Partially Exchangeable Random Partitions

By

Jim Pitman*

Technical Report No. 343
May 1992

*Research supported by NSF Grant MCS91-07531

Department of Statistics
University of California
Berkeley, California 94720

Partially Exchangeable Random Partitions

Jim Pitman *
Department of Statistics
U.C. Berkeley CA 94720

May 13, 1992

Abstract

A random partition of the positive integers is called *partially exchangeable* if for each finite sequence of positive integers $n_1, \ldots, n_k$, the probability that the partition breaks the first $n_1 + \cdots + n_k$ integers into k particular classes, of sizes $n_1, \ldots, n_k$ in order of their first elements, has the same value $p(n_1, \ldots, n_k)$ for every possible choice of classes subject to the sizes constraint. A random partition is exchangeable iff it is partially exchangeable for a symmetric function $p(n_1, \ldots, n_k)$. A representation is given for partially exchangeable random partitions that is similar to Kingman's representation in the exchangeable case. These representations are viewed as variations of de Finetti's representation of exchangeable sequences, and as identifications of the Martin boundary of associated Markov chains. In the exchangeable case, information is provided about the joint distribution of the proportions of classes in order of their appearance. This gives a constraint on the finite dimensional distributions of a random discrete probability distribution on the positive integers that is equivalent to invariance under size-biased random permutation. The results are illustrated by the two-parameter generalization of Ewens' partition structure.

*Research supported by N.S.F. Grant MCS91-07531

1 Introduction

Kingman[28] introduced the concept of a *partition structure*, that is to say a seqence $(\mathbb{P}_n,\ n = 1, 2, \ldots)$ of distributions for random partitions π_n of n, which is *consistent* in the following sense: if n objects are partitioned into subsets with sizes given by π_n, and an object is deleted uniformly at random, independently of π_n, the partition of the $n-1$ remaining objects has component sizes distributed according to $\mathbb{P}_{n-1}$. Kingman[30] and Aldous[1] interpreted this concept in terms of exchangeable partitions Π_n of the sets $\mathbb{N}_n := \{1, \ldots, n\}$. Given π_n, the partition of n, the partition Π_n of the set $\mathbb{N}_n$ is defined by assuming all partitions with component sizes dictated by π_n are equally likely. The distributions $\mathbb{P}_n$ for π_n are consistent iff the restriction of Π_n to $\mathbb{N}_{n-1}$ has the same distribution as Π_{n-1}. By Kolmogorov's extension theorem, there is then a one to one correspondence between partition structures $(\mathbb{P}_n)$ and distributions for an *exchangeable random partition of* $\mathbb{N} := \{1, 2, \ldots\}$, that is a random partition $\Pi = (\Pi_n)$ of $\mathbb{N}$ whose restriction Π_n to $\mathbb{N}_n$ is exchangeable for every n. The *distribution* of Π then refers to the collection of distributions of the Π_n. In other terms, this is Kingman's idea of an *exchangeable random equivalence relation* on $\mathbb{N}$.

Given a sequence of random variables $(X_1, X_2, \ldots)$, let $\Pi(X_1, X_2, \ldots)$, a random partition of $\mathbb{N}$, be induced by the random equivalence relation

$$i \sim j \Leftrightarrow X_i = X_j.$$

Clearly, if $(X_1, X_2, \ldots)$ is exchangeable then $\Pi(X_1, X_2, \ldots)$ is an exchangeable partition of $\mathbb{N}$. Conversely, starting from Π, an exchangeable partition of $\mathbb{N}$, Aldous [1] constructs exchangeable random variables $X_1, X_2, \ldots$ with values in $[0, 1]$ so that $\Pi = \Pi(X_1, X_2, \ldots)$ almost surely, by letting $X_i = U_j$ if i belongs to the jth component of Π, where the U_j are uniform $[0, 1]$ random variables, independent of each other and of Π. As Aldous points out, de Finetti's theorem then implies the X_i have limiting empirical distribution P_∞, a random probability measure on $[0, 1]$, and the conditional distribution of Π given P_∞ is the distribution of $\Pi(X_1, X_2, \ldots)$ where $X_1, X_2, \ldots$ are i.i.d. according to P_∞. This is Kingman's representation theorem: the most general partition structure $(\mathbb{P}_n)$ is obtained by letting $\mathbb{P}_n$ be the distribution of the partition of n induced by the partition $\Pi(X_1, \ldots, X_n)$ of $\mathbb{N}_n$, where given P_∞ the $X_1, X_2, \ldots$ are i.i.d. with distribution P_∞, and P_∞ is allowed to be random.

In principle, the distribution $I\!P_n$ of the partition of n is determined by the distribution of the random probability measure P_∞. The only feature of the distribution of P_∞ that is relevant to computation of $I\!P_n$ is the joint distribution of the sizes of the ranked atoms of P_∞, say

$$P_{(1)} \geq P_{(2)} \geq \ldots \geq 0. \tag{1}$$

Thus Kingman's representation sets up a one to one correspondence between partition structures $(I\!P_n)$ and joint distributions for a sequence of random variables $(P_{(1)}, P_{(2)}, \ldots)$ satisfying the order constraint (1) and $\sum_{i=1}^{\infty} P_{(i)} \leq 1$. If $\Pi = (\Pi_n)$ is an exchangeable random partition of $I\!N$ such that the partition of n induced by Π_n has distribution $I\!P_n$, then corresponding $P_{(i)}$ can be expressed as

$$P_{(i)} = \lim_{n\to\infty} \frac{N_{(i)n}}{n} \ a.s.,$$

where $N_{(i)n}$ is the size of the ith largest component in Π_n. See Kingman[28, 30], Aldous[1] for further discussion.

Two difficulties arise in working with ranked sizes of components in a partition structure. Firstly, the joint distribution of the limiting ranked proportions $P_{(i)}$ turn out to be rather complicated, even for the simplest partition structures, such as those corresponding to Ewens' sampling formula. See for instance Shepp and Lloyd [43], Vershik & Schmidt [45], Watterson [47]. Secondly, formulae for the distribution $I\!P_n$ of π_n in terms of the joint distribution of the $P_{(i)}$ involve infinite sums of expectations of products of the $P_{(i)}$, which are not easy to evaluate. For instance the probability that π_n consists of two components of different sizes n_1 and n_2 is

$$\binom{n_1+n_2}{n_1} \sum_{i\neq j} E[P_{(i)}^{n_1} P_{(j)}^{n_2}].$$

In the case corresponding to Ewens' sampling formula it is well known that these difficulties are avoided if the size-biased permutation of the atoms of P_∞ is considered instead of the rank ordering (Hoppe [22, 23, 24], Donelly[8], Ewens[16]). See Donnelly and Joyce [10] for a general discussion of relations between the ranked and size-biased presentations of a random discrete distribution. But while size-biasing the atoms simplifies many distributional computations, it complicates the formulation of Kingman's representation. In the *proper* case, when P_∞ is discrete almost surely, the problem is to

describe which random discrete probability distributions (P_n) are invariant under size-biased permutation, a condition that appears at first to be essentially infinite dimensional. Still, it turns out that this condition can be presented as a conjunction of simple constraints on the finite dimensional distributions of (P_n). See Corollary 16.

The present paper offers a broad view of these matters by a variation of Kingman's representation for a larger class of random partitions of $\mathbb{N}$, called *partially exchangeable*. The terminology is consistent with the general concept of partial exchangeability due to de Finetti [6]. See Diaconis and Freedman [7] for a recent survey. The representation provided by Theorem 15 of this paper fits perfectly into the Diaconis-Freedman framework of extreme point descriptions for models defined by a sequence of sufficient statistics. See also Martin-Löf [36], Lauritzen [32, 33, 34] and Dynkin [13] for development of similar frameworks. However the proof of the representation is based on direct application of de Finetti's theorem rather than any general extreme point theory.

When specialized to the exchangeable case, the main representation theorem provides information not readily availiable from Kingman's representation. The main theme of the paper is that there are several results involving size-biased sampling of components from an exchangeable random partition which seem best understood as the restriction to the exchangeable case of corresponding results for partially exchangeable partitions.

Section 2 records formulae relating various standard codings of a random partition of n to a single basic function derived from the associated exchangeable partition of $\mathbb{N}_n$, the *exchangeable partition probability function* (EPF). Partially exchangeable partions of $\mathbb{N}_n$ are introduced in Section 3, along with their probability functions (PEPF's). (One letter P is omitted in the acronyms to avoid excessive alliteration). The representation theorem for partially exchangeable partitions of $\mathbb{N}$ is established in Section 4. There is a close parallel between the results of Sections 3 and 4 and certain formulations of de Finetti's theorem for exchangeable sequences of random variables with a countable number of values, considered in Section 5. Section 6 makes the connection with the Doob-Hunt theory of Martin boundaries for Markov chains. Section 7 considers the family of partially exchangeable partitions of $\mathbb{N}$ derived from residual allocation models with independent factors. In particular, a two-parameter family of such models with beta distributed factors corresponds to the two-parameter generalization of Ewens' partition struc-

ture studied in Pitman[41]. It is shown in Pitman [40] that apart from some trivial examples, all residual allocation models that are invariant under size biased sampling belong to this two-parameter family.

2 Preliminaries

Let $n \in \mathbb{N} = \{1, 2, \ldots\}$. A *partition of* n is an unordered collection of n positive integers with sum n. There are two common ways to code a partition of n:
(i) by the finite sequence of ranked integer component sizes, say

$$n_{(1)} \geq n_{(2)} \geq \ldots \geq n_{(k)} \text{ with } \Sigma_{i=1}^{k} n_{(i)} = n.$$

(ii) by the collection of counts of component sizes

$$m_j = \#\{i : n_{(i)} = j\},\ j = 1, \ldots, n.$$

The *number of components* in the partition is $\sum m_j = k$, while $\sum j m_j = n$. A *random partition of* n, is a random variable π_n with values in the set of all partitions of n. Let K_n denote the random number of parts of π_n. The distribution of π_n, call it $\mathbb{P}_n$, is a probability distribution on the set of all unordered partitions of n. Such a distribution$\mathbb{P}_n$ is typically presented either (i) via the joint distribution of the counts of components of different sizes; or (ii) via the joint distribution of K_n and the K_n component sizes presented in some order.

In case (ii) the order could be ranked order, purely random order, size-biased random order, or an order with some more complex dependence on the partition of n.

It is clear in principle that any one these presentations of$\mathbb{P}_n$ determines each of the others. In the literature of models for random partitions, formulae for different presentations of particular random partitions have been derived from each other in many special cases. See for example Watterson[47], Donnelly and Tavaré[12]. The general form of these connections is made obvious by relating each presentation to yet another coding of the distribution of π_n. This involves Π_n, the random partition of the set $\mathbb{N}_n = \{1, 2, \ldots, n\}$, induced as follows by the random partition π_n of the integer n: given π_n with $K_n = k$, Π_n is equally likely to be any of the unordered partitions of $\mathbb{N}_n$ into disjoint

subsets $\{A_i\}_1^k$ with sizes dictated by π_n This random partition Π_n of the set $\mathbb{N}_n$ is an *exchangeable random partition (EP) of* $\mathbb{N}_n$ as defined by Aldous[1]. It is clear that for any particular partition of $\mathbb{N}_n$ into non-empty subsets $\{A_i\}_1^k$ of sizes n_i,

$$P(\Pi_n = \{A_i\}_1^k) = p(n_i), \quad \text{where} \quad (n_i) \to p(n_i) = p(n_1, \ldots, n_k) \tag{2}$$

is some symmetric function of sequences of positive integers

$$(n_i) = (n_1, \ldots, n_k) \text{ with } \Sigma_1^k n_i = n. \tag{3}$$

A random partition Π_n of $\mathbb{N}_n$ is exchangeable iff (2) holds for some symmetric function $p(n_i)$. Then call $p(n_i)$ an *exchangeable partition probability function* (EPF).

The following proposition expresses various presentations of the distribution of a random partition of n in terms of the corresponding EPF. The point is that any of the basic presentations can be expressed in terms of the EPF and simple combinatorial factors. This suggests that the EPF should be regarded as the fundamental descriptor both for an EP of $\mathbb{N}_n$, and for the corresponding random partition of n. This point of view is maintained throughout the paper.

Proposition 1 *Let π_n be a random partition of n with K_n components. Let Π_n be the exchangeable random partition $\mathbb{N}_n$ associated with π_n, and let $p(n_1, \ldots, n_k)$ as in (2) be the EPF that gives the probability that Π_n is any particular partition of $\mathbb{N}_n$ into k subsets A_i of sizes n_i, $i = 1, \ldots, k$.*
Component size counts: *The joint distribution of $M_1, \ldots, M_n$, where M_j is the number of parts of π_n of size j, is given by*

$$P(M_j = m_j, 1 \le j \le n) = N(m_j)\tilde{p}(m_j) \tag{4}$$

where

$$N(m_j) := \frac{n!}{\prod_{j=1}^n (j!)^{m_j} m_j!} \tag{5}$$

is the number of partitions of $\mathbb{N}_n$ into m_j classes of size j, $1 \le j \le n$, and $\tilde{p}(m_j)$ is the common value of the symmetric function $p(n_i)$ for every (n_i) with

$$\#(i : n_i = j\} = m_j, \ 1 \le j \le n. \tag{6}$$

Ranked order: *The joint distribution* $N_{(1)}, \ldots, N_{(K_n)}$, *the ranked component sizes, is given by*

$$P(K_n = k,\ N_{(i)} = n_i, 1 \le i \le k) = N(m_j)p(n_i) \tag{7}$$

where the m_j *are defined by* (6) .
Exchangeable order: *If given* $K_n = k$ *and* $N_{(1)}, \ldots, N_{(k)}$, *the sequence* $N_1^{EX}, \ldots, N_k^{EX}$ *presents the component sizes in exchangeable random order (all* $k!$ *permutations equally likely),*

$$P(K_n = k, N_i^{EX} = n_i, 1 \le i \le k) = \frac{n!}{k! \prod_{i=1}^k n_i!} p(n_i) \tag{8}$$

Size-biased order: *If given* $K_n = k$ *and* $N_{(1)}, \ldots, N_{(k)}$, *the sequence* $N_1^{SB}, \ldots, N_k^{SB}$ *is a size-biased random permutation of* $N_{(1)}, \ldots, N_{(k)}$, (*in particular if the* N_i^{SB} *are the component sizes of* Π_n *ordered by their first elements*)

$$P(K_n = k, N_i^{SB} = n_i, 1 \le i \le k) = \#(n_i)p(n_i) \tag{9}$$

where

$$\#(n_i) = \frac{n!}{n_k(n_k + n_{k-1}) \cdots (n_k + \cdots + n_1)\ \prod_{i=1}^k (n_i - 1)!} \tag{10}$$

is the number of ways to arrange n_1 *values of one type,* n_2 *of a second, and so on, subject to the constraint that the first value is of the first type, the next distinct value of the second type, and so on.*
Arbitrary order: *If* $N_1, \ldots, N_{K_n}$ *is a presentation of the components of* π_n *in some order with arbitrary dependence on* π_n, *and*

$$P(K_n = k, N_1 = n_1, \ldots, N_k = n_k) = p_0(n_1, \ldots, n_k),$$

then

$$p(n_1, \ldots, n_k) = \frac{\prod_{i=1}^k n_i!}{n!} \sum_\sigma p_0(n_{\sigma(1)}, \ldots, n_{\sigma(k)}) \tag{11}$$

where the sum is over all $k!$ *permutations* σ *of* $\{1, \ldots, k\}$.

Proof. These formulae all follow from the definition and symmetry of $p(n_1, \ldots, n_k)$ by elementary counting arguments. To derive (11), randomize the order of components and use (8). □

Corollary 2 *The most general possible distribution for π_n, a random partition of n, corresponds via the formulae of Proposition 1 to an arbitrary non-negative symmetric function $p(n_1, \ldots, n_k)$ with*

$$\sum_{(n_i)} \frac{n!}{k! \prod_{i=1}^{k} n_i!} p(n_i) = 1, \tag{12}$$

where the sum is over all finite sequences of positive integers (n_i) with $\sum_i n_i = n$.

Proof. This is immediate from formula (8). □

Alternative expressions for the sum on the left side of (12) can be derived from (7) and (9). See also Corollary 7 for a simple recursive way to check (12).

3 Partitions of the first n integers

Let $\mathbb{N}^* = \bigcup_{k=1}^{\infty} \mathbb{N}^k$, the set of finite sequences of positive integers. Denote a generic element of $\mathbb{N}^*$ by $(n_i) = (n_1, \ldots, n_k)$. Write $\sum n_i$ for $\sum_{i=1}^{k} n_i$. Let $p(n_i)$ be a non-negative function of (n_i).

Definition 3 A random partition Π_n of $\mathbb{N}_n$ is a *partially exchangeable partition* (PEP) if for some $p(n_i)$ defined for

$$(n_i) \text{ with } \Sigma n_i = n, \tag{13}$$

for every partition $\{A_i\}$ of $\mathbb{N}_n$ into non-empty subsets of sizes n_i that satisfy the *order constraint*: $1 \in A_1$, the first element not in A_1 is in A_2, and so on,

$$P(\Pi_n = \{A_i\}) = p(n_i). \tag{14}$$

Formula (14) sets up a one to one correspondence between distributions of a PEP of Π_n, and non-negative functions $p(n_i)$ with domain (13) such that

$$\sum_{(n_i)} \#(n_i) p(n_i) = 1. \tag{15}$$

where the sum is over the domain (13), and $\#(n_i)$ as in (10) is the number of partitions of $\mathbb{N}_n$ satisfying the order constraint with the given component

sizes. Less intimidating conditions equivalent to (15) are provided below. Call $p(n_i)$ the *partially exchangeable partition probability function* (PEPF) corresponding to Π_n. Note that Π_n is exchangeable iff Π_n is a PEP with PEPF that is a symmetric function of (n_i). The PEPF is then an EPF as in Section 2.

The meaning of the partial exchangeability condition is clarified by reformulating the definition in various ways. First some notation. Let K_n be the number of components of Π_n, and let N_i be the size of the ith component of Π_n, when components are ordered by their first elements. Let

$$N_{*n} := (N_1, \ldots, N_{K_n}), \tag{16}$$

a random element of $\mathbb{N}^*$. Then N_{*n} is a sufficient statistic for distributions of PEP's Π_n. More formally:

Proposition 4 *A random partition Π_n of $\mathbb{N}_n$ is a PEP iff given $N_{*n} = (n_i)$ the partition Π_n is uniformly distributed over the $\#(n_i)$ distinct partitions of $\mathbb{N}_n$ that meet the order constraint for the given sizes (n_i). Then the distribution of N_{*n} is related to the PEPF p of Π_n by*

$$P(N_{*n} = (n_i)) = \#(n_i)p(n_i). \tag{17}$$

Proof. This is immediate from the definitions. □

The most general partially exchangeable distribution of Π_n is thus obtained by prescribing an arbitrary distribution for N_{*n} over (n_i) with $\sum n_i = n$, then declaring that the distribution of Π_n is uniform given N_{*n}. The rest of this section develops various other descriptions of PEP's of $\mathbb{N}_n$. The reader primarily interested in the representation theorem for PEP's of $\mathbb{N}$ can skip to the next section and refer to this one only as necessary for proofs.

The following construction, easily seen to yield the most general PEP of $\mathbb{N}_n$, is the finite "sampling without replacement" version of the construction used in the next section to create the most general PEP of $\mathbb{N}$.

Construction 5 Let $\mathcal{A}_1, \ldots, \mathcal{A}_{K_n}$ denote the random subsets of $\mathbb{N}_n$ defined by the components of Π_n ordered by the first elements. Let N_1, the size of $\mathcal{A}_1$, have distribution

$$P(N_1 = n_1) = P(n_1),\ 1 \leq n_1 \leq n,$$

where $P(\cdot)$ is some arbitrary probability distribution on $\{1,\ldots,n\}$. Given $N_1 = n_1$, let $\mathcal{A}_1$ consist of 1 and a uniformly distributed random subset of $n - n_1 - 1$ elements of $\{2,\ldots,n\}$.

Inductively: Given that $\mathcal{A}_1,\ldots,\mathcal{A}_i$ have been defined, with $N_j = n_j$, $1 \le j \le i$, such that $\sum_{j=1}^{i} n_i < n$, let N_{i+1} have distribution

$$P(N_{i+1} = n_{i+1}|\mathcal{A}_1,\ldots,\mathcal{A}_i) = P(n_{i+1}|n_1,\ldots,n_i) \tag{18}$$

where $P(\cdot\,|n_1,\ldots,n_i)$ is some arbitrary distribution on $\{1,\ldots,n-\sum_{j=1}^{i} n_j\}$. And given $\mathcal{A}_1\ldots,\mathcal{A}_i$ and $N_{i+1} = n_{i+1}$, let $\mathcal{A}_{i+1}$ comprise the first element of $\mathbb{N}_n - \bigcup_1^i \mathcal{A}_j$ together with a uniformly distributed random subset of $n_{i+1} - 1$ elements of the remaining $n - \sum_1^i n_j - 1$ elements of $\mathbb{N}_n$.

The random partition Π_n so constructed is partially exchangeable, with PEPF

$$p(n_i) = \frac{P(n_1)}{\binom{n-1}{n_1-1}} \frac{P(n_2|n_1)}{\binom{n-n_1-1}{n_2-1}} \cdots \frac{P(n_k|n_1,\ldots,n_{k-1})}{\binom{n-n_1-\cdots-n_{k-1}-1}{n_k-1}}.$$

In the preceding descriptions of a PEP of $\mathbb{N}_n$, the size n of the set partitioned was thought of as fixed. A more dynamic view develops in terms of growing the sequence of partitions Π_m, where Π_m is the restriction of Π_n to $\mathbb{N}_m$ for $m = 1,\ldots,n$. It is easily seen that if Π_n is a PEP, then so is Π_m for $m = 1,\ldots,n$. This leads to the characterization of PEPF's provided by the next proposition.

Notation For $(n_i) = (n_1,\ldots,n_k) \in \mathbb{N}^*$, define $(n_i^{j+}) \in \mathbb{N}^*$ by incrementing n_j by 1:

$$(n_i^{j+}) := (n_i + 1(i = j)) \ \in \mathbb{N}^* \tag{19}$$

for $j = 1,\ldots,k+1$, where it is understood that $(n_i^{(k+1)+})$ is obtained by appending a 1 to (n_i) at place $k+1$.

In the exchangeable case, the consistency condition of the following proposition is the expression in terms of EPF's of Kingman's notion of consistency between the distributions of a partition of m and a partition of $m+1$. This is a simpler expression of consistency than Kingman's [27] formula in terms of counts of component sizes. But the two conditions are the same due to formula (4).

Proposition 6 *Fix $n \in \mathbb{N}$. Let $p(n_i)$ be a non-negative function defined for $(n_i) \in \mathbb{N}^*$ with $\sum n_i = n$. Define the* consistent extension *of p to $(n_i) \in \mathbb{N}^*$ with $\sum n_i \leq n$ by inductive application of the* consistency condition*:*

$$p(n_i) = \sum_{j=1}^{k+1} p(n_i^{j+}) \tag{20}$$

to $(n_i) = (n_1, \ldots, n_k)$ with $\sum_i n_i = m$, for $m = n-1, n-2, \ldots, 1$. Then the following are equivalent:
(i) *$p(n_i)$ for $(n_i) \in \mathbb{N}^*$ with $\sum n_i = n$ is the PEPF for some PEP of $\mathbb{N}_n$, say Π_n.*
(ii) *The consistent extension of $p(n_i)$ to (n_i) with $\sum n_i \leq n$ is such that $p(1) = 1$.*
(iii) *For each $m = 1, \ldots, n$, the consistent extension of $p(n_i)$ restricted to (n_i) with $\sum n_i = m$ is the PEPF of Π_m, the PEP of $\mathbb{N}_m$ that is the restriction of Π_n to $\mathbb{N}_m$.*

Proof. The obvious probabilistic interpretation of the consistency condition shows that (i) $\Rightarrow$ (iii) $\Rightarrow$ (ii). To see that (ii) $\Rightarrow$ (i), suppose that $p(n_i)$ solves (20) for $\sum n_i < n$, with $p(1) = 1$. Let Σ be the sum on the left side of (15). It must be shown that $\Sigma = 1$. Clearly $\Sigma > 0$, so $q(n_i) := p(n_i)/\Sigma$ is a PEPF. Application of (i) $\Rightarrow$ (ii), for q instead of p, shows that $p(1)/\Sigma = q(1) = 1$. Thus $\Sigma = p(1) = 1$. □

Corollary 7 *A PEP of $\mathbb{N}_n$ is exchangeable iff its PEPF is a symmetric function. Let Π_n be the exchangeable random partition associated with π_n, a random partition of n. Then the formula $P(\Pi_n = \{A_i\}) = p(n_i)$ sets up a one to one correspondence between distributions $\mathbb{P}_n$ for π_n, and non-negative symmetric functions $p(n_i)$, defined for all sequences of positive integers (n_i) with $\sum n_i \leq n$, that satisfy the consistency condition* (20) *and $p(1) = 1$.*

Proof. This follows immediately from the definitions and the preceding proposition. □

Recall from Section 2 that the symmetric PEPF corresponding to an exchangeable partition of $\mathbb{N}_n$ derived from a random partition of n, is called the *exchangeable partition probability function (EPF)* for the random partition of n.

Proposition 8 *Let Π_n be a PEP of $\mathbb{N}_n$ with PEPF $p(n_1, \ldots, n_k)$. Then the EPF $p^*(n_1, \ldots, n_k)$ associated with the partition of n induced by Π_n is*

$$p^*(n_1, \ldots, n_k) = \frac{\Pi_{i=1}^k n_i!}{n!} \sum_\sigma \#(n_{\sigma(1)}, \ldots, n_{\sigma(k)}) p(n_{\sigma(1)}, \ldots, n_{\sigma(k)}) \qquad (21)$$

where the sum is over all $k!$ permutations σ of $\{1, \ldots, k\}$.

Proof. This follows immediately from formulae (17) and (11). □

Note. As the following example makes clear, $p^*(n_i)$ is *not* the usual symmetrization of $p(n_i)$.

Example 9 Define the PEPF for a PEP Π_3 of $\{1, 2, 3\}$ by $p(n_i) \geq 0$ where $p(3) = p(1, 1, 1) = 0$ and $p(1, 2)$ and $p(2, 1)$ are arbitrary subject to the constraint (15) : $p(1, 2) + 2p(2, 1) = 1$. So

$$P(\Pi_3 = \{\{1\}, \{2, 3\}\}) = p(1, 2), \qquad (22)$$

$$P(\Pi_3 = \{\{1, 2\}, \{3\}\}) = P(\Pi_3 = \{\{1, 3\}, \{2\}\}) = p(2, 1) \qquad (23)$$

The corresponding exchangeable random partition of $\{1, 2, 3\}$ places equal probabilities on these three partitions:

$$p^*(1, 2) = p^*(2, 1) = 1/3, \qquad (24)$$

so

$$p^*(i, j) \neq \frac{1}{2}[p(i, j) + p(j, i)] \qquad (25)$$

unless $p(1, 2) = p(2, 1) = 1/3$.

Corollary 10 *For a PEP Π_n, with components of sizes $(N_1, \ldots, N_{K_n})$ in order of their first elements, the following four statements are equivalent:*

- *Π_n is exchangeable.*
- *$p(n_i)$ is a symmetric function of (n_i) with $\sum n_i = n$.*
- *$p^*(n_i) = p(n_i)$ for all (n_i) with $\sum n_i = n$.*
- *The distribution of $(N_1, \ldots, N_{K_n})$ is invariant under size-biased random permutation.*

Proof. This follows immediately from the preceding proposition and Proposition 1. □

For $m \leq n$, with Π_m the restriction to $\mathbb{N}_m$ of Π_n, let

$$N_{*m} := (N_{1,m}, \ldots, N_{K_m,m}), \tag{26}$$

the element of $\mathbb{N}^*$ defined by the sizes of components of Π_m in order of their first elements. As m increases, N_{*m} develops at each step either by incrementing one of its components by 1, or by adding 1 as a new component at the right end. Which component is incremented between m and $m+1$ shows how element $m+1$ is attached by Π_{m+1} to the classes already defined by Π_m. Thus $(N_{*1}, \ldots, N_{*n})$ is a bijective transformation of partitions Π_n of $\mathbb{N}_n$ into sequences of length n of elements of $\mathbb{N}^*$ subject to the obvious constraints that $N_{*1} = (1)$, and that if $N_{*m} = (n_i) = (n_1, \ldots, n_k)$, say, then $N_{*m+1} = (n_i^{j+})$ for some $j = 1, \ldots, k+1$, for every $m = 1, \ldots n-1$. If Π_n is a random partition, then $(N_{*1}, \ldots, N_{*n})$ is an $\mathbb{N}^*$ valued random process. The distribution of Π_n determines that of $(N_{*1}, \ldots, N_{*n})$, and vice versa.

Proposition 11 *Let $(N_{*1}, \ldots, N_{*n})$ be the $\mathbb{N}^*$ valued random process associated as above with Π_n, a random partition of $\mathbb{N}_n$. The following three conditions are equivalent:*

(i) *Π_n is a PEP with PEPF $p(n_i)$.*

(ii) *$(N_{*1}, \ldots, N_{*n})$ is a Markov chain with co-transition probabilities*

$$P(N_{*m-1} = (n_i^{j-}) \mid N_{*m} = (n_i)) \;=\; \frac{\#(n_i^{j-})}{\#(n_i)}, \quad j = 1, \ldots k, \tag{27}$$

for $(n_i) = (n_1, \ldots n_k)$ with $\sum n_i = m$, $m = 2, \ldots n$, where

$$(n_i^{j-}) := (n_i - 1(i = j)) \tag{28}$$

for $j = 1, \ldots, k$, and $\#(n_i)$ is defined by (10) for $(n_i) \in \mathbb{N}^*$, and defined to be zero otherwise.

(iii) *$(N_{*1}, \ldots, N_{*n})$ is a Markov chain with transition probabilities of the special form*

$$P(N_{*m+1} = (n_i^{j+}) \mid N_{*m} = (n_i)) \;=\; \frac{p(n_i^{j+})}{p(n_i)}, \quad j = 1, \ldots k+1, \tag{29}$$

for $(n_i) = (n_1, \ldots n_k)$ with $\sum n_i = m$ and $p(n_i) > 0$, $m = 1, \ldots n-1$, and zero otherwise, for some non-negative function $p(n_i)$ defined for all $(n_i) \in \mathbb{N}^*$ with $\sum n_i \leq n$, such that $p(1) = 1$.

Proof.

That (i) $\Rightarrow$ (ii) follows from Proposition 4. For it is easily seen that $(N_{*1}, \ldots, N_{*n})$ is Markovian with the stated co-transition probabilities in case Π_n is exchangeable, and according to Proposition 4, the conditional distribution of Π_n given N_{*n}, hence also that of $(N_{*1}, \ldots, N_{*n})$ given N_{*n}, is the same for Π_n partially exchangeable as for Π_n exchangeable.

That (ii) $\Rightarrow$ (i) follows immediately from (i) $\Rightarrow$ (ii) and the fact that Π_n is determined by $(N_{*1}, \ldots, N_{*n})$.

Next, (ii) $\Rightarrow$ (iii) by Bayes' rule and the definition of the PEPF. Finally, (iii) $\Rightarrow$ (i) by the implication (ii) $\Rightarrow$ (i) of Proposition 6. The required consistency of $p(n_i)$ is due to the implicit assumption in (iii) that formula (29) defines a transition probability matrix. □

An application of the above proposition is provided in Section 5 by Proposition 24.

4 Partitions of the positive integers

Definition. A random partition Π of $\mathbb{N}$ is *partially exchangeable* iff for every n the restriction Π_n of Π to $\mathbb{N}_n$ is a PEP of $\mathbb{N}_n$.

Proposition 12 (i) *The formula* $P(\Pi_n = \{A_i\}) = p(n_i)$ *sets up a one to one correspondence between distributions of PEP's* (Π_n) *of* $\mathbb{N}$*, and their PEPF's, that is to say non-negative functions* $p(n_i)$ *defined for all* $(n_i) \in \mathbb{N}^*$*, with the consistency property* (20) *for all* (n_i)*, and* $p(1) = 1$.
(ii) *In particular, there is a one to one correspondence between Kingman's partition structures* $(\mathbb{P}_n,\ n = 1, 2, \ldots)$*, and EPF's, that is to say symmetric PEPF's defined for all* $(n_i) \in \mathbb{N}^*$.

Proof. This is immediate from Proposition 6. □

The variation of Kingman's "paintbox" construction required for the partially exchangeable case is the following generalization of a construction due to Hoppe [24] in the exchangeable case corresponding to Ewens' sampling formula. Think of painting integers in $\mathcal{A}_1$ a first color, then those in $\mathcal{A}_2$ a second color, and so on.

Construction 13 Given an arbitrary joint distribution for a sequence of random variables $(W_1, W_2, \ldots)$ with values $W_i \in [0,1]$, define a random partition Π of $\mathbb{N}$ into random subsets $\mathcal{A}_1, \mathcal{A}_2, \ldots$ as follows. Let $(X_{ni}, n = 1,2,\ldots, i = 1,2,\ldots)$ be indicator variables with

$$P(X_{ni} = 1 | W_1, W_2, \ldots) = W_i.$$

Let $\mathcal{A}_1 = \{1\} \cup \{n : X_{n1} = 1\}$. Given that $\mathcal{A}_1 \neq \mathbb{N}$ (or, what is the same, $W_1 < 1$) let $\mathcal{C}_1 = \mathbb{N} - \mathcal{A}_1$, let $\mathcal{A}_2 = \{\min\{\mathcal{C}_1\}\} \cup \{n : n \in \mathcal{C}_1 \text{ and } X_{n2} = 1\}$, and so on. Let $\mathcal{C}_i = \mathbb{N} - (\mathcal{A}_1 \cup \ldots \cup \mathcal{A}_i)$. Given $\mathcal{C}_i$ is non-empty (or, what is the same, $\Pi_{j=1}^{i}(1 - W_i) > 0$), let

$$\mathcal{A}_{i+1} = \{\min\{\mathcal{C}_i\} \cup \{n : n \in \mathcal{C}_i \text{ and } X_{n,i+1} = 1\}$$

Note that by construction the $\mathcal{A}_i$ are in order of their first elements. It is easily seen that an alternative, sequential description of this random partition of $\mathbb{N}$ can be given as follows.

Proposition 14 *Let*

$$P_i = \bar{W}_1 \ldots \bar{W}_{i-1} W_i,$$

for W_i as in the previous construction, where $\bar{w} = 1 - w$. Then $\Pi_1 = \{1\}$, and for each $n \in \mathbb{N}$, conditionally given $\Pi_n = \{\{A_i\}_1^k\}$, Π_{n+1} is an extension of Π_n in which element $n+1$ attaches to class A_i with probability P_i, $1 \leq i \leq k$, and forms a new class with probability $1 - \sum_1^i P_j$.

It follows at once from this sequential description that $\Pi = (\Pi_n)$ so created is a PEP of $\mathbb{N}$ with PEPF

$$p(n_1, \ldots, n_k) = E\left[\left(\prod_{i=1}^{k} P_i^{n_i - 1}\right) \prod_{i=1}^{k-1} (1 - \sum_{1}^{i} P_j)\right] \tag{30}$$

$$= E\left[\prod_{i=1}^{k} W_i^{n_i - 1} \bar{W}_i^{n_{i+1} + \cdots + n_k}\right]. \tag{31}$$

The P_i appear in Π as almost sure limits, due to the strong law of large numbers:

$$\lim_{n \to \infty} \frac{\#(\mathcal{A}_i \cap \mathbb{N}_n)}{n} = P_i \text{ a.s., } i = 1,2,\ldots \tag{32}$$

The analog of Kingman's representation theorem for PEP's of $\mathbb{N}$ is the following:

Theorem 15 *Every PEP of $\mathbb{N}$ has the same distribution as one of the kind described by Construction 13 and Proposition 14. Formula* (30) *sets up a one to one correspondence between PEPF's for a PEP of $\mathbb{N}$, as described in Proposition* 12 *, and joint distributions for a sequence of random variables $(P_1, P_2, \ldots)$ with $P_i \geq 0$ and $\sum P_i \leq 1$. If $\Pi = \{\mathcal{A}_1, \mathcal{A}_2, \ldots\}$ is a PEP of $\mathbb{N}$, where the $\mathcal{A}_i$ are ordered by their first elements, then there exist almost sure limits (P_i) as in* (32)*, and the joint distribution of these (P_i) is related to the PEPF of Π via* (30)*.*

Remark. In keeping with the general theory of partial exchangeability induced by a sequence of sufficient statistics, the set of all distributions of PEP's of $\mathbb{N}$ is an abstract simplex. The theorem identifies the set of extreme points of this simplex as the set of distributions of PEP's constructed as in Proposition 14 from deterministic $P_i = p_i$, say. The corresponding PEPF's are of the form

$$p(n_1, \ldots, n_k) = \prod_{i=1}^{k} p_i^{n_i - 1} \prod_{i=1}^{k-1} (1 - \Sigma_1^i p_j). \tag{33}$$

for a deterministic sequence (p_i) with $p_i \geq 0$ and $\sum p_i \leq 1$. This is the PEPF for the unique PEP such that the asymptotic proportion in the ith class to appear is p_i.

Proof. Let $\Pi = \{\mathcal{A}_1, \mathcal{A}_2, \ldots\}$ be a PEP corresponding to the PEPF $p(n_i)$. Define positive integer valued random variables $T_1, T_2, \ldots$ by $T_m = i$ if $m \in \mathcal{A}_i$. Call T_m the *type* of m. Then for every sequence of positive integers $(i_m, 1 \leq m \leq n)$ with $i_m = 1$, $i_{m+1} \leq i_m + 1$, and $\#\{m : i_m = i\} = n_i$, $1 \leq i \leq k = \max_{1 \leq m \leq n} i_m$

$$P(T_m = i_m, 1 \leq m \leq n) = p(n_i) \tag{34}$$

Let

$$N_{in} = \#\{m \leq n : T_m = i\} = \#(\mathcal{A}_i \cap \mathbb{N}_n)$$

Now (34) implies the sequence of indicator random variables

$$(1(T_n = 1),\ n = 2, 3, \ldots) \tag{35}$$

is exchangeable. By de Finetti's theorem,

$$\lim_{n \to \infty} \frac{N_{1n}}{n} = P_1 \text{ a.s.}$$

for some random variable P_1 with $0 \leq P_1 \leq 1$. Moreover, conditionally given P_1, the $1(T_n = 1)$, $n = 2, 3, \ldots$ are independent Bernoulli (P_1) random variables. Let

$$\nu_i := \inf\{n : N_{in} = 1\} = \inf\{n : T_n \geq i\}.$$

Given that $\nu_2 = m$, it can be seen from (34) that the partial exchangeability implies the random variables

$$T_n 1(T_n \leq 2),\ n = m+1, m+2, \ldots \tag{36}$$

are exchangeable, hence

$$\lim_{n\to\infty} \frac{N_{2n}}{n} = P_2 \ a.s.$$

for some P_2 with $P_2 \geq 0$ and $0 \leq P_1 + P_2 \leq 1$. Moreover given $\nu_2 = m$, and the values of P_1 and P_2, the random variables $T_n 1(T_n \leq 2)$ for $n \geq m+1$ are independent, with distribution according to the table

value	0	1	2
probability	$1 - P_1 - P_2$	P_1	P_2

From the previous analysis of the sequence (35)

$$P(\nu_2 = m|P_1) = P_1^{m-2}(1 - P_1),\ m = 2, 3, \ldots \tag{37}$$

But for $n \geq m$

$$P(\nu_2 = m|N_{1n}) = P(\nu_2 = m|N_{*n}), = P(\nu_2 = m|N_{*r}, r \geq n)$$

where the first equality is clear from Construction 5, and the second is due to the Markov property of $(N_{*r}, r = 1, 2, \ldots)$ described in Proposition 11. It follows that (37) can be sharpened by reversed martingale convergence to

$$P(\nu_2 = m|P_1, P_2) = P_1^{m-2}(1 - P_1),\ m = 2, 3, \ldots$$

It is clear that this argument can be continued indefinitely by induction, to establish the existence of a.s. limits P_i for N_{in}/n as $n \to \infty$, such that $P_i \geq 0$ and $\sum_{i=1}^{\infty} P_i \leq 1$, and conditionally given all the (P_i), the joint law of the (T_n) is as follows: Given $T_1, \ldots, T_n$ with $\max_{1\leq m\leq n} T_m = k$

$$T_{n+1} = \begin{cases} i \text{ for } 1 \leq i \leq k, \text{ with probability } P_i, \\ k+1 \text{ with probability } 1 - P_1 - \ldots - P_k. \end{cases}$$

But by definition of the T_m, this is precisely the sequential description of Π_n as in Proposition 14. □

When specialized to exchangeable partitions, as in the following corollary, the above representation is seen to be closely related but not identical to Kingman's.

Corollary 16 *Let (P_i) be a sequence of random variables such that $P_i \geq 0$ and $\sum_{i=1}^{n} P_i \leq 1$ a.s. for all n. The following statements are equivalent:*

(i) *There exists an exchangeable random partition Π of $\mathbb{N}$ such that if P_i' is the a.s. limiting proportion of the ith class of Π to appear, then (P_i') has the same joint distribution as (P_i).*

(ii) *For each $k = 2, 3, \ldots$ the function p of k-tuples of positive integers defined by*

$$p(n_1, \ldots, n_k) := E\left[\left(\prod_{i=1}^{k} P_i^{n_i - 1}\right) \prod_{i=1}^{k-1} (1 - \sum_{1}^{i} P_j)\right], \tag{38}$$

is a symmetric function of $(n_1, \ldots, n_k)$.

(iii) *for each $k = 2, 3, \ldots$, the measure G_k on $\mathbb{R}^k$ defined by*

$$G_k(dp_1, \ldots, dp_k) = P(P_1 \in dp_1, \ldots, P_k \in dp_k) \prod_{i=1}^{k-1} (1 - \sum_{1}^{i} p_j). \tag{39}$$

is symmetric with respect to permutation of the coordinates in $\mathbb{R}^k$.

In case these equivalent conditions hold, $p(n_1, \ldots, n_k)$ in (ii) *is the consistent EPF for the partition structure $(\mathbb{P}_n)$ associated with the random partition Π constructed from (P_i) as in Proposition* 14. *Assuming either $P_1 > 0$ a.s., or $\sum_i P_i = 1$, (conditions that are equivalent in case* (i)-(iii) *hold) a fourth statement equivalent to* (i)-(iii) *is*

(iv) *$\sum_i P_i = 1$ and (P_i) is invariant under size-biased random permutation.*

Proof. (i) $\Leftrightarrow$ (ii). According to Theorem 15, $p(n_1, \ldots, n_k)$ defined by (38) is the PEPF associated with the unique partially exchangeable random partition of $\mathbb{N}$ whose asymptotic class proportions are $(P_1, P_2, \ldots)$. Since a

partially exchangeable random partition is exchangeable iff the PEPF is symmetric, the equivalence of (i) and (ii) is clear.

(ii) $\Leftrightarrow$ (iii): This is immediate from the definition of G_k, and the fact that polynomials are dense in the space of continuous function on $[0,1]^k$.

(i) $\Leftrightarrow$ (iv). This follows easily from Kingman's representation and Corollary 10. See the proof of the next corollary, which is just a restatement of Kingman's representation. □

See Donnelly [9] for quite a different characterization of random discrete distributions invariant under size-biased permutation.

Corollary 17 Kingman [28, 30] . *Let* Π *be a partially exchangeable random partition of* $\mathbb{N}$, P_i *the a.s. limiting relative frequency the ith class of* Π *to appear. Given the* (P_i), *let* P_∞ *be a probability distribution on* $[0,1]$ *with atoms of weights* P_i *and continuous component of weight* $R := 1 - \sum_i P_i$, *for example*

$$P_\infty = \sum_i P_i \delta_{U_i} + R\mu$$

where δ_{U_i} *is a unit mass at* U_i, *and the* U_i *are i.i.d. random variables with uniform distribution* μ *on* $[0,1]$. *Given* P_∞, *let* $X_1, X_2, \ldots$ *be i.i.d with distribution* P_∞. *Then* Π *has the same distribution as* $\Pi(X_1, X_2, \ldots)$ *iff* Π *is exchangeable.*

Proof. The argument will be indicated just in the *proper* case $\sum P_i = 1$. (Only in this case does Kingman's representation follow easily from the partially exchangeable representation. See the remark below.) According to Theorem 15, given $(P_1, P_2, \ldots)$ the partitions Π_n obtained by restriction of Π to $\mathbb{N}_n$ admit the sequential description of Proposition 14. But by construction, the partitions Π'_n generated by $X_1, \ldots, X_n$ as above admit the same description with $(P'_1, P'_2, \ldots)$, a size-biased permutation of $(P_1, P_2, \ldots)$, instead of $(P_1, P_2, \ldots)$. Thus $\Pi \stackrel{d}{=} \Pi'$ iff $(P_1, P_2, \ldots)$ is invariant under size-biased permutation. In case Π is exchangeable this invariance follows easily from its discrete analog stated in Corollary 10 by passage to the limit as $n \to \infty$. The converse is obvious.

Remark. For the most general possible joint distribution for the limiting proportions (P_i) of the classes of an exchangeable random partition in order of their first elements, allowing $\sum := \sum_i P_i$ to be less than 1, the description

analogous to condition (iv) in Theorem 15 is tricky to state, though easily derived from Kingman's representation. The description is as follows, in terms of an arbitrary distribution F on $[0,1]$ for $\sum$, and a family of proper joint distributions (Q_s), indexed measurably by $0 < s \leq 1$, where each Q_s is a joint distribution for (P_i) with $P_i \geq 0$ and $\sum_i P_i = 1$ that is invariant under size-biased permutation. Let $\sum$ have distribution F. Given $\sum = s$ for $0 < s \leq 1$, let Q_s be the conditional distribution of $(P_i^*/\sum, i = 1,2,\ldots)$, where (P_i^*) is the subsequence of strictly positive terms extracted from (P_i), with the convention $P_i^* = 0$ if there are fewer than i strictly positive terms. Given $\sum$ and (P_i^*) the distribution of (P_i) is described by the following process of insertion of zeros into the sequence (P_i^*): Let the number of zeros in (P_i) preceding P_1^* have geometric $(\sum)$ distribution on $\{0,1,2,\ldots\}$. Given this number of zeros, the number of consecutive zeros following P_1^* has geometric $(\sum - P_1^*)$ distribution; given the numbers of zeros before and after P_1^*, and assuming the latter number is finite (i.e. $P_2^* > 0$), the number of consecutive zeros following P_2^* has geometric $(\sum - P_1^* - P_2^*)$ distribution, and so on.

To conclude this section, here is another immediate corollary of Theorem 15. This corollary contains Theorem 4 of Hoppe[24] as the case when the W_i are i.i.d. beta$(1,\theta)$, so the partition of n is governed by Ewens' sampling formula (see Proposition 24 below)). In that case, and whenever Π is exchangeable, $(N_{1n},\ldots,N_{nn})$ is a size-biased presentation of the partition of n.

Corollary 18 *Let $\Pi = \{\mathcal{A}_i\}$ be a PEP of $\mathbb{N}$, P_i the almost sure limit as $n \to \infty$ of N_{in}/n, where $N_{in} = \#(\mathcal{A}_i \cap \mathbb{N}_n)$. For each $i \geq 0$, given $(P_1, P_2, \ldots)$ and (for $i \geq 1$) given also $N_{1n},\ldots,N_{in}$ with $\sum_1^i N_{jn} < n$, the random variable $N_{i+1,n} - 1$ has binomial $(n - \sum_1^i N_{jn} - 1, W_{i+1})$ distribution, where $W_{i+1} = P_{i+1}/(1 - \sum_1^i P_j)$.*

5 Variations on a theorem of de Finetti

According to Theorem 15, formula (30) sets up a one to one correspondence between consistent non-negative functions $p(n_1,\ldots,n_k)$ with $p(1) = 1$ and joint distributions for a sequence of random variables (P_i) with $P_i \geq 0$, $\sum_i P_i \leq 1$. This parallels the following variation of de Finetti's theorem:

Proposition 19 *The formula*

$$p(n_i) = E\left(\prod_i P_i^{n_i}\right) \tag{40}$$

sets up a one to one correspondence between non-negative functions p of sequences of non-negative integers (n_i) with $\sum_i n_i < \infty$, that satisfy $p(0,0,\ldots) = 1$ and

$$p(n_i) = \sum_{j=1}^{\infty} p(n_i^{j+}), \quad \text{where } n_i^{j+} = n_i + 1(i=j), \tag{41}$$

and joint distributions for a sequence of random variables (P_i) with $P_i \geq 0$, $\sum_i P_i = 1$.

Proof. For $(x_1,\ldots,x_n) \in \mathbb{N}^n$ define (n_i) with $\sum_i n_i = n$ by the frequencies $n_i = \#\{j : x_j = i\}$. It is then easily verified that (41) holds iff the formula $P(x_1,\ldots,x_n) = p(n_i)$ defines a consistent family of exchangeable distributions on $\mathbb{N}^n$. The proposition is now seen to be a reformulation of de Finetti's theorem for $\mathbb{N}$ valued exchangeable sequences (Hewitt and Savage[21], Aldous[1]). □

Proposition 19 generalizes Theorem III of Blackwell and Kendall [3], which is the special case when for some $k \geq 2$ it is assumed that $p(n_i) = 0$ for all (n_i) with $n_i > 0$ for some $i > k$. Then $P_i = 0$ for $i > k$, and $\sum_1^k P_i = 1$. In particular, for $k = 2$, this result can be stated in simpler notation as follows:

Corollary 20 (Blackwell and Kendall) *The formula*

$$m(r,s) = \int_0^1 x^r (1-x)^s \mu(dx) \tag{42}$$

defines a one to one correspondence between non-negative bounded Borel measures μ on $[0,1]$, and non-negative solutions of the recurrence relation

$$m(r,s) = m(r+1,s) + m(r,s+1), \; r,s = 0,1,2,\ldots \tag{43}$$

This is a restatement of the classical result of Hausdorff[20] that a sequence $(m(r), r = 0,1,\ldots)$ is the moment sequence of some positive measure μ on $[0,1]$ iff

$$m(r,s) := (-1)^s(\Delta^s m)(r) \tag{44}$$

is non-negative for all $r = 0, 1, 2 \ldots$, $s = 0, 1, 2 \ldots$, where Δ^s is the s-fold iterate of the finite difference operator Δ. The above proof of Proposition 19 in the case for Corollary 20 just reverses the well known derivation of de Finetti's theorem from Hausdorff's result (Feller[17], Section VII.3). Blackwell and Kendall proved Corollary 20 by application of Martin boundary theory. See next section for references to further literature with this point of view.

Another variation of de Finetti's theorem is the analog of Proposition 11 in the setting of Proposition 19. The one-dimensional analog of (i) $\Leftrightarrow$ (iii) in Proposition 11 is stated as the next proposition. As explained in the next section, this proposition is a paraphrase of the expression of de Finetti's theorem by identifying the Martin boundary of an associated space-time random walk. (The corresponding analog of (i) $\Leftrightarrow$ (ii) in Proposition 11 well known characterization of exchangeable sequences in terms of sampling without replacement). The general formulation and proof of this result in the setting of Proposition 11 are straightforward and left to the reader.

Proposition 21 *Let $X_1, \ldots, \ldots X_n$ be random a sequence of zeros and ones, $S_n = X_1 + \cdots X_n$. Then the following are equivalent*
(i) $X_1, \ldots, \ldots X_n$ is exchangeable .
(i) $S_1, \ldots, \ldots S_n$ is an inhomogeneous Markov chain, with transition probabilities of the special form

$$P(S_{m+1} = t | S_m = s) = \frac{p(m+1, t)}{p(m, s)}, \tag{45}$$

for $0 \leq m < n$, $p(m, s) > 0$ and $t = s$ or $s + 1$, and zero otherwise, for some non-negative function $p(m, s)$ defined for all (m, s) with $0 \leq s \leq m$. In that case, the unique such function with $p(0, 0) = 1$ is related to the distribution of S_m for every $1 \leq m \leq n$ by

$$P(S_m = s) = \binom{m}{s} p(m, s), \quad 0 \leq s \leq m \tag{46}$$

In particular, S_n derived from an infinite random sequence of zeros and ones is Markovian with such transition probabilities for all m iff S_n/n converges a.s. to a random variable Y with

$$E(Y^s(1 - Y)^{m-s}) = p(m, s), \quad 0 \leq s \leq m, \tag{47}$$

and given $Y = p$ the X_i are independent Bernoulli(p) random variables.

The following corollary is a special case of the the above proposition, stated here for ease of reference in Section 7. This is just a statement of well known properties of Pólya's urn scheme (see e.g. Freedman [18] Theorems 2.1 and 2.2).

Corollary 22 *Let $X_1, X_2 \ldots, \ldots$ be random a sequence of zeros and ones, $S_0 = 0$, $S_n = X_1 + \cdots X_n$. Let $a > 0, b > 0$. Then the following are equivalent (i) (S_n) is an inhomogeneous Markov chain, with*

$$P(S_{n+1} = s + 1 | S_n = s) = \frac{a+s}{a+b+n}. \tag{48}$$

(ii) S_n/n converges a.s. to a random variable Y with beta(a,b) distribution on $[0,1]$, and given $Y = p$ the X_i are independent Bernoulli(p) random variables.

6 Martin Boundaries

As noted in the last section, it is well known that de Finetti's[5] representation theorem for exchangeable sequences of zeros and ones is equivalent to Hausdorff's[20] characterization of moment sequences of a probability distribution on $[0,1]$. It is also well known that Hausdorff's result follows from the identification with $[0,1]$ of the Martin boundary of a space-time random walk whose steps are defined by a suitable infinite exchangeable sequence of zeros and ones $(X_n, n = 1, 2, \ldots)$. See e.g. Watanabe[46], Spitzer[44], Williams[48], who make the random walk steps by fair coin tossing, and Blackwell and Kendall[3], who draw the steps from Pólya's urn. Freedman [19] considers both cases in consecutive examples. These authors all identify the boundary of the space-time walk with $[0,1]$ using the Doob-Hunt boundary theory, without reference to de Finetti's theorem. But as pointed out by Martin Löf [36], this identificaton of the boundary, and the implied integral representation of harmonic functions of the space-time walk, is essentially just a reframing of de Finetti's theorem. See Lauritzen and Küchler [31] for recent extensions and developments of this idea.

The following elementary proposition is a restatement of Proposition 21 in terms of the Doob-Hunt theory.

Proposition 23 *Let Q denote the transition matrix of the space-time walk $((S_n, n), n = 0, 1, \ldots)$, where $S_n = X_1 + \ldots + X_n$ is derived from an infinite*

exchangeable sequence of zeros and ones $(X_n, n = 1, 2, \ldots)$ *such that* $P(S_2 = 1) > 0$. *A random sequence of zeros and ones* (X'_n) *is exchangeable iff the corresponding space-time process* $((S'_n, n))$ *is the Doob h-transform of* (S_n) *for some non-negative Q-harmonic function h with* $h(0,0) = 1$, *which is then given by*

$$h(s,n) = \frac{P(S'_n = s)}{P(S_n = s)}. \tag{49}$$

In particular (X'_n) that is a sequence of independent Bernoulli (p) random variables is obtained from $h = h_p$ defined by

$$h_p(s,n) = \frac{\binom{n}{s} p^s (1-p)^{n-s}}{P(S_n = s)}$$

Due to formula (49), convex combinations of harmonic functions h correspond to mixtures of laws for exchangeable sequences (X'_n). Thus de Finetti's theorem, that

> *the law of every exchangeable sequence of zeros and ones is a unique integral mixture of laws of independent Bernoulli* (p) *sequences,*

amounts to:

> *every non-negative Q-harmonic h with* $h(0,0) = 1$ *is a unique integral mixture of the extreme Q-harmonic functions* h_p.

Easily then

> $p \leftrightarrow h_p$ *is a homeomorphism between* $[0,1]$ *and the extreme Q-harmonic functions that define the Martin boundary of the space-time walk.*

The above story can be retold to the last detail with the two valued exchangeable reference process (X_n) replaced by any exchangeable reference process with values in a finite or countable set I, such that the support of the joint distribution of (X_1, X_2) is $I \times I$, with the space-time walk replaced by the empirical frequency process (F_n),

$$F_n = (F_{ni}, i \in I), \text{ where } F_{ni} = \#\{j : 1 \leq j \leq n, X_j = i\}.$$

Then (F_n) is a Markov chain, and the Doob h-processes derived from (F_n) for harmonic functions h are the empirical frequency processes of exchangeable I-valued sequences (X'_n). Using analytic methods based on Choquet theory, Blackwell and Kendall[3] identified the Martin boundary of the frequencies process (F_n), derived from a generalization of Pólya's scheme to balls of k colors, as $[0,1]^k$. The above observation regarding the h-processes of (F_n) shows that their result amounts to de Finetti's theorem for k-valued exchangeable processes. Similarly, for a countable value set I, and (F_n) derived from an exchangeable (X_n) as above, Proposition 19 identifies the Martin boundary of (F_n) with the infinite simplex $\{(p_1, p_2, \ldots) : p_i \geq 0, \Sigma_i p_i = 1\}$.

Theorem 15 can be viewed similarly as identifying the Martin boundary of the chain of frequencies (N_{*m}) derived as in (26) from any non-degenerate partially exchangeable random partition Π of $\mathbb{N}$, where non-degenerate means that the limiting proportions (P_i) are such that $P_i > 0$ for all i. The boundary points are identified as in Remark 33 with $\{(p_1, p_2, \ldots) : p_i \geq 0, \Sigma_i p_i \leq 1\}$.

Kingman's representation identifies the Martin boundary of the Markov chain (π_n) derived from a non-degenerate exchangeable partition of $\mathbb{N}$, where π_n is the partition of n induced by the restriction of Π to $\mathbb{N}_n$. Yet another variation on the theme is provided by the representation of consistent ordered sampling distributions due to Donnelly and Joyce [11]. In all these examples, the boundary theory provides a common framework, but no recipe for identifying the extreme points. The simplest way to identify the extreme points in these examples is by application of some more elementary form of de Finetti's theorem. See also Lauritzen [33, 34], Diaconis and Freedman [7] for discussion of similar problems.

7 Residual Allocation Models

The product model $P_i = \bar{W}_1 \ldots \bar{W}_{i-1} W_i$ for a random discrete distribution (P_i), with independent W_i, is known as a *residual allocation model.* See Patil and Taillie [37]. Let $m_i(r,s) = E[W_i^r \bar{W}_i^s]$. Then from (31), the formula

$$p(n_1, \ldots, n_k) = \prod_{i=1}^{k} m_i(n_i - 1, n_{i+1} + \ldots + n_k) \tag{50}$$

defines the PEPF corresponding to a PEP of $\mathbb{N}$ such that the asymptotic frequency P_i of the ith class to appear is $P_i = \bar{W}_1 \ldots \bar{W}_{i-1} W_i$.

Given a sequence of distributions F_i for W_i on $[0,1]$, it is not obvious by inspection of formula (50) whether $p(n_1,\ldots,n_k)$ is symmetric in $(n_1,\ldots,n_k)$, that is to say whether the random partition of $\mathbb{N}$ is exchangeable. Apart from some rather trivial examples, it turns out that the only possible distributions for the W_i are as described in the next proposition. See Pitman [40] for details.

The entire circle of ideas presented in this paper is really a development of the following proposition, which contains many known results as special cases and corollaries. An attempt to provide due credits is made after the proof.

Proposition 24 *For each pair of real parameters α and θ, such that*

$$\text{either } 0 \le \alpha < 1 \text{ and } \theta > -\alpha, \tag{51}$$

$$\text{or } \alpha < 0 \text{ and } \theta = -m\alpha \text{ for some } m \in \mathbb{N} \tag{52}$$

an exchangeable random partition $\Pi = (\Pi_n)$ of $\mathbb{N}$ can be constructed as follows: $\Pi_1 = \{1\}$; for each $n \in \mathbb{N}$, conditionally given $\Pi_n = \{\{A_i\}_1^k\}$, for any particular partition of $\mathbb{N}_n$ into k subsets A_i of sizes n_i, $i = 1,\ldots,k$ the partition Π_{n+1} of $\mathbb{N}_{n+1}$ is an extension of Π_n such that $n+1$ attaches to class A_i with probability $(n_i - \alpha)/(n+\theta), 1 \le i \le k$, and $n+1$ forms a new class with probability $(k\alpha+\theta)/(n+\theta)$. The corresponding EPF is

$$p(n_1,\ldots,n_k) = \frac{[\theta+\alpha]_{k-1;\alpha}}{[\theta+1]_{n-1}} \prod_{i=1}^{k} [1-\alpha]_{n_i-1} \tag{53}$$

where for real numbers x and a and non-negative integer m

$$[x]_{m;a} = \begin{cases} 1 & \text{for } m = 0 \\ x(x+a)\ldots(x+(m-1)a) & \text{for } m = 1,2,\ldots \end{cases}$$

and $[x]_m = [x]_{m;1}$. Let $(N_{in}, 1 \le i \le K_n)$ be the sizes of the classes of Π_n in order of their first elements. The almost sure limits $P_i = \lim_{n\to\infty} N_{in}/n$ are such that $P_i = \bar{W}_1 \ldots \bar{W}_{i-1} W_i$, where the W_i are independent random variables with beta$(1-\alpha, \theta+i\alpha)$ distributions (with the convention in case (52) that $W_m = 1$ and W_i is undefined for $i > m$). In all cases $\sum_i P_i = 1$ a.s., and this random discrete distribution (P_i) is invariant under size-biased permutation.

Proof. It is easily checked that the transition probabilities are of the form required by condition (iii) of Proposition 29 for the given $p(n_i)$, which is obviously symmetric. The form of the joint distribution of the P_i can be checked either from (50) by computation of moments derived from the beta distributions, or by repeated application of Corollary 22, following the argument of Hoppe [24] in the case $\alpha \leq 0$. That $\sum_i P_i = 1$ a.s., and the invariance of (P_i) under size-biased permutation, follow from condition (iv) of Corollary 16. □

For $\alpha = 0$, formula (53) for the EPF appears in Kingman [30], Theorem 4. As Kingman observes, the corresponding formula (4) for the joint distribution of the counts of component sizes is Ewens' [15] sampling formula. See Hoppe [24] and Ewens [16] for a variety of developments and applications of the case $\alpha = 0$ to population genetics. Antoniak [2] found Ewens' formula using the sequential description of the random partition as above, which he derived from the Blackwell-McQueen [4] urn scheme description of sampling from a Dirichlet prior distribution. It was analysis of the consistency feature of Ewens' formula as n varies which led Kingman [26, 27, 28, 29] to the concept of a partition structure. The fact that the corresponding residual allocation model with beta$(1, \theta)$ factors is invariant under size-biased permutation was known already to McCloskey [35], who showed this is the only residual allocation model with i.i.d. factors invariant size-biased permutation. See Pitman [40] for a similar characterization of the two parameter scheme described above among residual allocoation models with independent non-identically distributed factors.

The case $\alpha < 0$ corresponds to the partition generated by random sampling from a random discrete distribution on m points with symmetric Dirichlet prior. The sampling formula (4) corresponding to the EPF (53) for $\alpha < 0$ appears in Watterson [47], who used it to derive Ewens' formula by passage to the limit as $m \to \infty$ for fixed θ. The residual allocation model in this case was noted by Patil and Taillie [37], and the sequential description of the random partition appears in Hoppe [24]. See also Rothman and Templeton [42], and Keener, Rothman and Starr [25] for further study of this case.

In the case $0 < \alpha < 1$, the residual allocation model was considered by Engen [14], who showed that a single size-biased pick from (P_i) has the same distribution as P_1. The full invariance of (P_i) under size-biased permutation in this case follows from the work of Perman, Pitman and Yor [38], who showed how this random discrete distribution can be obtained by size-biased sampling of the normalized jumps of a stable subordinator with index α. See

Pitman [39] for further details of this connection. The above sequential description of the random partition and the formula (53) in this case seem to be new. See Pitman [41] for further study of this two-parameter generalization of Ewens' formula.

References

[1] D.J. Aldous. Exchangeability and related topics. In P.L. Hennequin, editor, *Ecole d'Été de Probabilités de Saint-Flour XII, Springer Lecture Notes in Mathematics, Vol. 1117.* Springer-Verlag, 1985.

[2] C. Antoniak. Mixtures of Dirichlet processes with application to Bayesian nonparametric problems. *Ann. Statist.*, 2:1152–1174, 1974.

[3] D. Blackwell and D. Kendall. The Martin boundary for Pólya's urn scheme, and an application to stochastic population growth. *J. Appl. Prob.*, 1:284–296, 1964.

[4] D. Blackwell and J.B. MacQueen. Ferguson distributions via Pólya urn schemes. *Ann. Statist.*, 1:353–355, 1973.

[5] B. de Finetti. La prévision: Ses lois logiques ses sources subjectives. *Ann. de l'institut Henri Poincaré*, 7:1–68, 1937. Translated in *Studies in Subjective Probability*, H. Kyburg and H. Smokler, editors, Wiley, New York, 1964.

[6] B. de Finetti. Sur la condition d'equivalence partielle. *Actualites Scientifiques et Industrielles*, 739, 1938. Herman and Cie: Paris. Translated in *Studies in Inductive Logic and Probability*, II, R. Jeffrey, editor, University of California Press: Berkeley, 1980.

[7] P. Diaconis and D. Freedman. Partial exchangeability and sufficiency. In J. K. Ghosh and J. Roy, editors, *Statistics Applications and New Directions; Proceedings of the Indian Statistical Institute Golden Jubilee International Conference; Sankhya A.* Indian Statistical Institute, 205-236, 1984.

[8] P. Donnelly. Partition structures, Pólya urns, the Ewens sampling formula, and the ages of alleles. *Theoretical Population Biology*, 30:271 – 288, 1986.

[9] P. Donnelly. The heaps process and size biased permutations. *Preprint*, 1991.

[10] P. Donnelly and P. Joyce. Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex. *Stochastic Processes and their Applications*, 31:89 – 103, 1989.

[11] P. Donnelly and P. Joyce. Consistent ordered sampling distributions: characterization and convergence. *Adv. Appl. Prob.*, 23:229–258, 1991.

[12] P. Donnelly and S. Tavaré. The ages of alleles and a coalescent. *Adv. Appl. Probab.*, 18:1–19 & 1023, 1986.

[13] E.B. Dynkin. Sufficient statistics and extreme points. *Ann. Probability*, 6:705–730, 1978.

[14] S. Engen. *Stochastic Abundance Models with Emphasis on Biological Communities and Species Diversity.* Chapman and Hall Ltd., 1978.

[15] W.J. Ewens. The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.*, 3:87 – 112, 1972.

[16] W.J. Ewens. Population genetics theory - the past and the future. In S. Lessard, editor, *Mathematical and Statistical Problems in Evolution.* University of Montreal Press, Montreal, 1988.

[17] W. Feller. *An Introduction to Probability Theory and its Applications, Vol 2.* Wiley, 1966.

[18] D. Freedman. Bernard Friedman's Urn. *Ann. Math. Statist*, 36:956–970, 1965.

[19] D. Freedman. *Markov Chains.* Holden-Day, San Francisco, 1971.

[20] F. Hausdorff. Summationsmethoden und Momentfolgen I. *Mathematische Zeitschrift*, 9:74–109, 1921.

[21] E. Hewitt and L.J. Savage. Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.*, 80:470–501, 1955.

[22] F. M. Hoppe. Pólya-like urns and the Ewens sampling formula. *Journal of Mathematical Biology*, 20:91 – 94, 1984.

[23] F. M. Hoppe. Size-biased filtering of Poisson-Dirichlet samples with an application to partition structures in genetics. *Journal of Applied Probability*, 23:1008 – 1012, 1986.

[24] F. M. Hoppe. The sampling theory of neutral alleles and an urn model in population genetics. *Journal of Mathematical Biology*, 25:123 – 159, 1987.

[25] R. Keener, E. Rothman, and N. Starr. Distributions on partitions. *Annals of Statistics*, 15:1466 – 1481, 1987.

[26] J. F. Kingman. The population structure associated with the Ewens sampling formula. *Theor. Popul. Biol.*, 11:274–283, 1977.

[27] J. F. Kingman. Random partitions in population genetics. *Proc. R. Soc. Lond. A.*, 361:1–20, 1978.

[28] J. F. Kingman. The representation of partition structures. *J. London Math. Soc.*, 18:374–380, 1978.

[29] J. F. Kingman. *The Mathematics of Genetic Diversity.* SIAM, 1980.

[30] J. F. Kingman. The coalescent. *Stochastic Processes and their Applications*, 13:235–248, 1982.

[31] U. Küchler and S. L. Lauritzen. Exponential families, extreme point models and minimal space-time invariant functions for stochastic processes with stationary and independent increments. *Scandinavian Journal of Statistics*, 16:237 – 261, 1989.

[32] S.L. Lauritzen. Sufficiency, prediction and extreme models. *Scand. J. Statist.*, 1:128–134, 1974.

[33] S.L. Lauritzen. Extreme point models in statistics. *Scandinavian Journal of Statistics*, 11:65 – 91, 1984.

[34] S.L. Lauritzen. *Extremal Families and Systems of Sufficient Statistics.* Springer-Verlag, Lecture Notes in Statistics, 49, 1988.

[35] J. W. McCloskey. A model for the distribution of individuals by species in an environment. Ph. D. thesis, Michigan State University, 1965.

[36] P. Martin-Löf. Repetetive structures and the relation between canonical and micro-canonical distributions in statistics and statistical mechanics. In *Proceedings of Conference of Foundational Questions in Statistical Inference.* Aarhus, 1974. O. Barndorff-Nielsen, P. Blaesild, G. Schon, editors.

[37] G. P. Patil and C. Taillie. Diversity as a concept and its implications for random communities. *Bull. Int. Stat. Inst.*, XLVII:497 – 515, 1977.

[38] M. Perman, J. Pitman, and M. Yor. Size-biased sampling of Poisson point processes and excursions. To appear in *Probability and Related Fields*, 1992.

[39] J. Pitman. Partition structures derived from Brownian motion and stable subordinators. Technical Report 346, Dept. Statistics, U.C. Berkeley, 1992.

[40] J. Pitman. Random discrete distributions invariant under size-biased permutation. Technical Report 344, Dept. Statistics, U.C. Berkeley, 1992.

[41] J. Pitman. The two-parameter generalization of Ewens' random partition structure. Technical Report 345, Dept. Statistics, U.C. Berkeley, 1992.

[42] E.D. Rothman and A.R. Templeton. On a class of models for selectively neutral alleles. *Theoret. Population Biology*, 18:135–150, 1980.

[43] L.A. Shepp and S.P. Lloyd. Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc.*, 121:340–357, 1966.

[44] F. Spitzer. *Principles of Random Walk.* Van Nostrand, Princeton, N.J., 1964.

[45] A.M. Vershik and A.A. Shmidt. Limit measures arising in the theory of groups, I. *Theor. Prob. Appl.*, 22:79–85, 1977.

[46] T. Watanabe. A probabilistic method in Hausdorff moment problem and Laplace-Stieltjes transform. *J. Math. Soc. Japan*, 12:192–206, 1960.

[47] G. A. Watterson. The stationary distribution of the infinitely-many neutral alleles diffusion model. *J. Appl. Probab.*, 13:639–651, 1976.

[48] D. Williams. *Diffusions, Markov Processes, and Martingales, Vol. 1: Foundations.* Wiley, Chichester, New York, 1979.