *Ann. Statist.* **31**, 493–511.

[22] Lijoi, A., and Prünster, I. (2004). A note on the problem of heaps. *Sankhya* **66**, 234–242.

[23] Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates: I Density estimates. *Ann. Statist.* **12**, 351–357.

[24] Lo, A. Y. and Weng, C.–S. (1989). On a class of Bayesian nonparametric estimates. II. Hazard rate estimates. *Ann. Inst. Statist. Math.* **41**, 227—245.

[25] Nieto-Barajas, L.E. and Walker, S.G. (2004). Bayesian nonparametric survival analysis via Lévy driven Markov processes. *Statist. Sinica* **14**, 1127–1146.

[26] Pitman, J. (2006). *Combinatorial stochastic processes.* Ecole d'Et de Probabilits de Saint-Flour XXXII–2002. Lecture Notes in Mathematics N. 1875, Springer, New York.

[27] Prünster, I. (2002). *Random probability measures derived from increasing additive processes and their application to Bayesian statistics.* Ph.d thesis, University of Pavia.

[28] Regazzini, E., Lijoi, A. & Prünster, I. (2003). Distributional results for means of random measures with independent increments. *Ann. Statist.* **31**, 560–585.

[29] Walker, S. and Damien, P. (1998). A full Bayesian non-parametric analysis involving a neutral to the right process. *Scand. J. Statist.*, **25**, 669–680.

[30] Walker, S. and Muliere, P. (1997). Beta-Stacy processes and a generalization of the Pólya-urn scheme. *Ann. Statist.*, **25**, 1762–1780.

[31] Wolpert, R. and Ickstadt, K. (1998). Poisson/gamma random field models for spatial statistics. *Biometrika* **85**, 251–267.