Combinatorics, Probability and Computing (2004) 13, 185–193. © 2004 Cambridge University Press
DOI: 10.1017/S0963548303005996 Printed in the United Kingdom

Three Sampling Formulas

ALEXANDER V. GNEDIN

Department of Mathematics, Utrecht University, Postbus 80010, 3508 TA Utrecht, The Netherlands
(e-mail: gnedin@math.uu.nl)

Received 22 October 2002; revised 18 August 2003

Sampling formulas describe probability laws of exchangeable combinatorial structures like partitions and compositions. We give a brief account of two known parametric families of sampling formulas for compositions and add a new family to the list.

1. Introduction

By an integer composition of weight n and length ℓ we shall mean an ordered collection of positive integer parts $\lambda = (\lambda_1, \dots, \lambda_\ell)$; we write $\lambda \vdash n$ for $\sum \lambda_j = n$. It will also be convenient to use variables $\Lambda_j = \lambda_j + \cdots + \lambda_\ell$, $j \leqslant \ell$, so that $\lambda_j = \Lambda_j - \Lambda_{j+1}$ with the convention $\Lambda_{\ell+1} = 0$.

A *composition structure* is a nonnegative function f on compositions such that, for each n, the values $\{f(\lambda) : \lambda \vdash n\}$ comprise a probability distribution, say f_n, and the $\{f_n\}$ satisfy the following sampling consistency condition. Imagine an ordered series of randomly many nonempty boxes filled in randomly with balls, so that the distribution of occupancy numbers from left to right is f_n. The condition requires that, if some $k < n$ balls are uniformly at random, then the distribution of the reduced occupancy numbers in nonempty boxes (in the same order) must be exactly f_{n-k} (without loss of generality we can take $k = 1$).

Ignoring the order of boxes yields Kingman's partition structure [9] (see [1] and [12] for systematic development of the theory of partition structures, their relation to exchangeability and many references). But the relation cannot be uniquely inverted, because for a given partition structure there are many ways to introduce the order in a consistent fashion.

Gnedin [5] showed that all composition structures can be uniquely represented by a version of Kingman's paintbox construction [9]. Let U be a *paintbox* – a random open subset of $[0,1]$. With a paintbox we associate an ordered partition of $[0,1]$ comprising the intervals of U and the individual elements of the complement U^c, with the order of

https://doi.org/10.1017/S0963548303005996 Published online by Cambridge University Press

blocks induced by the order on reals. Suppose n independent uniform random points are sampled from $[0,1]$ independently of U. The sample points group somehow within the partition blocks and we obtain a random composition by writing the nonzero occupation numbers from left to right. With probability one there is no tie among the sample points and the consistency for various sample sizes follows from exchangeability in the sample.

From a topological viewpoint, the representation establishes a homeomorphism between the space of extreme composition structures and the compact space of open subsets of $[0,1]$ (endowed with a weak topology), and also identifies the generic composition structure with a unique mixture of extremes. Thus already the set of extremes is intrinsically infinite-dimensional, not to mention the mixtures. It is therefore a question of interest to find smaller parametric families which admit a reasonably simple description.

In this note we discuss briefly three such families: one is an ordered modification, due to Donnelly and Joyce [3], of the ubiquitous Ewens sampling formula (corresponding to the $(0,\theta)$-partition structure from the Ewens–Pitman two-parametric family [11]); another one, due to Pitman [10], is an ordered (symmetric) version of the (α,α)-partition structure; and the third composition structure is new. Despite the fact that the new composition structure is, in a sense, constructed from the beta distributions, like the first two, the corresponding partition structure does not fit into the Ewens–Pitman family. All three belong to a large infinite-dimensional family of regenerative compositions introduced and characterized in [8] and all three are mixed, *i.e.*, generated by genuinely random paintboxes.

2. Ordered ESF

In their encyclopaedic exposition of the multivariate Ewens distribution, Ewens and Tavaré presented the ordered version of ESF (see [4, equation 41.6]), that is,

$$e(\lambda) = \frac{\theta^{\ell}\, n!}{[\theta]_n} \prod_{j=1}^{\ell} \frac{1}{\Lambda_j} \qquad \theta > 0, \tag{2.1}$$

in connection with a size-biased permutation of the Ewens partition structure (henceforth, $[\theta]_n = \theta(\theta+1)\cdots(\theta+n-1)$ is the Pochhammer factorial). The special case $\theta = 1$ is well known to combinatorialists as the distribution of cycle lengths in a uniform random permutation, provided the cycles are ordered by increasing their least elements.

Donnelly and Joyce [3] observed that the formula also defines a composition structure, *i.e.*, that (2.1) determines a consistent sequence of random exchangeable partitions taken together with an *intrinsic* ordering of blocks, based neither on the sizes of blocks nor on labelling of 'balls in boxes'. They argued that the ordered structure is of some significance for biological applications, and proved the following paintbox representation of e.

Let Z_j be independent random variables with beta density

$$\omega(\mathrm{d}z) = \theta\,(1-z)^{\theta-1}\,\mathrm{d}z \qquad 0 < z < 1.$$

Let U_e be the open set complementary to the stick-breaking sequence

$$1 - \prod_{j=1}^{k}(1-Z_j) \qquad k = 1,2,\ldots$$

https://doi.org/10.1017/S0963548303005996 Published online by Cambridge University Press

taken together with the endpoints of $[0,1]$. Then, rephrasing Theorem 10 from [3], we have the following result.

Theorem 2.1. *The composition structure e can be derived from the paintbox U_e.*

The proof of this result given in [3] relied on the twin fact about weak convergence of the size-biased permutation of ESF. Next is a direct argument which offers some more insight and exemplifies the approach taken in this paper.

Proof. Introduce the binomial moments

$$\begin{aligned} w(n:m) &= \binom{n}{m} \int_0^1 z^m (1-z)^{n-m}\, \omega(\mathrm{d}z) \\ &= \theta \binom{n}{m} \mathrm{B}(m+1, n-m+\theta). \end{aligned} \tag{2.2}$$

For I the leftmost interval of U_e (adjacent to 0), the size of I equals Z_1. Letting $\hat{e}$ denote the composition structure derived from U_e, we aim to show that $\hat{e} = e$.

The argument is based on two facts. Firstly, suppose n uniform points have been sampled from $[0,1]$ and I happened to contain m sample points. Then, conditionally given I, the configuration of other $n-m$ points is as if it were a uniform sample from I^c. The second fact is that, given I, the set $U_e \setminus I$ is a scaled distributional copy of U_e, as is clear from the definition of the paintbox via stick-breaking.

Suppose composition $(\lambda_1, \ldots, \lambda_\ell)$ appears. In this case the leftmost of the intervals of U_e induced by the sample contains exactly λ_1 sample points, and this is only possible if the number of sample points in I equals either λ_1 or 0. The chance that λ_1 points fall in I is $w(n:\lambda_1)$ and in this event the composition derived from the piece of U_e to the right from I must be $(\lambda_2, \ldots, \lambda_\ell)$. If none of the points falls in I then all n points group within $U_e \setminus I$ in accordance with λ. Combining these events we get the equation

$$\hat{e}(\lambda_1, \ldots, \lambda_\ell) = w(n:\lambda_1)\, \hat{e}(\lambda_2, \ldots, \lambda_\ell) + w(n:0)\, \hat{e}(\lambda_1, \ldots, \lambda_\ell),$$

leading to the recursion

$$\hat{e}(\lambda_1, \ldots, \lambda_\ell) = \frac{w(n:\lambda_1)}{1 - w(n:0)}\, \hat{e}(\lambda_2, \ldots, \lambda_\ell),$$

which is solved by

$$\hat{e}(\lambda) = \prod_{j=1}^{\ell} q(\Lambda_j : \lambda_j), \tag{2.3}$$

where

$$\begin{aligned} q(n:m) &:= \frac{w(n:m)}{1 - w(n:0)} \\ &= \frac{\theta}{n} \frac{n!}{(n-m)!} \frac{[\theta]_{n-m}}{[\theta]_n}. \end{aligned} \tag{2.4}$$

Plugging this into (2.3) and simplifying, we arrive at (2.1), and thus $\hat{e} = e$. □

https://doi.org/10.1017/S0963548303005996 Published online by Cambridge University Press

There is a canonical correspondence between composition structures and probability distributions of exchangeable compositions of an infinite set $\{\underline{1},\underline{2},\ldots\}$ (see [5]). In terms of the paintbox representation, the composition derived from U is obtained by sampling infinitely many uniform points and then assigning objects $\underline{i}$ and $\underline{j}$ to distinct blocks if the closed interval spanned on the ith and the jth sample points has a nonempty intersection with U^c.

The infinite composition associated with e, say $\mathscr{E}$, has a simply ordered collection of blocks, and the law of large numbers says that the asymptotic frequencies of the blocks (in a growing sample) coincide with the sizes of stick-breaking residuals, from left to right. When we view $\mathscr{E}$ from the perspective of restrictions $\mathscr{E}_n$ on finite sets $\{\underline{1},\ldots,\underline{n}\}$, the collection of blocks *stabilizes* with probability one. As n grows, new blocks, which always start as singletons, will eventually appear only in the 'tail' of the sequence of blocks of $\mathscr{E}_n$, rather than in the 'head' of any given length. Saying this more formally, for any fixed k and n sufficiently large, each of the first k blocks of $\mathscr{E}_{n'}$, for $n' > n$, will have a representative in the analogous block of $\mathscr{E}_n$.

Compositions with the simply ordered collection of blocks, whose asymptotic frequencies are necessarily positive and sum up to 1, are the 'representable' compositions studied in [3]. A family of such compositions generated by the general stick-breaking paintbox was characterized in [8] and further studied in [7].

3. PSF

Pitman's composition structure is given by [10, equation (30)]:

$$p(\lambda) = \frac{n!\,\alpha^{\ell}}{[\alpha]_n} \prod_{j=1}^{\ell} \frac{[1-\alpha]_{\lambda_j - 1}}{\lambda_j!} \qquad 0 < \alpha < 1. \tag{3.1}$$

This sampling formula was derived from the following paintbox representation.

Theorem 3.1. *The paintbox U_p for p is the union of excursion intervals of the Bessel bridge of dimension $2-2\alpha$.*

Equivalently, the complement U_p^c is the set of zeros of the Bessel bridge on $[0,1]$. The case $\alpha = 1/2$ corresponds to the Brownian bridge.

In fact, p is a conditional version of another Pitman's composition structure p' derived from the paintbox complementary to the set of zeros of a Bessel process (which has a final meander interval adjacent to the right point of $[0,1]$). Pitman obtained a formula for p' akin to (3.1) (see [10, equation (28)]) using self-similarity of the Bessel process and distribution of the length of the meander interval. The partition structure corresponding to p' is of the Ewens–Pitman type $(\alpha, 0)$, and a relation between the two composition structures is that

$$p(\lambda) = \mathrm{const}(n)\, p'(\lambda, 1) \qquad \lambda \vdash n,$$

where $(\lambda, 1) = (\lambda_1, \ldots, \lambda_\ell, 1)$ for $\lambda = (\lambda_1, \ldots, \lambda_\ell)$.

https://doi.org/10.1017/S0963548303005996 Published online by Cambridge University Press

Gnedin and Pitman [8] give a characterization of p related to the observation that this composition structure is also of the product form (similar to (2.3)) with

$$q(n:m) = -\frac{\binom{\alpha}{m}\binom{-\alpha}{n-m}}{\binom{-\alpha}{n}} = -\binom{n}{m}\frac{[\alpha-m+1]_m}{[-\alpha-n+1]_m}.$$

For ℓ fixed, p is a symmetric function of the parts. This is reflected in that U_p is *symmetric*, that is, U_p has component intervals 'in random order' (in [1] the open sets with this kind of invariance are called 'exchangeable interval partitions'). Summing $p(\lambda)$ over distinct permutations of parts yields a function on integer partitions which is the (α,α)-partition structure from the Ewens–Pitman family. It follows that p could be obtained from the partition structure by permuting the parts in uniform random order (this is the general device allowing us to derive symmetric composition structures and symmetric open sets from their unordered relatives [6]).

Blocks of the Pitman's composition $\mathscr{P}$ of an infinite set $\{\underline{1},\underline{2},\dots,\}$ are ordered like the set of rational numbers and as such cannot be labelled by integers consistently with their intrinsic order. This happens each time a composition has infinitely many blocks (almost surely) and is symmetric. A consequence is that the infinite composition $\mathscr{P}$ has no definite first, second, *etc.*, or last block: in particular, the first (hence kth) block in $\mathscr{P}_n$ does not stabilize as n grows, and the total frequency of the k first blocks approaches zero as n grows.

4. A new sampling formula

Here is a new composition structure,

$$g(\lambda) = \frac{n!}{[\theta]_n}\prod_{j=1}^{\ell}\frac{1}{\lambda_j\, h_\theta(\Lambda_j)} \qquad \theta>0, \tag{4.1}$$

where

$$h_\theta(n) = \sum_{k=1}^{n}\frac{1}{\theta+k-1}$$

are the generalized harmonic numbers which coincide with the partial sums of the harmonic series when $\theta=1$.

To explain the genesis of the formula consider stick-breaking with the general beta density

$$\omega(\mathrm{d}z) = \text{const}\cdot\, z^{\alpha-1}(1-z)^{\theta-1}\,\mathrm{d}z \qquad \alpha,\theta>0. \tag{4.2}$$

The resulting paintbox generates a composition structure given by the right-hand side of (2.3) with

$$q(n:m) = \binom{n}{m}\frac{[\alpha]_m\,[\theta]_{n-m}}{[\alpha+\theta]_n-[\theta]_n}, \tag{4.3}$$

where (4.3) is obtained like (2.4) from the binomial moments of the beta density (4.2) (to see this just follow the proof of Theorem 2.1). Parameters α,θ in (4.2) should not be confused with twin parameters of the Ewens–Pitman family of partition structures.

https://doi.org/10.1017/S0963548303005996 Published online by Cambridge University Press

For general α and θ the induced composition structure cannot be expressed by a simple product formula, because the denominator has no good factorization. One notable exception is the ESF appearing when $\alpha = 1$. Another exception is the case $\alpha = 0$ giving rise to g; but this should be interpreted properly because measure ω becomes infinite.

Theorem 4.1. *When $\alpha \downarrow 0$ the stick-breaking composition structure directed by the beta density* (4.2) *converges to g.*

Proof. Expansions in powers of α start with

$$[\alpha]_m = \alpha\,(m-1)! + \cdots, \qquad [\alpha+\theta]_n - [\theta]_n = \alpha\, h_\theta(n)[\theta]_n + \cdots.$$

Therefore, when α approaches 0, we get

$$q(n:m) = \frac{n!}{(n-m)!}\,\frac{[\theta]_{n-m}}{[\theta]_n}\,\frac{1}{m\,h_\theta(n)}. \tag{4.4}$$

which yields g by computing the product as in (2.3). □

For n fixed, (4.4) is the distribution of the first part of a random composition of weight n. The sampling formula (4.1) becomes especially simple for $\theta = 1$ when $q(n:\cdot)$ gives a weight proportional to m^{-1} to each $m = 1, \ldots, n$.

To determine the paintbox representation for g we will extend the classical stick-breaking procedure by embedding the process into continuous time, and allowing infinitely many breaks within any time interval. Note that defining a composition structure via the right-hand side of (2.3), through the binomial moments of some measure ω and

$$q(n:m) := \frac{w(n:m)}{w(n:1) + \cdots + w(n:n)},$$

we need not require that the measure ω be finite, and need only impose the condition

$$\int_0^1 z\,\omega(\mathrm{d}z) < \infty$$

to have all binomial moments finite for $1 \leqslant m \leqslant n < \infty$.

In particular, our g appears when we take improper density

$$\omega(\mathrm{d}z) = z^{-1}(1-z)^{\theta-1}\,\mathrm{d}z \tag{4.5}$$

(see [8] for more examples). For this ω consider a Poisson point process (PPP) in the infinite strip $[0,\infty] \times [0,1]$ with Lebesgue $\times\,\omega$ as intensity measure. The PPP has countably many atoms (τ_j, ξ_j) (we adopt the conventional fake labelling of atoms, which is not intended to say that τ_j or ξ_j is a definite random variable for particular j), and each location on the abscissa is a concentration point for the set of atoms. Define a pure-jump process with increasing right-continuous paths

$$S_t = 1 - \prod_{(\tau_j,\xi_j):\,\tau_j<t} (1-\xi_j),$$

where the product is over all PPP atoms to the left from t. For any t the product converges because $z\,\omega(\mathrm{d}z)$ is a finite measure. The process (S_t) is a *geometric subordinator*:

https://doi.org/10.1017/S0963548303005996 Published online by Cambridge University Press

for $t' > t$ the ratio $(1 - S_{t'})/(1 - S_t)$ is independent of the partial path on $[0, t]$ and has same distribution as $1 - S_{t'-t}$.

(The reader feeling more comfortable breaking sticks from right to left should translate paintbox formulas using the involution $z \leftrightarrow 1 - z$ and also mirror the sampling formulas.)

Theorem 4.2. *The paintbox U_g representing g is the complement to the closure of the random set $\{S_t : t > 0\}$, which is the range of the geometric subordinator.*

The result can be derived from known facts about conventional subordinators (positive Lévy processes) and we refer to [8], where this approach is taken. To keep exposition of the present paper self-contained we shall give an independent fairly general argument, which relies solely on a variant of the Lévy–Khintchine formula for the Mellin transform of a geometric subordinator with zero drift:

$$\mathbf{E}(1 - S_t)^v = \exp(-t\, W(v)), \qquad v > 0, \tag{4.6}$$

where

$$W(v) := \int_0^1 (1 - (1 - z)^v)\, \omega(\mathrm{d}z).$$

The formula follows easily from its classical counterpart; see [2, p. 72]. Almost tautologically, for $\beta(v)$ a random variable independent of (S_t) with distribution Beta$(1, v)$, $v > 0$, formula (4.6) reads as

$$\mathbf{P}(S_t < \beta(v)) = \exp(-t\, W(v))$$

because $\mathbf{P}(\beta(v) > z) = (1 - z)^v$.

Proof. Fix $\lambda \vdash n$ and consider a uniform sample of size n. We use notation $S_\bullet = (S_t)$ and label sample points $\epsilon_1, \ldots, \epsilon_n$ in increasing order. The composition λ appears when for some t a component interval of $U_g \cap [0, S_t]$ contains exactly λ_1 first sample points and the composition on the remaining $(n - \lambda_1)$ points is $(\lambda_2, \ldots, \lambda_\ell)$. From the properties of uniform distribution and because PPP is ruled by a product measure, it follows that the composition structure induced by U_g is of the product form as in (2.3) and we only need to justify formula (4.4) for the distribution of the first part of composition of n.

To that end, let A be the event that there exists $t > 0$ such that $S_{t-} < \epsilon_1$ and $S_t \in [\epsilon_m, \epsilon_{m+1}]$ (with the convention that $\epsilon_{n+1} = 1$). If such a t does exist it is unique, and we say that A occurs at time t. Defining $\pi(t)$ to be the probability that A occurs *after* t, we aim to determine $\pi(0) = q(n : m)$.

For $t, t' > 0$, if event A occurs after time $t + t'$ then, of course, $S_t < \epsilon_1$, but conditionally given S_t and $S_t < \epsilon_1$, the subordinator after t is a scaled copy of $S_\bullet$ and $\epsilon_1, \ldots, \epsilon_n$ are distributed as n uniform order statistics in $[S_t, 1]$; thus, by (4.6),

$$\pi(t + t') = \mathbf{P}(S_t < \epsilon_1)\, \pi(t') = \exp(-t\, W(n))\, \pi(t'),$$

https://doi.org/10.1017/S0963548303005996 Published online by Cambridge University Press

since distribution of ϵ_1 is Beta$(1,n)$. It is not hard to see that π is continuous, and thus we can send $t' \downarrow 0$ to obtain

$$\pi(t) = \exp(-t\, W(n))\, \pi(0).$$

Differentiating at 0 yields $-\pi'(0)/W(n) = \pi(0)$, and thus the problem is reduced to computing the derivative $\pi'(0)$.

Let us compute the infinitesimal probability that A occurs before t for $t \to 0$. We write $\doteq$ for 'equal up to $o(t)$ as $t \to 0$'. First of all, $\mathbf{P}(S_t > \epsilon_1) \doteq t\, W(n)$, by expanding (4.6), and this implies that the probability that *two* or more jump intervals of $S_\bullet$ on $[0,t]$ cover some sample points is $o(t)$. It follows that

$$\pi(0) - \pi(t) \doteq \mathbf{P}(S_t \in [\epsilon_m, \epsilon_{m+1}]).$$

Now, since the sample is uniform we can write

$$\mathbf{P}(S_t \in [\epsilon_m,\, \epsilon_{m+1}]) = \int_0^1 \binom{n}{m} z^m (1-z)^{n-m}\, \mathbf{P}(S_t \in \mathrm{d}z),$$

and substituting $z^m = (1-(1-z))^m$ and expanding in powers of $(1-z)$, we see that this is equal to

$$\int_0^1 \binom{n}{m} \sum_{j=0}^{m} (-1)^j \binom{m}{j} (1-z)^{n-m+j}\, \mathbf{P}(S_t \in \mathrm{d}z)$$

which, by (4.6), is equal to

$$\binom{n}{m} \sum_{j=0}^{m} (-1)^j \binom{m}{j} \exp(-t\, W(n-m+j))$$

and for small t becomes

$$\mathbf{P}(S_t \in [\epsilon_m,\, \epsilon_{m+1}]) \doteq \binom{n}{m} \sum_{j=0}^{m} (-1)^j \binom{m}{j} (1 - t\, W(n-m+j)) = w(n:m)\, t,$$

by application of the formulas

$$W(n) = \sum_{m=1}^{n} w(n:m), \quad w(n:m) = \binom{n}{m} \sum_{j=0}^{m} (-1)^{j+1} \binom{m}{j} W(n-m+j),$$

$$\sum_{j=0}^{m} (-1)^j \binom{m}{j} = 0.$$

Finally, $-\pi'(0) = w(n:m)$ and therefore $\pi(0) = w(n:m)/W(n)$, as wanted. □

The infinite composition $\mathscr{G}$ associated with g has infinitely many blocks, and the set of blocks is order isomorphic to the set of rational numbers. Unlike Pitman's $\mathscr{P}$ it is not symmetric, *i.e.*, g is not invariant under permutation of parts λ_j when $\ell > 1$, and the representing paintbox U_g is not an 'exchangeable interval partition'. A combinatorialist might find it natural to view g as a function on Young diagrams $(\Lambda_1, \ldots, \Lambda_\ell)$ with strictly decreasing parts.

https://doi.org/10.1017/S0963548303005996 Published online by Cambridge University Press

Ignoring the order in $\mathscr{G}$ yields a novel partition structure. This partition does not belong to the Ewens–Pitman two-parameter family for any value of θ, and that family covered practically all explicit sampling formulas known to date. The distinction can be seen immediately by comparing the probability of one-block partition, our $g(n) = q(n : n)$ given by (4.4), versus the analogous quantity computed via [11, equation (16)] (the formulas do not match for $n > 4$, whatever the values of parameters).

Taking other integer values of α in (4.2) leads to formulas involving products of stereotypic polynomial factors, *e.g.*, for $\alpha = 2$ we obtain yet another composition structure

$$g_2(\lambda) = \frac{n!\,\theta^{\ell}\,(1+\theta)^{\ell}}{[\theta]_n} \prod_{j=1}^{\ell} \frac{\lambda_j + 1}{\Lambda_j(\Lambda_j + 2\theta + 1)}.$$

The resulting infinite compositions have simply ordered blocks and thus are more in line with $\mathscr{E}$.

Acknowledgement

The author is indebted to a referee for numerous improvements and helpful remarks, and especially for the advice to include a complete proof of Theorem 4.2.

References

[1] Aldous, D. J. (1985) Exchangeability and related topics. In *École d'Éte de Probabilités de Saint Flour XII*, Vol. 1117 of *Lecture Notes in Mathematics*, Springer, New York.
[2] Bertoin, J. (1996) *Lévy Processes*, Cambridge University Press.
[3] Donnelly, P. and Joyce, P. (1991) Consistent ordered sampling distributions: Characterization and convergence. *Adv. Appl. Probab.* **23** 229–258.
[4] Ewens, W. J. and Tavare, S. (1995) The Ewens sampling formula. In *Multivariate Discrete Distributions* (N. S. Johnson, *et al.*, eds), Wiley, New York.
[5] Gnedin, A. V. (1997) The representation of composition structures. *Ann. Probab.* **25** 1437–1450.
[6] Gnedin, A. V. (1998) On the Poisson–Dirichlet limit. *J. Multivariate Analysis* **67** 90–98.
[7] Gnedin, A. V. (2003) Bernoulli sieve. To appear in *Bernoulli.*
[8] Gnedin, A. V. and Pitman, J. (2003) Regenerative composition structures. UC Berkeley Dept. Stats. Technical Report No. 644.
[9] Kingman, J. F. C. (1978) The representation of partition structures. *J. London Math. Soc.* **18** 374–380.
[10] Pitman, J. (1997) Partition structures derived from Brownian motion and stable subordinators. *Bernoulli* **3** 79–96.
[11] Pitman, J. (1995) Exchangeable and partially exchangeable random partitions. *Probab. Theory Rel. Fields* **102** 145–158.
[12] Pitman, J. (2002) Combinatorial stochastic processes. Lecture Notes for St. Flour course, July 2002. UC Berkeley Dept. Stats. Technical Report No. 621.

https://doi.org/10.1017/S0963548303005996 Published online by Cambridge University Press