

# RESIDUAL FRACTIONS OF SIZE-BIASED PERMUTATIONS OF DISCRETE PRIOR ASSOCIATED WITH GIBBS PARTITIONS

Yamato, Hajime  
Kagoshima University : Emeritus Professor

<https://doi.org/10.5109/1434310>

---

出版情報 : Bulletin of informatics and cybernetics. 43, pp.41-52, 2011-12. Research Association  
of Statistical Sciences

バージョン :

権利関係 :

**RESIDUAL FRACTIONS OF SIZE-BIASED PERMUTATIONS OF  
DISCRETE PRIOR ASSOCIATED WITH GIBBS PARTITIONS**

by

**Hajime YAMATO**

---

*Reprinted from the Bulletin of Informatics and Cybernetics  
Research Association of Statistical Sciences, Vol.43*

---

FUKUOKA, JAPAN  
2011

# RESIDUAL FRACTIONS OF SIZE-BIASED PERMUTATIONS OF DISCRETE PRIOR ASSOCIATED WITH GIBBS PARTITIONS

By

Hajime YAMATO\*

## Abstract

A prior distribution is considered over all discrete distributions on positive integers. The sample from this prior yields a random partition of integers. We consider the case in which the distribution of the random partition is described by the Gibbs form. We give the distributions of the residual fractions of residual allocation model based on the size-biased permutation of a prior distribution.

*Key Words and Phrases:* Prior distribution, Gibbs partition, residual allocation model, size-biased permutation.

## 1. Introduction

Let  $\mathbb{N} = \{1, 2, \dots\}$  and  $\mathbb{N}_n = \{1, 2, \dots, n\}$  for  $n \in \mathbb{N}$ . We denote a random discrete distribution on  $\mathbb{N}$  by  $\mathcal{P}$  and the random sample of size  $n$  from  $\mathcal{P}$  by  $X_1, X_2, \dots, X_n$ . Thus,  $X_1, X_2, \dots, X_n$  are independent and identically distributed according to the discrete distribution  $\mathcal{P}$  given  $\mathcal{P}$ . We say  $i$  and  $j$  are equivalent and write  $i \sim j$  if  $X_i = X_j$  for  $i, j \in \mathbb{N}_n$ . This equivalence relation  $\sim$  generates a random partition  $\Pi_n$  of  $\mathbb{N}_n$ . The number  $K_n$  of non empty subsets in  $\Pi_n$  is also a random variable. Now we denote the elements of  $\mathcal{P}$  by  $P_1, P_2, \dots$ . Since  $X_1, X_2, \dots, X_n$  are independent and identically distributed given  $\mathcal{P}$ , we have for each particular partition  $(A_1, A_2, \dots, A_k)$  of  $\mathbb{N}_n$  and  $n_i = |A_i| (i = 1, \dots, k)$

$$P(\Pi_n = (A_1, A_2, \dots, A_k), K_n = k) = \sum_{(j_1, \dots, j_k)} E \left[ \prod_{i=1}^k P_{j_i}^{n_i} \right]$$

where the sum is over all permutations of  $k$  positive integers. The right-hand side is symmetric function of positive integers  $n_1, n_2, \dots, n_k$ . We denote the right-hand side by  $p(n_1, n_2, \dots, n_k)$ , which is called exchangeable partition probability function (EPPF).

For the elements of  $\mathcal{P}$ ,  $P_1, P_2, \dots$ , their size-biased permutation is denoted by  $\tilde{P}_1, \tilde{P}_2, \dots$  (for the size-biased permutation, see Appendix 3.1.1).

Using the size-biased permutation, one representation of EPPF  $p$  is given by

$$p(n_1, \dots, n_k) = E \left[ \tilde{P}_1^{n_1-1} (1 - \tilde{P}_1) \tilde{P}_2^{n_2-1} \dots (1 - \tilde{P}_1 - \dots - \tilde{P}_{k-1}) \tilde{P}_k^{n_k-1} \right]. \quad (1)$$

---

\* Emeritus Kagoshima University, Take 3-32-1-708 Kagoshima 890-0045 Japan. tel +81-99-250-7536

We put

$$W_1 = \tilde{P}_1, \quad W_j = \frac{\tilde{P}_j}{1 - \tilde{P}_1 - \cdots - \tilde{P}_{j-1}} \quad (j = 2, 3, \dots), \quad (2)$$

then  $0 \leq W_1, W_2, \dots \leq 1$  and we can write

$$\tilde{P}_1 = W_1, \quad \tilde{P}_j = (1 - W_1) \cdots (1 - W_{j-1}) W_j \quad (j = 2, 3, \dots).$$

This is a residual allocation model, where the residual fractions  $W_1, W_2, \dots$  may not be independent. The function  $p(n_1, \dots, n_k)$  can be written as

$$p(n_1, \dots, n_k) = E[W_1^{n_1-1} (1 - W_1)^{n-n_1} W_2^{n_2-1} (1 - W_2)^{n-n_1-n_2} \times \cdots \\ \times W_{k-1}^{n_{k-1}-1} (1 - W_{k-1})^{n-n_1-\cdots-n_{k-1}} W_k^{n_k-1}]$$

(See, for example, Pitman (1995, 2003)).

The sequence of random partitions  $\Pi_n$  of  $\mathbb{N}_n$  ( $n \geq 1$ ), which are stated above, is consistent and exchangeable. That is, the distribution of the random partition of  $\mathbb{N}_n$  obtained from  $\Pi_{n+1}$  by discarding  $n+1$  is equal to the one of  $\Pi_n$  and the distribution of  $\Pi_n$  is invariant under permutation of  $\mathbb{N}_n$ , for  $n = 1, 2, \dots$ .

As a representation of EPPF  $p$ , we consider Gibbs form

$$p(n_1, \dots, n_k) = V_{n,k} \prod_{j=1}^k w_{n_j} \quad (3)$$

for positive integers  $n_1, \dots, n_k$  satisfying  $n_1 + \cdots + n_k = n$  and some nonnegative weights  $w_j$  and  $V_{n,k}$  ( $k = 1, \dots, n$ ).

By the consistency and exchangeability of  $\Pi_n$ ,  $n \geq 1$ , for  $j = 1, 2, \dots$ ,  $w_j$  of (3) is given by

$$w_j = (1 - \alpha)^{[j-1]} \quad \text{for } -\infty < \alpha < 1, \quad (4)$$

where  $x^{[r]} = x(x+1) \cdots (x+r-1)$ . The weights  $V_{1,1} = 1$  and  $V_{n,k}$  satisfies the backward recursion

$$V_{n,k} = (n - \alpha k) V_{n+1,k} + V_{n+1,k+1}, \quad (-\infty < \alpha < 1).$$

Or  $w_j = 1$  ( $j = 1, 2, \dots$ ),  $V_{1,1} = 1$  and  $V_{n,k}$  satisfies the backward recursion

$$V_{n,k} = k V_{n+1,k} + V_{n+1,k+1}$$

(Gnedin and Pitman (2006)). In the latter case of  $w_j = 1$  ( $j = 1, 2, \dots$ ), the EPPF  $p$  does not depend on the values of the  $n_1, \dots, n_k$ . Therefore we exclude this case and consider the case of  $-\infty < \alpha < 1$  hereafter.

If we take  $V_{n,k} = \theta^k / \theta^{[n]}$  ( $\theta > 0$ ), then EPPF  $p$  gives Ewens' sampling formula

$$p(n_1, \dots, n_k) = \frac{\theta^k}{\theta^{[n]}} \prod_{j=1}^k (n_j - 1)!. \quad (5)$$

$W_1, W_2, \dots$  are independent and have the same beta distribution  $Be(1, \theta)$ . The distribution of the corresponding  $\tilde{P}_1, \tilde{P}_2, \dots$  is well-known as GEM distribution (see, for

example, Johnson et al. (1997)).

If we take  $V_{n,k} = \theta^{[k;\alpha]}/\theta^{[n]}$  ( $0 \leq \alpha < 1$ ,  $\theta > -\alpha$ ), then EPPF  $p$  gives Pitman's sampling formula

$$p(n_1, \dots, n_k) = \frac{\theta^{[k;\alpha]}}{\theta^{[n]}} \prod_{j=1}^k (1-\alpha)^{[n_j-1]}, \quad (6)$$

where  $\theta^{[k;\alpha]} = \theta(\theta+\alpha)(\theta+2\alpha)\dots(\theta+(k-1)\alpha)$ .  $W_1, W_2, \dots$  are independent and each  $W_j$  has the beta distribution  $Be(1-\alpha, \theta+j\alpha)$ . The distribution of the corresponding  $\tilde{P}_1, \tilde{P}_2, \dots$  is known as the two-parameter GEM distribution (see, for example, Pitman (2003) and Yamato et al. (2001)).

The characterization of the Gibbs form (3) with (4) is given by Gnedin and Pitman (2006): Each exchangeable Gibbs partition of a fixed type  $\alpha \in (-\infty, 1)$  given by (3) with  $w_j = (1-\alpha)^{[j-1]}$  is a unique probability mixture of extreme partitions of this type, which are

- (I) PD( $\alpha, m|\alpha$ ) partitions with  $m = 0, 1, \dots, \infty$  for  $\alpha \in (-\infty, 0)$ ,
- (II) PD( $\theta$ ) partitions with  $\theta \in [0, \infty)$  for  $\alpha = 0$ ,
- (III) PK( $\rho_\alpha|t$ ) partitions with  $t \in [0, \infty)$  for  $\alpha \in (0, 1)$ .

PD( $\alpha, m|\alpha$ ) partitions of (I) is a random partition based on a sample from the  $m$ -dimensional Dirichlet distribution. PD( $\theta$ ) partitions of (II) is Ewens' sampling formula given by (5). The  $\rho_\alpha$  ( $0 < \alpha < 1$ ) of (III) is Lévy density  $\rho_\alpha(x) = \alpha x^{-\alpha-1}/\Gamma(1-\alpha)$  and Pitman's sampling formula (6) is a special case of (III). The matters related with PD and PK are explained in Appendix.

The purpose of this paper is to give the distributions of the residual fractions  $W_1, W_2, \dots$  associated with  $\tilde{P}_1, \tilde{P}_2, \dots$  which is given by (2) for the three cases of (I), (II) and (III), which are given in the next section.

## 2. Properties of Residual Fractions

We discuss the distribution of  $(W_1, W_2, \dots)$  given by (2) for the three cases of (I), (II) and (III) of Section 1.

**Case (I)** The Poisson-Dirichlet PD( $\alpha, m|\alpha$ ) is the symmetric Dirichlet distribution with parameter  $|\alpha|$  on the  $m$ -dimensional simplex  $\{(x_1, \dots, x_{m+1}) : x_i \geq 0, \sum_{i=1}^{m+1} x_i = 1\}$ . In case of  $m = 0$ , PD( $\alpha, m|\alpha$ ) degenerates to 1. Thus for  $m = 0$ , the residual fraction  $W_1$  is equal to 1. For  $m \geq 1$ ,  $W_1, W_2, \dots, W_{m+1}$  are independent,  $W_j$  ( $j = 1, \dots, m$ ) has the beta distribution  $Be(|\alpha|+1, (m-j+1)|\alpha|)$  and  $W_{m+1} = 1$  (see, Appendix 3.1.2). Under the mixing distribution  $\gamma$  on  $\mathbb{N}$ ,  $W_1, W_2, \dots$  take the following form;

- (i) with probability  $\gamma(0)$ ,  $W_1 = 1$
- (ii) with probability  $\gamma(m)$  ( $m = 2, 3, \dots$ ),  $W_1$  has  $Be(|\alpha|+1, m|\alpha|)$ ,  $W_2$  has beta  $(|\alpha|+1, (m-1)|\alpha|)$ , ...,  $W_m$  has  $Be(|\alpha|+1, |\alpha|)$  and  $W_{m+1} = 1$ .

We show the distributions of  $W_1, W_2, \dots$  by the form of table as follows.

Table: Distributions of  $W'$ s

$PD$	$W_1$	$W_2$	$W_3$	$W_4$	$\dots$	$Prob.$
$P( \alpha , 0 \alpha)$	1					$\gamma(0)$
$P( \alpha , 1 \alpha)$	$Be( \alpha  + 1,  \alpha )$	1				$\gamma(1)$
$P( \alpha , 2 \alpha)$	$Be( \alpha  + 1, 2 \alpha )$	$Be( \alpha  + 1,  \alpha )$	1			$\gamma(2)$
$P( \alpha , 3 \alpha)$	$Be( \alpha  + 1, 3 \alpha )$	$Be( \alpha  + 1, 2 \alpha )$	$Be( \alpha  + 1,  \alpha )$	1		$\gamma(3)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

The marginal distribution of  $W_1$  is as follows;  $W_1$  is equal to 1 with probability  $\gamma(0)$  and has the beta distribution  $Be(|\alpha| + 1, j|\alpha|)$  with probability  $\gamma(j)$  for  $j = 1, 2, \dots$ . That is, with the probability  $1 - \gamma(0)$ ,  $W_1$  has the density

$$g_{W_1}(w_1) = \sum_{j=1}^{\infty} \frac{w_1^{|\alpha|} (1 - w_1)^{j|\alpha| - 1}}{B(|\alpha| + 1, j|\alpha|)} \frac{\gamma(j)}{1 - \gamma(0)},$$

where  $B$  is the beta function such that  $B(\nu, \omega) = \int_0^1 u^{\nu-1} (1-u)^{\omega-1} du$ . For  $r \geq 2$ , the marginal distribution of  $W_r$  is as follows;  $W_r$  is equal to 1 with probability  $\gamma(r-1)$ , and has the beta distribution  $Be(|\alpha| + 1, (j-r+1)|\alpha|)$  with probability  $\gamma(j)$  for  $j = r, r+1, \dots$ . That is, with probability  $\gamma^*(r) = \gamma(r) + \gamma(r+1) + \dots$ ,  $W_r$  has the density

$$g_{W_r}(w_r) = \sum_{j=r}^{\infty} \frac{w_r^{|\alpha|} (1 - w_r)^{(j-r+1)|\alpha| - 1}}{B(|\alpha| + 1, (j-r+1)|\alpha|)} \frac{\gamma(j)}{\gamma^*(r)}.$$

With probability  $\gamma(0) + \dots + \gamma(r-1)$ ,  $W_r$  is not defined.

Especially, we consider the case in which  $\alpha = -1$  and  $\gamma$  is Poisson distribution with parameter  $\lambda (> 0)$ .  $W_1$  is equal to 1 with probability with  $e^{-\lambda}$ . With probability  $1 - e^{-\lambda}$ ,  $W_1$  has the density function

$$g_{W_1}(w_1) = \frac{[2 + \lambda(1 - w_1)]\lambda w_1 e^{-\lambda w_1}}{1 - e^{-\lambda}}, \quad 0 < w_1 < 1.$$

$W_2$  is equal to 1 with probability with  $\lambda e^{-\lambda}$ . With probability  $1 - (1 + \lambda)e^{-\lambda}$ ,  $W_2$  has the density function

$$g_{W_2}(w_2) = \frac{\lambda w_2 e^{-\lambda w_2}}{1 - (1 + \lambda)e^{-\lambda}}, \quad 0 < w_2 < 1,$$

and  $W_2$  is not defined with probability  $e^{-\lambda}$ .  $f_{W_2}^*(w_2)$  is a truncated Gamma distribution (Johnson (1994), p.380).  $W_3$  is equal to 1 with probability with  $\lambda^2 e^{-\lambda}/2$ . With probability  $1 - (1 + \lambda + \lambda^2/2)e^{-\lambda}$ ,  $W_3$  has the density function for  $0 < w_3 < 1$

$$g_{W_3}(w_3) = \frac{[\lambda^2(1 - w_3)^2 - 2\lambda(1 - w_3) + 2 - 2e^{-\lambda(1-w_3)}]w_3(1 - w_3)^{-3}e^{-\lambda w_3}}{1 - (1 + \lambda + \lambda^2/2)e^{-\lambda}},$$

and  $W_3$  is not defined with probability  $(1 + \lambda)e^{-\lambda}$ .

**Case (II)** In this case the distribution of  $(P_1, P_2, \dots)$  is given by mixing the Poisson-Dirichlet distribution  $PD(0, \theta)$  with the distribution  $\nu(\theta)$  on  $[0, \infty)$ . Since  $PD(\theta)$

partitions is Ewens' sampling formula, given  $\theta$ ,  $W_1, W_2, \dots$  are independent identically distributed with the beta distribution  $\beta(1, \theta)$ ,  $\theta > 0$ . The conditional distribution function of  $W_i$  given  $\theta$ ,  $G_W(t|\theta) = P(0 \leq W_i \leq t|\theta) = 1 - (1-t)^\theta$  ( $0 < t < 1$ ). Thus, in this case (II), the distribution function and density of  $W_i$  ( $n = 1, 2, \dots$ ) are given by

$$G_W(w) = 1 - \int_0^\infty (1-w)^\theta \nu(d\theta) \quad \text{and} \quad g_W(w) = \int_0^\infty \theta(1-w)^{\theta-1} \nu(d\theta) \quad (0 < w < 1) \quad (7)$$

Similarly, for  $0 < w_i < 1$  ( $i = 1, \dots, n$ ) the joint distribution function and density of  $W_1, \dots, W_n$  are given by

$$\begin{aligned} P(0 \leq W_1 \leq w_1, \dots, 0 \leq W_n \leq w_n) &= E_\theta [P(0 \leq W_1 \leq w_1|\theta) \cdots P(0 \leq W_n \leq w_n|\theta)] \\ &= E_\theta [(1 - (1 - w_1)^\theta) \cdots (1 - (1 - w_n)^\theta)] \\ &= \int_0^\infty (1 - (1 - w_1)^\theta) \cdots (1 - (1 - w_n)^\theta) \nu(d\theta) \end{aligned}$$

and

$$\begin{aligned} g_{W_1, \dots, W_n}(w_1, \dots, w_n) &= \int_0^\infty g_{W_1|\theta}(w_1|\theta) \cdots g_{W_n|\theta}(w_n|\theta) \nu(d\theta) \\ &= \int_0^\infty \theta^n [(1 - w_1) \cdots (1 - w_n)]^{\theta-1} \nu(d\theta), \end{aligned}$$

respectively. Thus we have

**PROPOSITION 2.1.** (Case (II)) *The distribution function and density of  $W_i$  ( $i = 1, 2, \dots$ ) are given by (10). For  $0 < w_i < 1$  ( $i = 1, \dots, n$ ), the joint distribution function and density of  $W_1, \dots, W_n$  are given by*

$$P(0 \leq W_1 \leq w_1, \dots, 0 \leq W_n \leq w_n) = \int_0^\infty (1 - (1 - w_1)^\theta) \cdots (1 - (1 - w_n)^\theta) \nu(d\theta),$$

and

$$g_{W_1, \dots, W_n}(w_1, \dots, w_n) = \int_0^\infty \theta^n [(1 - w_1) \cdots (1 - w_n)]^{\theta-1} \nu(d\theta),$$

respectively.

We note that these distributions of the residual fractions  $W$ 's gives also the distributions of residual fractions  $W$ 's associated with size-biased permutation of random proportions of mixtures of Dirichlet process.  $(P_1, P_2, \dots)$  having  $\text{PD}(0, \theta)$  ( $=\text{PD}(\theta)$ ) gives random probabilities of Dirichlet process with parameter  $\theta H$ ,  $H$  is any continuous distribution (Ferguson (1973)). Therefore, mixing  $\text{PD}(0, \theta)$  with respect to  $\theta$  by a distribution  $\nu(\theta)$  on  $[0, \infty)$  corresponds to mixing Dirichlet process of parameter  $\theta H$  with respect to the total mass  $\theta$  by a distribution  $\nu(\theta)$  on  $[0, \infty)$ , which yields the mixture of Dirichlet process (Antoniak (1974), Cerquetti (2008)). The random proportions of mixtures of Dirichlet process depend only on the total mass  $\theta$ . Thus, the distributions of  $W$ 's of the Proposition 2.1 gives the distributions of residual fractions  $W$ 's associated with size-biased permutation of random proportions of mixtures of Dirichlet process.

As an example, for the distribution of  $\theta$  we choose the gamma distribution having the density  $\nu(dx) = (x/b)^{c-1}e^{-x/b}/b\Gamma(c)dx$ , where  $x > 0$  and  $b, c > 0$ . Its Laplace transform is  $E_\theta e^{-\theta s} = 1/(1+bs)^c$ ,  $s > -1$ , where the expectation is taken with respect to the random variable  $\theta$ . Then, the distribution function of  $W_i$  is

$$G_W(t) = 1 - E_\theta e^{-\theta[-\log(1-t)]} = 1 - [1 - b\log(1-t)]^{-c}, \quad 0 < t < 1.$$

We have also

$$P(b\log(1-W) \leq x) = 1 - P(W \leq 1 - e^{x/b}) = (1-x)^{-c}, \quad x < 0.$$

The density function of  $W_i$  is

$$g_W(t) = \frac{bc}{[1 - b\log(1-t)]^{(c+1)}(1-t)}, \quad 0 < t < 1.$$

In case of  $b = c = 1$ , that is the exponential distribution  $e(1)$ , for  $0 < t < 1$

$$G_W(t) = 1 - [1 - \log(1-t)]^{-1}, \quad g_W(t) = \frac{1}{[1 - \log(1-t)]^2(1-t)}.$$

The joint distribution function of any  $W_i$  and  $W_j$  ( $i \neq j$ ) is for  $0 < t_1 < 1, 0 < t_2 < 1$

$$\begin{aligned} P(0 \leq W_i \leq t_1, 0 \leq W_j \leq t_2) &= E_\theta [P(0 \leq W_i \leq t_1 | \theta) P(0 \leq W_j \leq t_2 | \theta)] \\ &= E_\theta [(1 - (1 - t_1)^\theta)(1 - (1 - t_2)^\theta)] \\ &= 1 - \frac{1}{[1 - b\log(1 - t_1)]^c} - \frac{1}{[1 - b\log(1 - t_2)]^c} \\ &\quad + \frac{1}{[1 - b\log(1 - t_1)(1 - t_2)]^c}. \end{aligned}$$

For any  $j_1 \neq \dots \neq j_l$  ( $l = 2, 3, \dots$ ), we have also

$$\begin{aligned} P(b\log(1 - W_{j_1}) \leq x_1, \dots, b\log(1 - W_{j_l}) \leq x_l) \\ = E_\theta e^{-\theta(-[x_1 + \dots + x_l]/b)} = (1 - x_1 - \dots - x_l)^{-c}, \quad x_1, \dots, x_l < 0. \end{aligned}$$

**Case (III)** Here, we use the notations of Appendix 3.2. We suppose that  $T$  have stable density  $f_\alpha$  and the Lévy density is given by  $\rho_\alpha(x) = \alpha x^{-\alpha-1}/\Gamma(1-\alpha)$ . As the mixing distribution on  $[0, \infty)$ , we consider a continuous distribution  $\gamma$ . From (14) of Appendix, the density of  $W_1 (= \tilde{P}_1)$  is given by

$$g_{W_1}(w_1) = \int_0^\infty \frac{\alpha(w_1 t)^{-\alpha}}{\Gamma(1-\alpha)} \frac{f_\alpha((1-w_1)t)}{f_\alpha(t)} \gamma(dt).$$

In general, from (12) of Appendix, the joint density of  $(T, W_1, \dots, W_n)$  is given by

$$\begin{aligned} g_{T, W_1, \dots, W_n}(t, w_1, \dots, w_n) &= \left( \frac{\alpha}{\Gamma(1-\alpha)} \right)^n t^{-n\alpha} \\ &\quad \times f_\alpha((1-w_1) \dots (1-w_n)t) \times [w_1(1-w_1)^{n-1} w_2(1-w_2)^{n-2} \dots w_n]^{-\alpha}. \end{aligned}$$



Therefore, given  $T = t$ , the conditional density of  $(W_1, \dots, W_n)$  is given by

$$g_{W_1, \dots, W_n | T}(w_1, \dots, w_n | t) = \left( \frac{\alpha}{\Gamma(1 - \alpha)} \right)^n t^{-n\alpha} [w_1(1-w_1)^{n-1} w_2(1-w_2)^{n-2} \dots w_n]^{-\alpha} \\ \times \frac{f_\alpha((1-w_1) \dots (1-w_n)t)}{f_\alpha(t)}.$$

Thus we get the following.

**PROPOSITION 2.2.** *In case of Case (III), for  $n = 1, 2, \dots$ , the joint density of  $(W_1, \dots, W_n)$  is given by*

$$g_{W_1, \dots, W_n}(w_1, \dots, w_n) = \left( \frac{\alpha}{\Gamma(1 - \alpha)} \right)^n [w_1(1-w_1)^{n-1} w_2(1-w_2)^{n-2} \dots w_n]^{-\alpha} \\ \times \int_0^\infty t^{-n\alpha} \frac{f_\alpha((1-w_1) \dots (1-w_n)t)}{f_\alpha(t)} \gamma(dt). \quad (8)$$

**Example 1** (Pitman's sampling formula) As the mixing distribution, we take

$$\gamma(t) = C_{\alpha, \theta}^{-1} t^{-\theta} f_\alpha(t),$$

where  $C_{\alpha, \theta} = E_\alpha(T^{-\theta}) = \Gamma(\frac{\theta}{\alpha} + 1)/\Gamma(\theta + 1)$ . From this expectation, we get the relation  $E_\alpha(T^{-(\theta+2\alpha)}) = \Gamma(\frac{\theta+2\alpha}{\alpha} + 1)/\Gamma(\theta + 2\alpha + 1)$ . For the right-hand side of the integral (7) with  $n = 2$ , we do the change of variable  $x = (1-w_1)(1-w_2)t$  and use this relation. Thus, in this example, we get

$$g_{W_1, W_2}(w_1, w_2) = \frac{1}{B(1 - \alpha, \theta + \alpha)} w_1^{-\alpha} (1-w_1)^{\theta + \alpha - 1} \\ \times \frac{1}{B(1 - \alpha, \theta + 2\alpha)} w_1^{-\alpha} (1-w_1)^{\theta + 2\alpha - 1}$$

which shows that  $W_1$  has  $\beta(1 - \alpha, \theta + \alpha)$ , and  $W_2$  is independent of  $W_1$  and has  $\beta(1 - \alpha, \theta + 2\alpha)$ . In general, it is well-known that  $W_1, W_2, \dots$  are independent and  $W_j$  has  $\beta(1 - \alpha, \theta + j\alpha)$  ( $j = 1, 2, \dots$ ), as stated in Section 1.

**Example 2** As the mixing distribution, we consider

$$\gamma(t) = f_\alpha^\lambda(t) = f_\alpha(t) \exp\{\lambda^\alpha - \lambda t\}, \quad \lambda \geq 0. \quad (9)$$

From (8), in this example, the joint density of  $(W_1, \dots, W_n)$  ( $n = 1, 2, \dots$ ) is given by

$$g_{W_1, \dots, W_n}(w_1, \dots, w_n) = \left( \frac{\alpha}{\Gamma(1 - \alpha)} \right)^n e^{\lambda^\alpha} [w_1(1-w_1)^{n-1} w_2(1-w_2)^{n-2} \dots w_n]^{-\alpha} \\ \times \int_0^\infty t^{-n\alpha} e^{-\lambda t} f_\alpha((1-w_1) \dots (1-w_n)t) dt.$$

Especially, for  $\alpha = 1/2$ ,

$$f_{1/2}(t) = \frac{1}{2\sqrt{\pi}} t^{-\frac{3}{2}} e^{-\frac{1}{4t}}.$$

Using the Bessel function of the 3rd kind

$$K_\lambda(z) = \frac{1}{2} \left(\frac{z}{2}\right)^\lambda \int_0^\infty t^{-\lambda-1} \exp\left\{-t - \frac{z^2}{4t}\right\} dt \quad (z > 0)$$

we have

$$\begin{aligned} & g_{W_1, \dots, W_n}(w_1, \dots, w_n) \\ &= \frac{\lambda^{\frac{n+1}{4}}}{2^{\frac{n-1}{2}} \pi^{\frac{n+1}{2}}} e^{\lambda \frac{1}{2}} (\bar{w}_1 \cdots \bar{w}_n)^{-\frac{n+5}{4}} \left(\frac{w_1 \cdots w_n}{\bar{w}_1 \bar{w}_2^2 \cdots \bar{w}_n^n}\right)^{-\frac{1}{2}} K_{\frac{n+1}{2}} \left(\sqrt{\frac{\lambda}{\bar{w}_1 \cdots \bar{w}_n}}\right), \end{aligned}$$

where  $\bar{w} = 1 - w$ . Thus we have the following Proposition.

**PROPOSITION 2.3.** *We take the distribution  $\gamma$  given by (9) as the mixing distribution. Then, the joint density of the residual fractions  $W_1, W_2, \dots, W_n$  ( $n = 1, 2, \dots$ ) is given by*

$$\begin{aligned} g_{W_1, \dots, W_n}(w_1, \dots, w_n) &= \left(\frac{\alpha}{\Gamma(1-\alpha)}\right)^n e^{\lambda \alpha} [w_1(1-w_1)^{n-1} w_2(1-w_2)^{n-2} \cdots w_n]^{-\alpha} \\ &\quad \times \int_0^\infty t^{-n\alpha} e^{-\lambda t} f_\alpha((1-w_1) \cdots (1-w_n)t) dt. \end{aligned}$$

*Epecially, for  $\alpha = 1/2$ , we have*

$$\begin{aligned} & g_{W_1, \dots, W_n}(w_1, \dots, w_n) \\ &= \frac{\lambda^{\frac{n+1}{4}}}{2^{\frac{n-1}{2}} \pi^{\frac{n+1}{2}}} e^{\lambda \frac{1}{2}} (\bar{w}_1 \cdots \bar{w}_n)^{-\frac{n+5}{4}} \left(\frac{w_1 \cdots w_n}{\bar{w}_1 \bar{w}_2^2 \cdots \bar{w}_n^n}\right)^{-\frac{1}{2}} K_{\frac{n+1}{2}} \left(\sqrt{\frac{\lambda}{\bar{w}_1 \cdots \bar{w}_n}}\right). \quad (10) \end{aligned}$$

We note that (10) with  $n = 1$  gives

$$f(w_1) = \frac{\sqrt{\lambda}}{\pi} e^{\sqrt{\lambda}} \frac{1}{\sqrt{w_1(1-w_1)}} K_1 \left(\sqrt{\frac{\lambda}{1-w_1}}\right),$$

which is equal to the density of  $\tilde{P}_1$  given by (64) of Pitman (2003).

We state the relation between this proposition and the random probabilities obtained by normalizing the generalized Gamma process. For the mixing distribution  $\gamma(t)$ , it holds that

$$\text{PK}(\rho_\alpha, \gamma) = \text{PK}(\rho_\alpha^\lambda), \quad \rho_\alpha^\lambda(s) = \frac{\alpha}{\Gamma(1-\alpha)} s^{-1-\alpha} e^{-\lambda s}$$

(Pitman (2003) and Cerquetti (2007)). It means that Poisson-Kingman distribution  $\text{PK}(\rho_\alpha, \gamma)$  considered in Example 2 is equal to Poisson-Dirichlet distribution  $\text{PK}(\rho_\alpha^\lambda)$ . This distribution is given by the Poisson process with the Lévy density  $\rho_\alpha^\lambda$  and therefore corresponds to the jumps of the generalized Gamma process with the intensity  $\rho_\alpha^\lambda$  (Lijoi et al. (2008)).

Thus we know that for the random probabilities  $P$  obtained by normalizing the generalized Gamma process with the intensity  $\rho_\alpha^\lambda$ , the distributions of their residual fraction  $W_1, W_2, \dots$  are given by the above proposition. Especially, (10) gives the distributions of the residual fraction  $W_1, W_2, \dots$  for Normalized inverse Gaussian process (Lijoi et al. (2005)).

### 3. Appendix

#### 3.1. Size-biased permutation

We give the definition of size-biased permutation and explain size-biased permutation of symmetric Dirichlet distribution.

##### 3.1.1. Size-biased permutation

For the elements of  $\mathcal{P}$ ,  $P_1, P_2, \dots$ , we consider their size-biased permutation by  $\tilde{P}_1, \tilde{P}_2, \dots$ . Let  $N_1, N_2, \dots$  be a sequence of random variables which take the value in  $\mathbb{N} \cup \{\infty\}$  and satisfy the following property:

$$P(N_1 = i | \mathcal{P} = (P_1, P_2, \dots)) = P_i, \quad i = 1, 2, \dots.$$

For any  $r \in \mathbb{N}$  and distinct  $j_1, \dots, j_r \in \mathbb{N}$ , if  $P_1, P_2, \dots$  have at least  $r$  positive elements,

$$P(N_r = j_r | \mathcal{P} = (P_1, P_2, \dots), N_1 = j_1, \dots, N_{r-1} = j_{r-1}) = \frac{P_{j_r}}{1 - P_1 - \dots - P_{j_{r-1}}}$$

or, if  $P_1, P_2, \dots$  have less than  $r$  positive elements,

$$P(N_r = \infty | \mathcal{P} = (P_1, P_2, \dots)) = 1.$$

The size-biased permutation  $(\tilde{P}_1, \tilde{P}_2, \tilde{P}_3, \dots)$  of  $(P_1, P_2, P_3, \dots)$  is given by

$$(\tilde{P}_1, \tilde{P}_2, \tilde{P}_3, \dots) = (P_{N_1}, P_{N_2}, P_{N_3}, \dots),$$

where we let  $P_\infty = 0$  (see, for example, Gneden (1998)).

##### 3.1.2. Size-biased permutation of symmetric Dirichlet distribution

Suppose that  $(P_1, P_2, \dots, P_{m+1})$  have  $m + 1$ -dimensional Dirichlet distribution with parameter  $\beta (> 0)$ , where  $P_1 + P_2 + \dots + P_{m+1} = 1$ . On the residual fractions  $W_1, W_2, \dots, W_m, W_{m+1} (= 1)$  of their size-biased permutation  $(\tilde{P}_1, \tilde{P}_2, \dots, \tilde{P}_{m+1})$ , it is well-known that  $W_1, W_2, \dots, W_m$  are independent and  $W_j$  has the beta distribution  $Be(\beta + 1, (m - j + 1)\beta)$  for  $j = 1, \dots, m$  (see, for example, Patil and Taillie (1977), Kingman 1993)).

We explain an outline of this fact. Using  $P_{N_1} = \tilde{P}_1$  of Appendix 3.1.1, we can write  $(P_{N_1}, P_1, \dots, P_{N_1-1}, P_{N_1+1}, \dots, P_{m+1}) = (P_{N_1}, (1 - P_{N_1})\mathcal{P}')$ , where

$$\mathcal{P}' = \left( \frac{P_1}{1 - P_{N_1}}, \dots, \frac{P_{N_1-1}}{1 - P_{N_1}}, \frac{P_{N_1+1}}{1 - P_{N_1}}, \dots, \frac{P_{m+1}}{1 - P_{N_1}} \right),$$

$P_{N_1} = W_1$  and  $\mathcal{P}'$  are independent.  $P_{N_1} = W_1$  has the beta distribution  $Be(\beta + 1, m\beta)$  and  $\mathcal{P}'$  has  $m$ -dimensional Dirichlet distribution with parameter  $\beta$ . We have the same discussion for  $\mathcal{P}'$ . By repeating the same discussion, we get the result.

#### 3.2. Poisson-Kingman distribution

We quote about the Poisson-Kingman distribution and the related matters from Pitman (2003) and Perman et. al (1992).

### 3.2.1. Poisson-Kingman distribution

We denote the elements of  $\mathcal{P}$  by  $P_1, P_2, \dots$  with the descending orders  $P_1 \geq P_2 \geq \dots$ . We give the distribution of  $(P_1, P_2, \dots)$ , using an inhomogeneous Poisson process with Lévy measure  $\Lambda$ . Suppose that  $\int_0^1 x\Lambda(dx) < \infty$  and  $\Lambda[1, \infty) < \infty$ . Let  $J_1 \geq J_2 \geq \dots \geq 0$  be the random length of the ordered points of the process and put  $T = \sum_{i=1}^{\infty} J_i$ . We assume that the Lévy measure  $\Lambda$  has a density  $\rho(x)$  and  $T$  has a strictly positive and continuous density on  $(0, \infty)$ . Then the Laplace transform of  $T$  can be written as

$$E(e^{-\lambda T}) = \int_0^{\infty} e^{-\lambda t} f(t) dt = e^{-\psi(\lambda)},$$

where  $\psi(\lambda) = \int_0^{\infty} (1 - e^{-\lambda x}) \rho(x) dx$ . We put  $P_i = J_i/T$  ( $i = 1, 2, \dots$ ). The distribution of  $(P_1, P_2, \dots)$  is called the Poisson-Kingman distribution with Lévy density  $\rho$  and denoted  $\text{PK}(\rho)$ . For the conditional distribution  $\text{PK}(\rho|t)$  of  $(P_1, P_2, \dots)$  given  $T = t$  and a probability distribution  $\gamma$  on  $(0, \infty)$ ,

$$\text{PK}(\rho, \gamma) = \int_0^{\infty} \text{PK}(\rho|t) \gamma(dt)$$

is called the Poisson-Kingman distribution with Lévy density  $\rho$  and mixing distribution  $\gamma$  and denoted  $\text{PK}(\rho, \gamma)$ .

### 3.2.2. Joint distribution of $T$ and $W$ 's

For the joint distribution of  $T$  and  $W_1, W_2, \dots$  associated with  $P_1, P_2, \dots$ , the following are obtained by Theorem 2.1 of Perman et al. (1992): Let  $T$  have the density  $f$  and Lévy measure  $\Lambda$  has the density  $\rho$ . The joint density of  $(T, W_1, W_2)$  is

$$g_{T, W_1, W_2}(t, w_1, w_2) = \Theta(w_1 t) \Theta(\bar{w}_1 w_2 t) f(\bar{w}_1 \bar{w}_2 t), \quad (\bar{w} = 1 - w, \Theta(x) = x\rho(x))$$

and for every  $n \geq 1$  there are similar product formula for the  $n + 1$  dimensional joint densities of  $(T, W_1, \dots, W_n)$ .

Thus, if  $T$  have the stable density  $f$  and Lévy density is given by  $\rho_\alpha = \alpha x^{-\alpha-1}/\Gamma(1-\alpha)$ , then the joint density of  $(T, W_1, \dots, W_n)$  is

$$f_{T, W_1, \dots, W_n}(t, w_1, \dots, w_n) = \left( \frac{\alpha}{\Gamma(1-\alpha)} \right)^n t^{-n\alpha} \\ \times f_\alpha((1-w_1) \cdots (1-w_n)t) \times [w_1(1-w_1)^{n-1} w_2(1-w_2)^{n-2} \cdots w_n]^{-\alpha}. \quad (11)$$

### 3.2.3. Poisson-Dirichlet distribution

In case of  $\rho(x) = \theta x^{-1} e^{-x}$  ( $\theta > 0$ ),  $T$  has a gamma distribution whose density given by  $f(t) = t^{\theta-1} e^{-t}/\Gamma(\theta)$ . The  $\text{PK}(\rho)$  with this  $\rho$  is the Poisson-Dirichlet distribution with parameter  $\theta$ ,  $\text{PD}(0, \theta)$  or  $\text{PD}(\theta)$ .

Now we consider  $T$  with stable distribution whose Laplace transform is given by  $E[e^{-\lambda T}] = \int_0^{\infty} e^{-\lambda t} f_\alpha(x) dx = \exp(-\lambda^\alpha)$  ( $0 < \alpha < 1$ ), where  $f_\alpha$  is the density of  $T$ . The Lévy density corresponding to this Laplace transform is

$$\rho_\alpha(x) = \frac{\alpha x^{-\alpha-1}}{\Gamma(1-\alpha)}. \quad (12)$$

In the  $\text{PK}(\rho_\alpha|t)$ , the distribution of  $\tilde{P}_1$  has the density

$$\tilde{f}_\alpha(p|t) = \frac{\alpha(pt)^{-\alpha}}{\Gamma(1-\alpha)} \frac{f_\alpha((1-p)t)}{f_\alpha(t)} \quad (0 < p < 1). \quad (13)$$

For parameter  $0 < \alpha < 1$ ,  $\theta > -\alpha$  and the above Lévy density  $\rho_\alpha(x)$ ,

$$\begin{aligned} V_{n,k}(t) &= \frac{\alpha^k t^{-n}}{\Gamma(n-k\alpha)f_\alpha(t)} \int_0^t s^{n-k\alpha-1} f_\alpha(t-s) ds \\ &= \frac{\alpha^k t^{-k\alpha}}{\Gamma(n-k\alpha)f_\alpha(t)} \int_0^1 y^{n-k\alpha-1} f_\alpha(t(1-y)) dy. \end{aligned} \quad (14)$$

If we take the continuous distribution  $\gamma$  on  $[0, \infty)$  as the mixing distribution,  $V_{n,k}$  can be written as

$$V_{n,k} = \int_0^\infty \frac{\alpha^k t^{-n}}{\Gamma(n-k\alpha)f_\alpha(t)} \int_0^t s^{n-k\alpha-1} f_\alpha(t-s) ds \gamma(dt). \quad (15)$$

We consider the mixing distribution  $\gamma$  given by  $\gamma_{\alpha,\theta}(dt) = C_{\alpha,\theta}^{-1} t^{-\theta} f_\alpha(t) dt$ . The  $\text{PK}(\rho_\alpha, \gamma_{\alpha,\theta})$  is the Poisson-Dirichlet distribution with parameter  $(\alpha, \theta)$ ,  $\text{PD}(\alpha, \theta)$ .

### 3.3. Another representation of Lévy density

We consider as the Lévy density of 3.3.1, instead of  $\rho_\alpha(x) = \alpha x^{-\alpha-1}/\Gamma(1-\alpha)$ , positive  $(\alpha, \delta)$ -stable density

$$\rho_{\alpha,\delta}(x) = \delta 2^\alpha \rho_\alpha(x), \quad \delta > 0$$

(see, Barndorff-Nielson Shephard (2001) and Cerquetti (2007)). For this Lévy density  $\rho_{\alpha,\delta}$  and the mixing distribution  $\gamma$ , the Poisson-Kingman distribution is given by

$$\text{PK}(\rho_{\alpha,\delta}, \gamma) = \int_0^\infty \text{PK}(\delta 2^\alpha \rho_\alpha|t) \gamma(dt).$$

Using the relation  $\text{PK}(c\rho_\alpha|t) = \text{PK}(\rho_\alpha|c^{-1/\alpha}t)$  ((55) of Pitman (2003)) to the above equation, we have

$$\text{PK}(\rho_{\alpha,\delta}, \gamma) = \int_0^\infty \text{PK}(\rho_\alpha|\tau) G(d\tau), \quad G(x) = \gamma((\delta 2^\alpha)^{1/\alpha} x).$$

Therefore, the Poisson-Kingman distribution with Lévy density  $\rho_{\alpha,\delta}$  and mixing distribution  $\gamma$  is equal to the Poisson-Kingman distribution with Lévy density  $\rho_\alpha$  and mixing distribution  $G$ , that is  $\text{PK}(\rho_{\alpha,\delta}, \gamma) = \text{PK}(\rho_\alpha, G)$ .

### Acknowledgement

This work was supported by Grant-in-Aid for Scientific Research (B) (No. 22300097), Japan Society for the Promotion of Science.

## References

- Antoniak, C. E. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Statist.* **2**, 1142–1174.
- Barndorff-Nielsen, O. E. and Shepard, N. (2001). Normal modified stable processes. *Th. Probab. Math. Statist.*, **65**, 1–19
- Cerquetti, A. (2008). On a Gibbs characterization of normalized generalized Gamma process. *Statist. & Probab. Lett.* **78**: 3123–3128.
- Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, 209–230.
- Gnedin, A. (1998). On convergence and extensions of size-biased permutations, *J. Appl. Prob.* **35**, 642–650.
- Gnedin, A. and Pitman, J. (2006). Exchangeable Gibbs partitions and Stirling triangles. *J. of Mathematical Sciences* **138**, 5674–5685.
- Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997a). *Continuous univariate distributions* Vol. 1 2nd ed. New York: John Wiley & Sons.
- Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997b). *Discrete multivariate distributions*. New York: John Wiley & Sons.
- Kingman, J. F. C. (1975). Random discrete distributions. *J. R. Statist. Soc.* **B 37**, 1–15
- Kingman, J. F. C. (1993). *Poisson processes*. New York: Oxford University Press.
- Lijoi, A., Mena, R. H. and Prünster, I. (2005). Hierarchical mixture modelling with normalized inverse Gaussian priors. *J. Am. Statist. Assoc.* **100**, 1278–1291.
- Lijoi, A., Prünster, I. and Walker G.W. (2008). Investigating nonparametric priors with Gibbs structure. *Statistica Sinica* **18**: 1653–1668.
- Patil, G. P. and Taillie, C. (1977). Diversity as a concept and its implications for random communities. *Bull. Internat. Statist. Institute* **41**, 497–515.
- Perman, M., Pitman, J. and Yor, M. (1992). Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Relat. Fields* **92**, 21–39
- Pitman, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Relat. Fields* **102**, 145–158.
- Pitman, J. (2003). Poisson-Kingman Partitions, In *Science Statistics : A Festschrift for Terry Speed* (Edited by D. R. Goldstein), vol. 40 of Lecture Notes, Monograph Series, 1–34. Hayward: IMS. *Statistica Sinica* **18**: 1653–1668.
- Yamato, H., Sibuya, M. and Nomachi, T. (2001). Ordered sample from two-parameter GEM distribution, *Stat. & Prob. Lett* **55**: 19–27.

Received February 3, 2011

Revised October 10, 2011