Probab. Theory Relat. Fields 102, 145 – 158 (1995)

Probability Theory and Related Fields

Exchangeable and partially exchangeable random partitions

Jim Pitman*

Department of Statistics, U.C. Berkeley, CA 94720, USA

Received: 5 May 1992/Accepted: 25 November 1994

Summary. Call a random partition of the positive integers *partially exchangeable* if for each finite sequence of positive integers $n_1,\ldots,n_k$, the probability that the partition breaks the first $n_1+\ldots+n_k$ 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 which provides a useful variation of Kingman's representation in the exchangeable case. Results are illustrated by the two-parameter generalization of Ewens' partition structure.

Mathematics Subject Classification : 60G09, 60C05, 60J50

1 Introduction

For a positive integer n, a *partition of n* is an unordered collection of positive integers with sum n. There are two common ways to code a partition of n:

(i) by the decreasing sequence of terms, say

$$n_{(1)} \geqq n_{(2)} \geqq \cdots \geqq n_{(k)} \quad \text{with } \sum n_{(i)} = n \; ;$$

(ii) by the numbers of terms of various sizes, say

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

where $\Sigma m_j = k$, and $\Sigma j m_j = n$. A *random partition* of n is a random variable π_n with values in the set of all partitions of n. Motivated by applications in genetics, Kingman [20, 21] developed the concept of a *partition structure*, that

* Research supported by N.S.F. Grants MCS91-07531 and DMS-9404345

is a sequence $P_1, P_2, \ldots$ of distributions for $\pi_1, \pi_2, \ldots$ which is *consistent* in the following sense: if n objects are partitioned into classes 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 class sizes distributed according to P_{n-1}.

Let $N_n := \{1, \ldots, n\}$, $:= \{1, 2, \ldots\}$. A *partition of* N_n is an unordered collection of disjoint non-empty subsets of N_n, say $\{A_i\}$, with $\cup_i A_i = N_n$. The A_i will be called *classes* of the partition. Given a partition $\{A_i\}$ of N_n, for $m < n$ the *restriction of* $\{A_i\}$ *to* N_m is the partition of N_m whose classes are the non-empty members of $\{A_i \cap N_m\}$. A *random partition of* N_n is a random variable Π_n with values in the finite set of all partitions of N_n. A *random partition of* N is a sequence $\Pi = (\Pi_n)$ of random partitions of N_n defined on a common probability space, such that for $m < n$ the restriction of Π_n to N_m is Π_m. Permutations of N_n act in a natural way on partitions of N_n, and on distributions of a random partition of N_n. Following Kingman [22] and Aldous [1], Π_n is called *exchangeable* if the distribution of Π_n is invariant under the action of all such permutations. And $\Pi = (\Pi_n)$ is exchangeable if Π_n is exchangeable for every n. As shown by Kingman, (P_n) is a partition structure iff there exists an exchangeable random partition $\Pi = (\Pi_n)$ of N such that P_n is the distribution of the partition of n induced by the class sizes of Π_n.

For a sequence of random variables $(X_1, X_2, \ldots)$, let $\Pi(X_1, X_2, \ldots)$ be the random partition of N defined by equivalence classes for the random equivalence relation $i \sim j \Leftrightarrow X_i = X_j$. According to *Kingman's representation* every exchangeable random partition Π of N has the same distribution as $\Pi(X_1, X_2, \ldots)$ where $X_1, X_2, \ldots$ are conditionally i.i.d. according to P_∞ given some random probability distribution P_∞. See Aldous [1] for a quick proof. The distribution P_n of the class sizes of Π_n is determined by the joint distribution of the sizes of the ranked atoms of P_∞, denoted

$$P_{(1)} \geqq P_{(2)} \geqq \ldots \geqq 0\,, \tag{1}$$

where $P_{(i)} = 0$ if P_∞ has fewer than i atoms. Moreover such $P_{(i)}$ can be recovered from Π as

$$P_{(i)} = \lim_{n \to \infty} \frac{N_{(i)n}}{n} \quad \text{a.s.}, \tag{2}$$

where $N_{(i)n}$ is the size of the ith largest class in Π_n. See [21, 22, 1] for further details.

Two difficulties arise in working with this representation of partition structures. First, the joint distribution of the limiting ranked proportions $P_{(i)}$ turns out to be rather complicated, even for the simplest partition structures, such as those corresponding to Ewens' sampling formula, when the joint distribution of the $P_{(i)}$ is the Poisson–Dirichlet distribution [30, 19, 17]. Second, the expression for the distribution P_n of the partition of n in terms of the joint distribution of the $P_{(i)}$, given by formulae (2.10) and (5.1) of Kingman [20], involves infinite sums of expectations of products of the $P_{(i)}$, which are not easy to evaluate. In the case corresponding to Ewens' sampling formula, it is well known [6, 12, 14–16] that there is a much simpler description of the joint distribution of the sequence $(P_1, P_2, \ldots)$ obtained by presenting the ranked sequence $(P_{(1)}, P_{(2)}, \ldots)$ in the random order in which the corresponding classes appear in the random partition Π.

To be precise, write

$$\Pi = \{\mathscr{A}_1, \mathscr{A}_2, \ldots\}\,, \tag{3}$$

where $\mathscr{A}_i$ is the random subset of N defined as the *ith class of Π to appear.* That is to say $\mathscr{A}_1$ is the class containing 1, $\mathscr{A}_2$ is the class containing the first element of N $- \mathscr{A}_1$, and so on. For convenience, let $\mathscr{A}_i = \emptyset$ if Π has fewer than i classes. Then P_i is defined to be the long run relative frequency of $\mathscr{A}_i$:

$$P_i := \lim_{n\to\infty} \frac{\#(\mathscr{A}_i \cap \mathrm{N}_n)}{n} \quad \text{a.s. } i = 1, 2, \ldots \tag{4}$$

The $P_{(i)}$ in (2) are obtained by ranking the P_i, and the existence of either collection of limits (2) or (4) follows easily from the other. See e.g. [1, Lemma 11.8], which implies also that if $\sum_i P_{(i)} = 1$ a.s. then $(P_1, P_2, \ldots)$ is a *size-biased random permutation* of the ranked sequence $(P_{(1)}, P_{(2)}, \ldots)$ as studied in [8, 27].

The main purpose of this paper is to answer the following questions which arise naturally from the above development:

Question 1 *What is the most general possible distribution of the sequence (P_i) of limiting relative frequencies of classes in order of appearance for an exchangeable random partition Π of* N ?

Question 2 *How is the distribution of the sequence (P_i) related to the corresponding partition structure* P_n ?

Question 3 *What is the conditional distribution of Π given* (P_i) ?

These questions about exchangeable random partitions are answered in Sect. 2 by a variation of Kingman's representation which holds for larger class of random partitions of N, called *partially exchangeable.* The terminology is consistent with the general concept of partial exchangeability due to de Finetti [4]. Both Kingman's representation, and the present representation of partially exchangeable random partitions, fit the general framework of Diaconis–Freedman [5] for extreme point descriptions of models defined by a sequence of sufficient statistics. From another point of view, these results identify the Martin boundaries of associated Markov chains. But while the general extreme point or boundary theory provides a common framework, it offers no recipe for identifying the extreme points. Like Aldous' proof of Kingman's representation, the proof of the representation of partially exchangeable random partitions, provided in Sect. 4, is based on a direct application of de Finetti's theorem rather than any general extreme point theory. Section 5 considers partitions of N derived from residual allocation models. In particular, a two-parameter family of such models with beta distributed factors, presented at the end of Sect. 2, corresponds to a two-parameter generalization of Ewens' partition structure.

2 Results

Definition 4 *Let* $\mathrm{N}^* = \bigcup_{k=1}^{\infty} \mathrm{N}^k$, *the set of finite sequences of positive integers. Denote a generic element of* N^* *by* $\mathbf{n} = (n_1,\ldots,n_k)$, *and write* $\Sigma(\mathbf{n})$ *for* $\sum_{i=1}^{k} n_i$. *Call a random partition* Π_n *of* N_n **partially exchangeable** (**PE**) *if for every partition* $\{A_1,\ldots,A_k\}$ *of* N_n, *where the* $A_1,\ldots A_k$ *are in* **order of appearance**, *i.e.* $1 \in A_1$, *and for each* $2 \leqq i \leqq k$ *the first element of* $\mathrm{N}_n - (A_1 \cup \ldots \cup A_{i-1})$ *belongs to* A_i,

$$P(\Pi_n = \{A_1,\ldots,A_k\}) = p(\#(A_1),\ldots,\#(A_k)) \tag{5}$$

for some function $p(\mathbf{n}) = p(n_1,\ldots,n_k)$ *defined for* $\mathbf{n} \in N^*$ *with* $\Sigma(\mathbf{n}) = n$. *Then call* $p(\mathbf{n})$ *a* **partially exchangeable probability function** (*PEPF*).

Say a random partition Π of N is PE if $\Pi = (\Pi_n)$ with Π_n a PE partition of N_n for every n. Then a corresponding PEPF $p(\mathbf{n})$ is defined for all $\mathbf{n} \in \mathrm{N}^*$. The following elementary proposition follows from the above definition and the discussion on page 85 of Aldous [1]:

Proposition 5 *A random partition of* N_n (*or of* N) *is exchangeable iff it is partially exchangeable with a PEPF* $p(\mathbf{n})$ *which is a symmetric function of its arguments, i.e.*

$$p(n_1,\ldots,n_k) = p(n_{\sigma(1)},\ldots,n_{\sigma(k)})$$

for every permutation σ *of* N_k, $k = 2,3,\ldots$

When Π is exchangeable, call the symmetric PEPF derived from Π an *exchangeable probability function* (EPF). The main result of this paper is the representation for PE partitions of N stated in the following theorem, which is proved in Sect. 4.

Theorem 6 *Let* $\Pi = \{\mathscr{A}_1,\mathscr{A}_2,\ldots\}$ *be a random partition of* N, *with* $\mathscr{A}_i$ *the ith class of* Π *to appear. Let* Π_n *be the restriction of* Π *to* N_n. *The following conditions are equivalent*:

(i) Π *is partially exchangeable*;
(ii) *there is a sequence of random variables* $(P_i,\ i = 1,2,\ldots)$ *with* $P_i \geqq 0$ *and* $\sum_i P_i \leqq 1$ *such that the conditional distribution of* Π *given the whole sequence* (P_i) *is as follows*: *for each* $n \in \mathrm{N}$, *conditionally given* (P_i) *and* $\Pi_n = \{A_1,\ldots,A_k\}$, *where the* A_i *are in order of appearance*, Π_{n+1} *is an extension of* Π_n *in which* $n+1$

$$\begin{aligned} &\textit{attaches to class } A_i \textit{ with probability } P_i, 1 \leqq i \leqq k\ ,\\ &\textit{forms a new class with probability } 1 - \textstyle\sum_1^k P_j\ . \end{aligned} \tag{6}$$

If Π *is partially exchangeable then the* P_i *in* (ii) *are a.s. unique and equal to almost sure limiting relative frequencies of the classes* $\mathscr{A}_i$ *as in* (4).

This theorem answers Question 3: all exchangeable random partitions Π of N share a common conditional distribution given (P_i) defined in (ii); moreover a random partition of N is partially exchangeable iff it admits such relative frequencies (P_i) and shares this conditional distribution given the (P_i). The following corollary is an immediate consequence of Theorem 6:

Corollary 7 *The formula*

$$p(n_1,\ldots,n_k) = E\left[\left(\prod_{i=1}^{k} P_i^{n_i-1}\right)\prod_{i=1}^{k-1}\left(1-\sum_{j=1}^{i} P_j\right)\right] \tag{7}$$

sets up a one to one correspondence between PEPF's p: $\mathbb{N}^* \to [0, 1]$ and joint distributions for a sequence of random variables $(P_1, P_2, \ldots)$ with $P_i \geqq 0$ and $\Sigma_i P_i \leqq 1$.

To spell out this correspondence: given the PEPF $p(\mathbf{n})$ of a PE partition Π of $\mathbb{N}$, the P_i are recovered as the limiting relative frequencies of the classes of Π in order of appearance. And given a distribution for (P_i), a PE partition Π of $\mathbf{N}$ with the corresponding PEPF $p(\mathbf{n})$ is created by the sequential prescription (ii) of Theorem 6. See also Construction 16 in Sect. 5 for an alternative construction of Π from such (P_i). The next corollary answers Question 1 by specializing the previous corollary to the exchangeable case.

Corollary 8 *Let (P_i) be a sequence of random variables such that $P_i \geqq 0$ and $\sum_i P_i \leqq 1$ a.s.. The following statements are equivalent :*

(i) *There exists an exchangeable random partition Π of $\mathbb{N}$ whose sequence of limiting relative frequencies of classes, in order of appearance, has the same distribution as (P_i).*
(ii) *For each $k = 2, 3, \ldots$ the function $p : \mathbb{N}^k \to [0, 1]$ defined by (7) 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*

$$G_k(dp_1,\ldots,dp_k) = P(P_1 \in dp_1,\ldots,P_k \in dp_k)\prod_{i=1}^{k-1}\left(1-\sum_{j=1}^{i} p_j\right) \tag{8}$$

is symmetric with respect to permutation of the coordinates in $\mathbb{R}^k$.
Then $p(n_1, \ldots, n_k)$ defined by (7) is the EPF of Π.

Pitman [27] obtained this result assuming either $P_1 > 0$ a.s., or $\sum_i P_i = 1$ (conditions that are equivalent in cases (i)–(iii) hold) and showed that then the following condition is also equivalent to (i)–(iii):

$$(P_i) \textit{ is invariant under size-biased random permutation.} \tag{9}$$

See also [7] for yet another characterization of such random discrete distributions.

Corollary 8 also answers Question 2. For if Π is exchangeable it follows from (7) by a simple counting argument that the distribution P_n of the corresponding partition of n is as follows. Let $\mathbf{M}_n = (M_1, \ldots, M_n)$ where M_j is the number of classes of Π_n of size j. Then for any vector of non-negative integer counts $\mathbf{m} := (m_1, \ldots, m_n)$

$$\mathrm{P}_n(\mathbf{m}) := P(\mathbf{M}_n = \mathbf{m}) = N(\mathbf{m})\tilde{p}(\mathbf{m}), \tag{10}$$

where

$$N(\mathbf{m}) := \frac{n!}{\prod_{j=1}^{n} (j!)^{m_j} m_j!} \tag{11}$$

is the number of partitions of N_n into m_j classes of size j, $1, \leqq j \leqq n$, and $\tilde{p}(\mathbf{m})$ is the common value of the symmetric function $p(\mathbf{n})$ for all $\mathbf{n}$ with

$$\Sigma(\mathbf{n}) = n \quad \text{and } \#\{i: n_i = j\} = m_j, \quad 1 \leqq j \leqq n . \tag{12}$$

The formulation of results in this paper was guided by the example presented in the following proposition, which contains many known results as special cases and corollaries, and is proved in Sect. 5.

Proposition 9 *For each pair of real parameters α and θ, such that*

$$\text{either } 0 \leqq \alpha < 1 \textit{ and } \theta > -\alpha , \tag{13}$$

$$\text{or } \alpha < 0 \textit{ and } \theta = -m\alpha \textit{ for some } m \in \mathrm{N} \tag{14}$$

an exchangeable random partition $\Pi = (\Pi_n)$ of N can be constructed as follows: for each $n \in \mathrm{N}$, conditionally given $\Pi_n = \{A_1, \ldots, A_k\}$, for any particular partition of N_n into k subsets A_i of sizes n_i, $i = 1, \ldots, k$, the partition Π_{n+1} is an extension of Π_n such that $n + 1$

$$\begin{aligned} &\textit{attaches to class } A_i \textit{ with probability } \frac{n_i - \alpha}{n + \theta}, \quad 1 \leqq i \leqq k , \\ &\textit{forms a new class with probability } \frac{k\alpha + \theta}{n + \theta} . \end{aligned} \tag{15}$$

The corresponding EPF is

$$p(n_1, \ldots, n_k) = \frac{[\theta + \alpha]_{k-1;\alpha}}{[\theta + 1]_{n-1}} \prod_{i=1}^{k} [1 - \alpha]_{n_i - 1} , \tag{16}$$

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

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

and $[x]_m = [x]_{m;1}$. The probability that Π_n has m_j classes of size j, for $(m_1, \ldots, m_n)$ with $\Sigma m_j = k$, and $\Sigma j m_j = n$, is

$$n! \frac{[\theta + \alpha]_{k-1;\alpha}}{[\theta + 1]_{n-1}} \prod_{j=1}^{n} \left(\frac{[1 - \alpha]_{j-1}}{j!} \right)^{m_j} \frac{1}{m_{j1}} \tag{17}$$

The a.s. limiting relative frequencies P_i of the classes of Π in order of appearance are such that

$$P_i = (1 - W_1) \ldots (1 - W_{i-1}) W_i , \tag{18}$$

where the W_i are independent random variables with beta $(1 - \alpha, \theta + i\alpha)$ distributions, with the convention in case (14) that $W_m = 1$ and W_i is undefined for $i > m$.

The above proposition is known for $\alpha \leqq 0$. For $\alpha = 0$, Antoniak [2] derived the sequential construction (15) from the Blackwell–McQueen [3] urn scheme description of sampling from a Dirichlet prior distribution, and used it to deduce (16) and (17). Formula (17) in this case is Ewens' [11] sampling formula. See also [22], Theorem 4, for a similar derivation. A variation of the sequential construction for $\alpha = 0$, devised by Dubins and Pitman to explain the cycle structure of random permutations, is the "Chinese restaurant process" described on page 92 of [1]. See [16, 6, 12] for developments and applications of these ideas to population genetics, and further references. The model (18) for generating a random discrete distribution (P_i) from independent W_i is known as a *residual allocation model.* The consequence of the above proposition for $\alpha = 0$ and (9), that the RAM with identically distributed beta$(1, \theta)$ factors is invariant under size-biased permutation, was known already to McCloskey [23], who showed this is the only RAM with i.i.d. factors invariant under size-biased permutation. See also [27] for a similar characterization of the two parameter scheme by the RAM for the (P_i).

The case $\alpha = -\kappa < 0$ and $\theta = m\kappa$ corresponds to the partition generated by random sampling from a symmetric Dirichlet prior with weight κ on each of m points. This model for species sampling was proposed by Fisher [13], who considered also the passage to the limit as $m \to \infty$ and $\kappa \to 0$ for fixed $\theta = m\kappa$, which leads to Ewens' partition structure. Watterson [30] found the sampling formula (17) for Fisher's model with finite m and deduced Ewens' formula by passage to the limit. The sequential construction in this case follows immediately from the well known urn scheme description of sampling from a Dirichlet prior, which dates back to Johnson [18]. The corresponding RAM was considered by Patil and Taillie [24], Engen [10] and Hoppe [16].

For $0 < \alpha < 1$, the RAM was considered by Engen [10], 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–Yor [25], 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. The sequential construction of the random partition, and the formulae (16) and (17) in this case, seem to be new. See [26, 28, 29] for further study. Motivated by philosophical aspects of species sampling [31], Zabell [32] gives a characterization of the two-parameter scheme based on the form of the sequential construction.

3 Partially exchangeable partitions

Recall that $\mathrm{N}^* = \bigcup_{k=1}^{\infty} \mathrm{N}^k$, and $\Sigma(\mathbf{n}) = \sum_{i=1}^{k} n_i$ for $\mathbf{n} = (n_1, \ldots, n_k) \in \mathrm{N}^*$. Let $\mathrm{N}_n^* := \{\mathbf{n} \in \mathrm{N}^* : \Sigma(\mathbf{n}) = n\}$. It is immediate from Definition 4 that (5) sets up a one to one correspondence between distributions of a PE partition Π_n of N_n, and non-negative functions $p(\mathbf{n}) : \mathrm{N}_n^* \to [0, 1]$ such that

$$\sum_{\mathbf{n} \in \mathrm{N}_n^*} \#(\mathbf{n}) p(\mathbf{n}) = 1 \,, \tag{19}$$

where

$$\#(\mathbf{n}) = \frac{n!}{n_k(n_k + n_{k-1}) \cdots (n_k + \cdots + n_1) \prod_{i=1}^{k} (n_i - 1)!} \tag{20}$$

is the number of partitions of N_n whose class sizes in order of appearance are given by $\mathbf{n} = (n_1, \ldots, n_k)$. Let $\mathbf{N}_n$ be the random element of N_n^* representing the class sizes of Π_n in order of appearance. Then $\mathbf{N}_n$ is a *sufficient statistic* for distributions of PE partitions Π_n. That is to say, Π_n is PE iff given $\mathbf{N}_n = \mathbf{n}$ for every $\mathbf{n} \in \mathrm{N}_n^*$ the partition Π_n is uniformly distributed over the $\#(\mathbf{n})$ distinct partitions of N_n with class sizes in order of appearance given by $\mathbf{n}$. (The corresponding description of exchangeable partitions of N_n, with $\mathbf{N}_n$ replaced by the decreasing rearrangement of $\mathbf{N}_n$ which encodes the induced partition of n, appears on page 85 of [1]). The distribution of $\mathbf{N}_n$ for a PE partition Π_n is related to the PEPF p of Π_n by

$$P(\mathbf{N}_n = \mathbf{n}) = \#(\mathbf{n}) p(\mathbf{n}) \, . \tag{21}$$

Assuming Π_n is PE, it can be seen that Π_n is exchangeable if and only if $\mathbf{N}_n$ is a size-biased random ordering of the partition of n, as defined in [9, Sect. 7]. For the exchangeable partition of N_n derived from Ewens' partition structure, with p defined by (16) for $\alpha = 0$, formula (21) reduces to formula (7.2) of [9].

Proposition 10 *For* $1 \leqq m \leqq n$ *let* Π_m *be the restriction to* N_m *of a PE partition of* N_n *with PEPF* $p(\mathbf{n})$ *defined for* $\mathbf{n} \in \mathrm{N}_n^*$. *Then*

(i) Π_m *is PE, with PEPF* $p(\mathbf{n})$ *defined for* $\mathbf{n} \in \mathrm{N}_m^*$ *by repeated application for* $m = n-1, n-2, \ldots, 1$ *of the consistency relation*:

$$p(\mathbf{n}) = \sum_{j=1}^{k+1} p(\mathbf{n}^{j+}) \tag{22}$$

where $\mathbf{n}^{j+} \in \mathrm{N}_{m+1}^*$ *is derived from* $\mathbf{n} = (n_1, \ldots, n_k) \in \mathrm{N}_m^*$ *by incrementing* n_j *by* 1 *if* $1 \leqq j \leqq k$, *and by appending a* 1 *to* $\mathbf{n}$ *at place* $k+1$ *if* $j = k+1$.

(ii) $(\mathbf{N}_1, \cdots, \mathbf{N}_n)$ *is a Markov chain with transition probabilities*

$$P(\mathbf{N}_{m+1} = \mathbf{n}^{j+} \mid \mathbf{N}_m = \mathbf{n}) = \frac{p(\mathbf{n}^{j+})}{p(\mathbf{n})}, \quad j = 1, \ldots, k+1 \tag{23}$$

for p: $\bigcup_1^n \mathrm{N}_m^* \to [0, 1]$ *defined as in* (i).

Proof. (i) Argue inductively as follows for $m = n-1, n-2, \ldots, 1$. Suppose that Π_{m+1} is PE, with PEPF p: $\mathrm{N}_{m+1}^* \to [0, 1]$. Consider the probability of the event that Π_m is a particular partition of N_m with k classes, with sizes in order of appearance given by $\mathbf{n}$. By decomposing this event according to the $k+1$ different possibilities for Π_{m+1}, this probability equals the right hand expression in (22), which depends on the choice of partition of N_m only through $\mathbf{n}$. That is to say, Π_m is PE with PEPF $p(\mathbf{n})$ defined by (22) for all $\mathbf{n} \in \mathrm{N}_m^*$.

(ii) The above argument shows that (23) holds with the conditioning event $(\mathbf{N}_m = \mathbf{n})$ replaced by the event that Π_m is any particular partition of N_m with

$\mathbf{N}_m = \mathbf{n}$. Since $\mathbf{N}_1, \ldots, \mathbf{N}_m$ are all functions of Π_m, this proves (23) along with the Markov property of $(\mathbf{N}_1, \ldots, \mathbf{N}_n)$. □

In the exchangeable case, the EPF and the distribution of the corresponding partition of n are related by formula (10). The above consistency relation then becomes the expression in terms of EPF's of Kingman's notion of consistency of partitions of n.

Corollary 11 *A function* p: $\mathbb{N}_n^* \to [0, \infty)$ *is a PEPF iff* p: $\bigcup_1^n \mathbb{N}_m^* \to [0, \infty)$ *defined by repeated application of* (22) *is such that* $p(1) = 1$.

Proof. The "only if" assertion follows from the preceding proposition. To check the "if" assertion, let S be the sum on the left side of (19). It must be shown that $S = 1$. Clearly $S > 0$, so $q(\mathbf{n}) := p(\mathbf{n})/S$ is a PEPF. Apply the "only if" part of the corollary to q instead of p to see that $p(1)/S = q(1) = 1$. So $S = p(1) = 1$. □

The following variation of the previous corollary is the key to the proof of Proposition 9 given in Sect. 5.

Corollary 12 *If* $(\mathbf{N}_1, \ldots, \mathbf{N}_n)$ *is a Markov chain with transition probabilities of the special form* (23) *for some* p: $\bigcup_1^n \mathbb{N}_m^* \to [0, \infty)$ *with* $p(1) = 1$, *then* Π_n *is PE with PEPF given by* p.

Proof. This follows from the previous corollary, because Π_n can be recovered from the sequence $(\mathbf{N}_1, \cdots, \mathbf{N}_n)$. □

As a complement to part (ii) of Proposition 10, it is easily seen that Π_n is PE iff the reversed sequence $(\mathbf{N}_n, \cdots, \mathbf{N}_1)$ is a Markov chain with co-transition probabilities

$$P(\mathbf{N}_{m-1} = \mathbf{n}^{j-} \mid \mathbf{N}_m = \mathbf{n}) = \frac{\#(\mathbf{n}^{j-})}{\#(\mathbf{n})} \tag{24}$$

for every $1 \leqq j \leqq k$ such that $\mathbf{n}^{j-} \in \mathbb{N}_{m-1}^*$, where $\mathbf{n} = (n_1, \ldots, n_k)$ and $\mathbf{n}^{j-}$ is defined by decrementing the jth component of $\mathbf{n}$ by 1.

To conclude this section, here is a simple construction that is easily seen to yield the most general PE partition of $\mathbb{N}_n$. This is a finite "sampling without replacement" version of Construction 16 in the next section for PE partitions of N.

Construction 13 Let $\mathscr{A}_1, \ldots, \mathscr{A}_{K_n}$ denote the random subsets of $\mathbb{N}_n$ defined by the classes of Π_n in order of appearance. Let N_1, the size of $\mathscr{A}_1$, have distribution

$$P(N_1 = n_1) = P(n_1), \quad 1 \leqq n_1 \leqq n\,,$$

where $\mathrm{P}(\cdot)$ is some arbitrary probability distribution on $\{1, \ldots, n\}$. Given $\mathrm{N}_1 = n_1$, let $\mathscr{A}_1$ consist of 1 and a uniformly distributed random subset of $n_1 - 1$ elements of $\{2, \ldots, n\}$. Inductively: given that $\mathscr{A}_1, \ldots, \mathscr{A}_i$ have been defined, with $\mathrm{N}_j = n_j$, $1 \leqq j \leqq i$, such that $\sum_{j=1}^{i} n_j < n$, let N_{i+1} have distribution

$$P(N_{i+1} = n_{i+1} \mid \mathscr{A}_1, \ldots, \mathscr{A}_i) = P(n_{i+1} \mid n_1, \ldots, n_i)\,, \tag{25}$$

where $\mathrm{P}(\cdot \mid n_1,\ldots,n_i)$ is some arbitrary distribution on $\{1,\ldots,n-\sum_{j=1}^{i} n_j\}$. And given $\mathscr{A}_1,\ldots,\mathscr{A}_i$ and $\mathrm{N}_{i+1}=n_{i+1}$, let $\mathscr{A}_{i+1}$ comprise the first element of $\mathrm{N}_n - \bigcup_1^i \mathscr{A}_j$ 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 N_n. The random partition Π_n so constructed is partially exchangeable, with PEPF

$$p(\mathbf{n}) = \#(\mathbf{n})^{-1}P(n_1)P(n_2 \mid n_1)\cdots P(n_k \mid n_1,\ldots,n_{k-1}).$$

4 Proof of Theorem 6

A random partition $\Pi=\{\mathscr{A}_1,\mathscr{A}_2,\ldots\}$ of N, with classes $\mathscr{A}_i$ in order of appearance, is conveniently recoded by the sequence of N-valued random variables $(Y_1,Y_2,\ldots)$, where $Y_n=i$ if $n\in\mathscr{A}_i$. Call Y_n the *class of* n, and call (Y_n) the *classification process* of Π. Clearly, a sequence of N-valued random variables (Y_n) is the classification process of some random partition Π of N iff $Y_1=1$ and

$$Y_{n+1} \leqq \max_{1\leqq m\leqq n} Y_m + 1 \quad \text{for all } n\in \mathrm{N}\,.$$

Also, the restriction Π_n of Π to N_n is PE with PEPF p iff for every sequence of positive integers $(i_m, m\in \mathrm{N}_n)$ with

$$i_1=1,\quad i_{m+1}\leqq \max_{1\leqq \ell\leqq m} i_\ell + 1,\quad 1\leqq m\leqq n-1\,,$$

we have

$$P(Y_m = i_m \text{ for all } m\in \mathrm{N}_n) = p(n_1,\ldots,n_k)\,, \tag{26}$$

where

$$n_i = \#\{m\in \mathrm{N}_n:\ i_m=i\} \quad \text{for } 1\leqq i\leqq k=\max_{1\leqq m\leqq n} i_m\,.$$

The key to the proof of Theorem 6 is the following lemma:

Lemma 14 *Suppose that Π_n is PE. Then for each m with $1\leqq m\leqq n-2$, conditionally given Π_m such that Π_m has j classes, the random variables*

$$Y_r 1(Y_r\leqq j)\,,\quad r=m+1,\ldots,n,\quad \text{are exchangeable.} \tag{27}$$

Proof. Since a probabilistic mixture of exchangeable sequences is exchangeable, it suffices to prove the assertion conditionally given also $\mathbf{N}_n=\mathbf{n}$ for arbitrary $\mathbf{n}\in \mathrm{N}_n^*$. But given $\mathbf{N}_n=\mathbf{n}$ and Π_m, the partition Π_n is uniformly distributed over all partitions of N_n subject to the constraints imposed by $\mathbf{N}_n$ and Π_m. Suppose the sequence of class sizes of Π_m is given by $\mathbf{N}_m=\mathbf{m}$ for some $\mathbf{m}=(m_1,\ldots,m_j)\in \mathrm{N}_m^*$. The sequence $(Y_r1(Y_r\leqq j),\ r=m+1,\cdots,n)$ is then constrained to have n_i-m_i terms equal to i for each $1\leqq i\leqq j$, and the remaining $n_{j+1}+\cdots+n_k$ terms equal to 0. The exchangeability claim therefore amounts to the following: every possible sequence of values for $(Y_r1(Y_r\leqq j),\ r=m+1,\ldots,n)$, say $(v_{m+1},\ldots,v_n)$, subject to these constraints, has the same conditional probability given such Π_m and $\mathbf{N}_n$. But given Π_m with $\mathbf{N}_m=\mathbf{m}$ and given $\mathbf{N}_n=\mathbf{n}$, the partition Π_n is determined by $(v_{m+1},\ldots,v_n)$ combined with the way that those r with $v_r=0$ are assigned by Π_n to classes $j+1,\ldots,k$. Since the number of ways in which the latter assignment can be made is the same, namely $\#(n_{j+1},\ldots,n_k)$ as in (20), no matter which subset of $n_{j+1}+\cdots+n_k$ elements of $\{m+1,\ldots,n\}$ is the set of r such that $v_r=0$, the exchangeability claim follows. □

Suppose now that Π is a PE partition of N. Let

$$\nu_j := \inf\{n:\ Y_n = j\}\,.$$

By application of the above lemma, for every $m \in \mathrm{N}$ with $\mathrm{P}(\nu_j = m) > 0$, conditionally given $\nu_j = m$, the random variables

$$Y_r 1(Y_r \leqq j)\,, \quad r = m+1, m+2, \ldots \text{ are exchangeable.} \tag{28}$$

Since $\#(\mathscr{A}_i \cap \mathrm{N}_n) = \#\{m \in \mathrm{N}_n : Y_m = i\}$ the existence of almost sure limiting relative frequencies P_i for $\mathscr{A}_i$ as in (4) follows easily a consequence of de Finetti's law of large numbers for exchangeable sequences of zeros and ones. Clearly, $\mathrm{P}_i \geqq 0$ and $\sum_i P_i \leqq 1$. The claim now is that condition (ii) of Theorem 6 holds with these P_i. That is to say, for every $m \in \mathrm{N}$, conditionally given all the (P_i), and given Π_m with j classes,

$$Y_{m+1} = \begin{cases} i & \text{for } 1 \leqq i \leqq j, \text{ with probability } \mathrm{P}_i, \\ j+1 & \qquad\qquad\quad \text{with probability } 1 - P_1 - \cdots - P_j\,. \end{cases}$$

This follows by passage to the limit as $n \to \infty$ from the following combinatorial analog, which is an easy consequence of Lemma 14:

Lemma 15 *If Π_n is PE then for every $1 \leqq m \leqq n-1$, conditionally given $\mathbf{N}_n = (n_1,\ldots,n_k)$ and given Π_m with j classes such that $\mathbf{N}_m = (m_1,\ldots,m_j)$,*

$$Y_{m+1} = \begin{cases} i & \textit{for } 1 \leqq i \leqq j, \textit{ with probability } \dfrac{n_i - m_i}{n-m}\,, \\ j+1 & \qquad\qquad\quad \textit{with probability } \dfrac{n - n_1 - \cdots - n_j}{n-m}. \end{cases}$$

Due to the Markov property of the sequence $(\mathbf{N}_n,\ n = 1,2,\ldots)$ described in Proposition 10, for a partially exchangeable partition Π of N these conditional probabilities are also valid given the counts $\mathbf{N}_{n+1}, \mathbf{N}_{n+2}, \ldots$ as well as $\mathbf{N}_n$. So the passage to the limit is justified by reversed martingale convergence.

The converse implication in Theorem 6, (ii) $\Rightarrow$ (i), is straightforward. It follows from (ii) by first multiplying conditional probabilities and then taking expectation that Π_n defined by (ii) meets the requirements of Definition 4 with $p(\mathbf{n})$ given by (7). Finally, to check the last sentence of the theorem, observe that for Π satisfying (ii), the limiting relative frequency of $\mathscr{A}_i$ exists and equals P_i by the law of large numbers for independent Bernoulli trials. □

5 Residual allocation models

For $0 \leqq w \leqq 1$ let $\overline{w} = 1 - w$. The *residual allocation model* (RAM)

$$P_i = \overline{W}_i \cdots \overline{W}_{i-1} W_i \tag{29}$$

for a random discrete distribution (P_i), with independent W_i, has been proposed in a number of contexts. See [27] for a survey. Relaxing the assumption of independent factors allows any sequence of random variables (P_i) with $P_i \geqq 0$ and $\sum_i P_i \leqq 1$ to be represented in the form (29). Call this expression for (P_i) a *generalized residual allocation model* (GRAM). Thinking in terms of the residual fractions W_i instead of the P_i suggests the following construction of a PE partition Π of N with prescribed limiting frequencies P_i for the classes in order of appearance. This generalizes Hoppe's [16] construction of the exchangeable partition of N governed by Ewens' sampling formula.

Construction 16 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 N into random subsets $\mathscr{A}_1, \mathscr{A}_2, \ldots$ as follows. Let

$$(X_{ni}, n = 1,2,\ldots, i = 1,2,\ldots)$$

be indicator variables that are conditionally independent given $(W_1, W_2, \ldots)$ with

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

Let $\mathscr{A}_1 = \{1\} \cup \{n \in \mathrm{N} \colon\ X_{n1} = 1\}$. Inductively:

$$\text{for } i \geqq 1 \quad \text{let } \mathscr{C}_i = \mathrm{N} - (\mathscr{A}_1 \cup \ldots \cup \mathscr{A}_i) \,;$$

given $\mathscr{C}_i$ is non-empty (or, what is the same, $\Pi^i_{j=1}(1 - W_i) > 0$), let

$$\mathscr{A}_{i+1} = \{\min\{\mathscr{C}_i\}\} \cup \{n \in \mathscr{C}_i \colon\ X_{n,i+1} = 1\} \,.$$

It is easily seen directly that Π is PE. By construction the $\mathscr{A}_i$ are in order of appearance, with limiting frequencies P_i as in (29), by repeated application of the law of large numbers. It can also be seen directly that the conditional distribution of Π given (P_i) is as in (ii) of Theorem 6. So the most general possible distribution for a PE partition of N can be obtained by the above construction. As an easy consequence of this construction, there is the following corollary of Theorem 6. This corollary contains Theorem 4 of Hoppe [16] in the exchangeable case governed by Ewens' sampling formula.

Corollary 17 *Let $\Pi = \{\mathscr{A}_i\}$ be a PE partition of* N*, P_i the almost sure limit as $n \to \infty$ of N_{in}/n, where $N_{in} = \#(\mathscr{A}_i \cap \mathrm{N}_n)$. For each $i \geqq 0$, given $(P_1, P_2, \ldots)$ and (for $i \geqq 1$) 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 $W_{i+1} = P_{i+1}/(1 - \sum_1^i P_j)$.*

Using the GRAM (29), expression (7) for the PEPF becomes

$$p(\mathbf{n}) = E\left[\prod_{i=1}^{k} W_i^{n_i - 1}\overline{W}_i^{n_{i+1}+\cdots+n_k}\right] \,. \tag{30}$$

Let $m_i(r,s) = E[W_i^r \overline{W}_i^s]$. Assuming independent factors W_i, (30) becomes

$$p(n_1, \ldots, n_k) = \prod_{i=1}^{k} m_i(n_i - 1, n_{i+1} + \cdots + n_k) \tag{31}$$

Given a sequence of distributions for W_i on [0, 1], it is not obvious by inspection of formula (31) whether $p(n_1,\ldots,n_k)$ is symmetric in $(n_1,\ldots,n_k)$, that is to say whether the random partition of N is exchangeable. See [27] for a construction of all such sequences of distributions. The main example is provided by the sequence of beta distributions for W_i described in Proposition 9.

Proof of Proposition 9 It is easily checked that the sequential construction of Π_n defines transition probabilities for $\mathbf{N}_1, \mathbf{N}_2, \ldots$ that are of the form (23) for the $p(\mathbf{n})$ defined by (16), which satisfies $p(1) = 1$ and is obviously symmetric. So the partition of N is exchangeable with EPF $p(\mathbf{n})$ by Corollary 12 and Proposition 5. Formula (17) follows from (16) by (10). The form of the joint distribution of the P_i can be checked either from (31) by computation of moments derived from the beta distributions, or a variation of the argument of Hoppe [16] in the case $\alpha \leqq 0$. □

Acknowledgement. Thanks to David Aldous, Persi Diaconis, Warren Ewens and Simon Tavaré for many stimulating conversations.

References

1. Aldous, D.J.: Exchangeability and related topics. In: Hennequin, P.L. (ed.) École d'Été de Probabilités de Saint-Flour XII. (Lect. Notes Math. vol. 1117) Berlin Heidelberg New York: Springer 1985
2. Antoniak, C.: Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. Ann. Stat. **2**, 1152–1174 (1974)
3. Blackwell, D., MacQueen, J.B.: Ferguson distributions via Pólya urn schemes. Ann. Stat. **1**, 353–355 (1973)
4. de Finetti, B. : Sur la condition d'équivalence partielle. Actualités Scientifiques et Industrielles. 739 (1938). Herman and Cie: Paris. Translated In: Studies in Inductive Logic and Probability, II. Jeffrey, R. (ed.) University of California Press: Berkeley 1980
5. Diaconis, P., Freedman, D.: Partial exchangeability and sufficiency. In: Ghosh, J.K., Roy, J. (eds.) Statistics Applications and New Directions; Proceedings of the Indian Statistical Institute Golden Jubilee International Conference; Sankhya A. pp. 205–236 Indian Statistical Institute 1984
6. Donnelly, P.: Partition structures, Pólya urns, the Ewens sampling formula, and the ages of alleles. Theoret. Population Biology **30**, 271–288 (1986)
7. Donnelly, P.: The heaps process, libraries and size biased permutations. J. Appl. Probab. **28**, 322–335 (1991)
8. Donnelly, P., Joyce, P.: Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex. Stochast. Processes Appl. **31**, 89–103 (1989)
9. Donnelly, P., Tavaré, S.: The ages of alleles and a coalescent. Adv. Appl. Probab. **18**, 1–19, 1023 (1986)
10. Engen, S.: Stochastic abundance models with emphasis on biological communities and species diversity. London: Chapman and Hall Ltd., 1978
11. Ewens, W.J.: The sampling theory of selectively neutral alleles. Theor. Popul. Biol. **3**, 87–112 (1972)
12. Ewens, W.J.: Population genetics theory – the past and the future. In: Lessard, S. (ed.), Mathematical and statistical problems in evolution. Montreal: University of Montreal Press 1988
13. Fisher, R.A., Corbet, A.S., Williams, C.B.: The relation between the number of species and the number of individuals in a random sample of an animal population. J. Animal Ecol. **12**, 42–58 (1943)

14. Hoppe, F.M.: Pólya-like urns and the Ewens sampling formula. J. Math. Biol. **20**, 91–94 (1984)
15. Hoppe, F.M.: Size-biased filtering of Poisson–Dirichlet samples with an application to partition structures in genetics. J. Appl. Probab. **23**, 1008–1012 (1986)
16. Hoppe, F.M.: The sampling theory of neutral alleles and an urn model in population genetics. J. Math. Biol. **25**, 123–159 (1987)
17. Ignatov, T.: On a constant arising in the theory of symmetric groups and on Poisson–Dirichlet measures. Theory Probab. Appl. **27**, 136–147 (1982)
18. Johnson, W.E.: Probability: the deductive and inductive problems. Mind **49**, 409–423 (1932)
19. Kingman, J.F.: The population structure associated with the Ewens sampling formula. Theor. Popul. Biol. **11**, 274–283 (1977)
20. Kingman, J.F.: Random partitions in population genetics. Proc. R. Soc. Lond. A. **361**, 1–20 (1978)
21. Kingman, J.F.: The representation of partition structures. J. London Math. Soc. **18**, 374–380 (1978)
22. Kingman, J.F.: The coalescent. Stochast. Processes Appl. **13**, 235–248 (1982)
23. McCloskey, J.W.: A model for the distribution of individuals by species in an environment. Ph. D. Thesis, Michigan State University (1965)
24. Patil, G.P., Taillie, C.: Diversity as a concept and its implications for random communities. Bull. Int. Stat. Inst. **XLVII**, 497–515 (1977)
25. Perman, M., Pitman, J., Yor, M.: Size-biased sampling of Poisson point processes and excursions. Probab. Related Fields **92**, 21–39 (1992)
26. Pitman, J.: Partition structures derived from Brownian motion and stable subordinators. Technical Report 346, Dept. Statistics, U.C. Berkeley Preprint (1992)
27. Pitman, J.: Random discrete distributions invariant under size-biased permutation. Technical Report 344, Dept. Statistics, U.C. Berkeley (1992) To appear in J. Appl. Probab.
28. Pitman, J.: The two-parameter generalization of Ewens' random partition structure. Technical Report 345, Dept. Statistics, U.C. Berkeley, (1992)
29. Pitman, J., Yor, M.: The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. Preprint (1994)
30. Watterson, G.A.: The stationary distribution of the infinitely-many neutral alleles diffusion model. J. Appl. Probab. **13**, 639–651 (1976)
31. Zabell, S.L.: Predicting the unpredictable. Synthese **90**, 205–232 (1992)
32. Zabell, S.L.: The continuum of inductive methods revisited. Preprint (1994)