Statistica Sinica **18**(2008), 1653-1668

INVESTIGATING NONPARAMETRIC PRIORS WITH GIBBS STRUCTURE

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

[1] *Universit'a di Pavia and CNR-IMATI Milano,*
[2] *Universit'a di Torino, Collegio Carlo Alberto and ICER*
and [3] *University of Kent*

Abstract: This paper investigates nonparametric priors that induce infinite Gibbs–type partitions; such a feature is desirable both from a conceptual and a mathematical point of view. Recently it has been shown that Gibbs–type priors, with $\sigma \in (0,1)$, are equivalent to σ–stable Poisson–Kingman models. By looking at solutions to a recursive equation arising through Gibbs partitions, we provide an alternative proof of this fundamental result. Since practical implementation of general σ–stable Poisson–Kingman models is difficult, we focus on a related class of priors, namely normalized random measures with independent increments; these are easily implementable in complex Bayesian models. We establish the result that the only Gibbs–type priors within this class are those based on a generalized gamma random measure.

Key words and phrases: Bayesian nonparametrics, Gibbs exchangeable partitions, generalized gamma process, normalized random measures with independent increments, recursive equation, stable distribution.

1. Introduction

Recently a number of generalizations of the Dirichlet process (Ferguson (1973)) appeared in the literature. Among them we mention species sampling models (Pitman (1996)), stick–breaking priors (Hjort (2000) and Ishwaran and James (2001)), normalized random measures with independent increments (Regazzini, Lijoi and Prünster (2003)), Poisson–Kingman models (Pitman (2003)), spatial neutral to the right models (James (2006)), neutral to the right species sampling models (James (2007)). The last two works are based on neutral to the right processes due to Doksum (1974). All these classes of priors share with the Dirichlet process the so–called "discreteness" property, i.e., they select almost surely discrete distributions. Special cases have been successfully exploited in the context of Bayesian semiparametric models. See Hjort (2003) and Müller and Quintana (2004) for a review of complex Bayesian models involving a nonparametric component. The richer clustering structure induced by these new priors turns out to have some advantages over the widely used Dirichlet process mixtures (Lo (1984)) in certain applications. See, e.g., Ishwaran and James (2001, 2003), Lijoi, Mena and Prünster (2005a,b, 2007a). Moreover, novel applications to data sequences with ties have also been proposed: for instance, Goldwater, Griffiths, and Johnson (2006) and Teh (2006a,b) apply them to language models, and Lijoi, Mena and Prünster (2007b) to genomic libraries. Based on the usefulness they have demonstrated in applications, the structural properties of these novel priors deserve thorough investigations.

By close inspection of the alternatives to the Dirichlet process, one can see that many specific tractable priors lead to a system of predictive distributions of the type

$$\mathbb{P}\left[X_{n+1} \in A \,\middle|\, X_1,\dots,X_n\right] = g_0(n,k)P_0(A)+g_1(n,k)\sum_{j=1}^{k}(n_j-\sigma)\delta_{X_j^*}(A) \quad (1.1)$$

for any $A \in \mathscr{X}$, $n \geq 1$, $\sigma \in [0,1)$, where $X_1^*,\dots,X_k^*$ are the k distinct observations in the n–sample with frequencies n_j, and P_0 represents the prior guess at the shape of the random distribution. Such random probability measure can be a termed *Gibbs–type* random probability measure since it generates a sequence $(\tilde{\mathscr{P}}_n)_{n\geq 1}$ of exchangeable random partitions of *Gibbs type*, as defined by Gnedin and Pitman (2005). See also Pitman (2006). Note that when $\sigma = 0$, (1.1) reduces to the predictive distributions of the Dirichlet process. It is well–known that such predictives characterize the Dirichlet process prior; see Regazzini (1978) and Lo (1991). Other examples of Gibbs–type priors are the normalized stable process (Kingman (1975)), the two–parameter Poisson Dirichlet process (Pitman (1996)), and the normalized inverse Gaussian process (Lijoi, Mena and Prünster (2005a)). In the following we focus on Gibbs–type priors with $\sigma \in (0,1)$.

From (1.1) the mathematical tractability of Gibbs–type priors is apparent; indeed, the prediction rule can be seen as resulting from a two step procedure: the $(n+1)$th observation X_{n+1} is either "new" (i.e., not coinciding with any of the previously observed X_i^*'s) or "old", with probability depending on n and k but not on the frequencies n_i's. Given X_{n+1} is "new", it is sampled from P_0. Given X_{n+1} is "old" (i.e., X_{n+1} is equal to one of the already sampled X_i^*'s), it will coincide with a particular X_j^* with probability $(n_j-\sigma)/(n-k\sigma)$. The fact that the assignment to "new" or "old" does not depend on the frequencies n_i's is crucial in simplifying both analytical calculations and the derivation of suitable sampling schemes in complex models. See Ishwaran and James (2003) for connections with Chinese restaurant process sampling schemes.

In this paper we investigate Gibbs–type priors by finding solutions to a specific recursive equation. With this approach we first provide an alternative proof of the characterization of Gibbs–type priors due to Gnedin and Pitman (2005); this deep result essentially states that Gibbs-type priors are Poisson–Kingman models based on the stable subordinator. Their proof resorts to combinatorial techniques, in particular, generalized Stirling triangles asymptotics; see also Pitman (2003, 2006). Given this result, we face the problem of determining which normalized random measures with independent increments (NRMI) are also Gibbs–type priors. The answer turns out to be that a NRMI belongs to the class of Gibbs–type priors if and only if it is a normalized generalized gamma process (whose definition is provided in Section 2.2). From a statistical perspective this result is of great relevance; typically, when defining new specific priors, one first aims at deriving the posterior distribution and then obtains the predictive distributions as the expected value of the posterior distribution. In the case of NRMI, which are uniquely characterized by the Poisson intensity measure, the posterior and predictive distributions have been derived in James, Lijoi and Prünster (2008) as function of the Poisson intensity. Hence, it is natural to ask which choices of the Poisson intensity lead to a NRMI that is also of Gibbs–type. Indeed, given the analytical and computational advantages of a Gibbs structure, a precise answer to this question, as we give in this paper, is of great importance. As a consequence of our result, if one is looking for a new tractable prior with Gibbs structure, attention has to be focused on Poisson–Kingman models based on the stable process; unfortunately, this class, in its generality, is much less tractable than NRMI and, hence, the search for new specific examples of Gibbs–type priors which admit explicit representations will be a difficult task. For appreciating the complexity of the posterior distributions of Poisson–Kingman models, the reader is referred to the characterizations provided in James (2002).

In Section 2 we set the framework for our analysis, define the classes of priors with which we will deal, and review some relevant results and concepts. Section 3 states the fundamental result of Gnedin and Pitman (2005) and characterizes the normalized generalized gamma process as the only NRMI of Gibbs type. In the Appendix we provide an alternative derivation of the result of Gnedin and Pitman (2005).

2. Gibbs–Type Priors and Related Random Distributions

Throughout the paper we consider the following setup. Let $(X_n)_{n\geq 1}$ be a sequence of exchangeable $\mathbb{X}$–random variables defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$, where $\mathbb{X}$ is a separable and complete metric space and $\mathscr{X}$ the corresponding Borel σ–field. In other terms, we suppose there exists some random probability measure, say $\tilde{P}$, such that

$$\mathbb{P}\left[X_1 \in A_1, \ldots, X_n \in A_n \,\middle|\, \tilde{P}\right] = \prod_{i=1}^{n} \tilde{P}(A_i)$$

for any $n \geq 1$ and $A_1, \ldots, A_n \in \mathscr{X}$, the law of $\tilde{P}$ acting as a nonparametric prior. For any $n \geq 1$, the joint law of $X^{(n)} = (X_1, \ldots, X_n)$ is then

$$\mathbb{P}\left[X_1 \in \mathrm{d}x_1, \ldots, X_n \in \mathrm{d}x_n\right] = \mathbb{E}\left[\tilde{P}(\mathrm{d}x_1) \cdots \tilde{P}(\mathrm{d}x_n)\right]. \tag{2.1}$$

In the sequel, we consider $\tilde{P}$ which are, almost surely, discrete and with non–atomic prior guess at the shape $P_0(A) := \mathbb{E}[\tilde{P}(A)]$ for any $A \in \mathscr{X}$. Given the discreteness, we expect ties in the sample, namely that $X^{(n)}$ contains $k \leq n$ distinct observations $X_1^*, \ldots, X_k^*$ with frequencies $n_1, \ldots, n_k$, respectively. By integrating (2.1) with respect to all n–samples with k distinct observations having frequencies $n_1, \ldots, n_k$, we obtain the joint distribution of the number of distinct observations K_n and of the vector of frequencies $(N_{1,n}, \ldots, N_{k,n})$

$$\mathbb{P}\left[\{K_n = k\} \cap \{N_{j,n} = n_j,\ j = 1, \ldots, k\}\right] = \Pi_k^{(n)}(n_1, \ldots, n_k),$$

known as *exchangeable partition probability function* (EPPF) and extensively studied in Pitman (1996, 2003, 2006). Note that the EPPF satisfies the addition rule

$$\Pi_k^{(n)}(n_1 \ldots, n_k) = \sum_{j=1}^{k} \Pi_k^{(n+1)}(n_1, \ldots, n_j + 1, \ldots, n_k) + \Pi_{k+1}^{(n+1)}(n_1, \ldots, n_k, 1).$$

Moreover the EPPF identifies the law of an exchangeable random partition $\tilde{\mathscr{P}}$ of the set of integers $\mathbb{N}$. Here by random partition of $\mathbb{N}$ we mean a sequence $(\tilde{\mathscr{P}}_n)_{n\geq 1}$, where $\tilde{\mathscr{P}}_n$ is a random partition of the set of integers $[n] = \{1, \ldots, n\}$, for each $n \geq 1$. The sequence must be consistent, in the sense that $\tilde{\mathscr{P}}_n$ is the partition resulting from $\tilde{\mathscr{P}}_{n+1}$, after dropping the integer $n+1$. For a stimulating account on random partitions, the reader is referred to Pitman (2006).

2.1. Gibbs–type priors

We focus attention on random probability measures inducing *Gibbs–type random partitions* (Gnedin and Pitman (2005)) that are characterized by EPPF's of the form

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = V_{n,k} \prod_{j=1}^{k} (1 - \sigma)_{n_j - 1} \tag{2.2}$$

for some $\sigma \in (0, 1)$ and some set of non–negative weights $\{V_{n,k} :\ n \geq 1,\ 1 \leq k \leq n\}$. Here $(c)_n = \Gamma(c + n)/\Gamma(c)$ stands for the Pochhammer symbol.

Definition 1. A discrete random probability measure governing a sequence of exchangeable observations $(X_n)_{n\geq 1}$ is a *Gibbs–type prior* if it can be expressed as (2.2).

The predictive distributions associated to a Gibbs–type prior $\tilde{P}$ are, indeed, of the form (1.1), and the weights can be expressed as

$$g_0(n,k) = \frac{V_{n+1,k+1}}{V_{n,k}} \qquad\qquad g_1(n,k) = \frac{V_{n+1,k}}{V_{n,k}}.$$

Crucial for our treatment is the fact, shown in Gnedin and Pitman (2005), that the $V_{n,k}$'s have to satisfy the recursive equation

$$V_{n,k} = (n - \sigma k)V_{n+1,k} + V_{n+1,k+1} \tag{2.3}$$

for any $k = 1, \ldots, n$ and $n \geq 1$ with $V_{1,1} = 1$. This recursive equation arises from the fact that probabilities sum to one; that is given n observations, which produce k groups of sizes $(n_1, \ldots, n_k)$, the probability that the next observation provides a new group is $V_{n+1,k+1}/V_{n,k}$ and the probability the next observation is from the group of size n_j is $(n_j - \sigma)V_{n+1,k}/V_{n,k}$. Then it must be that $1 = V_{n+1,k+1}/V_{n,k} + \sum_{j=1}^{k}(n_j - \sigma)V_{n+1,k}/V_{n,k}$, which is tantamount to (2.3). From Lemma 2 in Gnedin and Pitman (2005) one has that a Gibbs type prior is uniquely characterized by a set of weights $\{V_{n,k} : \; n \geq 1, \; 1 \leq k \leq n\}$ and a prior guess P_0 at the shape of $\tilde{P}$.

In the following we explore which priors induce a Gibbs–type partition structure: we will see that Gibbs priors are closely connected with Poisson–Kingman models and NRMI. To this end we first define these two classes of priors. Let us start from the latter, which was introduced in Regazzini, Lijoi and Prünster (2003), and extended to general spaces in James (2002). Further results and characterizations have been provided in James, Lijoi and Prünster (2006, 2008).

2.2. NRMIs

Let $\mathbb{X}$ be some complete separable metric space equipped with the usual Borel σ–field $\mathscr{X}$. A *completely random measure* (Kingman (1967)) on $\mathbb{X}$ defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is a random measure $\tilde{\mu}$ such that

(i) $\tilde{\mu}(\emptyset) = 0$ a.s.-$\mathbb{P}$
(ii) for any collection of disjoint sets $A_1, A_2, \ldots$, the random variables $\tilde{\mu}(A_1)$, $\tilde{\mu}(A_2), \ldots$ are mutually independent and $\tilde{\mu}(\cup_{j\geq 1} A_j) = \sum_{j\geq 1} \tilde{\mu}(A_j)$ holds true a.s.-$\mathbb{P}$.

The Laplace transform of $\tilde{\mu}(C)$ for any measurable C is given by

$$\mathbb{E}[\mathrm{e}^{-\lambda\tilde{\mu}(C)}] = \mathrm{e}^{-\int_{\mathbb{R}^+ \times C}[1-\mathrm{e}^{-\lambda v}]\nu(\mathrm{d}v,\mathrm{d}x)}, \qquad \lambda \geq 0, \tag{2.4}$$

where ν stands for the Poisson intensity measure which uniquely characterizes $\tilde{\mu}$. Moreover,

$$\psi_C(\lambda) := \int_{\mathbb{R}^+ \times C} \left[1 - \mathrm{e}^{-\lambda s}\right] \nu(\mathrm{d}s, \, \mathrm{d}x) \tag{2.5}$$

is the so–called Laplace exponent of $\tilde{\mu}(C)$. See Kingman (1993) for a detailed account of completely random measures. According to Regazzini, Lijoi and Prünster (2003), a NRMI can be introduced as follows.

Definition 2. The measure

$$\tilde{P}(\,\cdot\,) = \frac{\tilde{\mu}(\,\cdot\,)}{\tilde{\mu}(\mathbb{X})}. \tag{2.6}$$

is a *NRMI* on $(\mathbb{X}, \mathscr{X})$ if $\tilde{\mu}$ is a completely random measure whose intensity measure ν is such that the Laplace exponent $\psi(\lambda) := \psi_{\mathbb{X}}(\lambda)$ is finite, for any positive λ, and $\nu(\mathbb{R}^+ \times \mathbb{X}) = +\infty$.

The finiteness of $\psi_{\mathbb{X}}(\lambda)$ and $\nu(\mathbb{R}^+ \times \mathbb{X}) = +\infty$ are required in order to ensure that the total mass $\tilde{\mu}(\mathbb{X})$ is finite and positive, almost surely. For details see Regazzini, Lijoi and Prünster (2003).

According to the particular decomposition of the intensity measure ν, one can distinguish two subclasses of NRMIs. Letting $H = \theta\, P_0$, where $0 < \theta < \infty$ and P_0 is a non–atomic probability distribution on $\mathbb{X}$, we have:

(a) if $\nu(\mathrm{d}v, \mathrm{d}x) = \rho(\mathrm{d}v)\, H(\mathrm{d}x)$, for some measure ρ on $\mathbb{R}^+$, then $\tilde{\mu}$ and the induced NRMI $\tilde{P}$ are *homogeneous*;
(b) if $\nu(\mathrm{d}v, \mathrm{d}x) = \rho(\mathrm{d}v|x)\, H(\mathrm{d}x)$, where $\rho(\,\cdot\,|x)$ is a measure on $\mathbb{R}^+$ for any x in $\mathbb{X}$, then $\tilde{\mu}$ and the induced NRMI $\tilde{P}$ are *non–homogeneous*.

Recall that, as shown e.g., in James (2003), NRMI select almost surely discrete distributions. James, Lijoi and Prünster (2008) obtained the EPPF of a NRMI as

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \frac{1}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1}\, \mathrm{e}^{-\psi(u)} \Big[\prod_{j=1}^k \int_{\mathbb{X}} \tau_{n_j}(u|x)\, H(\mathrm{d}x) \Big]\, \mathrm{d}u, \tag{2.7}$$

where, for $i = 1, \ldots, k$,

$$\tau_{n_i}(u|Y_i) = \int_{\mathbb{R}^+} v^{n_i}\, \mathrm{e}^{-uv}\, \rho(\mathrm{d}v|Y_i). \tag{2.8}$$

The EPPF of a homogeneous NRMI, derived in Pitman (2003), is of the form

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \frac{\theta^k}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1}\, \mathrm{e}^{-\psi(u)} \Big[\prod_{j=1}^k \tau_{n_j}(u) \Big]\, \mathrm{d}u. \tag{2.9}$$

An important case of NRMI, relevant for subsequent developments, is the *generalized gamma NRMI*, which is derived from the generalized gamma process studied in Brix (1999). Such a NRMI is homogeneous and characterized by a Poisson intensity measure of the type

$$\nu(\mathrm{d}v, \mathrm{d}x) = \frac{\mathrm{e}^{-\tau v}}{\Gamma(1-\sigma)\, v^{1+\sigma}} \mathrm{d}v H(\mathrm{d}x), \tag{2.10}$$

where $\tau \in (0,\infty)$ and $\sigma \in (0,1)$. Note that it includes, as a special case, the stable NRMI due to (Kingman (1975)), which arises by setting $\tau = 0$ in (2.10). The EPPF of the generalized gamma NRMI is

$$\Pi_k^{(n)}(n_1,\ldots,n_k) = \frac{\mathrm{e}^{\beta}\, \sigma^{k-1}}{\Gamma(n)} \sum_{i=0}^{n-1} \binom{n-1}{i} (-1)^i \beta^{\frac{i}{\sigma}}\, \Gamma\Big(k - \frac{i}{\sigma}; \beta\Big) \prod_{i=1}^{k} (1-\sigma)_{n_i-1}, \tag{2.11}$$

where $\beta = \theta\tau^{\sigma}/\sigma$ and $\Gamma(\,\cdot\,;\,\cdot\,)$ is the incomplete Gamma function. By setting $\tau = 0$, (2.11) simplifies to

$$\Pi_k^{(n)}(n_1,\ldots,n_k) = \frac{\sigma^{k-1}\Gamma(k)}{\Gamma(n)} \prod_{i=1}^{k} (1-\sigma)_{n_i-1}, \tag{2.12}$$

the EPPF of the stable NRMI (Pitman (1996)). By comparing (2.11) with (2.2), it is apparent that generalized gamma NRMI are of Gibbs–type.

2.3. Poisson–Kingman models

Let us now review the essentials of Poisson–Kingman (PK) models introduced in Pitman (2003). To give the general definition of PK model let $\tilde{P} = \tilde{\mu}/T$, where $T = \tilde{\mu}(\mathbb{X})$, be a homogeneous NRMI, and take $f : \mathbb{X} \to \mathbb{R}$ to be any function such that $\tilde{P}(|f|) = \int_{\mathbb{X}} |f(x)|\, \tilde{P}(\mathrm{d}x) < \infty$.

Definition 3. If γ is a probability distribution on $\mathbb{R}^+$, a *Poisson–Kingman model* is the random probability measure $\tilde{Q}$ characterized by

$$\mathbb{E}\Big[\exp\Big\{-\tilde{Q}(f)\Big\}\Big] = \int_{\mathbb{R}^+} \mathbb{E}\Big[\exp\Big\{-\frac{\tilde{\mu}(f)}{T}\Big\}\,\Big|\, T\Big]\, \mathrm{d}\gamma, \tag{2.13}$$

where $\mathbb{E}[\,\mathrm{e}^{-\tilde{\mu}(f)/T}\,|T]$ is a regular version of the conditional expectation of $\mathrm{e}^{-\tilde{\mu}(f)/T}$ given T.

Intuitively, a PK model is a (homogeneous) NRMI conditional on the total mass $\tilde{\mu}(\mathbb{X}) = T$, and mixed with respect to any probability distribution on $\mathbb{R}^+$. Consequently, $\tilde{Q}$ in (2.13) is uniquely defined by the Poisson intensity underlying $\tilde{P}$ and the mixing distribution γ. If γ coincides with the distribution of $\tilde{\mu}(\mathbb{X})$, then (2.13) is a homogeneous NRMI. An important special case is the *σ–stable PK model* with $\sigma \in (0,1)$: such a model is constructed starting from a σ–stable

NRMI characterized by the Poisson intensity (2.10) with $\tau = 0$. Let $C \in \mathscr{X}$ be any set such that $H(C) = \sigma$ and denote by f_σ the density of $\tilde{\mu}(C)$; the Laplace transform is then given by $\int_{\mathbb{R}^+} \mathrm{e}^{-\lambda x} f_\sigma(x)\mathrm{d}x = \mathrm{e}^{-\lambda^\sigma}$. As shown in Pitman (2003), the EPPF of the σ–stable NRMI conditioned on its total mass $\tilde{\mu}(\mathbb{X})$ is

$$\Pi_k^{(n)}(n_1, \ldots, n_k|t) = \frac{\sigma^k t^{-n}}{\Gamma(n-k\sigma)\, f_\sigma(t)} \int_0^t s^{n-k\sigma-1} f_\sigma(t-s)\,\mathrm{d}s \prod_{i=1}^k (1-\sigma)_{n_i-1}. \tag{2.14}$$

Obviously, the EPPF of the σ–stable PK model is then given by

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \int_{\mathbb{R}^+} \frac{\sigma^k t^{-n}}{\Gamma(n-k\sigma) f_\sigma(t)} \int_0^t s^{n-k\sigma-1} f_\sigma(t-s)\mathrm{d}s\gamma(\mathrm{d}t) \prod_{i=1}^k (1-\sigma)_{n_i-1}. \tag{2.15}$$

Hence, a σ–stable PK model with any mixing distribution γ is a Gibbs–type prior. Finally, it is important to recall that the generalized gamma NRMI can be recovered from the σ–stable PK model by a suitable choice of the mixing distribution γ (Pitman (2003)).

3. Characterization of Gibbs–Type Priors

In this section we focus on the characterization of Gibbs priors; the aim is to identify all priors which lead to a simple predictive structure of the form (1.1). Gnedin and Pitman (2005), resorting to combinatorial arguments and generalized Stirling number asymptotics, have shown that Gibbs priors coincide with stable PK models.

Proposition 1. (Gnedin and Pitman (2005)) *A random probability measure $\tilde{P}$ is of Gibbs–type with parameter $\sigma \in (0,1)$ if and only if $\tilde{P}$ is a σ–stable PK model.*

In the Appendix we provide an alternative proof of this fundamental result. Essentially, we face the problem of finding a solution to the recursive equations (2.3) that characterize Gibbs priors.

Remark. Gnedin and Pitman (2005) also characterize Gibbs priors for $\sigma \in (-\infty, 0]$. For the case where $\sigma < 0$, they show that Gibbs priors have finite support and correspond to mixtures of symmetric Dirichlet distributions. For $\sigma = 0$, Gibbs priors are mixtures, with respect to the parameter θ, of a Dirichlet process. This latter result was already conjectured by Kerov (1989, 2003) and stated by Kingman (1980). An alternative proof can be derived using our analytic approach.

We now state the main result of the paper, we show that the only NRMI with Gibbs structure is the one based on the generalized gamma process. From a practical point of view, the consequence of such a result is clear. Proposition 1 implies that all Gibbs–type priors are σ–stable PK models. However, simulating from a general σ–stable PK model seems impossible to date; one would need either a closed form expression for the stable density involved in (2.15) that would allow one to implement a Blackwell–MacQueen–type urn scheme via numerical integration, or a way to simulate from a σ–stable process conditioned on its total mass. Neither strategy is exploitable. Instead, simulating from a NRMI is feasible by resorting to urn schemes as, e.g, in Ishwaran and James (2001), or to Ferguson–Klass algorithms as, e.g., in Nieto-Barajas, Prünster, and Walker (2004). Hence, it is natural to ask which NRMIs have a Gibbs structure.

Before stating the result it is worth showing that a σ–stable PK model with a particular mixing distribution leads to the generalized gamma NRMI characterized by the Poisson intensity at (2.10). Pitman (2003) shows this using scaling arguments, whereas we resort to analytic tools. Take the mixing distribution in (2.15) as $\gamma(\mathrm{d}t) = f_\sigma(t)\,\phi(t)\,\mathrm{d}t$ for some (for the moment unspecified) function ϕ. Then

$$\begin{aligned} V_{n,k} &= \frac{\sigma^k}{\Gamma(n-k\sigma)} \int_0^\infty \int_0^t t^{-n} f_\sigma(t-s) s^{n-k\sigma-1} \phi(t) \mathrm{d}s\, \mathrm{d}t \\ &= \frac{\sigma^k}{\Gamma(n-k\sigma)} \int_0^\infty \int_s^\infty t^{-n} f_\sigma(t-s) s^{n-k\sigma-1} \phi(t)\, \mathrm{d}t \mathrm{d}s \\ &= \frac{\sigma^k}{\Gamma(n-k\sigma)} \int_0^\infty \int_0^\infty (u+s)^{-n} f_\sigma(u) s^{n-k\sigma-1} \phi(s+u)\, \mathrm{d}s\, \mathrm{d}u. \end{aligned} \tag{3.1}$$

By the gamma identity we have

$$(\frac{\theta}{\sigma})^{\frac{n}{\sigma}} \Big[(\frac{\theta}{\sigma})^{\frac{1}{\sigma}} (u+s) \Big]^{-n} = \frac{1}{\Gamma(n)} \int_0^\infty x^{n-1} \exp\Big\{ -x(\frac{\theta}{\sigma})^{\frac{1}{\sigma}} (u+s) \Big\} \mathrm{d}x,$$

which, inserted into (3.1), leads to

$$V_{n,k} = \frac{\sigma^k (\frac{\theta}{\sigma})^{\frac{n}{\sigma}}}{\Gamma(n)\Gamma(n-k\sigma)} \int_0^\infty \int_0^\infty \int_0^\infty x^{n-1} e^{-x(\frac{\theta}{\sigma})^{\frac{1}{\sigma}}(s+u)} f_\sigma(u) s^{n-k\sigma-1} \phi(s+u) \mathrm{d}x \mathrm{d}u \mathrm{d}s.$$

We choose, for some $a, \tau > 0$, $\phi(t) = a \exp(-t\tau(\theta/\sigma)^{1/\sigma})$, and hence

$$\begin{aligned} V_{n,k} &= \frac{a(\frac{\theta}{\sigma})^{\frac{n}{\sigma}} \sigma^k}{\Gamma(n)\Gamma(n-k\sigma)} \int_0^\infty x^{n-1} \int_0^\infty \int_0^\infty e^{-(\frac{\theta}{\sigma})^{\frac{1}{\sigma}}(u+s)(\tau+x)} s^{n-k\sigma-1} f_\sigma(u) \mathrm{d}s \mathrm{d}u \mathrm{d}x \\ &= \frac{a\theta^k}{\Gamma(n)} \int_0^\infty \frac{x^{n-1}}{(\tau+x)^{n-k\sigma}} \int_0^\infty e^{-u(\frac{\theta}{\sigma})^{\frac{1}{\sigma}}(\tau+x)} f_\sigma(u) \mathrm{d}u \mathrm{d}x \\ &= \frac{a\theta^k}{\Gamma(n)} \int_0^\infty \frac{x^{n-1}}{(\tau+x)^{n-k\sigma}} e^{-(\frac{\theta}{\sigma})(\tau+x)^\sigma}\, \mathrm{d}x. \end{aligned}$$

In order for $V_{1,1} = 1$, it must be that $a = \exp(\theta\tau^\sigma/\sigma)$. Hence, we have

$$V_{n,k} = \frac{\theta^k \exp(\frac{\theta\tau^\sigma}{\sigma})}{\Gamma(n)} \int_0^\infty \frac{x^{n-1}}{(\tau + x)^{n-k\sigma}} e^{-(\tau+x)^\sigma} \, \mathrm{d}x, \tag{3.2}$$

which is precisely the EPPF of a homogeneous NRMI (2.9) computed for the generalized gamma intensity (2.10). By simple algebra from (3.2), one obtains (2.11) as in Lijoi, Mena and Prünster (2007a). We are now in a position to state the characterization of the generalized gamma NRMI.

Proposition 2. *Let $\tilde{P}$ be a NRMI. Then $\tilde{P}$ is of Gibbs–type, with $\sigma \in (0,1)$, if and only if it is a generalized gamma NRMI.*

Proof. First note that, combining the definition of a PK model with Proposition 1, a necessary condition for a NRMI to be of Gibbs–type is that it be homogeneous. Hence, $\tilde{P}$ has an EPPF of the form (2.9). From (2.2), it then follows that $\tilde{P}$ is of Gibbs–type if

$$V_{n,k} = \frac{\theta^k}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1} \, \mathrm{e}^{-\psi(u)} \Big[\prod_{j=1}^k \frac{\tau_{n_j}(u)}{(1-\sigma)_{n_j-1}} \Big] \, \mathrm{d}u \tag{3.3}$$

for every $n \geq 1$ and $k = 1, \ldots, n$. Since σ and the sequence of $V_{n,k}$'s uniquely identify the underlying Lévy intensity ρ, we first assume ρ is such that $\prod_{j=1}^k [\tau_{n_j}(u)/(1-\sigma)_{n_j-1}]$ depends only on n and k, for almost every u. Then we resort to a constructive approach in order to determine such a ρ and discover that it coincides with the Lévy intensity of a generalized gamma NRMI.

In accordance with our assumption, set $f_{n,k}(u) = \prod_{j=1}^k \tau_{n_j}(u)/(1-\sigma)_{n_j-1}$. If we now put $k = 2$, then it is easy to see that we must have

$$\frac{\tau_{n+1}(u)}{\tau_n(u)} = \frac{(n-\sigma)}{(1-\sigma)} \, \frac{\tau_2(u)}{\tau_1(u)}.$$

This follows by taking $n_1 = n+1$ and $n_2 = 1$, then $n_1 = n$ and $n_2 = 2$. Hence, writing $l(u) = \tau_2(u)/\tau_1(u)$, we have

$$\tau_{n+1}(u) = \frac{(n-\sigma)\cdots(1-\sigma)}{(1-\sigma)^n} l^n(u) \, \tau_1(u),$$

and so

$$\tau_n(u) = \frac{(1-\sigma)_{n-1}}{(1-\sigma)^{n-1}} \, l^{n-1}(u)\tau_1(u). \tag{3.4}$$

Consequently

$$\prod_{j=1}^k \tau_{n_j}(u) = \prod_{j=1}^k (1-\sigma)_{n_j-1} \, \frac{l^{n-k}(u)\tau_1^k(u)}{(1-\sigma)^{n-k}},$$

and it follows that

$$V_{n,k} = \frac{\theta^k}{(1-\sigma)^{n-k}\Gamma(n)} \int_0^\infty u^{n-1}\, \mathrm{e}^{-\psi(u)}\, l^{n-k}(u)\tau_1^k(u)\, \mathrm{d}u. \tag{3.5}$$

At this point we claim that

$$l'(u) = -\frac{l^2(u)}{1-\sigma}. \tag{3.6}$$

To see this, recall that $l(u) = \tau_2(u)/\tau_1(u)$, so that $l'(u) = -\tau_3(u)/\tau_1(u) + \tau_2^2(u)/\tau_1^2(u)$, since $\tau_n'(u) = -\tau_{n+1}(u)$. Using (3.4), we obtain $\tau_3(u) = \{(2-\sigma)/(1-\sigma)\}l^2(u)\tau_1(u)$ and, hence, (3.6) follows.

Now, the solution to (3.6) is given by

$$l(u) = \frac{1}{c + \frac{u}{1-\sigma}} = \frac{1-\sigma}{\tau + u}, \tag{3.7}$$

where c is some constant and $\tau = c(1-\sigma)$. Therefore, (3.5) becomes

$$V_{n,k} = \frac{1}{\Gamma(n)} \int_0^\infty u^{n-1}\, \mathrm{e}^{-\psi(u)}\, \frac{(\theta\tau_1(u))^k}{(\tau+u)^{n-k}}\, \mathrm{d}u. \tag{3.8}$$

Recall that $l(u) = \tau_2(u)/\tau_1(u)$ and that, by (2.8), one has $\tau_2(u) = -\mathrm{d}\tau_1(u)/\mathrm{d}u$. Consequently

$$\frac{-\frac{\mathrm{d}}{\mathrm{d}u}\tau_1(u)}{\tau_1(u)} = \frac{1-\sigma}{\tau+u},$$

which is solved by

$$\tau_1(u) = \frac{d}{(\tau+u)^{1-\sigma}} \tag{3.9}$$

for some constant d. Finally, note that $\mathrm{e}^{\psi(u)}(-\mathrm{d}\mathrm{e}^{-\psi(u)}/\mathrm{d}u) = -\theta\tau_1(u)$; solving this differential equation, with (3.9) for τ_1, yields

$$\mathrm{e}^{-\psi(u)} = a \exp\left\{-\frac{\theta}{\sigma}(\tau+u)^\sigma\right\} \tag{3.10}$$

for some constant $a > 0$. Inserting (3.9) and (3.10) into (3.8), we obtain

$$V_{n,k} = \frac{a d^k \theta^k}{\Gamma(n)} \int_0^\infty \frac{u^{n-1}}{(\tau+u)^{n-k\sigma}} \exp\left\{-\frac{\theta}{\sigma}(\tau+u)^\sigma\right\} \mathrm{d}u. \tag{3.11}$$

The condition $V_{1,1} = 1$ then requires $a = \exp\{(\theta/\sigma)\tau^\sigma\}$ and $d = 1$. Since (3.11) coincides with (3.2), it follows that the generalized gamma NRMI is the unique solution to (3.3).

Up to now, we have proved that, if $\prod_{j=1}^k [\tau_{n_j}(u)/(1-\sigma)_{n_j-1}]$ depends only on n and k, then a NRMI is of Gibbs–type if and only if it is a generalized gamma NRMI. Suppose there exists another NRMI such that $\prod_{j=1}^k [\tau_{n_j}(u)/(1-\sigma)_{n_j-1}]$ depends also on $(n_1,\ldots,n_k)$ and is of Gibbs–type. This implies that it still satisfies (3.3) but, since a Gibbs prior is characterized by σ and the $V_{n,k}$'s, it necessarily leads to a set of weights $V_{n,k}$ different from those induced by the generalized gamma NRMI. Denote this new set of weights by $\{Z_{n,k} : 1 \leq k \leq n, n \geq 1\}$. Then, letting Q be any partition–distribution of the integers $\{1,\ldots,n\}$ in k blocks, one can also write

$$\begin{aligned} Z_{n,k} &= \frac{\theta^k}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1}\,\mathrm{e}^{-\psi(u)} \sum_{\pi} Q(n_1,\ldots,n_k) \Big[\prod_{j=1}^k \frac{\tau_{n_j}(u)}{(1-\sigma)_{n_j-1}}\Big]\,\mathrm{d}u \\ &= \frac{\theta^k}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1}\,\mathrm{e}^{-\psi(u)}\, g_{n,k}^Q(u)\,\mathrm{d}u, \end{aligned}$$

where $\sum_\pi$ stands for the sum over partitions in k blocks. Since we are assuming that $\prod_{j=1}^k [\tau_{n_j}(u)/(1-\sigma)_{n_j-1}]$ depends on $(n_1,\ldots,n_k)$, it must be that $g_{n,k}^Q$ depends on Q, whereas $Z_{n,k}$ does not. Because of the arbitrariness of Q, this cannot happen unless $\prod_{j=1}^k [\tau_{n_j}(u)/(1-\sigma)_{n_j-1}]$ does not depend on the specific configuration $(n_1,\ldots,n_k)$, which in turn implies the NRMI is of generalized gamma type. Thus the proof is complete.

Apart from the practical implications we have already outlined, the result in Proposition 2 sheds some light on the theory of PK models. Indeed, it can be equivalently formulated as follows: NRMIs based on generalized gamma completely random measures are the only NRMIs representable, in distribution, as σ–stable PK models, with $\sigma \in (0,1)$.

4. Appendix: Alternative Proof of Proposition 1

.

In the following we exclude the trivial case of the one–block partition. Before entering the details, we outline a brief sketch of the proof. The main idea is to directly solve (2.3) with $V_{1,1} = 1$. The problem is first formulated in terms of an equivalent array of weights $\{W_{n,k} : k = 1,\ldots,n,\ n \geq 1\}$. Secondly, we suppose that $W_{n,1}$ can be expressed as a mixture of some function g, and a solution to the new recursive formula for $W_{n,k}$ is given. Hence, the function g is identified as a suitable function of the density of a σ–stable random variable. This leads to the EPPF as the one generated by a mixture of σ–stable PK models.

According to the above scheme, put

$$V_{n,k} = \frac{\sigma^k}{\Gamma(n-k\sigma)} W_{n,k}$$

so that the $W_{n,k}$ must solve

$$W_{n,k} - W_{n+1,k} = \frac{\sigma W_{n+1,k+1}\Gamma(n-k\sigma)}{\Gamma(n+1-(k+1)\sigma)}. \tag{4.1}$$

Note that, for any $\theta > 0$, the definition of the Beta function entails

$$\frac{\Gamma(n-k\sigma)}{\Gamma(n+1-(k+1)\sigma)} = \frac{\int_0^\theta (\theta-s)^{-\sigma} s^{n-k\sigma-1}\mathrm{d}s}{\Gamma(1-\sigma)\theta^{n+1-(k+1)\sigma-1}}$$

which, inserted into (4.1), yields

$$\theta^{n+1-(k+1)\sigma-1}\left(W_{n,k} - W_{n+1,k}\right) = \frac{\sigma W_{n+1,k+1}}{\Gamma(1-\sigma)}\int_0^\theta (\theta-s)^{-\sigma} s^{n-k\sigma-1}\mathrm{d}s.$$

Consider now some function $g_{n,t,\sigma}(\theta)$ such that $\int_0^\infty \theta^{n+1-(k+1)\sigma-1} g_{n,t,\sigma}(\theta)\,\mathrm{d}\theta < \infty$, where $t > 0$ is a parameter, and write

$$\begin{aligned}(W_{n,k} - W_{n+1,k}) &\int_0^\infty \theta^{n+1-(k+1)\sigma-1} g_{n+1,t,\sigma}(\theta)\mathrm{d}\theta \\ &= \frac{\sigma\ W_{n+1,k+1}}{\Gamma(1-\sigma)}\int_0^\infty \int_0^\theta (\theta-s)^{-\sigma} s^{n-k\sigma-1}\,\mathrm{d}s\ \ g_{n+1,t,\sigma}(\theta)\,\mathrm{d}\theta. \end{aligned} \tag{4.2}$$

The sequence $\{V_{n,1} : n \geq 1\}$ clearly defines the whole array $\{V_{n,k} :\ n \geq 1,\ 1 \leq k \leq n\}$ via (2.3). Hence, the same holds true for $\{W_{n,1} : n \geq 1\}$ and $\{W_{n,k} :\ n \geq 1,\ 1 \leq k \leq n\}$. We suppose that g is such that $W_{n,1} = \int_0^\infty \theta^{n-\sigma-1} g_{n,t,\sigma}(\theta)\,\mathrm{d}\theta$, and that it further satisfies

$$\frac{\sigma}{\Gamma(1-\sigma)}\int_s^\infty (\theta-s)^{-\sigma}\, g_{n+1,t,\sigma}(\theta)\mathrm{d}\theta = g_{n,t,\sigma}(s) - s g_{n+1,t,\sigma}(s). \tag{4.3}$$

Later we identify an appropriate g. We now show that, given the $\{W_{n,1} : n \geq 1\}$, the solutions $\{W_{n,k} :\ n \geq 1,\ 1 \leq k \leq n\}$ coincide with

$$W_{n,k} = \int_0^\infty \theta^{n-k\sigma-1} g_{n,t,\sigma}(\theta)\mathrm{d}\theta. \tag{4.4}$$

Indeed, from (4.2), (4.3), and for $W_{n,k}$ given by (4.4), we have that

$$\frac{W_{n,k} - W_{n+1,k}}{W_{n+1,k+1}} = \frac{\int_0^\infty s^{n-k\sigma-1} g_{n,t,\sigma}(s)\mathrm{d}s - \int_0^\infty s^{n+1-k\sigma-1} g_{n+1,t,\sigma}(s)\mathrm{d}s}{\int_0^\infty s^{n+1-(k+1)\sigma-1}\, g_{n+1,t,\sigma}(s)\mathrm{d}s}.$$

Hence, since the array $\{W_{n,k} :\ n \geq 1,\ 1 \leq k \leq n\}$ is uniquely determined given $\{W_{n,1} : n \geq 1\}$, solutions are of the kind (4.4), with g satisfying (4.3). At this point we have to identify g by solving (4.3). To this end, we consider

$$g_{n,t,\sigma}(\theta) = a\, t^{-n} \mathbb{I}_{(-\infty,t]}(\theta)\, h_\sigma(t-\theta) \tag{4.5}$$

for some function h_σ and some constant $a > 0$, with $\mathbb{I}_A$ the indicator function of the set A. For $s < t$, (4.3) leads to the identity

$$\frac{\sigma}{\Gamma(1-\sigma)\, t^{n+1}} \int_s^t (\theta - s)^{-\sigma} h_\sigma(t-\theta)\mathrm{d}\theta \, = (t^{-n} - st^{-(n+1)}) h_\sigma(t-s),$$

which is equivalent to

$$\frac{\sigma}{\Gamma(1-\sigma)} \int_s^t (\theta - s)^{-\sigma} h_\sigma(t-\theta)\mathrm{d}\theta = (t-s) h_\sigma(t-s). \tag{4.6}$$

Putting $u = t - s$, (4.6) can be written as

$$\frac{\sigma}{\Gamma(1-\sigma)} \int_s^{s+u} (\theta - s)^{-\sigma} h_\sigma(s+u-\theta)\mathrm{d}\theta = u h_\sigma(u).$$

Finally, set $v = \theta - s$ to obtain the integral equation

$$h_\sigma(u) = \frac{\sigma}{\Gamma(1-\sigma)\, u} \int_0^u v^{-\sigma} h_\sigma(u-v)\mathrm{d}v. \tag{4.7}$$

It is well–known that the density f of an absolutely continuous infinitely divisible random variable is uniquely characterized by its Lévy measure η via the integral equation $f(u) = \int_0^u v\, u^{-1} f(u-v)\, \eta(\mathrm{d}v)$; see, e.g., (4.5) in Pitman (2006). Now, let $\tilde{\mu}$ be a stable random measure whose Poisson intensity is given by $\Gamma(1-\sigma)v^{-1-\sigma}H(\mathrm{d}x)$, and let C be a set such that $H(C) = \sigma$. Then $\tilde{\mu}(C)$, with density f_σ, has Lévy measure $\sigma[\Gamma(1-\sigma)]^{-1}v^{-1-\sigma}\mathrm{d}v$ which, inserted into the previous integral equation, leads to (4.7). Hence, the unique solution to (4.7) is given by $h_\sigma = f_\sigma$. It is clear that any mixture of (4.5) will also work; that is, $g_{n,\sigma}(\theta) = \int_0^\infty g_{n,t,\sigma}(\theta)\gamma(\mathrm{d}t)$ for any distribution γ will also solve (4.3). Hence, we have shown that solutions to (2.3) are of the type

$$V_{n,k}^{(t)} = \frac{\sigma^k t^{-n}}{\Gamma(n-k\sigma) f_\sigma(t)} \int_0^t s^{n-k\sigma-1} f_\sigma(t-s)\, \mathrm{d}s, \tag{4.8}$$

where we have set $a = 1/f_\sigma(t)$ in (4.5) to ensure that $V_{1,1} = 1$, and that any mixture over t will also be a solution. But (4.8) and its mixtures over t identify precisely the $V_{n,k}$'s corresponding to the EPPF of a σ–stable PK model recalled in (2.14) and (2.15). The proof is complete.

Acknowledgements

The authors are grateful to two anonymous referees for their suggestions that led to an improvement in the presentation of the paper. Antonio Lijoi and Igor Prünster are partially supported by MUR grants 2006/134525 and 2006/133449, respectively. Stephen G. Walker is supported by an EPSRC Advanced Research Fellowship.

References

Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. Appl. Probab.* **31**, 929-953.

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

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, 209-230.

Gnedin, A. and Pitman, J. (2005). Exchangeable Gibbs partitions and Stirling triangles. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **325**, 83-102.

Goldwater, S., Griffiths, T. and Johnson, M. (2006). Interpolating between types and tokens by estimating power-law generators. *Advances in Neural Information Processing Systems* **18**.

Hjort, N. L. (2000). Bayesian analysis for a generalized Dirichlet process prior. *Statist. Research Report* **7**, University of Oslo.

Hjort, N. L. (2003). Topics in non-parametric Bayesian statistics. In *Highly Structured Stochastic Systems* (Edited by P. J. Green, et al.), 455-478, Oxford University Press, Oxford.

Ishwaran, H. and James, L. F. (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Stat. Assoc.* **96**, 161-173.

Ishwaran, H. and James, L. F. (2003). Generalized weighted Chinese restaurant processes for species sampling mixture models. *Statist. Sinica* **13**, 1211-1235.

James, L. F. (2002). Poisson Process Partition Calculus with applications to exchangeable models and Bayesian Nonparametrics. arXiv.org/math.PR/0205093.

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

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

James, L. F. (2007). Neutral-to-the-right species sampling mixture models. In *Advances in statistical modeling and inference. Essays in honor of Kjell A. Doksum* (Edited by V. Nair), 425-439. World Scientific, Hackensack, NJ.

James, L. F., Lijoi, A. and Prünster, I. (2006). Conjugacy as a distinctive feature of the Dirichlet process. *Scand. J. Statist.* **33**, 105-120.

James, L. F., Lijoi, A. and Prünster, I. (2008). Posterior analysis for normalized random measures with independent increments. *Scand. J. Statist.*, DOI: 10.1111/j.1467-9469.2008.00609.x

Kerov, S. V. (1989). Combinatorial examples in the theory of AF-algebras. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov* (LOMI), **172**, 55-67. Translated in *J. Soviet Math.* **59**, 1063-1071, 1992.

Kerov, S. V. (2003) *Asymptotic Representation Theory of the Symmetric Group and its Applications in Analysis*, Amer. Math. Soc., Providence, RI.

Kingman, J. F. C. (1967). Completely random measures. *Pacific J. Math.* **21**, 59-78.

Kingman, J. F. C. (1975). Random discrete distributions. *J. Roy. Stat. Soc. Ser. B* **37**, 1-22.

Kingman, J. F. C. (1980). *The Mathematics of Genetic Diversity.* SIAM.

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

Lijoi, A., Mena, R. H. and Prünster, I. (2005a). Hierarchical mixture modelling with normalized inverse Gaussian priors. *J. Amer. Statist. Assoc.* **100**, 1278-1291.

Lijoi, A., Mena, R. H. and Prünster, I. (2005b). Bayesian nonparametric analysis for a generalized Dirichlet process prior. *Statist. Inference Stoch. Process.* **8**, 283-309.

Lijoi, A., Mena, R. H. and Prünster, I. (2007a). Controlling the reinforcement in Bayesian nonparametric mixture models. *J. Roy. Statist. Soc. B* **69**, 715-740.

Lijoi, A., Mena, R. H. and Prünster, I. (2007b). Bayesian nonparametric estimation of the probability of discovering a new species. *Biometrika* **94**, 769-786.

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

Lo, A. Y. (1991). A characterization of the Dirichlet process. *Statist. Probab. Lett.* **12**, 185-187.

Müller, P. and Quintana, F. A. (2004). Nonparametric Bayesian data analysis. *Statist. Science* **19**, 95-110.

Nieto-Barajas, L. E., Prünster, I. and Walker, S. G. (2004). Normalized Random Measures driven by Increasing Additive Processes. *Ann. Statist.* **32**, 2343-2360.

Pitman, J. (1996). Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, Probability and Game Theory. Papers in honor of David Blackwell* **30** (Edited by T. S. Ferguson, et al.), 245–267. Lecture Notes, Monograph Series. IMS, Hayward.

Pitman, J. (2003). Poisson-Kingman partitions. In *Science and Statistics: A Festschrift for Terry Speed* **40** (Edited by D. R. Goldstein), 1-35. Lecture Notes, Monograph Series. IMS, Hayward.

Pitman, J. (2006). *Combinatorial stochastic processes.* Ecole d'Eté de Probabilités de Saint-Flour XXXII. Lecture Notes in Mathematics N. 1875, Springer.

Regazzini, E. (1978). Intorno ad alcune questioni relative alla definizione del premio secondo la teoria della credibilità. *Giornale dell' Istituto Italiano degli Attuari*, **41**, 77-89.

Regazzini, E., Lijoi, A. and Prünster, I. (2003). Distributional results for means of random measures with independent increments. *Ann. Statist.* **31**, 560-585.

Teh, Y. W. (2006a). A hierarchical Bayesian language model based on Pitman-Yor processes. *Proceedings of Coling–ACL 2006.*

Teh, Y. W. (2006b). A Bayesian interpretation of interpolated Kneser–Ney. Technical Report 2/06, School of Computing, National University of Singapore.

Dipartimento Economia Politica e Metodi Quantitativi, Università di Pavia, via San Felice 5, 27100 Pavia, and CNR-IMATI, Milano, Italy.

E-mail: lijoi@unipv.it

Dipartimento di Statistica e Matematica Applicata, Collegio Carlo Alberto and ICER, Università di Torino, Corso Unione Sovietica 218/bis, 10134 Torino, Italy.

E-mail: igor@econ.unito.it

Institute of Mathematics, Statistics and Actuarial Science, University of Kent, Kent CT2 7NZ, UK.

E-mail: S.G.Walker@kent.ac.uk

(Received September 2006; accepted May 2007)