2

Exchangeable random partitions

This chapter is a review of basic ideas from Kingman's theory of exchangeable random partitions [253], as further developed in [14, 347, 350]. This theory turns out to be of interest in a number of contexts, for instance in the study of population genetics, Bayesian statistics, and models for processes of coagulation and fragmentation. The chapter is arranged as follows.

2.1. Finite partitions This section introduces the *exchangeable partition probability function (EPPF)* associated with an exchangeable random partition Π_n of the set $[n] := \{1, \ldots, n\}$. This symmetric function of compositions $(n_1, \cdots, n_k)$ of n gives the probability that Π_n equals any *particular* partition of $[n]$ into k subsets of sizes n_1, n_2, $\ldots$,n_k, where $n_i \geq 1$ and $\Sigma_i n_i = n$. Basic examples are provided by *Gibbs partitions* for which the EPPF assumes a product form.

2.2. Infinite partitions A random partition Π_∞ of the set $\mathbb{N}$ of positive integers is called exchangeable if its restriction Π_n to $[n]$ is exchangeable for every n. The distribution of Π_∞ is determined by an EPPF which is a function of compositions of positive integers subject to an addition rule expressing the consistency of the partitions Π_n as n varies. Kingman [250] established a one-to-one correspondence between distributions of such exchangeable random partitions of $\mathbb{N}$ and distributions for a sequence of nonnegative random variables $P_1^\downarrow, P_2^\downarrow, \ldots$ with $P_1^\downarrow \geq P_2^\downarrow \geq \ldots$ and $\sum_k P_k^\downarrow \leq 1$. In Kingman's *paintbox representation*, the blocks of Π_∞ are the equivalence classes generated by the random equivalence relation $\sim$ on positive integers, constructed as follows from *ranked frequencies* $(P_k^\downarrow)$ and a sequence of independent random variables U_i with uniform distribution on $[0,1]$, where (U_i) and $(P_k^\downarrow)$ are independent: $i \sim j$ iff either $i = j$ or both U_i and U_j fall in I_k for some k, where the I_k are some disjoint random sub-intervals of $[0,1]$ of lengths $P_k^\downarrow$. Each $P_k^\downarrow$ with $P_k^\downarrow > 0$ is then the asymptotic frequency of some corresponding block of Π_∞, and if $\sum_k P_k^\downarrow < 1$ there is also a remaining subset of $\mathbb{N}$ with asymptotic frequency $1 - \sum_k P_k^\downarrow$, each of whose elements is a singleton block of Π_∞.

2.3. Structural distributions A basic property of every exchangeable random partition Π_∞ of $\mathbb{N}$ is that each block of Π_∞ has a limiting relative frequency almost surely. The *structural distribution* associated with Π_∞ is the probability distribution on $[0,1]$ of the asymptotic frequency of the block of Π_∞ that contains a particular positive integer, say 1. In terms of Kingman's representation, this is the distribution of a size-biased pick from the associated sequence of random frequencies $(P_k^\downarrow)$. Many important features of exchangeable random partitions and associated random discrete distributions, such as the mean number of frequencies in a given interval, can be expressed in terms of the structural distribution.

2.4. Convergence Convergence in distribution of a sequence of exchangeable random partitions Π_n of $[n]$ as $n \to \infty$ can be expressed in several equivalent ways: in terms of induced partitions of $[m]$ for fixed m, in terms of ranked or size-biased frequencies, and in terms of an associated process with exchangeable increments.

2.5. Limits of Gibbs partitions Limits of Gibbs partitions lead to exchangeable random partitions of $\mathbb{N}$ with ranked frequencies $(P_i^\downarrow, i \geq 1)$ distributed according to some mixture over s of the conditional distribution of ranked jumps of some subordinator $(T_u, 0 \leq u \leq s)$ given $T_s = 1$. Two important special cases arise when T is a gamma process, or a stable subordinator of index $\alpha \in (0,1)$. The study of such limit distributions is pursued further in Chapter 4.

2.1. Finite partitions

A random partition Π_n of $[n]$ is called *exchangeable* if its distribution is invariant under the natural action on partitions of $[n]$ by the symmetric group of permutations of $[n]$. Equivalently, for each partition $\{A_1, \ldots, A_k\}$ of $[n]$,

$$\mathbb{P}(\Pi_n = \{A_1, \ldots, A_k\}) = p(|A_1|, \ldots, |A_k|)$$

for some symmetric function p of compositions $(n_1, \ldots, n_k)$ of n. This function p is called the *exchangeable partition probability function (EPPF)* of Π_n. For instance, given two positive sequences $v_\bullet = (v_1, v_2, \ldots)$ and $w_\bullet = (w_1, w_2, \ldots)$, the formula

$$p(n_1, \ldots, n_k; v_\bullet, w_\bullet) := \frac{v_k \prod_{i=1}^k w_{n_i}}{B_n(v_\bullet, w_\bullet)} \tag{2.1}$$

where $B_n(v_\bullet, w_\bullet)$ is a normalization constant, defines the EPPF of a Gibbs partition determined by $v_\bullet$ and $w_\bullet$ as discussed in Section 1.5 . In most applications, it is the sizes of blocks of an exchangeable random partition Π_n which are of primary interest. The next three paragraphs present three different ways to encode these block sizes as a random composition of $[n]$, and show how the distributions of these encodings are determined by the EPPF p.

Decreasing order Let $(N_{n,1}^\downarrow, \ldots, N_{n,K_n}^\downarrow)$ denote the *partition of* n *induced by* Π_n, that is the random composition of n defined by the sizes of blocks of

Π_n with blocks in decreasing order of size. Then for each partition of n with component sizes (n_i) in decreasing order,

$$\mathbb{P}((N^{\downarrow}_{n,1}, \dots, N^{\downarrow}_{n,K_n}) = (n_1, \dots, n_k)) = \frac{n!}{\prod_{i=1}^{n} (i!)^{m_i} m_i!} \, p(n_1, \dots, n_k) \qquad (2.2)$$

where

$$m_i := \sum_{\ell=1}^{k} 1(n_\ell = i) \qquad (2.3)$$

is the number of components of size i in the given partition of n, and the combinatorial factor is the number of partitions of $[n]$ corresponding to the given partition of n. Let $|\Pi_n|_j$ denote the number of blocks of Π_n of size j. Due to the bijection between partitions of n and possible vectors of counts $(m_i, 1 \le i \le n)$, for (m_i) a vector of non-negative integers subject to $\sum_i m_i = k$ and $\sum_i i m_i = n$, the probability

$$\mathbb{P}(|\Pi_n|_i = m_i \text{ for } 1 \le i \le n), \qquad (2.4)$$

that is the probability that Π_n has m_i blocks of size i for each $1 \le i \le n$, is identical to the the probability in (2.2) for $(n_1, \dots, n_k)$ the decreasing sequence subject to (2.3).

Size-biased order of least elements Let $(\tilde{N}_{n,1}, \dots, \tilde{N}_{n,K_n})$ denote the random composition of n defined by the sizes of blocks of Π_n with blocks in order of appearance. Then for all compositions $(n_1, \dots, n_k)$ of n into k parts,

$$\mathbb{P}((\tilde{N}_{n,1}, \dots, \tilde{N}_{n,K_n}) = (n_1, \dots, n_k)) \qquad (2.5)$$

$$= \frac{n!}{n_k(n_k + n_{k-1}) \cdots (n_k + \cdots + n_1) \; \prod_{i=1}^{k} (n_i - 1)!} \, p(n_1, \dots, n_k) \qquad (2.6)$$

where the combinatorial factor is the number of partitions of $[n]$ with the prescribed block sizes in order of appearance [115]. Note that $(\tilde{N}_{n,1}, \dots, \tilde{N}_{n,K_n})$ is a *size-biased random permutation* of $(N^{\downarrow}_{n,1}, \dots, N^{\downarrow}_{n,K_n})$, meaning that given the decreasing rearrangement, the blocks appear in the random order in which they would be discovered in a process of simple random sampling without replacement.

Exchangeable random order It is often convenient to consider the block sizes of a random partition of $[n]$ in *exchangeable random order*, meaning that conditionally given $\Pi_n = \{A_1, \dots, A_k\}$, random variables $(N^{ex}_{n,1}, \dots, N^{ex}_{n,k})$ are constructed as $N^{ex}_{n,i} = |A_{\sigma(i)}|$ where σ is a uniformly distributed random permutation of $[k]$. Then

$$\mathbb{P}((N^{ex}_{n,1}, \dots, N^{ex}_{n,K_n}) = (n_1, \dots, n_k)) = \binom{n}{n_1, \dots, n_k} \frac{1}{k!} \, p(n_1, \dots, n_k). \qquad (2.7)$$

To see this, recall that $p(n_1, \dots, n_k)$ is the probability of any particular partition of $[n]$ with block sizes $(n_1, \dots, n_k)$ in some order. Dividing by $k!$ gives the probability of obtaining a particular ordered partition of $[n]$ after randomizing the order of the blocks, and the multinomial coefficent is the number of such ordered partitions consistent with $(n_1, \dots, n_k)$.

Partitions generated by sampling without replacement. Let

$$\Pi(x_1, \dots, x_n)$$

denote the partition of $[n]$ generated by a sequence $x_1, \dots, x_n$. That is the partition whose blocks are the equivalence classes for the random equivalence relation $i \sim j$ iff $x_i = x_j$. If $(X_1, \dots, X_n)$ is a sequence of exchangeable random variables, then $\Pi(X_1, \dots, X_n)$ is an exchangeable random partition of $[n]$. Moreover, the most general possible distribution of an exchangeable random partition of $[n]$ is obtained this way. To be more precise, there is the following basic result. See Figure 2.1 for a less formal statement.

Proposition 2.1. [14] *Let Π_n be an exchangeable random partition of $[n]$, and let $\pi_n := (N^{\downarrow}_{n,i}, 1 \le i \le K_n)$ be the corresponding partition of n defined by the decreasing rearrangement of block sizes of Π_n. Then the joint law of Π_n and π_n is that of $\Pi(X_1, \dots, X_n)$ and π_n, where conditionally given π_n the sequence $(X_1, \dots, X_n)$ is defined by simple random sampling without replacement from a list $x_1, \dots, x_n$ with $N^{\downarrow}_{n,i}$ values equal to i for each $1 \le i \le K_n$, say $(X_1, \dots, X_n) = (x_{\sigma(1)}, \dots, x_{\sigma(n)})$ where σ is a uniform random permutation of $[n]$.*

Exercises

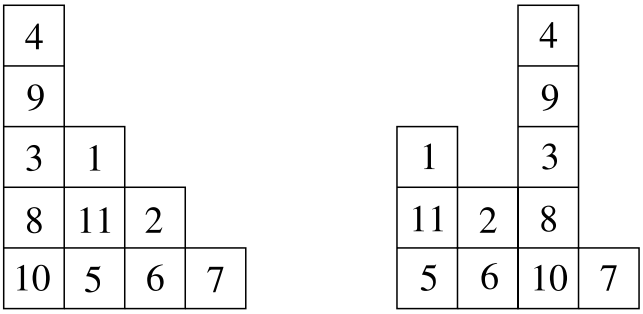
2.1.1. Prove Proposition 2.1.

2.1.2. Corresponding to each probability distribution Q on the set $\mathcal{P}_n$ of partitions of n, there is a unique distribution of an exchangeable partition Π_n of $[n]$ which induces a partition π_n of n with distribution Q: given π_n, let Π_n have uniform distribution on the set of all partitions of $[n]$ whose block sizes are consistent with π_n.

2.1.3. A function p defined on the set of compositions of n is the EPPF of some exchangeable random partition Π_n of $[n]$ if and only if p is non-negative, symmetric, and

$$\sum_{k=1}^{n} \sum_{(n_1,\dots,n_k)} \binom{n}{n_1, \dots, n_k} \frac{1}{k!} p(n_1, \dots, n_k) = 1,$$

where the second sum is over all compositions $(n_1, \dots, n_k)$ of n with k parts.



Block sizes in decreasing order of size : $(5, 3, 2, 1)$

Block sizes in the size-biased order of the least element : $(3, 2, 5, 1)$

Figure 2.1: A random partition Π_{11} of [11]. To state Proposition 2.1 less formally: if Π_n is exchangeable, then given that the block sizes of Π_n in decreasing order define some pattern of boxes, as above left for $n = 11$, known as a Ferrer's diagram, corresponding to a partition of the integer n, the partition of $[n]$ is recovered by filling the boxes with numbers sampled from $[n]$ without replacement, then taking the partition generated by the columns of boxes, to get e.g. $\Pi_n = \{\{4, 9, 3, 8, 10\}, \{1, 11, 5\}, \{2, 6\}, \{7\}\}$ as above.

2.1.4. (Serban Nacu [318]) . Let X_i be the indicator of the event that i is the least element of some block of an exchangeable random partition Π_n of $[n]$. Show that the joint law of the $(X_i, 1 \leq i \leq n)$ determines the law of Π_n.

2.1.5. (Problem) Characterize all possible laws of strings of 0's and 1's which can arise as in the previous exercise. Variants of this problem, with side conditions on the laws, are easier but still of some interest. Compare with Exercise 3.2.4 .

2.1.6. The EPPF of an exchangeable random partition Π_n of $[n]$ is $p(n_1, \ldots, n_k) := \mathbb{P}(\Pi_n = \Pi)$ for each particular partition $\Pi = \{A_1, \ldots, A_k\}$ of $[n]$ with $|A_i| = n_i$ for all $1 \leq i \leq n$. Let $q(n_1, \ldots, n_k)$ be the common value of $\mathbb{P}(\Pi_n \geq \Pi)$ for each such Π, where $\Pi_n \geq \Pi$ means that Π_n is *coarser than* Π, i.e. each block of Π_n is some union of blocks of Π. Each of the functions p and q determines the other via the formula

$$q(n_1, \ldots, n_k) = \sum_{j=1}^{k} \sum_{\{B_1, \ldots, B_j\}} p(n_{B_1}, \ldots, n_{B_j}) \tag{2.8}$$

where the second sum is over partitions $\{B_1, \ldots, B_j\}$ of $[k]$, and $n_B := \sum_{i \in B} n_i$.

2.2. Infinite partitions

For $1 \leq m \leq n$ let $\Pi_{m,n}$ denote the restriction to $[m]$ of Π_n, an exchangeable random partition of $[n]$. Then $\Pi_{m,n}$ is an exchangeable random partition of $[m]$ with some EPPF $p_n : \mathcal{C}_m \to [0,1]$, where $\mathcal{C}_m$ is the set of compositions of m. So for each partition $\{A_1, \ldots, A_k\}$ of $[m]$

$$\mathbb{P}(\Pi_{m,n} = \{A_1, \ldots, A_k\}) = p_n(|A_1|, \ldots, |A_k|)$$

where the definition of the EPPF of Π_n, that is $p_n : \mathcal{C}_n \to [0,1]$, is extended recursively to $\mathcal{C}_m$ for $m = n-1, n-2, \ldots, 1$, using the addition rule of probability. Thus the function $p = p_n$ satisfies the following *addition rule*: for each composition $(n_1, \ldots, n_k)$ of $m < n$

$$p(n_1, \ldots, n_k) = \sum_{j=1}^{k} p(\ldots, n_j + 1, \ldots) + p(n_1, \ldots, n_k, 1) \tag{2.9}$$

where $(\ldots, n_j + 1, \ldots)$ is derived from $(n_1, \ldots, n_k)$ by substituting $n_j + 1$ for n_j. For instance,

$$1 = p(1) = p(2) + p(1,1) \tag{2.10}$$

and

$$p(2) = p(3) + p(2,1); \quad p(1,1) = p(1,2) + p(2,1) + p(1,1,1) \tag{2.11}$$

where $p(1,2) = p(2,1)$ by symmetry of the EPPF.

Consistency [253, 14, 347] Call a sequence of exchangeable random partitions (Π_n) *consistent in distribution* if Π_m has the same distribution as $\Pi_{m,n}$ for every $m < n$. Equivalently, there is a symmetric function p defined on the set of all integer compositions (an *infinite EPPF*) such that $p(1) = 1$, the addition rule (2.9) holds for all integer compositions $(n_1, \ldots, n_k)$, and the restriction of p to $\mathcal{C}_n$ is the EPPF of Π_n. Such (Π_n) can then be constructed so that $\Pi_m = \Pi_{m,n}$ almost surely for every $m < n$. The sequence of random partitions $\Pi_\infty := (\Pi_n)$ is then called an *exchangeable random partition of* $\mathbb{N}$, or an *infinite exchangeable random partition*. Such a Π_∞ can be regarded as a random element of the set $\mathcal{P}_\mathbb{N}$ of partitions of $\mathbb{N}$, equipped with the σ-field generated by the restriction maps from $\mathcal{P}_\mathbb{N}$ to $\mathcal{P}_{[n]}$ for all n. One motivation for the study of exchangeable partitions of $\mathbb{N}$ is that if (Π_n) is any sequence of exchangeable partitions of $[n]$ for $n = 1, 2, \ldots$ which has a limit in distribution in the sense that $\Pi_{m,n} \xrightarrow{d} \Pi_{m,\infty}$ for each m as $n \to \infty$, then the sequence of limit partitions $(\Pi_{m,\infty}, m = 1, 2, \ldots)$ is consistent in distribution, hence constructible as an exchangeable partition of $\mathbb{N}$. This notion of weak convergence of random partitions is further developed in Section 2.4.

Partitions generated by random sampling Let (X_n) be an infinite exchangeable sequence of real random variables. According to de Finetti's theorem, (X_n) is obtained by *sampling from some random probability distribution* F. That is to say there is a random probability distribution F on the line, such that conditionally given F the X_i are i.i.d. according to F. To be more explicit, if

$$F_n(x) := \frac{1}{n} \sum_{i=1}^{n} 1(X_i \leq x)$$

is the empirical distribution of the first n values of the sequence, then by combination of de Finetti's theorem [122, p. 269] and the Glivenko-Cantelli theorem [122, p. 59]

$$F(x) = \lim_n F_n(x) \text{ uniformly in } x \text{ almost surely.} \tag{2.12}$$

Let Π_∞ be the exchangeable random partition of $\mathbb{N}$ generated by (X_n), meaning that the restriction Π_n of Π_∞ to $[n]$ is the partition generated by $(X_1, \ldots, X_n)$, as defined above Proposition 2.1. The distribution of $\Pi_\infty := (\Pi_n)$ is determined by the distribution of $(P_i^\downarrow, i \geq 1)$, where $P_i^\downarrow$ is the magnitude of the ith largest atom of F. Note that $1 - \sum_i P_i^\downarrow$ is the magnitude of the continuous component of F, which might be strictly positive, and that almost surely each i such that X_i is not an atom of F contributes a singleton component $\{i\}$ to Π_∞. To summarize this setup, say Π_∞ is *generated by sampling from a random distribution with ranked atoms* $(P_i^\downarrow, i \geq 1)$. According to the following theorem, every infinite exchangeable partition has the same distribution as one generated this way. This is the infinite analog of Proposition 2.1, according to which every finite exchangeable random partition can be generated by a process of random sampling without replacement from some random population.

Theorem 2.2. (Kingman's representation [251, 253]) *Let* $\Pi_\infty := (\Pi_n)$ *be an exchangeable random partition of* $\mathbb{N}$*, and let* $(N_{n,i}^\downarrow, i \geq 1)$ *be the decreasing rearrangement of block sizes of* Π_n*, with* $N_{n,i}^\downarrow = 0$ *if* Π_n *has fewer than i blocks. Then* $N_{n,i}^\downarrow / n$ *has an almost sure limit* $P_i^\downarrow$ *as* $n \to \infty$ *for each* i*. Moreover the conditional distribution of* Π_∞ *given* $(P_i^\downarrow, i \geq 1)$ *is as if* Π_∞ *were generated by random sampling from a random distribution with ranked atoms* $(P_i^\downarrow, i \geq 1)$*.*

Proof. (Sketch, following Aldous [14, p. 88]) Without loss of generality, it can be supposed that on the same probability space as Π_∞ there is an independent sequence of i.i.d. uniform $[0, 1]$ variables U_j. Let $X_n = U_j$ if n falls in the jth class of Π_∞ to appear. Then $(X_n, n = 1, 2, \ldots)$ is exchangeable. Hence Π_∞ is generated by random sampling from F which is the uniform almost sure limit of

$$F_n(u) := \frac{1}{n} \sum_{m=1}^{n} 1(X_m \leq u) = \sum_{i=1}^{\infty} \frac{N_{n,i}^\downarrow}{n} 1(\hat{U}_{n,i} \leq u)$$

for some Π_n-dependent rearrangement $\hat{U}_{n,i}$ of the U_j. By the almost sure uniformity (2.12) of convergence of F_n to F, the size $N^{\downarrow}_{n,i}/n$ of the ith largest atom of F_n has almost sure limit $P^{\downarrow}_i$ which is the size of the ith largest atom of F. □

Theorem 2.2 sets up a bijection (*Kingman's correspondence*) between probability distributions for an infinite exchangeable random partition, as specified by an infinite EPPF, and probability distributions of $(P^{\downarrow}_i)$ on the set

$$\mathcal{P}^{\downarrow}_{[0,1]} := \{(p_1, p_2, \ldots) : p_1 \geq p_2 \geq \cdots \geq 0 \text{ and } \sum_{i=1}^{\infty} p_i \leq 1\} \tag{2.13}$$

of ranked sub-probability distributions on $\mathbb{N}$.

Note that the set of all infinite EPPF's $p : \cup_{n=1}^{\infty} \mathcal{C}_n \to [0,1]$, with the topology of pointwise convergence, is compact.

Theorem 2.3. (Continuity of Kingman's correspondence [250, §5], [252, p. 45]) *Pointwise convergence of EPPF's is equivalent to weak convergence of finite dimensional distributions of the corresponding ranked frequencies.*

A similar result holds for the frequencies of blocks in order of appearance. See Theorem 3.1. Assuming for simplicity that Π_∞ has *proper frequencies*, meaning that $\sum_i P^{\downarrow}_i = 1$ a.s., Kingman's correspondence can be made more explicit as follows. Let (P_i) denote any rearrangement of the ranked frequencies $(P^{\downarrow}_i)$, which can even be a random rearrangement. Then

$$p(n_1, \ldots, n_k) = \sum_{(j_1, \ldots, j_k)} \mathbb{E}\left[\prod_{i=1}^{k} P_{j_i}^{n_i}\right] \tag{2.14}$$

where $(j_1, \ldots, j_k)$ ranges over all ordered k-tuples of distinct positive integers. This is easily seen from Kingman's representation for $(P_i) = (P^{\downarrow}_i)$. The formula holds also for any rearrangement of these frequencies, because the right side is the expectation of a function of $(P_1, P_2, \ldots)$ which is invariant under finite or infinite permutations of its arguments. In particular (P_i) could be the sequence $(\tilde{P}_i)$ of limit frequencies of classes of (Π_n) in order of appearance, which is a size-biased random permutation of $(P^{\downarrow}_i)$. A much simpler formula in this case is described later in Theorem 3.1.

Exercises

The first two exercises recall some forms of Pólya's urn scheme [151, VII.4], which allow explicit sequential constructions of exchangeable sequences and random partitions. See [300],[350] for more in this vein.

2.2.1. (Beta-binomial) Fix $a, b > 0$. Let $S_n := X_1 + \cdots + X_n$, where the X_i have values 0 or 1. Check that

$$\mathbb{P}(X_{n+1} = 1 \mid X_1, \ldots, X_n) = \frac{a + S_n}{a + b + n} \tag{2.15}$$

for all $n \geq 0$ if and only if the X_i are exchangeable and the almost sure limit of S_n/n has the beta(a, b) distribution.

2.2.2. (Dirichlet-multinomial) Fix $\theta_1, \ldots, \theta_m > 0$. Let $(X_n, n = 1, 2, \ldots)$ be a process with values in $\{1, \ldots, m\}$. If for each $n \geq 0$, given $(X_1, \ldots, X_n)$ with n_i values equal to i for each $1 \leq i \leq m$, where $n_1 + \cdots + n_m = n$,

$$X_{n+1} = i \text{ with probability } \frac{\theta_i + n_i}{\theta_1 + \cdots + \theta_m + n}$$

then (X_n) is exchangeable with asymptotic frequencies P_i with the *Dirichlet* $(\theta_1, \ldots, \theta_m)$ distribution (0.47), and conversely.

2.2.3. (Sampling from exchangeable frequencies) Let $p(n_1, \ldots, n_k)$ be the EPPF corresponding to some sequence of random ranked frequencies $(P_1^{\downarrow}, \ldots, P_m^{\downarrow})$ with $\sum_{i=1}^m P_i^{\downarrow} = 1$ for some $m < \infty$. Let $(P_1, \ldots, P_m)$ be the exchangeable sequence with $\sum_{i=1}^m P_i = 1$ obtained by putting these ranked frequencies in exchangeable random order. Then

$$p(n_1, \ldots, n_k) = (m)_{k\downarrow} \mathbb{E}\left[\prod_{i=1}^k P_i^{n_i}\right].$$

2.2.4. (Coupon Collecting) If $P_i^{\downarrow} = 1/m$ for $1 \leq i \leq m$ then

$$p(n_1, \ldots, n_k) = (m)_{k\downarrow}/m^n \text{ where } n := \sum_{i=1}^k n_i. \tag{2.16}$$

2.2.5. (Sampling from exchangeable Dirichlet frequencies) [428] If $(P_1, \ldots, P_m)$ has the symmetric Dirichlet distribution (0.47) with parameters $\theta_1 = \cdots = \theta_m = \kappa > 0$, then

$$p(n_1, \ldots, n_k) = (m)_{k\downarrow} \frac{\prod_{i=1}^k (\kappa)_{n_i \uparrow}}{(m\kappa)_{n\uparrow}}. \tag{2.17}$$

Note that the coupon collector's partition (2.16) is recovered in the limit as $\kappa \to \infty$.

2.2.6. (The Blackwell-MacQueen urn scheme) [68]. Fix $\theta > 0$. Let (X_n) with values in $[0, 1]$ be governed by the following prediction rule: $n \geq 0$,

$$\mathbb{P}(X_{n+1} \in \cdot \mid X_1, \ldots, X_n) = \frac{\theta\, \lambda(\cdot) + \sum_{i=1}^n 1(X_i \in \cdot)}{\theta + n} \tag{2.18}$$

where $\lambda(\cdot)$ is Lebesgue measure on $[0, 1]$. Then (X_n) is exchangeable, distributed as a sample from a Dirichlet(θ) process F_θ as in (0.46).

2.2.7. (The Ewens sampling formula) [144, 26, 145] As $m \to \infty$ and $\kappa \to 0$ with $m\kappa \to \theta$, the EPPF in (2.17) converges to the EPPF

$$p_{0,\theta}(n_1, \ldots, n_k) = \frac{\theta^k}{(\theta)_{n\uparrow}} \prod_{i=1}^{k} (n_i - 1)! \tag{2.19}$$

Such a partition is generated by $X_1, \ldots, X_n$ governed by the Blackwell-MacQueen urn scheme (2.18). The corresponding partition of n has distribution

$$\mathbb{P}(|\Pi_n|_i = m_i, 1 \leq i \leq n) = \frac{n!\,\theta^k}{(\theta)_{n\uparrow}} \prod_{i=1}^{n} \frac{1}{i^{m_i} m_i!} \tag{2.20}$$

for (m_i) as in (2.4). The corresponding ranked frequencies are the ranked jumps of the Dirichlet(θ) process. The frequencies in order of appearance are described in Theorem 3.2.

2.2.8. (Continuity of Kingman's correspondence) Prove Theorem 2.3.

Notes and comments

The theory of exchangeable random partitions described here, following [14] and [347], is essentially equivalent to Kingman's theories of *partition structures* [250, 251] and of *exchangeable random equivalence relations* [253]. The theory is simplified by describing the consistent sequence of distributions of partitions of $[n]$ by its EPPF, rather than by the corresponding sequence of distributions of integer partitions, which is what Kingman called a partition structure. Donnelly and Joyce [114] and Gnedin [172] developed a parallel theory of *composition structures*, whose extreme points are represented by open subsets of $[0, 1]$. See also [198, 199] for alternate approaches.

In the work of Kerov and Vershik on multiplicative branchings [241, 242, 244], each extreme infinite exchangeable partition corresponds to a real-valued character of the algebra of symmetric functions, with certain positivity conditions. See also Aldous [14] regarding exchangeable arrays, and Kallenberg [231] for paintbox representations of random partitions with general symmetries.

2.3. Structural distributions

Let (P_i) be a random discrete probability distribution with size-biased permutation $(\tilde{P}_j)$. So in particular

$$\tilde{P}_1 = P_{\sigma(1)} \text{ where } \mathbb{P}(\sigma(1) = i \,|\, P_1, P_2, \ldots) = P_i \qquad (i = 1, 2, \ldots). \tag{2.21}$$

The random variable $\tilde{P}_1$ may be called a *size-biased pick* from (P_i). Let $\tilde{\nu}$ denote the distribution of $\tilde{P}_1$ on $(0, 1]$. Following the terminology of Engen [132], $\tilde{\nu}$ is called the *structural distribution* associated with the random discrete distribution (P_i). Note that if a random partition Π_∞ is derived by sampling from

(P_i), then the size-biased permutation $(\tilde{P}_j)$ can be constructed as the sequence of class frequencies of Π_∞ in order of appearance. Then $\tilde{P}_1$ is the frequency of the class of Π_∞ that contains 1. It follows from (2.21) that for an arbitrary non-negative measurable function g,

$$\int \tilde{\nu}(dp)g(p) = \mathbb{E}[g(\tilde{P}_1)] = \mathbb{E}\left[\sum_i P_i g(P_i)\right]. \tag{2.22}$$

Hence, taking $g(p) = f(p)/p$, for arbitrary non-negative measurable function f there is the formula

$$\mathbb{E}\left[\sum_i f(P_i)\right] = \mathbb{E}\left[\frac{f(\tilde{P}_1)}{\tilde{P}_1}\right] = \int_0^1 \frac{f(p)}{p}\,\tilde{\nu}(dp). \tag{2.23}$$

Formula (2.23) shows that the structural distribution $\tilde{\nu}$ encodes much information about the entire sequence of random frequencies. Taking f in (2.23) to be the indicator of a subset B of $(0,1]$, formula (2.23) shows that the point process with a point at each $P_j \in (0,1]$ has mean intensity measure $p^{-1}\tilde{\nu}(dp)$. If $(P_i) = (P_i^{\downarrow})$ is in decreasing order, for $x > \frac{1}{2}$ there can be at most one $P_i^{\downarrow} > x$, so the structural distribution $\tilde{\nu}$ determines the distribution ν of $P_1^{\downarrow} = \max_i \tilde{P}_i$ on $(\frac{1}{2}, 1]$ via the formula

$$\mathbb{P}(P_1^{\downarrow} > x) = \nu(x,1] = \int_{(x,1]} p^{-1}\tilde{\nu}(dp) \qquad (x > \tfrac{1}{2}). \tag{2.24}$$

Typically, formulas for $\mathbb{P}(P_1^{\downarrow} > x)$ get progressively more complicated on the intervals $(\frac{1}{3}, \frac{1}{2}]$, $(\frac{1}{4}, \frac{1}{3}], \cdots$. See e.g. [339, 371]. Note that by (2.14) for $k = 1$ and $n_1 = n$ and (2.23)

$$p(n) = \mathbb{E}\left[\sum_i P_i^n\right] = \mathbb{E}[\,\tilde{P}_1^{n-1}\,] = \mu(n-1) \qquad (n = 1, 2, \cdots), \tag{2.25}$$

where $\mu(q)$ is the qth moment of the distribution $\tilde{\nu}$ of $\tilde{P}_1$ on $(0,1]$. From (2.10), (2.11), and (2.25) the following values of the EPPF of an infinite exchangeable random partition Π_∞ are also determined by the first two moments of the structural distribution:

$$p(1,1) = 1-\mu(1);\;\; p(2,1) = \mu(1)-\mu(2);\;\; p(1,1,1) = 1-3\mu(1)+2\mu(2). \tag{2.26}$$

So the distribution of Π_3 on partitions of the set $\{1,2,3\}$ is determined by the first two moments of $\tilde{P}_1$. The distribution of Π_n is not determined for all n by the structural distribution (Exercise 2.3.4). But moments of the structural distribution play a key role in the description of a number of particular models for random partitions. See for instance [362, 170].

Exercises

2.3.1. (Improper frequencies) Show how to modify the results of this section to be valid also for exchangeable random partitions of the positive integers with improper frequencies. Show that formula (2.14) is false in the improper case. Find the patch for that formula, which is not so pretty. See for instance Kerov [244, equation (1.3.1)].

2.3.2. (Mean number of blocks) Engen [132]. For an infinite exchangeable partition (Π_n) with $\tilde{P}_1$ the frequency of the block containing 1,

$$\mathbb{E}(|\Pi_n|) = \mathbb{E}[k_n(\tilde{P}_1)], \tag{2.27}$$

where $k_n(v) := (1-(1-v)^n)/v$ is a polynomial of degree $n-1$.

2.3.3. (Proper frequencies) [350] For an infinite exchangeable partition (Π_n) with frequencies $\tilde{P}_i$, the frequencies are proper, meaning $\sum_i \tilde{P}_i = 1$ almost surely, iff $\mathbb{P}(\tilde{P}_1 > 0) = 1$, and also iff $|\Pi_n|/n \to 0$ almost surely.

2.3.4. (The structural distribution does not determine the distribution of the infinite partition) Provide an appropriate example.

2.3.5. (Problem: characterization of structural distributions) What is a necessary and sufficient condition for a probability distribution F on $[0,1]$ to be a structural distribution? For some necessary and some sufficient conditions see [368].

2.4. Convergence

There are many natural combinatorial constructions of exchangeable random partitions Π_n of $[n]$ which are not consistent in distribution as n varies, so not immediately associated with an infinite exchangeable partition Π_∞. However, it is often the case that a sequence of combinatorially defined exchangeable partitions (Π_n) *converges in distribution as* $n \to \infty$ meaning that

$$\Pi_{m,n} \xrightarrow{d} \Pi_{m,\infty} \text{ for each fixed } m \text{ as } n \to \infty, \tag{2.28}$$

where $\Pi_{m,n}$ is the restriction to $[m]$ of Π_n, and $(\Pi_{m,\infty}, m = 1, 2 \ldots)$ is some sequence of limit random partitions. Let $p_n(n_1, \ldots, n_k)$ denote the EPPF of Π_n, defined as a function of compositions $(n_1, \ldots, n_k)$ of m for every $m \le n$, as discussed in Section 2.2. Then (2.28) means that for all integer compositions $(n_1, \ldots, n_k)$ of an arbitrary fixed m,

$$p_n(n_1, \ldots, n_k) \to p(n_1, \ldots, n_k) \text{ as } n \to \infty \tag{2.29}$$

for some limit function p. It is easily seen that any such limit p must be an infinite EPPF, meaning that the sequence of random partitions $\Pi_{m,\infty}$ in (2.28) can be constructed consistently to make an infinite exchangeable random partition

$\Pi_\infty := (\Pi_{m,\infty}, m = 1, 2, \ldots)$ whose EPPF is p. Let $(\tilde{P}_i)$ and $(P_i^\downarrow)$ denote the class frequencies of Π_∞, in order of appearance, and ranked order respectively. And let $(N_{n,i}, i \geq 1)$ and $(N_{n,i}^\downarrow, i \geq 1)$ denote the sizes of blocks of Π_n, in order of appearance and ranked order respectively, with padding by zeros to make infinite sequences. It follows from the continuity of Kingman's correspondence (Theorem 2.3) together with Proposition 2.1, and the obvious coupling between sampling with and without replacement for a sample of fixed size as the population size tends to ∞, that this notion (2.28)–(2.29) of convergence in distribution of Π_n to Π_∞ is further equivalent to

$$(N_{n,i}/n)_{i\geq 1} \xrightarrow{d} (\tilde{P}_i)_{i\geq 1} \tag{2.30}$$

meaning weak convergence of finite dimensional distributions, and similarly equivalent to

$$(N_{n,i}^\downarrow/n)_{i\geq 1} \xrightarrow{d} (P_i^\downarrow)_{i\geq 1} \tag{2.31}$$

in the same sense [173]. Let (U_i) be a sequence of independent and identically distributed uniform $(0, 1)$ variables independent of the Π_n. Another equivalent condition is that for each fixed $u \in [0, 1]$

$$\sum_{i=1}^{\infty} (N_{n,i}/n) 1(U_i \leq u) \xrightarrow{d} F(u) \tag{2.32}$$

for some random variable $F(u)$. According to Kallenberg's theory of processes with exchangeable increments [226], a limit process $(F(u), 0 \leq u \leq 1)$ can then be constructed as an increasing right-continuous process with exchangeable increments, with $F(0) = 0$ and $F(1) = 1$ a.s., and the convergence (2.32) then holds jointly as u varies, and in the sense of convergence in distribution on the Skorohod space $D[0, 1]$. To be more explicit,

$$F(u) = \sum_{i=1}^{\infty} P_i 1(U_i \leq u) + (1 - \Sigma_{i=1}^\infty P_i) u \tag{2.33}$$

where the U_i with uniform distribution on $[0, 1]$ are independent of the P_i, and either $(P_i) = (\tilde{P}_i)$ or $(P_i) = (P_i^\downarrow)$. The limit partition Π_∞ can then be generated by sampling from any random distribution such as F whose ranked atoms are distributed like $(P_i^\downarrow)$. The restriction $\Pi_{m,\infty}$ of Π_∞ to $[m]$ can then be generated for all $m = 1, 2, \ldots$ by sampling from F, meaning that i and j with $i, j \leq m$ lie in the same block of $\Pi_{m,\infty}$ iff $X_i = X_j$ where the X_i are random variables which conditionally given F are independent and identically distributed according to F:

$$\mathbb{P}(X_i \leq u \,|\, F) = F(u) \qquad (0 \leq u \leq 1).$$

This connects Kingman's theory of exchangeable random partitions to the theory of Bayesian statistical inference [350]. See also James [213, 212] for recent work in this vein.

2.5. Limits of Gibbs partitions

As an immediate consequence of (1.50), the decreasing arrangement of relative sizes of blocks of a $\text{Gibbs}_{[n]}(v_\bullet, w_\bullet)$ partition Π_n, say

$$(N^\downarrow_{n,1}/n, \ldots, N^\downarrow_{n,|\Pi_n|}/n) \tag{2.34}$$

has the same distribution as the decreasing sequence of order statistics of

$$(X_1/n, \ldots, X_K/n) \text{ given } S_K/n = 1$$

where the X_i have distribution (1.41) and K with distribution (1.42) is independent of the X_i, for some arbitrary $\xi > 0$ with $v(w(\xi)) < \infty$. Since the distribution of a $\text{Gibbs}_{[n]}(v_\bullet, w_\bullet)$ partition depends only on the v_j and w_j for $1 \le j \le n$, in this representation for fixed n the condition $v(w(\xi)) < \infty$ can always be arranged by setting $v_j = w_j = 0$ for $j > n$. It is well known [151, XVII.7] [230]that if $X_{n,1}, \ldots, X_{n,k_n}$ is for each n a sequence of independent and identically distributed variables of some non-random length k_n, with $k_n \to \infty$ as $n \to \infty$, then under appropriate conditions

$$\sum_{i=1}^{k_n} X_{n,i} \xrightarrow{d} T := \sum_{i=1}^{\infty} J_i^\downarrow$$

where $J_1^\downarrow \ge J_2^\downarrow \ge \ldots \ge 0$ are the points of a Poisson point process on $(0, \infty)$ with intensity measure $\Lambda(dx)$, for some Lévy measure Λ on $(0, \infty)$ with

$$\Psi(\lambda) := \int_0^\infty (1 - e^{-\lambda x})\Lambda(dx) < \infty \tag{2.35}$$

for all $\lambda > 0$. Then

$$\Lambda(x, \infty) = \lim_{n\to\infty} k_n \mathbb{P}(X_{n,1} > x)$$

for all continuity points x of Λ, and the Laplace transform of T is given by the Lévy-Khintchine formula

$$\mathbb{E}(e^{-\lambda T}) = \exp(-\Psi(\lambda)).$$

It is also known that if such a sum $\sum_{i=1}^{k_n} X_{n,i}$ has T as its limit in distribution as $n \to \infty$, then the convergence in distribution of $\sum_{i=1}^{k_n} X_{n,i}$ to T holds jointly with convergence in distribution of the k largest order statistics of the $X_{n,i}, 1 \le i \le k_n$ to the k largest points $J_1^\downarrow, \ldots, J_k^\downarrow$ of the Poisson process.

It is therefore to be anticipated that if a sequence of $\text{Gibbs}_{[n]}(v_\bullet, w_\bullet)$ partitions converges as $n \to \infty$ to some infinite partition Π_∞, where either $v_\bullet = v_\bullet^{(n)}$ or $w_\bullet = w_\bullet^{(n)}$ might be allowed to depend on n, and $v_\bullet^{(n)}$ is chosen to ensure that the distribution of the number of components K_n of Π_n grows to ∞ in a deterministic manner, say $K_n/k_n \to s > 0$ for some normalizing constants

k_n, then the distribution of ranked frequencies $(P_i^{\downarrow})$ of Π_∞ obtained from the convergence of finite-dimensional distributions

$$(N_{n,i}^{\downarrow}/n)_{i\geq 1} \xrightarrow{d} (P_i^{\downarrow})_{i\geq 1} \text{ with } K_n/k_n \to s \tag{2.36}$$

should be representable as

$$(P_i^{\downarrow})_{i\geq 1} \stackrel{d}{=} ((J_{s,i}^{\downarrow})_{i\geq 1} \,|\, T_s = 1) \tag{2.37}$$

for the ranked points $J_{s,i}^{\downarrow}$ of a Poisson point process with intensity $s\Lambda$, with $\sum_i J_{s,i}^{\downarrow} = T_s$. This Poisson process may be constructed as the jumps of $(T_u, 0 \leq u \leq s)$, where $(T_u, u \geq 0)$ is a subordinator with no drift and Lévy measure Λ. Then for $v_\bullet^{(n)}$ chosen so that K_n/k_n converges in distribution to S for some strictly positive random variable S, the limit law of $(P_i^{\downarrow})$ in (2.36) should be

$$\int_0^\infty \mathbb{P}((J_{s,i}^{\downarrow}) \in \cdot \,|\, T_s = 1)\mathbb{P}(S \in ds). \tag{2.38}$$

To make rigorous sense of this, it is first necessary to give a rigorous meaning to the law of $(J_{s,i}^{\downarrow})$ given $T_s = 1$, for instance by showing that for fixed s the law of $(J_{s,i}^{\downarrow})$ given $T_s = t$ can be constructed to be weakly continuous in t. Second, to justify weak convergence of conditional probability distributions it is necessary to establish an appropriate local limit theorem.

This program has been carried out in two cases of combinatorial significance. One case, treated in detail in [27], covers the class of logarithmic combinatorial assemblies:

Theorem 2.4. [189, 27] *Let $w_\bullet = (w_j)$ be a sequence of weights with*

$$w_j \sim \theta(j-1)!y^j \text{ as } j \to \infty$$

for some $\theta > 0$ and $y > 0$. Let Π_n be a Gibbs$_{[n]}(v_\bullet^{(n)}, w_\bullet)$ partition, either for $v_\bullet^{(n)} \equiv 1^\bullet$, or more generally for any array of weights $v_\bullet^{(n)}$ such that $|\Pi_n|/\log n$ converges in probability to θ as $n \to \infty$. Then Π_n converges in distribution to Π_∞ as $n \to \infty$, where Π_∞ is a $(0,\theta)$-partition with EPPF (2.19)*, whose Poisson-Dirichlet$(0,\theta)$ frequencies are the ranked jumps of a gamma process $(T_u, 0 \leq u \leq \theta)$ given $T_\theta = 1$.*

Sketch of proof. The case when $v_\bullet^{(n)} \equiv 1^\bullet$ can be read from the work of [27], where it is shown that in this case $|\Pi_n|/\log n$ converges in probability to θ as $n \to \infty$. The extension to more general $v_\bullet^{(n)}$ is quite straightforward. □

Two cases of Theorem 2.4 of special interest, discussed further in following chapters, are

- Π_n generated by the cycles of a uniform random permutation of $[n]$, when $w_j = (j-1)!, y = 1, \theta = 1$;

- Π_n generated by the basins of a uniform random mapping of $[n]$, with $w_j = (j-1)! \sum_{i=0}^{j-1} j^i/i!$ as in (1.61), $y = e, \theta = \frac{1}{2}$.

See [27] for many more examples. Note that mixtures over θ of $(0,\theta)$ partitions could arise by suitable choice of $v_\bullet^{(n)}$ so that $|\Pi_n|/\log n$ had a non-degenerate limit distribution, but this phenomenon does not seem to arise naturally in combinatorial examples.

Another case, treated by Pavlov [335, 336, 337], and Aldous-Pitman [17] covers a large number of examples involving random forests, where the limit involves the stable subordinator of index $\frac{1}{2}$. A more general result, where the limit partition is derived from a stable subordinator of index α for $\alpha \in (0,1)$, can be formulated as follows:

Theorem 2.5. *Let $w_\bullet = (w_j)$ be a sequence of weights with exponential generating function $w(\xi) := \sum_{j=1}^{\infty} \xi^j w_j / j!$ such that $w(\xi) = 1$ for some $\xi > 0$. Let $(p_j, j = 1, 2, \ldots)$ be the probability distribution defined by $p_j = \xi^j w_j / j!$ for ξ with $w(\xi) = 1$, and suppose that*

$$\sum_{j=i}^{\infty} p_j \sim \frac{c\, i^{-\alpha}}{\Gamma(1-\alpha)} \text{ as } i \to \infty \tag{2.39}$$

for some $\alpha \in (0,1)$. Let Π_n be a Gibbs$_{[n]}(v_\bullet^{(n)}, w_\bullet)$ partition, for any array of weights $v_\bullet^{(n)}$ such that $|\Pi_n|/n^\alpha$ converges in probability to s as $n \to \infty$. Then Π_n converges in distribution to Π_∞ as $n \to \infty$, where Π_∞ has ranked frequencies distributed like the

$$\text{ranked jumps of } (T_u, 0 \le u \le cs) \text{ given } T_{cs} = 1, \tag{2.40}$$

where $(T_u, u \ge 0)$ is the stable subordinator of index α with $\mathbb{E}\exp(-\lambda T_u) = \exp(-u\lambda^\alpha)$.

Sketch of proof. This was argued in some detail in [17] for the particular weight sequence $w_j = j^{j-1}$, corresponding to blocks with an internal structure specified by a rooted labeled tree. Then $\xi = e^{-1}$, $\alpha = \frac{1}{2}$, and the limiting partition can also be described in terms of the lengths of excursions of a Brownian motion or Brownian bridge, as discussed in Section 4.4 . The proof of the result stated above follows the same lines, appealing to the well known criterion for convergence to a stable law, and the local limit theorem of Ibragimov-Linnik [204]. □

In Section 4.3 the limiting partition Π_∞ appearing in Theorem 2.5 is called an $(\alpha|cs)$ partition. Mixtures of these distributions, obtained by randomizing s for fixed α, arise naturally in a number of different ways, as shown in Chapter 4.

Exercises

2.5.1. (Problem: Characterizing all weak limits of Gibbs partitions) Intuitively, the above discussion suggests that the only possible weak limits of Gibbs partitions are partitions whose ranked frequencies are mixtures over s of the law of ranked jumps of some subordinator $(T_u, 0 \leq u \leq s)$ given $T_s = 1$, allowing also the possibility of conditioning on the number of jumps in the compound Poisson case. Show that if the conditioning is well defined by some regularity of the distribution of T_s for all s, then such a partition can be achieved as a limit of Gibbs partitions, allowing both $v_\bullet$ and $w_\bullet$ to depend on n. But due to the difficulty in giving meaning to the conditioning when T_s does not have a density, it is not clear how to formulate a rigorous result. Can that be done? Does it make any difference whether or not $w_\bullet$ is allowed to depend on n?