J Theor Probab (2007) 20: 295–308
DOI 10.1007/s10959-007-0069-y

The Dirichlet Distribution and Process through Neutralities

Konstancja Bobecka · Jacek Wesołowski

Received: 23 June 2005 / Revised: 28 February 2006 /
Published online: 26 April 2007

Abstract A new characterization of the Dirichlet distribution, based on the notion of complete neutrality and a regression version of neutrality, is derived. It unifies earlier characterizations by James and Mosimann (Ann. Stat. **8**, 183–189, 1980) and by Seshadri and Wesołowski (Sankhyā, A **65**, 248–291, 2003). Also new results on identification of the Dirichlet process in the class of neutral-to-the-right processes are obtained. The proof of the main result makes an extensive use of the method of moments.

Keywords Neutrality · Complete neutrality · Neutrality-to-the-right · Dirichlet distribution · Dirichlet process · Beta distribution · Constancy of regression · Method of moments

1 Introduction

The notion of neutrality and complete neutrality was introduced by Connor and Mosimann [2] while studying the concepts of independence for random proportions.

Let $T_n = \{\mathbf{x} = (x_1, \ldots, x_n) \in (0, \infty)^n : \sum_{i=1}^n x_i < 1\}$. Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a T_n-valued random vector with $n \geq 2$.

This research was supported by the Ministry of Science and Higher Education of Poland (Project No. 1P03A022427).

K. Bobecka · J. Wesołowski (✉)
Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska, Warszawa, Poland
e-mail: wesolo@mini.pw.edu.pl

K. Bobecka
e-mail: bobecka@mini.pw.edu.pl

Definition 1 Let $k \in \{1, \dots, n-1\}$. A subvector $(X_1, \dots, X_k)$ is neutral in $\mathbf{X}$ if the vectors

$$(X_1, \dots, X_k) \quad \text{and} \quad \left(1 - \sum_{i=1}^{k} X_i\right)^{-1} (X_{k+1}, \dots, X_n),$$

are independent.

Let us define a mapping ψ on T_n by

$$\psi(\mathbf{x}) = \left(x_1, \frac{x_2}{1-x_1}, \dots, \frac{x_n}{1 - \sum_{k=1}^{n-1} x_k}\right), \qquad \mathbf{x} = (x_1, \dots, x_n) \in T_n. \tag{1}$$

Definition 2 $\mathbf{X}$ is completely neutral if the components of the random vector $\psi(\mathbf{X})$ are mutually independent.

Note that the order of the coordinates of $\mathbf{X}$ is important in the definition of complete neutrality and it will be important also in the related characterizations which are the main issue we address here.

A refined version of complete neutrality of a random vector is a notion of neutrality-to-the-right (ntr) for random distribution functions, which was introduced in Doksum [4] and has been studied intensively since then, mostly in the context of Bayesian statistics, see for instance [5, 6, 13, 16]. More precisely let $(F(t))_{t \in \mathbf{R}}$ be a stochastic process such that its trajectories are almost surely distribution functions of some probability measure. We say that $(F(t))$ is an ntr process if for any $n \geq 1$ and for any $t_1 < \dots < t_n$ the random vector $(F(t_1), F(t_2) - F(t_1), \dots, F(t_n) - F(t_{n-1}))$ is completely neutral. Dually, we say (following [4]) that $(F(t))$ is neutral-to-the-left (ntl) if $(1 - F(t_n), F(t_n) - F(t_{n-1}), \dots, F(t_2) - F(t_1))$ is completely neutral.

Recall that $\mathbf{X} = (X_1, \dots, X_n)$ has a Dirichlet distribution with parameters $(\alpha_1, \dots, \alpha_{n+1})$, $\mathbf{X} \sim \mathrm{Dir}(\alpha_1, \dots, \alpha_{n+1})$, if it has the density of the form

$$f(\mathbf{x}) = \frac{\Gamma(\sum_{i=1}^{n+1} \alpha_i)}{\prod_{i=1}^{n+1} \Gamma(\alpha_i)} \prod_{i=1}^{n} x_i^{\alpha_i - 1} \left(1 - \sum_{i=1}^{n} x_i\right)^{\alpha_{n+1}-1} I_{T_n}(\mathbf{x}),$$

where $\alpha_i > 0$, $i = 1, \dots, n+1$.

This is the basic example of a probability distribution on a unit simplex which has all neutrality and complete neutrality properties, while the Dirichlet process (see [8]) is essential for the ntr and ntl properties. The Dirichlet process $(F(t))$ on the real line driven by a finitely additive measure μ is a process such that finite dimensional distribution of its increments is Dirichlet

$$\begin{aligned}(F(t_1), F(t_2) - F(t_1), \dots, F(t_n) - F(t_{n-1})) \\ \sim \mathrm{Dir}(\mu((-\infty, t_1]), \mu((t_1, t_2]), \dots, \mu((t_{n-1}, t_n]))\end{aligned}$$

for any $n \geq 2$ and any $t_1 < \dots < t_n$.

Several characterizations of the Dirichlet distribution and Dirichlet process related to concepts of neutrality are known in the literature, see e.g. [3, 7, 11] (referred to by

JM in the sequel), [9, 12, 14]. For other characterizations of the Dirichlet distribution one may consult Chamayou and Letac [1] or a fairly recent review by Gupta and Richards [10].

In this paper we provide a new characterization of the Dirichlet distribution which makes use of the concept of complete neutrality combined with a weak version of neutrality expressed in terms of constancy of regressions. As it will be seen later on this result unifies JM characterization of the Dirichlet distribution and a characterization of the beta distribution given recently in Seshadri and Wesołowski [15]. The main result of the paper is discussed in Sect. 2, while its proof, based on the method of moments, is given in Sect. 4. Section 3 is devoted to related characterizations of the Dirichlet process. In particular we provide new results developed on the basis of the characterization of the Dirichlet distribution obtained in Sect. 2. Except of that we formulate and prove a corrected version of the characterization of the Dirichlet process stated in JM, providing a solution of the Doksum [4] conjecture that the only random distribution function which is both ntr and ntl is Dirichlet.

2 Complete Neutrality

The following characterization of the Dirichlet distribution is proved in JM:

Theorem 3 *Let* $\mathbf{X} = (X_1, \dots, X_n)$ *be a* T_n*-valued random vector with* $n \geq 2$. *If* $\mathbf{X}$ *is completely neutral and* X_n *is neutral in* $\mathbf{X}$ *then* $\mathbf{X}$ *has the Dirichlet distribution.*

We say that U has a beta distribution with parameters $(p, q) : U \sim \mathrm{B}(p, q)$ if it has the density of the form

$$f(u) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} u^{p-1}(1-u)^{q-1} I_{(0,1)}(u), \quad \text{with } p, q > 0.$$

It is easy to see that $\mathbf{X} \sim \mathrm{Dir}(\alpha_1, \dots, \alpha_{n+1})$ iff $\bar{U} = (U_1, \dots, U_n) = \psi(\mathbf{X})$, where ψ is defined by (1), has independent beta components, $U_k \sim \mathrm{B}(\alpha_k, \alpha_{k+1} + \cdots + \alpha_{n+1})$, $k = 1, \dots, n$. This observation, which will be used in the proof of our characterization in Sect. 4, follows, for instance, by direct computation since $\psi : T_n \to (0,1)^n$ is a bijection.

Below we give an extension of Theorem 3, which is the main result of this paper.

Theorem 4 *Let* $\mathbf{X} = (X_1, \dots, X_n)$ *be a* T_n*-valued random vector. If* $\mathbf{X}$ *is completely neutral and for some* $2 \leq i \leq n$

$$E\left(\left(1 - \sum_{k=i}^{n} X_k\right)^{-1} X_j \,\middle|\, X_i, \dots, X_n\right) = c_j \quad j = 1, 2, \dots, i-1, \tag{2}$$

$$E\left(\left(1 - \sum_{k=i}^{n} X_k\right)^{-1} X_1^2 \,\middle|\, X_i, \dots, X_n\right) = d, \tag{3}$$

then there exist positive constants $p, \alpha_i, \dots, \alpha_n$ *such that the distribution of* $\mathbf{X}$ *is Dirichlet*, $\mathrm{Dir}(\alpha_1, \dots, \alpha_{n+1})$, *where*

$$\alpha_1 = \frac{c_1(d-c_1)}{c_1^2 - d} > 0, \qquad \alpha_j = \frac{c_j}{c_1} p > 0, \quad j = 2, \dots, i-1,$$

$$\alpha_{n+1} = \frac{(1-C)(d-c_1)}{c_1^2 - d} > 0, \quad \text{with } C = \sum_{j=1}^{i-1} c_j.$$

Remark 5 Note that the regression conditions (2) and (3) are much weaker than neutrality of the subvector $(X_i, \dots, X_n)$ in $(X_1, \dots, X_n)$. So the above result is a generalization of the characterization of the Dirichlet distribution proved in JM.

It appears that Theorem 4 covers also a recent characterization of the beta distribution:

Remark 6 Seshadri and Wesołowski [15] proved the following result:

If X, Y are independent (0, 1)-valued random variables such that

$$E\left(\frac{1-Y}{1-XY} \,\middle|\, XY\right) = c, \qquad E\left(\left(\frac{1-Y}{1-XY}\right)^2 \,\middle|\, XY\right) = d,$$

then there exists $p > 0$ such that $X \sim \mathrm{B}(p, q)$, $Y \sim \mathrm{B}(p+q, r)$, where

$$q = \frac{(1-c)(d-c)}{c^2 - d} > 0, \qquad r = \frac{c(d-c)}{c^2 - d} > 0.$$

Note that by taking $X_1 = 1 - Y$ and $X_2 = XY$ we have:

$$X = \frac{X_2}{1 - X_1}, \qquad \frac{1-Y}{1-XY} = \frac{X_1}{1-X_2}.$$

Thus the above result is equivalent to Theorem 4 for $n = 2$ and $i = 2$.

3 Neutrality-to-the-Right

Doksum [4] conjectured that the only random distribution function which is both ntr and ltr is the Dirichlet process. JM claim that it is a consequence of Theorem 3. In a sense it is really the case, however their argument needs some clarification: apparently the condition stated in their Sect. 3: "$1 - F(t_n)$ is neutral in

$$(F(t_1), F(t_2) - F(t_1), \dots, F(t_{n-1}) - F(t_{n-2}), 1 - F(t_n))\text{"}$$

is not implied by neutrality to the left of the process $(F(t))$.

In this context it is more convenient to consider a version of Theorem 3 for $(n+1)$-dimensional random vector $(X_1, \dots, X_n, X_{n+1})$ such that $X_i > 0$, $i = 1, \dots, n+1$ and $\sum_{i=1}^{n+1} X_i = 1$. Note that complete neutrality of $(X_1, \dots, X_n)$ is equivalent to complete neutrality of $(X_1, \dots, X_n, X_{n+1})$ (observe that $X_{n+1}/(1 - \sum_{i=1}^{n} X_i) = 1$). Such a version of Theorem 3 has the form

Theorem 7 *Let* $\mathbf{X} = (X_1, \ldots, X_n, X_{n+1})$ *be a random vector with positive components such that their sum equals* 1. *If* $\mathbf{X}$ *is completely neutral and* X_{n+1} *is neutral in* $\mathbf{X}$ *then* $(X_1, \ldots, X_n)$ *has the Dirichlet distribution.*

Proof Note that complete neutrality of $\mathbf{X}$ is equivalent to the following set of independencies: for any $k = 1, \ldots, n-1$

$$(X_1, \ldots, X_k) \quad \text{and} \quad \left(1 - \sum_{j=1}^{k} X_j\right)^{-1} (X_{k+1}, \ldots, X_n, X_{n+1}) \tag{4}$$

are independent. On the other hand, neutrality of X_{n+1} in $\mathbf{X}$ is equivalent to

$$X_{n+1} \quad \text{and} \quad (1 - X_{n+1})^{-1}(X_1, \ldots, X_{n-1}, X_n) \tag{5}$$

are independent.

Define $Y_i = X_i$ for $i = 1, \ldots, n-1$, $Y_n = X_{n+1}$, $Y_{n+1} = X_n$. Then (4) is equivalent to complete neutrality of $(Y_1, \ldots, Y_n)$ and (5) is equivalent to neutrality of Y_n in $(Y_1, \ldots, Y_n)$. Thus by Theorem 3 it follows that $(Y_1, \ldots, Y_n)$ is Dirichlet. Consequently $(X_1, \ldots, X_n)$ is also Dirichlet. □

As a corollary to this result we have a corrected version of the characterization of the Dirichlet process stated in JM.

Theorem 8 *Let* $(F(t))_{t \in \mathbf{R}}$ *be a stochastic process such that its trajectories are a.s. distribution functions. Assume that* $(F(t))$ *is an ntr process. Let* $1 - F(t_n)$ *be neutral in*

$$(F(t_1), F(t_2) - F(t_1), \ldots, F(t_n) - F(t_{n-1}), 1 - F(t_n))$$

for any $n \geq 2$ *and for any* $t_1 < \cdots < t_n$.

Then (F_t) *is a Dirichlet process.*

Proof Fix $n \geq 2$ and an arbitrary set of points $t_1 < \cdots < t_n$. Denote then $X_1 = F(t_1)$, $X_k = F(t_k) - F(t_{k-1})$, $k = 2, \ldots, n$, $X_{n+1} = 1 - F(t_n)$. Now the assumptions of Theorem 7 are satisfied. Thus $(F(t_1), F(t_2) - F(t_1), \ldots, F(t_n) - F(t_{n-1}))$ has an n-dimensional Dirichlet distribution

$$\mathrm{Dir}(\alpha_{1,n}(t_1, \ldots, t_n), \ldots, \alpha_{n,n}(t_1, \ldots, t_n), \alpha_{n+1,n}(t_1, \ldots, t_n))$$

for $(n+1)$ positive functions $\alpha_{1,n}, \ldots, \alpha_{n,n}, \alpha_{n+1,n}$. The same holds true for any $n \geq 2$ and for any $t_1 < \cdots < t_n$. Now by the Kolmogorov consistency condition it follows that there exist a finitely additive measure μ on $\mathbf{R}$ such that $\alpha_{1,n}(t_1, \ldots, t_n) = \mu((-\infty, t_1])$, $\alpha_{k,n}(t_1, \ldots, t_n) = \mu((t_{k-1}, t_k])$, $k = 2, \ldots, n$, and $\alpha_{n+1,n}(t_1, \ldots, t_n) = \mu((t_n, \infty))$. □

Now we will explore implications of Theorem 4 for the theory of the Dirichlet process.

First, one can get a regression version of Theorem 7 from Theorem 4 with $i = n$.

Theorem 9 *Let* $\mathbf{X} = (X_1, \ldots, X_n, X_{n+1})$ *be a random vector with positive components such that their sum equals* 1. *Let* $\mathbf{X}$ *be completely neutral. Assume that*

$$E\left(\frac{X_j}{1 - X_{n+1}} \,\middle|\, X_{n+1}\right) = c_j, \quad j = 1, \ldots, n,$$

and

$$E\left(\left(\frac{X_1}{1 - X_{n+1}}\right)^2 \,\middle|\, X_{n+1}\right) = d$$

for some real constants c_j, $j = 1, \ldots, n$, *and* d.

Then $(X_1, \ldots, X_n)$ *has a Dirichlet distribution.*

Proof It follows by renumbering variables, similarly to the way Theorem 7 was obtained from Theorem 3. □

Thus Theorem 9 leads to a new characterization of the Dirichlet process which results in a further weakening the assumptions of the Doksum [4] conjecture. Namely, we have the following strengthened version of Theorem 8.

Theorem 10 *Let* $(F(t))_{t\in\mathbf{R}}$ *be a stochastic process whose trajectories are a.s. distribution functions. Assume that* $(F(t))$ *is an ntr process. Let*

$$E\left(\frac{F(t_j) - F(t_{j-1})}{F(t_n)} \,\middle|\, F(t_n)\right) = c_{j,n}(t_1, \ldots, t_n), \quad j = 1, 2, \ldots, n,$$

$$E\left(\left(\frac{F(t_1)}{F(t_n)}\right)^2 \,\middle|\, F(t_n)\right) = d_n(t_1, \ldots, t_n),$$

for any $n \geq 2$ *and for any* $-\infty = t_0 < t_1 < \cdots < t_n$, *where* $c_{j,n}$, $j = 1, \ldots, n$ *and* d_n *are some real functions.*

Then (F_t) *is a Dirichlet process.*

Proof It is an immediate consequence of the version of Theorem 9 with $(X_1, \ldots, X_n, X_{n+1})$ defined as in the proof of Theorem 8 and of the Kolmogorov consistency condition. □

This way of reasoning for getting another answer to the Doksum [4] conjecture does not apply to the extension of JM characterization of the Dirichlet distribution provided in Theorem 4 if $i < n$. The essential reason is that a version of Theorem 4 for the vector $(X_1, \ldots, X_n, X_{n+1})$, $\sum_{i=1}^{n+1} X_i = 1$, does not hold for i different than $n+1$, i.e. complete neutrality of $(X_1, \ldots, X_n, X_{n+1})$ and neutrality of $(X_i, \ldots, X_{n+1})$ in $(X_1, \ldots, X_n, X_{n+1})$ does not imply that $(X_1, \ldots, X_n)$ is Dirichlet, except the case $i = n + 1$. To see this consider the following example:

Example 11 Take $n = 3$. By Theorem 4 it follows that if

$$X_1 \quad \text{and} \quad \left(\frac{X_2}{1 - X_1}, \frac{X_3}{1 - X_1}, \frac{X_4}{1 - X_1}\right) \quad \text{are independent,} \tag{6}$$

$$(X_1, X_2) \quad \text{and} \quad \left(\frac{X_3}{1-X_1-X_2}, \frac{X_4}{1-X_1-X_2}\right) \quad \text{are independent,} \tag{7}$$

$$(X_2, X_3) \quad \text{and} \quad \left(\frac{X_1}{1-X_2-X_3}, \frac{X_4}{1-X_2-X_3}\right) \quad \text{are independent,} \tag{8}$$

then (X_1, X_2, X_3) is Dirichlet.

Now, to get a proper analogue of Theorem 4, i.e. for instance a version of Theorem 9 for $3 = i < n+1 = 4$, we would like to change (8) into

$$(X_3, X_4) \quad \text{and} \quad \left(\frac{X_1}{1-X_3-X_4}, \frac{X_2}{1-X_3-X_4}\right) \quad \text{are independent.} \tag{9}$$

However (6), (7) and (9) do not imply that (X_1, X_2, X_3) is Dirichlet: it suffices to take (X_1, X_2) to be Dirichlet and $X_3 = Y(1 - X_1 - X_2)$, where Y is any $(0,1)$-valued random variable independent of (X_1, X_2), and $X_4 = 1 - X_1 - X_2 - X_3$. Then (X_1, X_2, X_3, X_4) is completely neutral, i.e. (6) and (7) hold, (X_3, X_4) is neutral in (X_1, X_2, X_3, X_4), i.e. (9) holds, but (X_1, X_2, X_3) is not Dirichlet.

Finally, let us point out that Theorem 4 leads to another characterization of the Dirichlet process in the class of ntr processes, which does not fall into the scheme of jointly ntr and ltr processes proposed in Doksum [4].

Theorem 12 *Let $(F(t))_{t\in\mathbf{R}}$ be a stochastic process whose trajectories are a.s. distribution functions. Assume that $(F(t))$ is an ntr process. Assume that there exists $i > 1$ such that for any $n \geq i$*

$$\begin{aligned}
&E\left(\frac{F(t_j)-F(t_{j-1})}{1-F(t_n)+F(t_{i-1})} \,\middle|\, F(t_i)-F(t_{i-1}), \dots, F(t_n)-F(t_{n-1})\right)\\
&\quad = C_{j,n}(t_1, \dots, t_n), \quad j = 1, \dots, i-1, \quad \textit{and}\\
&E\left(\left(\frac{F(t_1)}{1-F(t_n)+F(t_{i-1})}\right)^2 \,\middle|\, F(t_i)-F(t_{i-1}), \dots, F(t_n)-F(t_{n-1})\right)\\
&\quad = D_n(t_1, \dots, t_n)
\end{aligned}$$

for any $-\infty = t_0 < t_1 < \cdots < t_n$, where $C_{j,n}$, $j = 1, \dots, i-1$ and D_n are some real functions.

Then $(F(t))_{t\in\mathbf{R}}$ is a Dirichlet process.

Proof It follows directly from Theorem 4 by taking $X_j = F(t_j) - F(t_{j-1})$, $j = 1, \dots, n$, that the random vector $(F(t_1), F(t_2) - F(t_1), \dots, F(t_n) - F(t_{n-1}))$ has a Dirichlet distribution for any $t_1 < \cdots < t_n$. An application of the Kolmogorov consistency condition to identify the parameters concludes the proof. □

4 Proof of Theorem 4

Proof Let $\mathbf{U} = (U_1, \dots, U_n) = \psi(\mathbf{X})$, where ψ is defined by (1). We have $\mathbf{X} = (U_1, (1-U_1)U_2, \dots, \prod_{i=1}^{n-1}(1-U_i)U_n)$.

Conditions (2) and (3) can be rewritten as:

$$\frac{1}{c_j} E\left(U_j \prod_{k=1}^{j-1}(1-U_k) \,\middle|\, U_i \prod_{k=1}^{i-1}(1-U_k), \dots, U_n \prod_{k=1}^{n-1}(1-U_k) \right) = 1 - \sum_{k=i}^{n} U_k \prod_{l=1}^{k-1}(1-U_l) \tag{10}$$

for $j = 1, \dots, i-1$, (where $\prod_{k=1}^{0} = 1$) and

$$E\left(U_1^2 \,\middle|\, U_i \prod_{k=1}^{i-1}(1-U_k), \dots, U_n \prod_{k=1}^{n-1}(1-U_k) \right) = d \left(1 - \sum_{k=i}^{n} U_k \prod_{l=1}^{k-1}(1-U_l) \right)^2. \tag{11}$$

Now for any natural $r_i, \dots, r_n$ the identity (10) implies

$$\frac{1}{c_j} E\left\{ U_j \prod_{k=1}^{j-1}(1-U_k) \left[U_i \prod_{k=1}^{i-1}(1-U_k) \right]^{r_i} \dots \left[U_n \prod_{k=1}^{n-1}(1-U_k) \right]^{r_n} \right\}$$
$$= E\left\{ \left(1 - \sum_{k=i}^{n} U_k \prod_{l=1}^{k-1}(1-U_l) \right) \left[U_i \prod_{k=1}^{i-1}(1-U_k) \right]^{r_i} \dots \left[U_n \prod_{k=1}^{n-1}(1-U_k) \right]^{r_n} \right\} \tag{12}$$

for $j = 1, \dots, i-1$.

For any non-negative integers k, r denote

$$M_l(k) = \frac{E(1-U_l)^{k+1}}{E(1-U_l)^k} \quad l = 1, \dots, i-1, \tag{13}$$

$$N_l(r,k) = \frac{E[U_l^r(1-U_l)^{k+1}]}{E[U_l^r(1-U_l)^k]}, \quad l = i, \dots, n-1 \tag{14}$$

and

$$O(k) = \frac{E[U_n^{k+1}]}{E[U_n^k]}. \tag{15}$$

Note that $0 < M_l(k) < 1$, $l = 1, \dots, i-1$, and $0 < O(k) < 1$ for any k.

Now we will prove that $U_k \sim \mathrm{B}(p_k, q_k)$, $k = 1, \dots, n$, and $q_k = p_{k+1} + q_{k+1}$ for $k = 1, \dots, n-1$. By the remark preceding the formulation of Theorem 4 it will follow that $\mathbf{X} \sim \mathrm{Dir}(\alpha_1, \dots, \alpha_{n+1})$ with $\alpha_k = p_k$ for $k = 1, \dots, n$, and $\alpha_{n+1} = q_n$.

The proof is divided now into five steps.

In the *first step* it will be shown that M_l for $l = 2, \dots, i-1$, is uniquely determined by M_1.

Since the right hand side of (10) does not depend on j and U_j's are independent then for any $j = 2, \dots, i-1$, we can write (after cancelling some of the terms which repeat on the both sides)

$$\frac{1}{c_1} E[U_1(1-U_1)^s] \left\{ \prod_{l=2}^{i-1} E[(1-U_l)^s] \right\}$$
$$= \frac{1}{c_j} \left\{ \prod_{l=1}^{j-1} E[(1-U_l)^{1+s}] \right\} E[U_j(1-U_j)^s] \left\{ \prod_{l=j+1}^{i-1} E[(1-U_l)^s] \right\},$$

with $s = \sum_{k=i}^{n} r_k$. Dividing both sides by $\prod_{l=1}^{i-1} E[(1-U_l)^s]$ and using (13) we can rewrite the above equation in a much shorter form as

$$\frac{1}{c_1}[1 - M_1(s)] = \frac{1}{c_j} \left[\prod_{l=1}^{j-1} M_l(s) \right] [1 - M_j(s)], \quad j = 2, \dots, i-1.$$

Thus

$$\prod_{l=1}^{j} M_l(s) = \prod_{l=1}^{j-1} M_l(s) - \frac{c_j}{c_1}[1 - M_1(s)], \quad j = 2, \dots, i-1.$$

Iterating we get

$$\prod_{l=1}^{j} M_l(s) = 1 - \frac{1 - M_1(s)}{c_1} \sum_{k=1}^{j} c_k, \quad j = 2, \dots, i-1. \tag{16}$$

Consequently M_j's for $j = 2, \dots, i-1$, are uniquely determined by M_1 as

$$M_j(s) = \frac{c_1 - [1 - M_1(s)] \sum_{k=1}^{j} c_k}{c_1 - [1 - M_1(s)] \sum_{k=1}^{j-1} c_k}, \quad j = 2, \dots, i-1. \tag{17}$$

In the *second step* we will express the product of N_j's in terms of M_1 and O. We will exploit here and also later on the following identity

$$\sum_{k=i}^{n} U_k \prod_{l=1}^{k-1} (1-U_l) = \prod_{l=1}^{i-1} (1-U_l) - \prod_{l=1}^{n} (1-U_l). \tag{18}$$

Applying (18) to (12) for $j = 1$ and using independence of U_j's we get

$$\frac{1}{c_1} E[U_1(1-U_1)^s] \left\{ \prod_{l=2}^{i-1} E[(1-U_l)^s] \right\} \left\{ \prod_{l=i}^{n} E[U_l^{r_l}(1-U_l)^{s_l}] \right\}$$

$$= \left\{\prod_{l=1}^{i-1} E[(1-U_l)^s]\right\}\left\{\prod_{l=i}^{n} E[U_l^{r_l}(1-U_l)^{s_l}]\right\} - \left\{\prod_{l=1}^{i-1} E[(1-U_l)^{1+s}]\right\}$$
$$\times \left\{\prod_{l=i}^{n} E[U_l^{r_l}(1-U_l)^{s_l}]\right\} + \left\{\prod_{l=1}^{i-1} E[(1-U_l)^{1+s}]\right\}\left\{\prod_{l=i}^{n} E[U_l^{r_l}(1-U_l)^{1+s_l}]\right\}$$

with $s_l = \sum_{k=l+1}^{n} r_k$, $l = i, \dots, n$, where $\sum_{k=n+1}^{n} = 0$. Dividing both sides by

$$\left\{\prod_{l=1}^{i-1} E[(1-U_l)^s]\right\}\left\{\prod_{l=i}^{n} E[U_l^{r_l}(1-U_l)^{s_l}]\right\} \tag{19}$$

and using (13), (14) and (15), we can rewrite the above equation in a much shorter form as

$$\frac{1}{c_1}[1 - M_1(s)] = 1 - \left[\prod_{l=1}^{i-1} M_l(s)\right]\left[1 - (1 - O(r_n))\prod_{l=i}^{n-1} N_l(r_l, s_l)\right].$$

Now, using (16) we arrive at

$$[1 - O(r_n)]\prod_{l=i}^{n-1} N_l(r_l, s_l) = \frac{(1-C)[1 - M_1(s)]}{c_1 - C[1 - M_1(s)]}, \quad \text{with } C = \sum_{k=1}^{i-1} c_k. \tag{20}$$

In the *third step* we will identify M_1 and O. Now for any natural $r_i, \dots, r_n$ the identity (11) implies

$$\frac{1}{d} E\left\{U_1^2\left[U_i\prod_{k=1}^{i-1}(1-U_k)\right]^{r_i} \dots \left[U_n\prod_{k=1}^{n-1}(1-U_k)\right]^{r_n}\right\}$$
$$= E\left\{\left[1 - \sum_{k=i}^{n} U_k\prod_{l=1}^{k-1}(1-U_l)\right]^2\left[U_i\prod_{k=1}^{i-1}(1-U_k)\right]^{r_i} \dots \left[U_n\prod_{k=1}^{n-1}(1-U_k)\right]^{r_n}\right\}.$$

Using (18) and independence of U_j's and then dividing both sides by (19) we get

$$\frac{1}{d}[1 - 2M_1(s) + M_1(s)M_1(s+1)]$$
$$= 1 - 2\prod_{l=1}^{i-1} M_l(s) + 2\left[\prod_{l=1}^{i-1} M_l(s)\right]\left[\prod_{l=i}^{n-1} N_l(r_l, s_l)\right][1 - O(r_n)] + \left[\prod_{l=1}^{i-1} M_l(s)\right]$$
$$\times \left[\prod_{l=1}^{i-1} M_l(s+1)\right] - 2\left[\prod_{l=1}^{i-1} M_l(s)\right]\left[\prod_{l=1}^{i-1} M_l(s+1)\right]\left[\prod_{l=i}^{n-1} N_l(r_l, s_l)\right]$$
$$\times [1 - O(r_n)]$$

$$+\left[\prod_{l=1}^{i-1} M_l(s)\right]\left[\prod_{l=1}^{i-1} M_l(s+1)\right]\left[\prod_{l=i}^{n-1} N_l(r_l,s_l)\right]\left[\prod_{l=i}^{n-1} N_l(r_l,1+s_l)\right]$$
$$\times [1-2O(r_n)+O(r_n)O(r_n+1)].$$

Plugging (17) and (20) in the above equation we obtain

$$\begin{aligned}
&\frac{1}{d}[1-2M_1(s)+M_1(s)M_1(s+1)] \\
&\quad = 1-2\left[1-\frac{C}{c_1}(1-M_1(s))\right]+2\frac{1-C}{c_1}(1-M_1(s))+\left[1-\frac{C}{c_1}(1-M_1(s))\right] \\
&\qquad \times \left[1-\frac{C}{c_1}(1-M_1(s+1))\right]-2\frac{1-C}{c_1}(1-M_1(s))\left[1-\frac{C}{c_1}(1-M_1(s+1))\right] \\
&\qquad + \frac{1-C}{c_1}(1-M_1(s))\frac{1-C}{c_1}(1-M_1(s+1))\frac{1-O(r_n)O(r_n+1)}{[1-O(r_n)][1-O(r_n+1)]}.
\end{aligned}$$

After some algebra we arrive at

$$\begin{aligned}
&\frac{c_1-Cd}{d}\left(\frac{1}{1-M_1(s+1)}-\frac{1}{1-M_1(s)}\right) \\
&\quad = \frac{d-c_1^2}{c_1 d}+\frac{(1-C)^2}{c_1}\left(\frac{1}{1-O(r_n+1)}-\frac{1}{1-O(r_n)}\right).
\end{aligned} \tag{21}$$

For $r_n=0$ we are getting

$$\frac{1}{1-M_1(k+1)}=\frac{1}{1-M_1(k)}+B_1, \quad k=0,1,\ldots, \tag{22}$$

where B_1 is a constant. Solving (22) we obtain

$$M_1(k)=\frac{q_1+k}{p_1+q_1+k}, \quad k=0,1,\ldots,$$

where

$$q_1=\frac{1}{B_1}\frac{1-EU_1}{EU_1}, \qquad p_1=\frac{1}{B_1}.$$

Since $0<M_1(k)<1$ we have

$$0<\frac{q_1+k}{p_1+q_1+k}<1, \quad k=0,1,\ldots.$$

We will prove that p_1 and q_1 are positive. Assume that $p_1<0$. Then for k large enough the denominator is positive. Thus $q_1+k<p_1+q_1+k$ which for negative p_1 is impossible. For $k=0$ we get $q_1=p_1\frac{1-EU_1}{EU_1}$ which is positive since $0<U_1<1$ a.s. Hence $1-U_1\sim \mathrm{B}(q_1,p_1)$, which implies $U_1\sim \mathrm{B}(p_1,q_1)$. (The same reasoning applies to the other U_j's considered later on in this proof.)

By (22) we conclude from (21) that

$$\frac{1}{1-O(k+1)} = \frac{1}{1-O(k)} + B_n, \quad k = 0, 1, \ldots,$$

where B_n is a constant. Thus

$$O(k) = \frac{p_n + k}{p_n + q_n + k}, \quad k = 0, 1, \ldots,$$

where

$$p_n = \frac{1}{B_n} \frac{EU_n}{1 - EU_n} > 0, \qquad q_n = \frac{1}{B_n} > 0,$$

and we obtain $U_n \sim \mathrm{B}(p_n, q_n)$.

In the *fourth step* we will identify U_j's for $j = 2, \ldots, n-1$. First, inserting M_1 into (17), we get for $j = 2, \ldots, i-1$

$$M_j(k) = \frac{q_j + k}{p_j + q_j + k}, \quad k = 0, 1, \ldots,$$

where

$$q_j = q_1 + \left(1 - \frac{1}{c_1}\sum_{l=1}^{j} c_l\right) p_1, \qquad p_j = \frac{c_j}{c_1} p_1.$$

Again, by the condition $0 < M_j < 1$, we conclude that $q_j > 0$, $p_j > 0$ and $U_j \sim \mathrm{B}(p_j, q_j)$ for $j = 2, \ldots, i-1$. Let us observe that

$$q_j = p_{j+1} + q_{j+1}, \quad j = 1, \ldots, i-2.$$

Now putting $r_n = 0$ into (20) we get

$$\begin{aligned} \prod_{l=i}^{n-1} N_l(r_l, s_l) &= \frac{1}{[1 - O(0)]} \frac{(1-C)[1 - M_1(s)]}{c_1 - C[1 - M_1(s)]} \\ &= \frac{\frac{p_n+q_n}{q_n}(1-C)\frac{p_1}{c_1}}{q_1 + p_1(1 - \frac{C}{c_1}) + s}, \end{aligned} \tag{23}$$

where $s = \sum_{k=i}^{n-1} r_k$, $s_l = \sum_{k=l+1}^{n-1} r_k$.

We first insert $r_i = r$ and $r_j = 0$ for $j = i+1, \ldots, n-1$ into (23), getting

$$N_i(r, 0) = \frac{q_i}{p_i + q_i + r}, \quad r = 0, 1, \ldots,$$

where

$$q_i = \frac{\frac{p_n+q_n}{q_n}(1-C)\frac{p_1}{c_1}}{\prod_{l=i+1}^{n-1} N_l(0,0)}$$

and $p_i = q_1 + p_1(1 - \frac{C}{c_1}) - q_i$. Thus, $p_i > 0$, $q_i > 0$ and $U_i \sim \mathrm{B}(p_i, q_i)$. Note that $q_{i-1} = p_i + q_i$.

Now we will show inductively that for $j = i+1, \dots, n-1$

$$N_j(0,r) = \frac{q_j + r}{q_{j-1} + r},$$

where $q_j = q_i \prod_{l=i+1}^{j} N_l(0,0)$. Observe that it means that $U_j \sim \mathrm{B}(p_j, q_j)$, with $p_j = q_{j-1} - q_j > 0$, $j = i+1, \dots, n-1$.

In the argument below the following fact will be used: if for some j we have $N_j(r,0) = \frac{q_j}{p_j + q_j + r}$ for any $r = 0, 1, \dots$, then $p_j > 0$, $q_j > 0$ and $U_j \sim \mathrm{B}(p_j, q_j)$ and consequently $N_j(0,r) = \frac{q_j + r}{p_j + q_j + r}$ for any $r = 0, 1, \dots$. To see this note that

$$\frac{EU_j^{r+1}}{EU_j^r} = 1 - N_j(r,0) = \frac{p_j + r}{p_j + q_j + r}, \quad r = 0, 1, \dots.$$

Thus $p_j > 0$, $q_j > 0$ and $U_j \sim \mathrm{B}(p_j, q_j)$. Hence $1 - U_j \sim \mathrm{B}(q_j, p_j)$ and

$$N_j(0,r) = \frac{E(1-U_j)^{r+1}}{E(1-U_j)^r} = \frac{q_j + r}{p_j + q_j + r}, \quad r = 0, 1, \dots.$$

Consider first (23) with $r_{i+1} = r$ and $r_j = 0$ for $j \neq i+1$. Then

$$N_{i+1}(r,0) = \frac{N_{i+1}(0,0)}{N_i(0,r)} \frac{q_i}{p_i + q_i + r} = \frac{q_i N_{i+1}(0,0)}{q_i + r}$$

proving the result for $i+1$.

Assume now that the induction assumptions hold for $j = i+1, \dots, k$. Putting $r_k = r$ and $r_j = 0$ for $j \neq k$ into (23) we obtain

$$N_{k+1}(r,0) = \frac{\prod_{l=i+1}^{k+1} N_l(0,0)}{\prod_{l=i}^{k} N_l(0,r)} \frac{q_i}{p_i + q_i + r}.$$

But the induction assumption implies

$$\prod_{l=i}^{k} N_l(0,r) = \frac{q_k + r}{p_i + q_i + r}.$$

Thus

$$N_{k+1}(r,0) = \frac{q_i \prod_{l=i+1}^{k+1} N_l(0,0)}{q_k + r} = \frac{q_{k+1}}{q_k + r}.$$

It implies that $U_{k+1} \sim \mathrm{B}(p_{k+1}, q_{k+1})$ with $p_{k+1} = q_k - q_{k+1} > 0$ and

$$N_{k+1}(0,r) = \frac{q_{k+1} + r}{q_k + r}.$$

In the *fifth step* we will show that $q_{n-1} = p_n + q_n$ and then we will identify the constants p_1 and q_n.

Inserting $r_n = r$ and $r_j = 0$ for $j = i, \ldots, n-1$ into (20) we obtain

$$\frac{q_{n-1}+r}{p_n+q_n+r} = \frac{(1-C)p_1}{c_1 q_n}, \quad r = 0, 1, \ldots.$$

Since the right hand side does not depend on r it follows that $q_{n-1} = q_n + p_n$ and $(1-C)p_1 = c_1 q_n$. On the other hand, (21) can be rewritten as

$$\frac{c_1 - Cd}{d}\frac{1}{p_1} = \frac{d-c_1^2}{c_1 d} + \frac{(1-C)^2}{c_1}\frac{1}{q_n}. \tag{24}$$

Since $1 - C > 0$, we get $p_1 = \frac{c_1}{1-C} q_n$ and plugging it into the above equation, after some easy algebra, we obtain

$$q_n = \frac{(1-C)(d-c_1)}{c_1^2 - d}, \quad p_1 = \frac{c_1(d-c_1)}{c_1^2-d}.$$

Taking expectations in (2) and (3) we have $c_1^2 < d < c_1$. Thus $q_n,\ p_1 > 0$. □

References

1. Chamayou, J.-F., Letac, G.: Transient random walk on stochastic matrices with Dirichlet distribution. Ann. Probab. **22**, 424–430 (1994)
2. Connor, J.R., Mosimann, J.E.: Concepts of independence for proportions with a generalization of the Dirichlet distribution. J. Am. Stat. Assoc. **64**, 194–206 (1969)
3. Darroch, J.N., Ratcliff, D.: A characterization of the Dirichlet distribution. J. Am. Stat. Soc. **66**, 641–643 (1971)
4. Doksum, K.: Tailfree and neutral random probabilities and their posterior distributions. Ann. Probab. **2**, 183–201 (1974)
5. Doksum, K., James, L.F.: On spatial neutral to the right processes and their posterior distributions. In: Mathematical Reliability: An Expository Perspective. Intern. Ser. Oper. Res. Manag. Sci., vol. 67, pp. 87–103. Kluwer, Boston (2004)
6. Epifani, I., Lijoi, A., Prünster, I.: Exponential functionals and means of neutral-to-the-right priors. Biometrika **90**, 791–808 (2003)
7. Fabius, J.: Two characterizations of the Dirichlet distribution. Ann. Stat. **1**, 583–587 (1973)
8. Ferguson, T.S.: A Bayesian analysis of some nonparametric problems. Ann. Stat. **1**, 209–230 (1973)
9. Geiger, D., Heckerman, D.: A characterization of the Dirichlet distribution through global and local parameter independence. Ann. Stat. **25**, 1344–1369 (1997)
10. Gupta, R.D., Richards, D.S.P.: The history of the Dirichlet and Liouville distributions. Intern. Stat. Rev. **69**, 433–446 (2001)
11. James, I.R., Mosimann, J.E.: A new characterization of the Dirichlet distribution through neutrality. Ann. Stat. **8**, 183–189 (1980)
12. Lo, A.Y.: A characterization of the Dirichlet process. Stat. Probab. Lett. **12**, 185–187 (1991)
13. Ramamoorthi, R.V., Draghici, L., Dey, J.: Characterizations of tailfree and neutral to the right priors. In: Advances on Theoretical and Methodological Aspects of Probability and Statistics, Hamilton, ON, 1998, pp. 305–316. Taylor & Francis, London (2002)
14. Rolin, J.-M.: Some useful properties of the Dirichlet process. Preprint, 1–19 (1992)
15. Seshadri, V., Wesołowski, J.: Constancy of regressions for beta distributions. Sankhyā A **65**, 284–291 (2003)
16. Walker, S., Muliere, P.: A characterization of a neutral to the right prior via an extension of Johnson's sufficientness postulate. Ann. Stat. **27**, 589–599 (1999)