Journal of Multivariate Analysis 129 (2014) 1–15

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Dirichlet distribution through neutralities with respect to two partitions

Agata Sakowicz *, Jacek Wesołowski

Politechnika Warszawska, Wydział Matematyki i Nauk Informacyjnych, ul. Koszykowa 75, 00-662 Warsaw, Poland

ARTICLE INFO

Article history:
Received 15 November 2013
Available online 18 April 2014

AMS 2010 subject classifications:
62H05
62E10
39B22

Keywords:
Neutrality with respect to partition
Dirichlet distribution
Method of moments
Characterization

ABSTRACT

Different concepts of neutrality have been studied in the literature in context of independence properties of vectors of random probabilities, in particular, for Dirichlet random vectors. Some neutrality conditions led to characterizations of the Dirichlet distribution. In this paper we provide a new characterization in terms of neutrality with respect to two partitions, which generalizes previous results. In particular, no restrictions on the size of the vector of random probabilities are imposed. In the proof we enhance the moments method approach proposed in Bobecka and Wesołowski (2009) [2] by combining it with some graph theoretic techniques.

1. Introduction

Let $\underline{X} = (X_1, \ldots, X_n)$ be a random vector of probabilities, i.e. $X_i \geq 0$, $i = 1, \ldots, n$, and $\sum_{i=1}^{n} X_i = 1$. A concept of neutrality was first introduced for such vectors by Connor and Mosimann in [5]. They indicated that given a vector of random probabilities $\underline{X} = (X_1, \ldots, X_n)$ it is desirable in some situations to eliminate one of the proportions, say X_1, and to analyse its effects on proportions of the form $X_2/(1 - X_1), \ldots, X_n/(1 - X_1)$. This led to the following definition (see [5]): X_1 is neutral in $\underline{X}$ whenever X_1 and the vector $(X_2/(1 - X_1), \ldots, X_n/(1 - X_1))$ are independent. These authors defined also neutrality of a subvector in a random vector of proportions and complete neutrality of a vector. Similar notions of neutrality to the right and neutrality to the left were defined in Doksum [7]. There were also other related notions of neutrality studied in the literature. All these notions embed in the notion of neutrality with respect to partition of an index set (introduced in [2]) which we recall below.

We say that $\pi = \{P_1, \ldots, P_K\}$ is a partition of a set E when $P_1, \ldots, P_K$ are nonempty pairwise disjoint subsets of E, whose union is E. The elements of π are called blocks.

Definition 1.1. Let $\pi = \{P_1, \ldots, P_K\}$ be a partition of $E = \{1, \ldots, n\}$. We say that a vector of random probabilities $\underline{X} = (X_1, \ldots, X_n)$ is neutral with respect to π (from here abbreviated nwrt π) if the following random vectors are mutually independent:

$$U = \left(\sum_{i \in P_1} X_i, \ldots, \sum_{i \in P_K} X_i\right),$$

$$W_{P_1} = \left(\frac{X_j}{\sum_{i \in P_1} X_i}, j \in P_1\right), \ldots, W_{P_K} = \left(\frac{X_j}{\sum_{i \in P_K} X_i}, j \in P_K\right).$$

* Corresponding author.
E-mail addresses: A.Sakowicz@mini.pw.edu.pl (A. Sakowicz), wesolo@mini.pw.edu.pl (J. Wesołowski).

http://dx.doi.org/10.1016/j.jmva.2014.04.004
0047-259X/

The notion of neutrality appeared to be a useful tool in studying independence properties of the Dirichlet distribution. In particular, the Dirichlet distribution, which can be imposed on a vector of random probabilities $\underline{X}$, is neutral with respect to all possible partitions of the corresponding index set. For a recent accounts on Dirichlet distributions, including relations to neutrality concepts see e.g. Ng, Tian and Tang [13] (in particular, Ch. 2.6) or Chang, Gupta and Richards [3].

Recall that a random vector $\underline{X} = (X_1, \dots, X_{n-1})$ has the Dirichlet distribution $\mathrm{Dir}(\alpha_1, \dots, \alpha_n)$ if its density is of the form

$$f(x_1, \dots, x_{n-1}) = \frac{\Gamma\left(\sum_{i=1}^{n} \alpha_i\right)}{\prod_{i=1}^{n} \Gamma(\alpha_i)} \prod_{i=1}^{n-1} x_i^{\alpha_i - 1} \left(1 - \sum_{i=1}^{n-1} x_i\right)^{\alpha_n - 1} \mathbb{1}_{T_{n-1}}(x_1, \dots, x_{n-1}),$$

where $\alpha_i > 0$, $i = 1, \dots, n$, and $T_{n-1} = \{(x_1, \dots, x_{n-1}) : x_i > 0, \sum_{i=1}^{n-1} x_i < 1\}$. In the sequel we will say that a vector of random probabilities $\underline{X} = (X_1, \dots, X_n)$ has the Dirichlet distribution if a subvector $(X_1, \dots, X_{n-1})$ has the density given by the above formula.

Characterizations of the Dirichlet distribution by different independence assumptions related to neutralities were discussed by several authors. All these results can be formulated in terms of neutrality with respect to partitions, although this notion had not been explicitly referred to. Darroch and Ratcliff proved in [6] a characterization of the Dirichlet distribution, using neutralities with respect to partitions $\pi_i = \{\{1\}, \dots, \{i-1\}, \{i+1\}, \dots, \{n-1\}, \{i, n\}\}$, $i = 1, \dots, n-1$. A result by Fabius, [8], concerned partitions $\pi_i = \{\{i\}, \{1, \dots, i-1, i+1, \dots, n\}\}$, $i = 1, \dots, n-1$. James and Mosimann presented in [10] a characterization by neutrality with respect to partitions $\pi_i = \{\{1\}, \dots, \{i\}, \{i+1, \dots, n\}\}$, $i = 1, \dots, n-2$, and $\pi_{n-1} = \{\{n-1\}, \{1, \dots, n-2, n\}\}$. Their result was further generalized in [1] by Bobecka and Wesołowski, where partitions $\pi_i = \{\{1\}, \dots, \{i\}, \{i+1, \dots, n\}\}$, $i = 1, \dots, n-2$, and $\pi_{n-1} = \{\{i_0\}, \{i_0+1\}, \dots, \{n-1\}, \{1, \dots, i_0-1, n\}\}$ for an arbitrary fixed i_0 were considered. Note that for a vector of size n all of these characterizations require exactly $n - 1$ partitions. Another result, requiring only 2 partitions for any n not being a prime number, was presented in [9] by Geiger and Heckerman. These authors proved, assuming the existence of a density, a characterization of an $L \times M$ Dirichlet random matrix, with one partition determined by its rows and another one by its columns. The proof was based on solving a functional equation for densities (further developed by Járai in [11], see also Chapter 23 of [12]). See also Heckerman, Geiger and Chickering [4] for a thorough description of the Bayesian networks context of this characterization. Bobecka and Wesołowski, [2], refined this result proving an analogous characterization by means of moments method and, consequently, without additional density assumption. They also generalized it to multi-way tables. The result of Geiger and Heckerman (and its extension) has been also recently proved within the Bayesian framework by Ramamoorthi and Sangalli in [15].

The aim of this paper is to present a new characterization of the Dirichlet distribution, which generalizes all the previous results when neutrality with respect to only two partitions is assumed. Actually we determine a set of all pairs of partitions such that neutrality with respect to both elements of the pair characterizes the Dirichlet distribution for the vector of random probabilities. We use the moments method as in [2], hence no density assumption is needed. In the proof we also rely heavily on graph theoretic techniques.

The paper is organized as follows. Section 2 contains some definitions and facts from graph theory which are used in the proof of the main result. In Section 3 we state and prove the characterization of the Dirichlet law, which is our main result. In the proof we use also an auxiliary result on a functional equation which is formulated in Section 3 and proved in the Appendix. In Section 4 we illustrate the characterization with several examples.

2. Facts from graph theory

In this section we present some definitions and facts from graph theory that will be used in the proof of our main result.

Definition 2.1. Let $\mathcal{G} = (V, E)$, where V is the set of vertices and E is the set of edges, be a connected graph. The vertex $v \in V$ is called a cut vertex if removing v from $\mathcal{G}$ disconnects the graph. Otherwise, we say that v is a non-cut vertex. Alternatively, v is non-cut if for any $u, w \in V \setminus \{v\}$ there exists a path between u and w that does not contain v.

Below we state a fact on non-cut vertices, which belongs to the graph theory folklore. Since we were not able to find an exact reference, a short proof is also given.

Lemma 2.2. *In every connected graph $\mathcal{G} = (V, E)$, $|V| \geq 2$, there exist at least two non-cut vertices.*

Proof. Let $\mathcal{T}$ denote a spanning tree of the graph $\mathcal{G}$. Since $|V| \geq 2$, there exist at least two leaves u, v in $\mathcal{T}$. As the removal of leaves does not disconnect the tree, u, v are non-cut in $\mathcal{T}$, and hence they are non-cut in $\mathcal{G}$. $\square$

For the purpose of this paper we introduce a notion of significance of a vertex.

Definition 2.3. Let C be a maximal clique in a graph $\mathcal{G}$ and $v \in C$. Denote by $N(v)$ the set of neighbours of v. We say that v is significant in C if $N(v) \cup \{v\} = C$.