Sankhyā : The Indian Journal of Statistics
2006, Volume 68, Part 3, pp. 461-487

Some Developments of the Normalized Random Measures with Independent Increments

Laura M. Sangalli
Università degli Studi di Pavia, Italy

Abstract

In this paper, we deal with sequences of exchangeable observations, which are conditionally independent and identically distributed with a common (random) distribution τ, where τ is defined by the conditioning of a normalized random measure with independent increments. By application of a multi-dimensional version of the Faà di Bruno formula, we derive explicit closed form expressions for the finite dimensional distributions and the predictive distributions of the sequence. The results are illustrated through suitable examples.

AMS (2000) *subject classification.* Primary 62F15; secondary 60G57.
Keywords and phrases. Bayesian nonparametric inference, normalized random measure with independent increments, species sampling sequence.

1 Introduction

In this paper, we deal with sequences of exchangeable observations, which are conditionally independent and identically distributed with a common (random) distribution τ, where τ is defined by suitable conditioning of a *normalized random measure with independent increments*. Random measures with independent increments are random measures whose values on disjoint events are stochastically independent random variables. The resulting normalized random measures, which are included in the class of random probabilities we consider, are studied, e.g., in Regazzini, Lijoi and Prünster (2003). The Ferguson-Dirichlet process is a special member of this class.

Such models have been extensively studied in the literature. The sequences of exchangeable observations we consider provide examples of generalized *species sampling sequences*. Species sampling sequences are defined in Pitman (1996b) and further analyzed in Hansen and Pitman (2000), whereas

Pitman (2003) concentrates on the special case of species sampling sequences corresponding to *Poisson-Kingam distributions*. See also Perman, Pitman and Yor (1992), Pitman (1995, 1996a) and Pitman and Yor (1997). Species sampling models stem from the theory of partition structures developed by Kingman (1978). In a Bayesian nonparametric context, these kinds of models have been studied, e.g., by James (2002, 2005) and James, Lijoi and Prünster (2005), by means of a closely related methodology, called *Poisson process partition calculus*, and inspired by the pioneering works of Lo (1984) and Lo and Weng (1989).

The laws of random measures with independent increments are characterized by measures $\nu(dx\,ds)$, as given in (2.1), (2.2) and (2.3). While the above-quoted literature always deals with the case where the measure $\eta(\cdot, s)$ in (2.2) is diffuse for each fixed s (or, correspondingly, the measure $\alpha(\cdot)$ in (2.3) is diffuse), in the present paper we allow this measure to have atoms. We study distributional properties which are of interest in Bayesian statistics. In particular, we provide explicit closed expressions for the finite-dimensional distributions and the predictive distributions of the sequence of observations. These are derived from the *Lévy-Khintchine representation* of the random measure with independent increments, through application of a recent *multidimensional version of the Faà di Bruno formula*, due to Constantine and Savits (1996). Condensed expressions are obtained in terms of *exponential Bell partition polynomials*, which lead to some computational advantage. Moreover, these special functions can be easily and quickly computed with a processor. The results are illustrated by resorting to suitable examples. In particular, we develop an example which includes as special cases the Poisson-Dirichlet with two parameters, Ferguson-Dirichlet and Kingman γ-stable models. We have not carried out a study of the posterior distributions. Under the additional hypothesis of diffuseness of $\eta(\cdot, s)$, characterizations of the posterior distributions are given, e.g., in James, Lijoi and Prünster (2005).

The paper is organized as follows. In Sections 2, we recall the basic elements of the theory of normalized random measures with independent increments and describe the conditioning argument. In Sections 3 and 4, we provide the expressions for the finite-dimensional distributions and the predictive distributions of the sequence of observations. In Section 5, we specialize all the results to the case of normalized random measures with independent increments. Finally, in Section 6, we briefly discuss a straightforward extension which covers common forms of hierarchical models. Detailed proofs are deferred to the Appendix.

2 Preliminaries

Let μ be a random measure with independent increments (RMI) on the real line $\mathbb{R}$, i.e., a random measure such that for any measurable collection $\{A_1, \ldots, A_n\}$ of pairwise disjoint subsets of $\mathbb{R}$, $n = 1, 2, \ldots$, the random variables $\mu(A_1), \ldots, \mu(A_n)$ are stochastically independent. A systematic account of these random measures is given, e.g., in Kingman (1967), where these measures are called *completely random measures*. It is well known that the distribution of μ is uniquely determined by a measure $\nu(dx\,ds)$ on $\mathbb{R} \times \mathbb{R}^+$, via the Laplace transform of $\mu(A)$, for any measurable subset A of $\mathbb{R}$,

$$E\left[e^{-u\,\mu(A)}\right] = \exp\left\{-\int_{A\times\mathbb{R}^+} (1 - e^{-us})\,\nu(dx\,ds)\right\} \quad (u \geq 0), \qquad (2.1)$$

where E denotes expectation with respect to P, P being the probability defined on the space which supports all the random elements considered through the present paper. This is just the Lévy-Khintchine representation mentioned in the previous section. Assume that

$$\nu(dx\,ds) = \eta(dx, s)\,\rho(ds), \qquad (2.2)$$

where ρ is a σ-finite measure on $\mathbb{R}^+$, and η is such that for each fixed s, $\eta(\cdot, s)$ is a measure on $\mathbb{R}$, and for each fixed measurable subset A of $\mathbb{R}$, $\eta(A, \cdot)$ is measurable. Moreover, assume that $\eta(\cdot, s)$ is uniformly σ-finite. See, e.g., Ash (2000), Section 2.6. An important special case is when $\eta(dx, s) = \alpha(dx)$ for any s in $\mathbb{R}^+$, α being some finite measure on $\mathbb{R}$, i.e.,

$$\nu(dx\,ds) = \alpha(dx)\,\rho(ds). \qquad (2.3)$$

Following Regazzini, Lijoi and Prünster (2003), one can determine a random probability φ by normalization of μ, i.e.,

$$\varphi(\cdot) := \frac{\mu(\cdot)}{\mu(\mathbb{R})}$$

which is the normalized random measure with independent increments (NRMI). This definition makes sense provided that the total mass of the NRM μ, $T := \mu(\mathbb{R})$, is finite and strictly positive almost surely, which in turn is true if and only if $\int_{\mathbb{R}\times\mathbb{R}^+}(1 - e^{-us})\,\nu(dx\,ds) < \infty$ for any positive u, and $\nu(\mathbb{R} \times \mathbb{R}^+) = +\infty$. In this framework, for instance, the Dirichlet process with parameter α (α being a finite measure on $\mathbb{R}$) appears when $\nu(dx\,ds) = \alpha(dx)\,s^{-1}\,e^{-s}\,ds$.

We are now in a position to introduce the conditioning argument. This is a classical argument, explicitly considered, e.g., in Kingman (1975) and Pitman (2003), and further developed in James (2002) and James, Lijoi and Prünster (2005). In our setting, it consists of a suitable deformation of the law of the total mass T of the RMI μ. The class of random probabilities, which can be defined by this argument, constitutes an extension of the class of NRMI's.

Given any collection $\{B_1, \ldots, B_n\}$ of measurable subsets of $\mathbb{R}$, $n = 1, 2, \ldots$, set

$$\begin{aligned} & q_{B_1,\ldots,B_n}(d\theta_1, \ldots, d\theta_n) \\ & := \int_{\mathbb{R}^+} Q\big\{\varphi(B_1) \in d\theta_1, \ldots, \varphi(B_n) \in d\theta_n \big| T = t\big\}\, h(t)\, Q\big\{T \in dt\big\} \end{aligned} \tag{2.5}$$

where Q is the probability distribution of μ, and h is any nonnegative measurable function on $\mathbb{R}^+$ such that

$$\int_{\mathbb{R}^+} h(t)\, Q\big\{T \in dt\big\} = 1. \tag{2.6}$$

The family $\{q_{B_1,\ldots,B_n} :\ B_1, \ldots, B_n \text{ measurable subsets of } \mathbb{R}, n = 1, 2, \ldots\}$ determines a unique probability distribution q for some random probability τ, and $q_{B_1,\ldots,B_n}$ represents the finite-dimensional distribution of τ at $\{B_1, \ldots, B_n\}$. The random probability τ can be called *conditioned NRMI*. It is obvious that one can recover the law of the NRMI φ by taking $h(t) = 1$ for every t.

It should be noted that the construction considered in Pitman (2003) corresponds to the case where the RMI μ has associated measure $\nu(dx\, ds)$ such as in (2.3), with $\alpha(dx)$ diffuse. James (2002) and James, Lijoi and Prünster (2005) deal with the more general case, where $\nu(dx\, ds)$ is as in (2.2), but make the analogous assumption of diffuseness of $\eta(dx, s)$ for each fixed s. Here, we allow $\eta(dx, s)$ to have atoms.

It is also possible to consider RMI's, and hence conditioned NRMI's, on more general Polish spaces. In fact, the arguments required by the general case are the same as the ones we will employ to deal with the real case.

3 Probability Law of the Observations

Designate the observations by the real-valued random variables $X_1, X_2, \ldots$ and assume that, conditionally on the hypothesis that τ is the true law of each observation, $X_1, X_2, \ldots$ are independent and identically distributed (i.i.d.) with common distribution τ. Then, by (2.5) and a straightforward application of Fubini's theorem,

$$\begin{aligned} P\{X_1 \in B_1, \ldots, X_n \in B_n\} &= \int_{[0,1]^n} \left[\prod_{i=1}^n \theta_i\right] q_{B_1,\ldots,B_n}(d\theta_1, \ldots, d\theta_n) \\ &= \int_{\mathbb{R}^+} h(t)\, \frac{1}{t^n}\, E\left[\prod_{i=1}^n \mu(B_i) \middle| T = t\right] Q\{T \in dt\} \end{aligned} \tag{3.1}$$

for any measurable $B_1, \ldots, B_n$ and $n = 1, 2, \ldots$.

We aim at providing a closed explicit expression for $P\{X_1 \in B_1, \ldots, X_n \in B_n\}$ based on the characterization (2.1) of the law of the random measure μ with independent increments.

To simplify the notation, set $f_T(\cdot) := Q\{T \in \cdot\}$. Moreover, set

$$\phi(u) := \exp\left\{-\int_{\mathbb{R}\times\mathbb{R}^+} (1 - e^{-us})\, \nu(dx\, ds)\right\} \quad (u \geq 0)$$

and, for any measurable subset B of $\mathbb{R}$, and any integer i,

$$\begin{aligned} \kappa_i(u; B) &:= \int_{B\times\mathbb{R}^+} e^{-us}\, s^i\, \nu(dx\, ds) = \int_{\mathbb{R}^+} e^{-us}\, s^i\, \eta(B, s)\, \rho(ds) \;\; (u \geq 0) \\ l_i(B)(ds) &:= s^i\, 1_{s>0}\, \eta(B, s)\, \rho(ds). \end{aligned}$$

In the special case corresponding to (2.3), one has

$$\begin{aligned} \phi(u) &= \exp\left\{-\alpha(\mathbb{R}) \int_{\mathbb{R}^+} (1 - e^{-us})\, \rho(ds)\right\}, \\ \kappa_i(u; B) &= \alpha(B)\, \kappa_i(u), \qquad\qquad l_i(B)(ds) = \alpha(B)\, l_i(ds), \end{aligned}$$

where

$$\kappa_i(u) := \int_{\mathbb{R}^+} s^i\, e^{-us}\, \rho(ds), \qquad\qquad l_i(ds) := s^i\, 1_{s>0}\, \rho(ds).$$

Finally, denote by $\lambda_1 * \ldots * \lambda_m$ the convolution of measure obtained using $\lambda_1, \ldots, \lambda_m$ ($m = 1, 2, \ldots$), and, given any vector of nonnegative integers $\boldsymbol{d} = (d_1, \ldots, d_n)$, let

$$\begin{aligned} l^{\boldsymbol{d}}(B) &:= \underbrace{l_1(B) * \ldots * l_1(B)}_{d_1} * \ldots * \underbrace{l_n(B) * \ldots * l_n(B)}_{d_n} \\ l^{\boldsymbol{d}} &:= \underbrace{l_1 * \ldots * l_1}_{d_1} * \ldots * \underbrace{l_n * \ldots * l_n}_{d_n}. \end{aligned}$$

LEMMA 3.1. *Let $X_1, X_2, \ldots$ be conditionally i.i.d. with common distribution τ, given the conditioned NRMI τ. For any measurable $B_1, \ldots, B_n$ and $n = 1, 2, \ldots$,*

$$\begin{aligned} & E\left[\prod_{i=1}^{n} \mu(B_i) \middle| T = t\right] f_T(dt) \\ & = \sum_{\boldsymbol{\pi}(n)} \left(f_T * l_{n_1}(\cap_{i\in\pi_1} B_i) * \cdots * l_{n_{q(\boldsymbol{\pi})}}(\cap_{i\in\pi_{q(\boldsymbol{\pi})}} B_i)\right)(dt), \end{aligned} \tag{3.2}$$

where the summation is over all partitions $\boldsymbol{\pi} = \{\pi_1, \ldots, \pi_{q(\boldsymbol{\pi})}\}$ of $\{1, \ldots, n\}$, and n_j denotes the number of elements in π_j. In particular, if (2.3) holds, then the right hand side of (3.2) reduces to

$$\sum_{\boldsymbol{\pi}(n)} \left(f_T * l_{n_1} * \cdots * l_{n_{q(\boldsymbol{\pi})}}\right)(dt) \prod_{j=1}^{q(\boldsymbol{\pi})} \alpha\left(\cap_{i\in\pi_j} B_i\right).$$

Moreover, if $A_1, \ldots, A_q$ are pairwise disjoint measurable sets, and $n_1, \ldots, n_q$ are nonnegative integers such that $n_1 + \ldots + n_q = n$, then

$$\begin{aligned} & E\left[\prod_{j=1}^{q} \mu(A_j)^{n_j} \middle| T = t\right] f_T(dt) \\ & = \sum_{\boldsymbol{d}(n_1,\ldots,n_q)} \left(f_T * l^{\boldsymbol{d}^{(1)}}(A_1) * \cdots * l^{\boldsymbol{d}^{(q)}}(A_q)\right)(dt) \prod_{j=1}^{q} b_{\boldsymbol{d}^{(j)}}, \end{aligned} \tag{3.3}$$

where the summation is over all nonnegative solutions $\boldsymbol{d}^{(j)} = (d_1^{(j)}, \ldots, d_{n_j}^{(j)})$ of the equations $d_1^{(j)} + 2d_2^{(j)} + \cdots + n_j d_{n_j}^{(j)} = n_j$, for $j = 1, \ldots, q$, and

$$b_{\boldsymbol{d}^{(j)}} := \frac{n_j!}{\prod_{i=1}^{n_j} d_i^{(j)}!\,(i!)^{d_i^{(j)}}}.$$

In particular, if (2.3) holds, then the right side of (3.3) reduces to

$$\sum_{\boldsymbol{d}(n_1,\ldots,n_q)} \left(f_T * l^{\boldsymbol{d}^{(1)}} * \cdots * l^{\boldsymbol{d}^{(q)}}\right)(dt) \prod_{j=1}^{q} b_{\boldsymbol{d}^{(j)}} \alpha(A_j)^{\bar{d}^{(j)}},$$

where $\bar{d}^{(j)} = d_1^{(j)} + \ldots + d_{n_j}^{(j)}$.

Thanks to Lemma 3.1 and (3.1), one immediately derives the following proposition.

PROPOSITION 3.1. *Let $X_1, X_2, \ldots$ be conditionally i.i.d. with common distribution τ, given the conditioned NRMI τ. For any measurable $B_1, \ldots, B_n$ and $n = 1, 2, \ldots$,*

$$P\{X_1 \in B_1, \ldots, X_n \in B_n\}$$
$$= \sum_{\boldsymbol{\pi}(n)} \int_{\mathbb{R}^+} h(t) \frac{1}{t^n} \left(f_T * l_{n_1}(\cap_{i\in\pi_1} B_i) * \cdots * l_{n_{q(\boldsymbol{\pi})}}(\cap_{i\in\pi_{q(\boldsymbol{\pi})}} B_i) \right) (dt).$$

In particular, if (2.3) holds, then the right side of the above equation reduces to

$$\sum_{\boldsymbol{\pi}(n)} \int_{\mathbb{R}^+} h(t) \frac{1}{t^n} (f_T * l_{n_1} * \cdots * l_{n_{q(\boldsymbol{\pi})}})(dt) \prod_{j=1}^{q(\boldsymbol{\pi})} \alpha \left(\cap_{i\in\pi_j} B_i \right).$$

Moreover, if $A_1, \ldots, A_q$ are pairwise disjoint measurable sets, $n_1, \ldots, n_q$ are nonnegative integers such that $n_1 + \ldots + n_q = n$, and $\widetilde{A}_1 \times \cdots \times \widetilde{A}_n$ stands for the Cartesian product of n_1 copies of $A_1, \ldots, n_q$ copies of A_q, taken in a fixed order, then

$$P\{(X_1, \ldots, X_n) \in \widetilde{A}_1 \times \cdots \times \widetilde{A}_n\}$$
$$= \sum_{\boldsymbol{d}(n_1,\ldots,n_q)} \int_{\mathbb{R}^+} h(t) \frac{1}{t^n} \left(f_T * l^{\boldsymbol{d}^{(1)}}(A_1) * \cdots * l^{\boldsymbol{d}^{(q)}}(A_q) \right) (dt) \prod_{j=1}^{q} b_{\boldsymbol{d}^{(j)}}.$$

In particular, if (2.3) holds, then the right side of the above equation reduces to

$$\sum_{\boldsymbol{d}(n_1,\ldots,n_q)} \int_{\mathbb{R}^+} h(t) \frac{1}{t^n} (f_T * l^{\boldsymbol{d}^{(1)}} * \cdots * l^{\boldsymbol{d}^{(q)}})(dt) \prod_{j=1}^{q} b_{\boldsymbol{d}^{(j)}} \, \alpha \left(A_j \right)^{\bar{d}^{(j)}}.$$

See also the expressions provided in Section 8 of James (2002) and in Section 8.3 of James, Lijoi and Prünster (2005), for the special case where $\eta(\cdot, s)$ is diffuse.

Before illustrating the above result with some example, we need to recall the definition of exponential Bell partition polynomials. Given any positive integer n, the exponential Bell partition polynomial of order n, B_n, is defined by

$$\mathrm{B}_n\,(w_1,\ldots,w_n) := \sum_{\boldsymbol{d}(n)} \frac{n!}{\prod_{i=1}^n d_i!(i!)^{d_i}} \prod_{i=1}^n w_i^{d_i},$$

where $\sum_{\boldsymbol{d}(n)}$ denotes the summation over all nonnegative integer solutions $\boldsymbol{d} = (d_1,\ldots,\, d_n)$ of the equation $d_1 + 2d_2 + \cdots + nd_n = n$. The definition is extended to $n = 0$ by setting $\mathrm{B}_0 := 1$. Given any positive integers n and k, with $k \le n$, the exponential *partial* Bell partition polynomial of order n and degree k, $\mathrm{B}_{n,d}$, is defined by

$$\mathrm{B}_{n,d}\,(w_1,\ldots,w_n) := \sum_{\boldsymbol{d}(n,k)} \frac{n!}{\prod_{i=1}^n d_i!(i!)^{d_i}} \prod_{i=1}^n w_i^{d_i},$$

where $\sum_{\boldsymbol{d}(n,k)}$ denotes the summation over all nonnegative integer solutions $\boldsymbol{d} = (d_1,\ldots,d_n)$ of the equations $d_1+2d_2+\cdots+nd_n = n$ and $d_1+\cdots+d_n = k$. As we shall see afterwards, the employment of these special functions yields some computational advantage, leading for example to a straightforward derivation of the predictive distributions of the sequence of observations. Moreover, these polynomials are of help in providing condensed expressions. Finally, thanks to recursive formulas, they can be quickly computed with a processor. See Regazzini (1998) for an example of their use in a closely related context. A systematic treatment of these special functions can be found, e.g., in Charalambides (2002).

We now introduce an example, which includes as special cases some classical models such as the Poisson-Dirichlet with two parameters, Ferguson-Dirichlet and Kingman γ-stable models. Consider the Kingman γ-stable model, corresponding to the measure $\nu(dx\,ds) = \alpha(dx)\,s^{-1-\gamma}ds$, where α is any finite measure and $0 < \gamma < 1$. From

$$\begin{aligned}\prod_{j=1}^{q(\boldsymbol{\pi})} \kappa_{n_j}(u) &= u^{\gamma q(\boldsymbol{\pi})-n} \prod_{j=1}^{q(\boldsymbol{\pi})} \Gamma(n_j-\gamma)\\ &= \int e^{-uw} \frac{w^{n-\gamma q(\boldsymbol{\pi})-1}}{\Gamma\,(n-\gamma q(\boldsymbol{\pi}))} 1_{w>0}\,dw \prod_{j=1}^{q(\boldsymbol{\pi})} \Gamma(n_j-\gamma),\end{aligned}$$

which holds in this special case, we obtain

$$\left(f_T{*}l_{n_1}{*}\cdots{*}l_{n_q}\right)(dt) = \left(\int_0^t \frac{(t-x)^{n-\gamma q(\boldsymbol{\pi})-1}}{\Gamma\,(n-\gamma\,q(\boldsymbol{\pi}))}\,f_T(dx)\right) dt \prod_{j=1}^{q(\boldsymbol{\pi})} \Gamma(n_j-\gamma).$$

Thus,

$$E\left[\prod_{i=1}^{n}\mu(B_i)\middle| T=t\right] f_T(dt)$$
$$=\sum_{\boldsymbol{\pi}(n)}\left(\int_0^t \frac{(t-x)^{n-\gamma q(\boldsymbol{\pi})-1}}{\Gamma\left(n-\gamma q(\boldsymbol{\pi})\right)} f_T(dx)\right) dt \prod_{j=1}^{q(\boldsymbol{\pi})}\Gamma(n_j-\gamma)\,\alpha\left(\cap_{i\in\pi_j}B_i\right)$$

for any $B_1,\dots,B_n$. Moreover, if $A_1,\dots,A_q$ are pairwise disjoint, and $n_1+\dots+n_q=n$, then

$$E\left[\prod_{j=1}^{q}\mu(A_j)^{n_j}\middle| T=t\right] f_T(dt)$$
$$=\sum_{\underline{d}(n_1,\dots,n_q)}\left(\int_0^t \frac{(t-x)^{n-\gamma\sum_{j=1}^q d^{(j)}-1}}{\Gamma\left(n-\gamma\sum_{j=1}^q d^{(j)}\right)} f_T(dx)\right) dt \prod_{j=1}^{q}\mathrm{B}_{n_j,d^{(j)}}\left(\boldsymbol{v}^{(j)}\right),$$

where $\sum_{\underline{d}(n_1,\dots,n_q)}$ denotes the summation over $d^{(j)}=1,\dots,n_j$, for $j=1,\dots,q$, and $\boldsymbol{v}^{(j)}:=(v_1^{(j)}\dots,v_{n_j}^{(j)})$ with $v_i^{(j)}:=\Gamma(i-\gamma)\,\alpha(A_j)$.

Now, following Pitman (2003), choose $h(t)\propto t^{-\beta}$, with $\beta\ge 0$. With some simple algebra, one can see that

$$h(t)=\frac{\Gamma(\beta+1)\left(\alpha(\mathbb{R})\,\Gamma(1-\gamma)/\gamma\right)^{\beta/\gamma}}{\Gamma(\beta/\gamma+1)}\,t^{-\beta} \tag{3.4}$$

satisfies (2.6). In this case, for any $B_1,\dots,B_n$,

$$P\{X_1\in B_1,\dots,X_n\in B_n\}$$
$$=\sum_{\boldsymbol{\pi}(n)}\frac{\prod_{m=1}^{q(\boldsymbol{\pi})-1}(\beta+m\,\gamma)}{[\beta+1]_{n-1}}\prod_{j=1}^{q(\boldsymbol{\pi})}[1-\gamma]_{n_j-1}\frac{\alpha\left(\cap_{i\in\pi_j}B_i\right)}{\alpha(\mathbb{R})}, \tag{3.5}$$

where, for any positive integer m, $[x]_m:=\prod_{j=1}^m(x+j-1)$, and $[x]_0:=1$. In particular, when α is diffuse and $\beta=\alpha(\mathbb{R})$, one has the Poisson-Dirichlet model with two parameters (α,γ) introduced by Pitman (1995). On the other hand, for a general α, setting $\beta=\alpha(\mathbb{R})$ and $\gamma=0$ in (3.5), gives the expression corresponding to the Ferguson-Dirichlet model with parameter α. Finally, when $\beta=0$, i.e., in the unconditioned case, one has the Kingman γ-stable model (see Section 5).

Let us call the model corresponding to (3.4) Poisson-Dirichlet model with three parameters (α, γ, β). If $A_1, \ldots, A_q$ are pairwise disjoint, and $n_1 + \ldots + n_q = n$, then

$$P\{(X_1, \ldots, X_n) \in \widetilde{A}_1 \times \cdots \times \widetilde{A}_n\} = \sum_{\underline{d}(n_1,\ldots,n_q)} \frac{\prod_{m=1}^{\sum_1^q d^{(j)}-1} (\beta + m\,\gamma)}{[\beta+1]_{n-1}} \prod_{j=1}^{q} \mathrm{B}_{n_j, d^{(j)}}\left(\boldsymbol{\omega}^{(j)}\right),$$

where $\boldsymbol{\omega}^{(j)} := (\omega_1^{(j)} \ldots, \omega_{n_j}^{(j)})$ and $\omega_i^{(j)} := [1-\gamma]_{i-1}\alpha(A_j)/\alpha(\mathbb{R})$. Using the properties of Bell polynomials one can easily check that, when $\beta = \alpha(\mathbb{R})$ and $\gamma = 0$, this reduces to the corresponding probabilities for the Ferguson-Dirichlet model with parameter α.

4 Predictive Distributions

Using the representation of $P\{(X_1, \ldots, X_n) \in \widetilde{A}_1 \times \cdots \times \widetilde{A}_n\}$ given in Proposition 3.1, one can easily derive expressions for predictive distributions. Even though this study can be carried out for conditioned NRMI associated with general measures $\nu(dx\,ds)$, we confine ourselves to considering the special case in which (2.3) holds.

Let us begin with predictive distributions based on grouped data, i.e., the conditional distribution of X_{n+1} given $\{(X_1, \ldots, X_n) \in \widetilde{A}_1 \times \cdots \times \widetilde{A}_n\}$. Given any set B, let B^c denote its complement.

Proposition 4.1. *Let $X_1, X_2, \ldots$ be conditionally i.i.d. with common distribution τ, given the conditioned NRMI τ, and assume that* (2.3) *holds. If $A_1, \ldots, A_q$ are pairwise disjoint measurable sets, and $n_1, \ldots, n_q$ are non-negative integers such that $n_1 + \ldots + n_q = n$. Then, for any measurable B,*

$$\begin{aligned} &P\{X_{n+1} \in B | (X_1, \ldots, X_n) \in \widetilde{A}_1 \times \cdots \times \widetilde{A}_n\} \\ &\quad = \sum_{j=1}^{q} \frac{\alpha(A_j \cap B)}{\alpha(A_j)} \frac{P\{(X_1, \ldots, X_n, X_{n+1}) \in A_1^{n_1} \times \cdots \times A_j^{n_j+1} \times \cdots \times A_q^{n_q}\}}{P\{(X_1, \ldots, X_n) \in A_1^{n_1} \times \cdots \times A_j^{n_j} \times \cdots \times A_q^{n_q}\}} \\ &\quad + \frac{P\{(X_1, \ldots, X_n, X_{n+1}) \in A_1^{n_1} \times \cdots \times A_q^{n_q} \times (B \cap (\cup_{j=1}^{q} A_j)^c)\}}{P\{(X_1, \ldots, X_n) \in A_1^{n_1} \times \cdots \times A_q^{n_q}\}}. \end{aligned}$$

Proposition 4.1 points out that predictive distributions based on grouped data can be represented as convex linear combinations, with weights given by the ratios of the probabilities of grouped data. In particular, the j-th weight ($j = 1, \ldots, q$) is obtained according to the hypothesis that the $(n+1)$-th observation belongs to A_j, and the $(q+1)$-th weight according to the hypothesis that the $(n+1)$-th observation does not belong to any A_j ($j = 1, \ldots, q$). See James, Lijoi and Prünster (2006) for another expression of these predictive distributions in the special case in which $h(t) = 1$ for each t.

Moreover, Proposition 4.1 is a first step towards the derivation of an analogous form for predictive distributions based on point data. Given any set B, denote its indicator function by $1(x \in B)$.

COROLLARY 4.1. *Let $X_1, X_2, \ldots$ be conditionally i.i.d. with common distribution τ, given the conditioned NRMI τ, and assume that* (2.3) *holds. If $x_1, \ldots, x_q$ are distinct real numbers, $n_1 + \ldots + n_q = n$, and $(\tilde{x}_1, \ldots, \tilde{x}_n)$ is any rearrangement of $(\underbrace{x_1, \ldots, x_1}_{n_1}, \ldots, \underbrace{x_q, \ldots, x_q}_{n_q})$, then, for any measurable B,*

$$P\{X_{n+1} \in B | X_1 = \tilde{x}_1, \ldots, X_n = \tilde{x}_n\} = \sum_{j=1,\ldots,q} p_j \, 1(x_j \in B) + p_{q+1} \frac{\alpha(B)}{\alpha(\mathbb{R})} \tag{4.1}$$

with $p_j \geq 0$, for $j = 1, \ldots, q+1$, and $\sum_{j=1}^{q+1} p_j = 1$. Moreover, if α is purely discrete, then

$$p_j = \frac{\displaystyle\sum_{\boldsymbol{d}(n_1,\ldots,n_j+1,\ldots,n_q)} \int_{\mathbb{R}^+} h(t) \frac{1}{t^{n+1}} \left(f_T * l^{\boldsymbol{d}^{(1)}} * \cdots * l^{\boldsymbol{d}^{(q)}}\right)(dt) \prod_{r=1}^{q} b_{\boldsymbol{d}^{(r)}} \, \alpha(\{x_r\})^{\bar{d}^{(r)}}}{\displaystyle\sum_{\boldsymbol{d}(n_1,\ldots,n_q)} \int_{\mathbb{R}^+} h(t) \frac{1}{t^{n}} \left(f_T * l^{\boldsymbol{d}^{(1)}} * \cdots * l^{\boldsymbol{d}^{(q)}}\right)(dt) \prod_{r=1}^{q} b_{\boldsymbol{d}^{(r)}} \, \alpha(\{x_r\})^{\bar{d}^{(r)}}} - \frac{\displaystyle\sum_{\boldsymbol{d}(n_1,\ldots,n_q)} \int_{\mathbb{R}^+} h(t) \frac{1}{t^{n+1}} \left(f_T * l^{\boldsymbol{d}^{(1)}} * \cdots * l^{\boldsymbol{d}^{(q)}} * l_1\right)(dt) \left[\prod_{r=1}^{q} b_{\boldsymbol{d}^{(r)}} \, \alpha(\{x_r\})^{\bar{d}^{(r)}}\right] \alpha(\{x_j\})}{\displaystyle\sum_{\boldsymbol{d}(n_1,\ldots,n_q)} \int_{\mathbb{R}^+} h(t) \frac{1}{t^{n}} \left(f_T * l^{\boldsymbol{d}^{(1)}} * \cdots * l^{\boldsymbol{d}^{(q)}}\right)(dt) \prod_{r=1}^{q} b_{\boldsymbol{d}^{(r)}} \, \alpha(\{x_r\})^{\bar{d}^{(r)}}}$$

for $j = 1, \ldots, q$, *and*

$$p_{q+1} = \frac{\displaystyle\sum_{\boldsymbol{d}(n_1,\ldots,n_q)} \int_{\mathbb{R}^+} h(t) \frac{1}{t^{n+1}} \Big(f_T * l^{\boldsymbol{d}^{(1)}} * \cdots * l^{\boldsymbol{d}^{(q)}} * l_1 \Big)(dt) \left[\prod_{r=1}^{q} b_{\boldsymbol{d}^{(r)}}\, \alpha(\{x_r\})^{\bar{d}^{(r)}} \right] \alpha(\mathbb{R})}{\displaystyle\sum_{\boldsymbol{d}(n_1,\ldots,n_q)} \int_{\mathbb{R}^+} h(t)\, \frac{1}{t^n} \left(f_T * l^{\boldsymbol{d}^{(1)}} * \cdots * l^{\boldsymbol{d}^{(q)}} \right)(dt) \prod_{r=1}^{q} b_{\boldsymbol{d}^{(r)}}\, \alpha\left(\{x_r\}\right)^{\bar{d}^{(r)}}}.$$

If α *is diffuse, then*

$$p_j = \frac{\displaystyle\int_{\mathbb{R}^+} h(t)\, \frac{1}{t^{n+1}}\, (f_T * l_{n_1} * \cdots * l_{n_j+1} * \cdots * l_{n_q})(dt)}{\displaystyle\int_{\mathbb{R}^+} h(t)\, \frac{1}{t^n}\, (f_T * l_{n_1} * \cdots * l_{n_j} * \cdots * l_{n_q})(dt)}$$

for $j = 1, \ldots, q$, *and*

$$p_{q+1} = \frac{\displaystyle\int_{\mathbb{R}^+} h(t)\, \frac{1}{t^{n+1}}\, (f_T * l_{n_1} * \cdots * l_{n_q} * l_1)(dt)}{\displaystyle\int_{\mathbb{R}^+} h(t)\, \frac{1}{t^n}\, (f_T * l_{n_1} * \cdots * l_{n_q})(dt)}\, \alpha(\mathbb{R}).$$

Combining the arguments used in the proof of Corollary 4.1, one can easily determine the expression of the weight p_j associated with a general measure α, which is a mixture of a diffuse and a discrete measure.

Exchangeable sequences having predictive distributions, which are convex linear combinations of a weighted empirical measure and a parameter measure, such as in (4.1), are natural models for species sampling problems. In the problem of species sampling, $X_1, X_2, \ldots$ form a random sample from a large population of individuals of various species, and the random variable X_i represents the species of the i-th individual sampled. According to the terminology introduced by Pitman (1996b), a species sampling sequence is an exchangeable sequence characterized by predictive distributions of the form (4.1) with the measure α diffuse. In this case, the weight p_j depends solely on the vector $(n_1, \ldots, n_q)$ of the counts of the various species. Hansen and Pitman (2000) note that one could also consider the case of a general measure α, with a discrete part, letting the weights p_j depend also on the set of atoms of α which have been visited by the sequence within the first n terms, $\{x \in \mathbf{B}(\mathbb{R}) : \alpha\{x\} > 0,\ X_i = x \text{ for some } i \le n\}$, i.e., on the actual species observed. Thus, the sequences of observations conditionally i.i.d. with common distribution τ, given the conditioned NRMI τ associated with a measure $\nu(dx\,ds)$ such as in (2.3), provide examples of generalized species sampling sequences — generalized in the sense that the measure α can have a discrete part. From Corollary 4.1, it is easy to derive other distributional properties such as the probability law of different clusterings of observations (the so-called *exchangeable partition probability function*), and of the number of distinct observations. These quantities are typically of interest in species sampling problems, where they relate to the partition of the population in groups of individuals having the same species, and to the number of different species. For the sake of brevity, we omit these here.

We now continue the example of Poisson-Dirichlet model with three parameters (α, γ, β). Let α be defined as a mixture of a diffuse and a discrete measure, and let $x_1, \ldots, x_q$ be such that

$$\begin{cases} \alpha(\{x_j\}) > 0, & j = 1, \ldots, q^* \\ \alpha(\{x_j\}) = 0, & j = q^* + 1, \ldots, q. \end{cases} \tag{4.2}$$

Using the notation introduced in the previous section, one has

$$p_j = \frac{\displaystyle\sum_{\underline{d}(n_1,\ldots,n_j+1,\ldots,n_{q^*})} \left[\prod_{m=1}^{\sum_{r=1}^{q^*} d^{(r)}+q-q^*-1} (\beta + m\,\gamma) \right] \prod_{r=1}^{q^*} \mathrm{B}_{n_r,d^{(r)}}\left(\boldsymbol{\omega}^{(r)}\right)}{\displaystyle(\beta+n) \sum_{\underline{d}(n_1,\ldots,n_{q^*})} \left[\prod_{m=1}^{\sum_{r=1}^{q^*} d^{(r)}+q-q^*-1} (\beta + m\,\gamma) \right] \prod_{r=1}^{q^*} \mathrm{B}_{n_r,d^{(r)}}\left(\boldsymbol{\omega}^{(r)}\right)} - \frac{\displaystyle\sum_{\underline{d}(n_1,\ldots,n_{q^*})} \left[\prod_{m=1}^{\sum_{r=1}^{q^*} d^{(r)}+q-q^*} (\beta + m\,\gamma) \right] \left[\prod_{r=1}^{q^*} \mathrm{B}_{n_r,d^{(r)}}\left(\boldsymbol{\omega}^{(r)}\right) \right] \omega_1^{(j)}}{\displaystyle(\beta+n) \sum_{\underline{d}(n_1,\ldots,n_{q^*})} \left[\prod_{m=1}^{\sum_{r=1}^{q^*} d^{(r)}+q-q^*-1} (\beta + m\,\gamma) \right] \prod_{r=1}^{q^*} \mathrm{B}_{n_r,d^{(r)}}\left(\boldsymbol{\omega}^{(r)}\right)}$$

whenever $j = 1, \ldots, q^*$. Moreover,

$$p_j = \frac{n_j - \gamma}{\beta + n}$$

whenever $j = q^* + 1, \ldots, q$. Finally,

$$p_{q+1} = \frac{\displaystyle\sum_{\underline{d}(n_1,\ldots,n_{q^*})} \left[\prod_{m=1}^{\sum_{r=1}^{q^*} d^{(r)}+q-q^*} (\beta + m\,\gamma) \right] \prod_{r=1}^{q^*} \mathrm{B}_{n_r, d^{(r)}}\Big(\boldsymbol{\omega}^{(r)}\Big)}{(\beta+n) \displaystyle\sum_{\underline{d}(n_1,\ldots,n_{q^*})} \left[\prod_{m=1}^{\sum_{r=1}^{q^*} d^{(r)}+q-q^*-1} (\beta + m\,\gamma) \right] \prod_{r=1}^{q^*} \mathrm{B}_{n_r, d^{(r)}}\Big(\boldsymbol{\omega}^{(r)}\Big)}.$$

In particular, if α is diffuse, one has

$$\begin{aligned} & P\{X_{n+1} \in B | X_1 = \tilde{x}_1, \ldots, X_n = \tilde{x}_n\} \\ & = \sum_{j=1,\ldots,q} \frac{n_j - \gamma}{\beta + n} 1(x_j \in B) + \frac{\beta + q\,\gamma}{\beta + n} \frac{\alpha(B)}{\alpha(\mathbb{R})}, \end{aligned}$$

which, for $\beta = \alpha(\mathbb{R})$, are the predictive distributions of the Poisson-Dirichlet models with two parameters. Whereas, for a general α, setting $\beta = \alpha(\mathbb{R})$ and $\gamma = 0$, one obtains the predictive distributions of the Ferguson-Dirichlet model.

5 A Special Case

In this section, we focus on the special case in which $h(t) = 1$ for every t, i.e., we consider sequences of observations which are conditionally i.i.d. with common distribution φ, given the NRMI φ.

The expressions for the law of the sequence simplifies as follows.

COROLLARY 5.1. *Let $X_1, X_2, \ldots$ be conditionally i.i.d. with common distribution φ, given the NRMI φ. For any measurable $B_1, \ldots, B_n$ and $n = 1, 2, \ldots$,*

$$\begin{aligned} & P\{X_1 \in B_1, \ldots, X_n \in B_n\} \\ & \quad = \sum_{\boldsymbol{\pi}(n)} \frac{1}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1} \phi(u) \left[\prod_{j=1}^{q(\boldsymbol{\pi})} \kappa_{n_j}(u; \cap_{i \in \pi_j} B_i) \right] du\,. \end{aligned} \tag{5.1}$$

In particular, if (2.3) *holds, then the right side of* (5.1) *reduces to*

$$\sum_{\boldsymbol{\pi}(n)} \frac{1}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1}\phi(u) \left[\prod_{j=1}^{q(\boldsymbol{\pi})} \kappa_{n_j}(u)\right] du \prod_{j=1}^{q(\boldsymbol{\pi})} \alpha\left(\cap_{i\in\pi_j} B_i\right).$$

Moreover, if $A_1, \ldots, A_q$ *are pairwise disjoint measurable sets, and* $n_1 + \cdots + n_q = n$, *then*

$$\begin{aligned} &P\{(X_1, \ldots, X_n) \in \widetilde{A}_1 \times \cdots \times \widetilde{A}_n\} \\ &\quad = \frac{1}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1}\phi(u) \left[\prod_{j=1}^{q} \mathrm{B}_{n_j}\left(\boldsymbol{\kappa}(u; A_j)\right)\right] du, \end{aligned} \tag{5.2}$$

where $\boldsymbol{\kappa}(u; A_j) := \big(\kappa_1(u; A_j), \ldots, \kappa_{n_j}(u; A_j)\big)$. *In particular, if (2.3) holds, then* $\boldsymbol{\kappa}(u; A_j) = \alpha(A_j) \times \boldsymbol{\kappa}^{(j)}(u)$, *where* $\boldsymbol{\kappa}^{(j)}(u) := \big(\kappa_1(u), \ldots, \kappa_{n_j}(u)\big)$.

It should be mentioned that James, Lijoi and Prünster (2005) derive (5.1) by means of Poisson process partition calculus methods, under the additional hypothesis of diffuseness of $\eta(\cdot, s)$. See also James (2002, 2005). For the special case where (2.3) holds, note that the quantity between braces in

$$\begin{aligned} &P\{X_1 \in B_1, \ldots, X_n \in B_n\} \\ &= \sum_{\boldsymbol{\pi}(n)} \left\{\frac{1}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1}\phi(u) \left[\prod_{j=1}^{q(\boldsymbol{\pi})} \kappa_{n_j}(u)\right] du\, \alpha(\mathbb{R})^{q(\boldsymbol{\pi})}\right\} \prod_{j=1}^{q(\boldsymbol{\pi})} \frac{\alpha\left(\cap_{i\in\pi_j} B_i\right)}{\alpha(\mathbb{R})} \end{aligned}$$

is the exchangeable partition probability function of Poisson-Kingman models provided by Pitman (2003, equation (36)).

Arguing as in the proof of Corollary 5.1, one can also consider the representations of predictive distributions based on point data, obtained in Corollary 4.1, in the special case of NRMI's. The expression of the predictive distributions for the case of an NRMI associated with a measure ν such as in (2.3), with α diffuse, were given in Pitman (2003), James (2002) and Prunster (2002).

COROLLARY 5.2. *Let* $X_1, X_2, \ldots$ *be conditionally i.i.d. with common distribution* φ, *given the NRMI* φ, *and assume that* (2.3) *holds. Then, the sequence* $X_1, X_2, \ldots$ *has predictive distributions of the form* (4.1) *with weights*

p_j's having the following expressions. If α is purely discrete, then

$$p_j = \frac{\displaystyle\int_{\mathbb{R}^+} u^n\,\phi(u)\left[\prod_{r\in\{1,\ldots,q\}\setminus\{j\}} \mathrm{B}_{n_r}\left(\boldsymbol{\kappa}(u;\{x_r\})\right)\right]\mathrm{B}_{n_j+1}\left(\boldsymbol{\kappa}(u;\{x_j\})\right)du}{\displaystyle n\int_{\mathbb{R}^+} u^{n-1}\phi(u)\left[\prod_{r=1}^{q}\mathrm{B}_{n_r}\left(\boldsymbol{\kappa}(u;\{x_r\})\right)\right]du}$$
$$-\frac{\displaystyle\int_{\mathbb{R}^+} u^n\,\phi(u)\left[\prod_{r=1}^{q}\mathrm{B}_{n_r}\left(\boldsymbol{\kappa}(u;\{x_r\})\right)\right]\kappa_1(u)\alpha(\{x_j\})du}{\displaystyle n\int_{\mathbb{R}^+} u^{n-1}\phi(u)\left[\prod_{r=1}^{q}\mathrm{B}_{n_r}\left(\boldsymbol{\kappa}(u;\{x_r\})\right)\right]du}$$

for $j=1,\ldots,q$, and

$$p_{q+1} = \frac{\displaystyle\int_{\mathbb{R}^+} u^n\,\phi(u)\left[\prod_{r=1}^{q}\mathrm{B}_{n_r}\left(\boldsymbol{\kappa}(u;\{x_r\})\right)\right]\kappa_1(u)\,\alpha(\mathbb{R})\,du}{\displaystyle n\int_{\mathbb{R}^+} u^{n-1}\phi(u)\left[\prod_{r=1}^{q}\mathrm{B}_{n_r}\left(\boldsymbol{\kappa}(u;\{x_r\})\right)\right]du}.$$

If α is diffuse, then

$$p_j = \frac{\displaystyle\int_{\mathbb{R}^+} u^n\phi(u)\left[\prod_{r\in\{1,\ldots,q\}\setminus\{j\}}\kappa_{n_r}(u)\right]\kappa_{n_j+1}(u)\,du}{\displaystyle n\int_{\mathbb{R}^+} u^{n-1}\phi(u)\left[\prod_{r=1}^{q}\kappa_{n_r}(u)\right]du}$$

for $j=1,\ldots,q$, and

$$p_{q+1} = \frac{\displaystyle\int_{\mathbb{R}^+} u^n\phi(u)\left[\prod_{r=1}^{q}\kappa_{n_r}(u)\right]\kappa_1(u)\,\alpha(\mathbb{R})\,du}{\displaystyle n\int_{\mathbb{R}^+} u^{n-1}\phi(u)\left[\prod_{r=1}^{q}\kappa_{n_r}(u)\right]du}.$$

Consider the Kingman γ-stable model. From Corollaries 5.1 and 5.2, or directly from the example developed in the previous sections, setting $\beta = 0$, one has the following expressions. For any $B_1, \ldots, B_n$,

$$P\{X_1 \in B_1, \ldots, X_n \in B_n\} = \sum_{\boldsymbol{\pi}(n)} \frac{\Gamma(q(\boldsymbol{\pi}))}{\gamma\,\Gamma(n)} \prod_{j=1}^{q(\boldsymbol{\pi})} \gamma\,[1-\gamma]_{n_j-1} \frac{\alpha\left(\cap_{i\in\pi_j} B_i\right)}{\alpha(\mathbb{R})}.$$

Analogously, if $A_1, \ldots, A_q$ are pairwise disjoint, and $n_1 + \ldots + n_q = n$,

$$P\{(X_1, \ldots, X_n) \in \widetilde{A}_1 \times \cdots \times \widetilde{A}_n\} = \sum_{\underline{d}(n_1,\ldots,n_q)} \frac{\Gamma\left(\sum_{k=1}^{q} d^{(j)}\right)}{\gamma\,\Gamma(n)} \prod_{j=1}^{q} \mathrm{B}_{n_j,d^{(j)}}\left(\gamma\boldsymbol{\omega}^{(j)}\right).$$

Concerning the predictive distributions, let α be a mixture of a diffuse and a discrete measure, and $x_1, \ldots, x_q$ be as in (4.2). Then,

$$p_j = \frac{\sum_{\underline{d}(n_1,\ldots,n_j+1,\ldots,n_{q^*})} \Gamma\left(\sum_{r=1}^{q^*} d^{(r)} + q - q^*\right) \prod_{r=1}^{q^*} \mathrm{B}_{n_r,d^{(r)}}\left(\gamma\boldsymbol{\omega}^{(r)}\right)}{n\sum_{\underline{d}(n_1,\ldots,n_{q^*})} \Gamma\left(\sum_{r=1}^{q^*} d^{(r)} + q - q^*\right) \prod_{r=1}^{q^*} \mathrm{B}_{n_r,d^{(r)}}\left(\gamma\boldsymbol{\omega}^{(r)}\right)} - \frac{\sum_{\underline{d}(n_1,\ldots,n_{q^*})} \Gamma\left(\sum_{r=1}^{q^*} d^{(r)} + q - q^* + 1\right) \left[\prod_{r=1}^{q^*} \mathrm{B}_{n_r,d^{(r)}}\left(\gamma\boldsymbol{\omega}^{(r)}\right)\right] \gamma\omega_1^{(j)}}{n\sum_{\underline{d}(n_1,\ldots,n_{q^*})} \Gamma\left(\sum_{r=1}^{q^*} d^{(r)} + q - q^*\right) \prod_{r=1}^{q^*} \mathrm{B}_{n_r,d^{(r)}}\left(\gamma\boldsymbol{\omega}^{(r)}\right)}$$

whenever $j = 1, \ldots, q^*$. Moreover,

$$p_j = \frac{n_j - \gamma}{n}$$

whenever $j = q^* + 1, \ldots, q$. Finally,

$$p_{q+1} = \frac{\sum_{\underline{d}(n_1,\ldots,n_{q^*})} \Gamma\left(\sum_{r=1}^{q^*} d^{(r)} + q - q^* + 1\right) \left[\prod_{r=1}^{q^*} \mathrm{B}_{n_r,d^{(r)}}\left(\gamma\boldsymbol{\omega}^{(r)}\right)\right] \gamma}{n\sum_{\underline{d}(n_1,\ldots,n_{q^*})} \Gamma\left(\sum_{r=1}^{q^*} d^{(r)} + q - q^*\right) \prod_{r=1}^{q^*} \mathrm{B}_{n_r,d^{(r)}}\left(\gamma\boldsymbol{\omega}^{(r)}\right)}.$$

When α is diffuse, this gives the well-known predictive distributions

$$P\{X_{n+1} \in B | X_1 = \tilde{x}_1, \ldots, X_n = \tilde{x}_n\} = \sum_{j=1,\ldots,q} \frac{n_j - \gamma}{\alpha(\mathbb{R}) + n} 1(x_j \in B) + \frac{\alpha(\mathbb{R}) + q\,\gamma}{\alpha(\mathbb{R}) + n} \frac{\alpha(B)}{\alpha(\mathbb{R})}$$

(see Pitman, 1995).

6 Hierarchical Models

The construction considered in Section 2 admits straightforward generalizations which cover hierarchical models such as the Dirichlet mixture models studied by Ferguson (1983) and Lo (1984), and increasingly popular since the first development of an appropriate Markov chain Monte Carlo technique by Escobar (1994) and Escobar and West (1995).

As an example, consider the random measure on $\mathbb{R}$ defined by

$$\varphi^K(\cdot) := \int_{\mathbb{R}} K(\cdot, x)\varphi(dx) \tag{6.1}$$

for an appropriate transition kernel K on $\mathbb{R} \times \mathbb{R}$. In this setting, the usual NRMI φ is recovered by taking $K(B, x) = 1(x \in B)$. Random measures such as the one defined in (6.1) are studied, e.g., by Nieto Barajas, Prünster and Walker (2004) and James, Lijoi and Prünster (2005). Now, for any measurable $B_1, \ldots, B_n$ and $n = 1, 2, \ldots$, set

$$\begin{aligned} &q^K_{B_1,\ldots,B_n}(d\theta_1, \ldots, d\theta_n) \\ &:= \int_{\mathbb{R}^+} Q\big\{\varphi^K(B_1) \in d\theta_1, \ldots, \varphi^K(B_n) \in d\theta_n \big| T = t\big\}\, h(t)\, Q\big\{T \in dt\big\}. \end{aligned}$$

The family $\{q^K_{B_1,\ldots,B_n}\colon B_1, \ldots, B_n$ measurable subsets of $\mathbb{R}, n = 1, 2, \ldots\}$ determines a unique probability distribution q^K for some random probability τ^K, and $q^K_{B_1,\ldots,B_n}$ represents the finite-dimensional distribution of τ^K at $\{B_1, \ldots, B_n\}$. It is obvious that when $h(t) = 1$ for every t, one recovers the law of the random measure φ^K.

Let $(Y_n)_n$ be a sequence of exchangeable observations which are conditionally i.i.d. with common random distribution τ^K, given the random measure τ^K. Then, as in (3.1), for any n and any measurable $B_1, \ldots, B_n$,

$$\begin{aligned} &P\{Y_1 \in B_1, \ldots, Y_n \in B_n\} \\ &\qquad = \int_{\mathbb{R}^+} h(t)\, \frac{1}{t^n}\, E\left[\prod_{i=1}^{n} \int_{\mathbb{R}} K(B_i, x)\mu(dx) \middle| T = t\right] f_T(dt). \end{aligned}$$

By a simple application of Fubini's theorem, one can check that it is possible to obtain a vector of exchangeable observations $(Y_1, \ldots, Y_n)$ having the above law, thanks to the following hierarchical structure: $X_1, \ldots, X_n$ are exchangeable variables, conditionally i.i.d. with common distribution τ, given the conditioned NRMI τ; given $X_1, \ldots, X_n$, the observations $Y_1, \ldots, Y_n$ are independent, with distributions $K(\cdot, X_1), \ldots, K(\cdot, X_n)$ respectively.

Set

$$l_i^K(B_r : r \in \pi_j)(ds) := s^i\, 1_{s>0} \left[\int_{\mathbb{R}} \left\{ \prod_{r \in \pi_j} K(B_r, x) \right\} \eta(dx, s) \right] \rho(ds).$$

By arguments same as those used to derive Proposition 3.1 and of Corollary 5.1, one gets the following proposition and corollary.

Proposition 6.1. *Let $Y_1, Y_2, \dots$ be conditionally i.i.d. with common distribution τ^K, given the random measure τ^K. For any measurable $B_1, \dots, B_n$ and $n = 1, 2, \dots,$*

$$P\{Y_1 \in B_1, \dots, Y_n \in B_n\}$$
$$= \sum_{\boldsymbol{\pi}(n)} \int_{\mathbb{R}^+} h(t)\, \frac{1}{t^n} \left(f_T * l_{n_1}^K (B_i : i \in \pi_1) * \cdots * l_{n_{q(\boldsymbol{\pi})}}^K \left(B_i : i \in \pi_{q(\boldsymbol{\pi})}\right)\right)(dt). \tag{6.2}$$

In particular, if (2.3) *holds, then the right hand side of* (6.2) *reduces to*

$$\sum_{\boldsymbol{\pi}(n)} \int_{\mathbb{R}^+} h(t)\, \frac{1}{t^n} \left(f_T * l_{n_1} * \cdots * l_{n_{q(\boldsymbol{\pi})}} \right)(dt) \prod_{j=1}^{q(\boldsymbol{\pi})} \int_{\mathbb{R}} \left[\prod_{i \in \pi_j} K(B_i, x) \right] \alpha(dx).$$

Corollary 6.1. *Let $Y_1, Y_2, \dots$ be conditionally i.i.d. with common distribution φ^K, given the random measure φ^K. For any measurable $B_1, \dots, B_n$ and $n = 1, 2, \dots,$*

$$P\{Y_1 \in B_1, \dots, Y_n \in B_n\} =$$
$$\sum_{\boldsymbol{\pi}(n)} \frac{1}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1} \phi(u) \left[\prod_{j=1}^{q(\boldsymbol{\pi})} \int_{\mathbb{R} \times \mathbb{R}^+} e^{-us} s^{n_j} \left\{ \prod_{i \in \pi_j} K(B_i, x) \right\} \nu(dx\, ds) \right] du. \tag{6.3}$$

In particular, if (2.3) *holds, then the right hand side of* (6.3) *reduces to*

$$\sum_{\boldsymbol{\pi}(n)} \frac{1}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1} \phi(u) \left[\prod_{j=1}^{q(\boldsymbol{\pi})} \kappa_{n_j}(u) \right] du \prod_{j=1}^{q(\boldsymbol{\pi})} \int_{\mathbb{R}} \left[\prod_{i \in \pi_j} K(B_i, x) \right] \alpha(dx).$$

James, Lijoi and Prünster (2005) derive (6.3) by means of Poisson process partition calculus methods, under the additional hypothesis of diffuseness of $\eta(\cdot, s)$. See also James (2002, 2005).

As an example, starting from the Poisson-Dirichlet model with three parameters (α, γ, β) with a transition kernel K, one has, for any $B_1, \ldots, B_n$,

$$P\{X_1 \in B_1, \ldots, X_n \in B_n\} = \sum_{\boldsymbol{\pi}(n)} \frac{\prod_{m=1}^{q(\boldsymbol{\pi})-1} (\beta + m\,\gamma)}{[\beta+1]_{n-1}} \prod_{j=1}^{q(\boldsymbol{\pi})} [1-\gamma]_{n_j-1} \frac{\int_{\mathbb{R}} \left[\prod_{i\in\pi_j} K(B_i, x)\right] \alpha(dx)}{\alpha(\mathbb{R})}.$$

Starting from the Kingman γ-stable model with a transition kernel K, for any $B_1, \ldots, B_n$,

$$P\{X_1 \in B_1, \ldots, X_n \in B_n\} = \sum_{\boldsymbol{\pi}(n)} \frac{\Gamma(q(\boldsymbol{\pi}))}{\gamma\,\Gamma(n)} \prod_{j=1}^{q(\boldsymbol{\pi})} \gamma\,[1-\gamma]_{n_j-1} \frac{\int_{\mathbb{R}} \left[\prod_{i\in\pi_j} K(B_i, x)\right] \alpha(dx)}{\alpha(\mathbb{R})}.$$

Following the seminal work of Escobar (1994) and Escobar and West (1995), various Markov chain Monte Carlo techniques have been developed for Dirichlet process hierarchical models. See e.g., the reviews in Neal (2000) and Papaspiliopoulos and Roberts (2004). The Gibbs sampling methods described by these authors make use of the predictive distributions of the Dirichlet process. The same approach can be adopted to deal with the more general case of hierarchical models based on conditioned NRMI's, by using the expressions of the predictive distributions provided in Corollary 4.1. Assume that $K(\cdot, x)$ has density $f(\cdot|x)$ with respect to some σ-finite measure λ for each x in $\mathbb{R}$. Then, to draw values from the posterior distribution of $\boldsymbol{X} = (X_1, \ldots, X_n)$, based on the sample $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$, one iteratively simulates values from the conditional distributions

$$P(X_i \in dx | \boldsymbol{X}^{(-i)} = \tilde{\boldsymbol{x}}^{(-i)}, \boldsymbol{Y} = \boldsymbol{y}) = \sum_{j=1,\ldots,q^{(-i)}} w_j^{(-i)}\, 1(x_j^{(-i)} \in dx) + w_{q^{(-i)}+1}^{(-i)} \frac{f(y_i|x)\alpha(dx)}{\int_{\mathbb{R}} f(y_i|x)\alpha(dx)},$$

where $w_j^{(-i)} \propto p_j^{(-i)} f(y_i|x_j^{(-i)})$ and $w_{q^{(-i)}+1}^{(-i)} \propto p_{q^{(-i)}+1}^{(-i)} \int_{\mathbb{R}} f(y_i|x)\alpha(dx)/\alpha(\mathbb{R})$, subject to the constraint $\sum_{j=1}^{q^{(-i)}+1} w_j^{(-i)} = 1$, and where the quantities with index $(-i)$ relate to the vector $\boldsymbol{X}^{(-i)} = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$. Lijoi, Mena and Prünster (2005a) and Lijoi, Mena and Prünster (2005b), e.g., apply this method to hierarchical models based on two concrete classes of NRMI's. Alternatively, it is possible to devise variations of the Chinese restaurant procedure, as the one set forth by Ishwaran and James (2003).

Acknowledgments. I am indebted to Eugenio Regazzini for introducing me to this line of research and providing countless helpful suggestions. I am also grateful to Gareth Roberts, Omiros Papaspiliopoulos, R.V. Ramamoorthi, Pietro Rigo and Federico Basetti for useful comments. Finally, I would like to thank the co-editor and the referee for many helpful comments.

Appendix

PROOF OF LEMMA 3.1. Note that

$$\int_{\mathbb{R}^+} e^{-ut} E\left[\prod_{i=1}^{n} \mu(B_i) \middle| T = t\right] f_T(dt)$$
$$= \frac{\partial^n}{\partial y_1 \cdots \partial y_n} E\left[\exp\left\{-uT + \sum_{i=1}^{n} y_i \mu(B_i)\right\}\right]\Bigg|_{\boldsymbol{y}=\mathbf{0}},$$

where $\boldsymbol{y} = (y_1, \ldots, y_n)$. Next, thanks to (2.1),

$$\frac{\partial^n}{\partial \boldsymbol{y}} E\left[\exp\left\{-uT + \sum_{i=1}^{n} y_i \mu(B_i)\right\}\right]\Bigg|_{\boldsymbol{y}=\mathbf{0}}$$
$$= \frac{\partial^n}{\partial \boldsymbol{y}} \exp\left\{-\int_{\mathbb{R}\times\mathbb{R}^+}\left[1 - \exp\left\{\left(-u + \sum_{i=1}^{n} y_i 1(x \in B_i)\right)s\right\}\right]\nu(dx\,ds)\right\}\Bigg|_{\boldsymbol{y}=\mathbf{0}}. \tag{A.1}$$

To compute the right side of (A.1), we resort to a multivariate generalization of the Faà di Bruno formula, provided by Constantine and Savits (1996), which in our case leads to

$$\frac{\partial^n}{\partial \boldsymbol{y}} \exp\left\{-\int_{\mathbb{R}\times\mathbb{R}^+}\left[1 - \exp\left\{\left(-u + \sum_{i=1}^{n} y_i 1(x \in B_i)\right)s\right\}\right]\nu(dx\,ds)\right\}\Bigg|_{\boldsymbol{y}=\mathbf{0}}$$
$$= \sum_{\boldsymbol{\pi}(n)} \phi(u) \prod_{j=1}^{q(\boldsymbol{\pi})} \kappa_{n_j}(u; \cap_{i\in\pi_j} B_i),$$

where the summation is extended over all partitions $\boldsymbol{\pi} = \{\pi_1, \ldots, \pi_{q(\boldsymbol{\pi})}\}$ of $\{1, \ldots, n\}$, and n_j denotes the number of elements in π_j. Since $\kappa_{n_j}(u; B)$ is the Laplace-Stieltjes transform of the measure $l_{n_j}(B)$ at u, and, by (2.1),

$\phi(u)$ is the Laplace-Stieltjes transform of T at u, we can write

$$\begin{aligned}
&\phi(u) \prod_{j=1}^{q(\boldsymbol{\pi})} \kappa_{n_j} \left(u; \cap_{i\in\pi_j} B_i\right) \\
&= \int_{\mathbb{R}^+} e^{-ut} \, \left(f_T * l_{n_1} \left(\cap_{i\in\pi_1} B_i\right) * \cdots * l_{n_{q(\boldsymbol{\pi})}} \left(\cap_{i\in\pi_{q(\boldsymbol{\pi})}} B_i\right)\right)(dt).
\end{aligned} \tag{A.2}$$

Thus, we get

$$\begin{aligned}
&\int_{\mathbb{R}^+} e^{-ut} \, E\left[\prod_{i=1}^{n} \mu(B_i) \middle| T = t\right] f_T(dt) \\
&= \int_{\mathbb{R}^+} e^{-ut} \sum_{\boldsymbol{\pi}(n)} \left(f_T * l_{n_1} \left(\cap_{i\in\pi_1} B_i\right) * \cdots * l_{n_{q(\boldsymbol{\pi})}} \left(\cap_{i\in\pi_{q(\boldsymbol{\pi})}} B_i\right)\right)(dt),
\end{aligned}$$

which gives (3.2). One derives (3.3) by the same arguments. □

Proof of Proposition 4.1. One has

$$\begin{aligned}
&P\{X_{n+1} \in B | (X_1, \ldots, X_n) \in \widetilde{A}_1 \times \cdots \times \widetilde{A}_n\} \\
&= \frac{P\{(X_1, \ldots, X_n, X_{n+1}) \in A_1^{n_1} \times \cdots \times A_q^{n_q} \times B\}}{P\{(X_1, \ldots, X_n) \in A_1^{n_1} \times \cdots \times A_q^{n_q}\}}.
\end{aligned}$$

Since $B = \left\{\cup_{j=1}^{q} (A_j \cap B)\right\} \cup \left\{B \cap (\cup_{j=1}^{q} A_j)^c\right\}$,

$$\begin{aligned}
&P\{(X_1, \ldots, X_{n+1}) \in A_1^{n_1} \times \cdots \times A_q^{n_q} \times B\} \\
&= \sum_{j=1}^{q} \sum_{z=0}^{n_j} \binom{n_j}{z} P\{(X_1, \ldots, X_{n+1}) \in A_1^{n_1} \times \cdots \times (A_j \cap B)^{z+1} \\
&\qquad \times (A_j \cap B^c)^{n_j - z} \times \cdots \times A_q^{n_q}\} \\
&\quad + P\left\{(X_1, \ldots, X_{n+1}) \in A_1^{n_1} \times \cdots \times A_q^{n_q} \times \left(B \cap (\cup_{r=1}^{q} A_r)^c\right)\right\}.
\end{aligned}$$

Thus, in order to prove Proposition 4.1, it suffices to show that

$$\sum_{z=0}^{n_j}\binom{n_j}{z}P\{(X_1,\ldots,X_{n+1})\in A_1^{n_1}\times\cdots\times(A_j\cap B)^{z+1}\times(A_j\cap B^c)^{n_j-z}\times\cdots\times A_q^{n_q}\}$$

$$=\frac{\alpha(A_j\cap B)}{\alpha(A_j)}\,P\{(X_1,\ldots,X_{n+1})\in A_1^{n_1}\times\cdots\times A_j^{n_j+1}\times\cdots\times A_q^{n_q}\}.$$

Thanks to (3.1), this is tantamount to verifying that

$$\sum_{z=0}^{n_j}\binom{n_j}{z}E\big[\mu(A_1)^{n_1}\cdots\mu(A_j\cap B)^{z+1}\mu(A_j\cap B^c)^{n_j-z}\cdots\mu(A_q)^{n_q}\big|T=t\big]\,f_T(dt)$$

$$=\frac{\alpha(A_j\cap B)}{\alpha(A_j)}\,E\big[\mu(A_1)^{n_1}\cdots\mu(A_j)^{n_j+1}\cdots\mu(A_q)^{n_q}\big|T=t\big]\,f_T(dt).$$

This can be checked by showing the equality of the Laplace-Stieltjes transform of the left and right sides. Now, by the arguments same as those used in the proof of Lemma 3.1, one has

$$\int_{\mathbb{R}^+}e^{-ut}\,E\left[\prod_{j=1}^{q}\mu(A_j)^{n_j}\middle|T=t\right]f_T(dt)=\phi(u)\prod_{j=1}^{q}\mathrm{B}_{n_j}\left(\boldsymbol{\kappa}(u;A_j)\right),$$

and hence

$$\int_{\mathbb{R}^+}e^{-ut}\sum_{z=0}^{n_j}\binom{n_j}{z}$$

$$E\big[\mu(A_1)^{n_1}\cdots\mu(A_j\cap B)^{z+1}\mu(A_j\cap B^c)^{n_j-z}\cdots\mu(A_q)^{n_q}\big|T=t\big]f_T(dt)$$

$$=\phi(u)\left[\prod_{r\in\{1,\ldots,q\}\setminus\{j\}}\mathrm{B}_{n_r}\left(\boldsymbol{\kappa}(u,A_r)\right)\right]$$

$$\times\left[\sum_{z=0}^{n_j}\binom{n_j}{z}B_{z+1}\left(\boldsymbol{\kappa}(u,A_j\cap B)\right)\mathrm{B}_{n_j-z}\left(\boldsymbol{\kappa}(u,A_j\cap B^c)\right)\right].$$

Next, by the properties of exponential partial Bell partition polynomials (see Charalambides, 2002, equation (11.10) and Section 11.7),

$$\sum_{z=0}^{n_j}\binom{n_j}{z}\mathrm{B}_{z+1}\left(\boldsymbol{\kappa}(u;A_j\cap B)\right)\,\mathrm{B}_{n_j-z}\left(\boldsymbol{\kappa}(u;A_j\cap B^c)\right)$$

$$=\frac{\alpha(A_j\cap B)}{\alpha(A_j)}\,\mathrm{B}_{n_j+1}\left(\boldsymbol{\kappa}(u;A_j)\right).$$

Thus,

$$\int_{\mathbb{R}^+} e^{-ut}\sum_{z=0}^{n_j}\binom{n_j}{z}$$
$$E\big[\mu(A_1)^{n_1}\cdots\mu(A_j\cap B)^{z+1}\mu(A_j\cap B^c)^{n_j-z}\cdots\mu(A_q)^{n_q}\big|T=t\big]f_T(dt)$$
$$=\phi(u)\Bigg[\prod_{r\in\{1,\ldots,q\}\setminus\{j\}}\mathrm{B}_{n_r}\left(\boldsymbol{\kappa}(u,A_r)\right)\Bigg]\frac{\alpha(A_j\cap B)}{\alpha(A_j)}\,B_{n_j+1}\left(\boldsymbol{\kappa}(u,A_j)\right)$$
$$=\int_{\mathbb{R}^+}e^{-ut}\frac{\alpha(A_j\cap B)}{\alpha(A_j)}\,E\big[\mu(A_1)^{n_1}\cdots\mu(A_j)^{n_j+1}\cdots\mu(A_q)^{n_q}\big|T=t\big]f_T(dt),$$

and this concludes the proof. □

Proof of Corollary 4.1. For δ positive, let ${}^{\delta}\!A_j=[x_j,x_j+\delta)$, $j=1,\ldots,q$. Then

$$P\{X_{n+1}\in B|X_1=\cdots=X_{n_1}=x_1,\ldots,X_{n_1+\ldots+n_{q-1}+1}=\cdots=X_n=x_q\}$$
$$=\lim_{\delta\to 0}P\{X_{n+1}\in B|(X_1,\ldots,X_n)\in {}^{\delta}\!A_1^{n_1}\times\cdots\times {}^{\delta}\!A_q^{n_q}\}.$$

Using the expression of the predictive distributions based on grouped data, given in Proposition 4.1, one can write

$$P\{X_{n+1}\in B|(X_1,\ldots,X_n)\in {}^{\delta}\!A_1^{n_1}\times\cdots\times {}^{\delta}\!A_q^{n_q}\}$$
$$=\sum_{j=1}^{q}{}^{\delta}p_j\frac{\alpha({}^{\delta}\!A_j\cap B)}{\alpha({}^{\delta}\!A_j)}+{}^{\delta}p_{q+1}\frac{\alpha(B)}{\alpha(\mathbb{R})},$$

where

$${}^{\delta}p_j:=\frac{P\{(X_1,\ldots,X_{n+1})\in {}^{\delta}\!A_1^{n_1}\times\cdots\times{}^{\delta}\!A_j^{n_j+1}\times\cdots\times{}^{\delta}\!A_q^{n_q}\}}{P\{(X_1,\ldots,X_n)\in {}^{\delta}\!A_1^{n_1}\times\cdots\times{}^{\delta}\!A_j^{n_j}\times\cdots\times{}^{\delta}\!A_q^{n_q}\}}$$
$$-\frac{\alpha({}^{\delta}\!A_j)}{\alpha\left(\mathbb{R}\cap(\cup_{r=1}^{q}{}^{\delta}\!A_r)^c\right)}$$
$$\cdot\frac{P\{(X_1,\ldots,X_{n+1})\in {}^{\delta}\!A_1^{n_1}\times\cdots\times{}^{\delta}\!A_q^{n_q}\times(\mathbb{R}\cap(\cup_{r=1}^{q}{}^{\delta}\!A_r)^c)\}}{P\{(X_1,\ldots,X_n)\in {}^{\delta}\!A_1^{n_1}\times\cdots\times{}^{\delta}\!A_j^{n_j}\times\cdots\times{}^{\delta}\!A_q^{n_q}\}}$$

for $j=1,\ldots,q$, and

$${}^{\delta}p_{q+1}:=\frac{\alpha(\mathbb{R})}{\alpha\left(\mathbb{R}\cap(\cup_{r=1}^{q}{}^{\delta}\!A_r)^c\right)}$$
$$\cdot\frac{P\{(X_1,\ldots,X_{n+1})\in {}^{\delta}\!A_1^{n_1}\times\cdots\times{}^{\delta}\!A_q^{n_q}\times(\mathbb{R}\cap(\cup_{r=1}^{q}{}^{\delta}\!A_r)^c)\}}{P\{(X_1,\ldots,X_n)\in {}^{\delta}\!A_1^{n_1}\times\cdots\times{}^{\delta}\!A_q^{n_q}\}}.$$

Note that, for any δ, $\sum_{j=1}^{q+1} {}^{\delta}p_j = 1$. Moreover, ${}^{\delta}p_{q+1} \geq 0$ and, using (11.10) in Charalambides (2002), it can be easily seen that ${}^{\delta}p_j \geq 0$ for any $j = 1, \dots, q$. Set $p_j := \lim_{\delta \to 0} {}^{\delta}p_j$, for $j = 1, \dots, q+1$. Then $\sum_{j=1}^{q+1} p_j = 1$ and $p_j \geq 0$ for $j = 1, \dots, q+1$. Now, the limit of $\alpha({}^{\delta}\!A_j \cap B)/\alpha({}^{\delta}\!A_j)$, as ${}^{\delta}\!A_j$ shrinks to $\{x_j\}$, is 1 if $x_j \in B$ and 0 if $x_j \notin B$, i.e., $\lim_{\delta \to 0} \alpha({}^{\delta}\!A_j \cap B)/\alpha({}^{\delta}\!A_j) = 1(x_j \in B)$. Finally, from Proposition 3.1, one immediately derives the expressions of the weights associated with a measure α which is purely discrete. On the other hand, when α is diffuse, is suffices to note that

$$\frac{P\{(X_1, \dots, X_n, X_{n+1}) \in {}^{\delta}\!A_1^{n_1} \times \cdots \times {}^{\delta}\!A_j^{n_j+1} \times \cdots \times {}^{\delta}\!A_q^{n_q}\}/\alpha({}^{\delta}\!A_1) \cdots \alpha({}^{\delta}\!A_q)}{P\{(X_1, \dots, X_n) \in {}^{\delta}\!A_1^{n_1} \times \cdots \times {}^{\delta}\!A_j^{n_j} \times \cdots \times {}^{\delta}\!A_q^{n_q}\}/\alpha({}^{\delta}\!A_1) \cdots \alpha({}^{\delta}\!A_q)}$$
$$= \frac{\int_{\mathbb{R}^+} h(t) \, \frac{1}{t^{n+1}} \, (f_T * l_{n_1} * \cdots * l_{n_j+1} * \cdots * l_{n_q})(dt) + G\left({}^{\delta}\!A_1, \dots, {}^{\delta}\!A_q\right)}{\int_{\mathbb{R}^+} h(t) \, \frac{1}{t^n} \, (f_T * l_{n_1} * \cdots * l_{n_j} * \cdots * l_{n_q})(dt) + F\left({}^{\delta}\!A_1, \dots, {}^{\delta}\!A_q\right)}$$

and

$$\frac{P\{(X_1, \dots, X_{n+1}) \in {}^{\delta}\!A_1^{n_1} \times \cdots \times {}^{\delta}\!A_q^{n_q} \times (\mathbb{R} \cap (\cup_{r=1}^q {}^{\delta}\!A_r)^c)\}/\alpha({}^{\delta}\!A_1) \cdots \alpha({}^{\delta}\!A_q)}{P\{(X_1, \dots, X_n) \in {}^{\delta}\!A_1^{n_1} \times \cdots \times {}^{\delta}\!A_q^{n_q}\}/\alpha({}^{\delta}\!A_1) \cdots \alpha({}^{\delta}\!A_q)}$$
$$= \frac{\int_{\mathbb{R}^+} h(t) \, \frac{1}{t^{n+1}} \, (f_T * l_{n_1} * \cdots * l_{n_q} * l_1)(dt) \, \alpha(\mathbb{R} \cap (\cup_{r=1}^q {}^{\delta}\!A_r)^c) + G'\left({}^{\delta}\!A_1, \dots, {}^{\delta}\!A_q\right)}{\int_{\mathbb{R}^+} h(t) \, \frac{1}{t^n} \, (f_T * l_{n_1} * \cdots * l_{n_q})(dt) + F\left({}^{\delta}\!A_1, \dots, {}^{\delta}\!A_q\right)}$$

with $G\left({}^{\delta}\!A_1, \dots, {}^{\delta}\!A_q\right)$, $G'\left({}^{\delta}\!A_1, \dots, {}^{\delta}\!A_q\right)$ and $F\left({}^{\delta}\!A_1, \dots, {}^{\delta}\!A_q\right)$ going to 0 as δ goes to 0, yielding the expressions of the weights in Corollary 4.1. □

PROOF OF COROLLARY 5.1. In order to derive the expressions in Corollary 5.1 from the corresponding in Proposition 3.1, it suffices to note that

$$\int_{\mathbb{R}^+} \frac{1}{t^n} \left(f_T * l_{n_1}(\cap_{i \in \pi_1} B_i) * \cdots * l_{n_{q(\boldsymbol{\pi})}}(\cap_{i \in \pi_{q(\boldsymbol{\pi})}} B_i) \right) (dt)$$
$$= \int_{\mathbb{R}^+} \left[\frac{1}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1} e^{-ut} du \right] \left(f_T * l_{n_1}(\cap_{i \in \pi_1} B_i) * \cdots * l_{n_{q(\boldsymbol{\pi})}}(\cap_{i \in \pi_{q(\boldsymbol{\pi})}} B_i) \right) (dt)$$
$$= \frac{1}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1} \int_{\mathbb{R}^+} e^{-ut} \left(f_T * l_{n_1}(\cap_{i \in \pi_1} B_i) * \cdots * l_{n_{q(\boldsymbol{\pi})}}(\cap_{i \in \pi_{q(\boldsymbol{\pi})}} B_i) \right) (dt) \, du$$
$$= \frac{1}{\Gamma(n)} \int_{\mathbb{R}^+} u^{n-1} \phi(u) \left[\prod_{j=1}^{q(\boldsymbol{\pi})} \kappa_{n_j}(u; \cap_{i \in \pi_j} B_i) \right] du,$$

where the first equality is due to the identity $t^{-n} = \int_{\mathbb{R}^+} u^{n-1} e^{-ut} du/\Gamma(n)$, the second one is due to Fubini's theorem, and the last one is due to (A.2).

This gives (5.1). By the same arguments and using exponential Bell partition polynomials, one has (5.2). □

References

ASH, R.B. (2000). *Probability and Measure Theory.* Harcourt/Academic Press, Burlington, MA, 2nd ed., with contributions by Catherine Doleans-Dade.

CHARALAMBIDES, C.A. (2002). *Enumerative combinatorics, CRC Press Series on Discrete Mathematics and its Applications*, Chapman & Hall/CRC, Boca Raton, FL.

CONSTANTINE, G.M. and SAVITS, T.H. (1996). A multivariate Faá di Bruno formula with applications. *Trans. Amer. Math. Soc.* **348**, 503–520.

ESCOBAR, M.D. (1994). Estimating normal means with a Dirichlet process prior. *J. Amer. Statist. Assoc.* **89**, 268–277.

ESCOBAR, M.D. and WEST, M. (1995). Bayesian density estimation and inference using mixtures. *J. Amer. Statist. Assoc.* **90**, 577–588.

FERGUSON, T.S. (1983). Bayesian density estimation by mixtures of normal distributions. In *Recent Advances in Statistics : Papers in honor of Herman Chernoff on his sixtieth birthday.* M.H. Rizvi, J.S. Rustagi and D. Siegmund, eds., Academic Press, New York, 287–302.

HANSEN, B. and PITMAN, J. (2000). Prediction rules for exchangeable sequences related to species sampling. *Statist. Probab. Lett.* **46**, 251–256.

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. Unpublished manuscript. Also at `http://arxiv.org/PS_cache/math/pdf/0205/0205093.pdf` (Mathematics ArXiv, math.PR/0205093).

JAMES, L.F. (2005). Bayesian Poisson process partition calculus with an application to Bayesian Levy moving averages. *Ann. Statist.* **33**, 1771–1799.

JAMES, L. F., LIJOI, A. and PRÜNSTER, I. (2005). Bayesian nonparamentric inference via classes of normalized random measures. Tech. Rep. ICER WP 2005/5, Università degli Studi di Torino and International Center for Economic Research. Also at `http://web.econ.unito.it/gma/icer-wps.htm`.

JAMES, L. F., LIJOI, A. and PRÜNSTER, I. (2006). Conjugacy as a distinctive feature of the Dirichlet process. *Scand. J. Statist.* **33**, 105–120.

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

KINGMAN, J.F.C. (1975). Random discrete distribution (with discussion). *J. Roy. Statist. Soc. Ser. B.* **37**, 1–22.

KINGMAN, J.F.C. (1978). The representation of partition structures. *J. London Math. Soc. (2).* **18**, 374–380.

LIJOI, A., MENA, R.H. and PRÜNSTER, I. (2005a). Bayesian nonparametric analysis for a generalized Dirichlet process prior. *Stat. Inference Stoch. Process.* **8**, 283–309.

LIJOI, A., MENA, R.H. and PRÜNSTER, I. (2005b). Hierarchical mixture modelling with normalized Gaussian priors. *J. Amer. Statist. Assoc.* **100**, 1278–1291.

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

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.

NEAL, R.M. (2000). Markov chain sampling methods for Dirichlet process mixture models. *J. Comput. Graph. Statist.* **9**, 249–265.

NIETO-BARAJAS, L.E., PRÜNSTER, I. and WALKER, S.G. (2004). Normalized random measures driven by increasing additive processes. *Ann. Statist.* **32**, 2343–2360.

PAPASPILIOPOULOS, O. and ROBERTS, G.O. (2004). Retrospective Markov chain Monte Carlo methods for Dirichlet process hierarchical models. Tech. rep. Lancaster University. Also at `http://www.maths.lancs.ac.uk/~papaspil/research.html`.

PERMAN, M., PITMAN, J. and YOR, M. (1992). Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Related Fields.* **92**, 21-39.

PITMAN, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields.* **102**, 145–158.

PITMAN, J. (1996a). Random discrete distributions invariant under sizebiased permutation. *Adv. in Appl. Probab.* **28**, 525–539.

PITMAN, J. (1996b). Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, Probability and Game Theory: Papers in honor of David Blackwell.* IMS Lecture Notes and Monograph Series **30**, T.S. Ferguson, L.S. Shapley and J.B. MacQueen, eds., Institute of Mathematical Statistics, Hayward, CA, 245–267.

PITMAN, J. (2003). Poisson-Kingman partitions. In *Statistics and Science: a Festschrift for Terry Speed.* IMS Lecture Notes and Monograph Series **40**, D.R. Goldstein, ed., Institute of Mathematical Statistics, Beachwood, OH, 1–34.

PITMAN, J. and YOR, M. (1997). The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25**, 855-900.

PRÜNSTER, I. (2002). *Random Probability Measures Derived from Increasing Additive Processes and Their Application to Bayesian Statitics.* PhD thesis, University of Pavia.

REGAZZINI, E. (1998). An example of the interplay between statistics and special functions. *Atti dei Convegni Lincei.* **147**, 303–320.

REGAZZINI, E., LIJOI, A. and PRÜNSTER, I. (2003). Distributional results for means of normalized random measures with independent increments. *Ann. Statist.* **31**, 560–585.

LAURA M. SANGALLI
DIPARTIMENTO DI MATEMATICA
UNIVERSITÀ DEGLI STUDI DI PAVIA
VIA FERRATA 1, 27100 PAVIA, ITALY
E-mail: laura.sangalli@unipv.it

Paper received January 2006; revised May 2006.