Electronic Journal of Probability

Vol. 14 (2009), Paper no. 4, pages 72–89.

Journal URL
http://www.math.washington.edu/~ejpecp/

De Finetti's-type results for some families of non identically distributed random variables

Ricardo Vélez Ibarrola * and Tomás Prieto-Rumeau
Statistics Deparment, UNED, Madrid, Spain

Abstract

We consider random selection processes of weighted elements in an arbitrary set. Their conditional distributions are shown to be a generalization of the hypergeometric distribution, while the marginal distributions can always be chosen as generalized binomial distributions. Then we propose sufficient conditions on the weight function ensuring that the marginal distributions are necessarily of the generalized binomial form. In these cases, the corresponding indicator random variables are conditionally independent (as in the classical De Finetti theorem) though they are neither exchangeable nor identically distributed.

Key words: De Finetti theorem; Exchangeability; Random assignment processes.

AMS 2000 Subject Classification: Primary 60G09.

Submitted to EJP on April 7, 2008, final version accepted December 11, 2008.

*Correspondence to: Ricardo Vélez Ibarrola. Departamento de Estadística, Facultad de Ciencias, UNED. Calle Senda del Rey 9, 28040 Madrid, Spain. Tel: (+34) 913 987 258. Fax: (+34) 913 987 261. e-mail: rvelez@ccia.uned.es

1 Introduction

The De Finetti theorem states that any infinite sequence of exchangeable random variables is conditionally i.i.d. This result was first proved for $\{0,1\}$-valued random variables (see, e.g., [4]), and it was then extended to sequences of exchangeable random variables taking values in more general domains; see, for instance, [5; 7; 8].

Our main result is the construction of a class of $\{0,1\}$-valued non identically distributed random variables that satisfy a De Finetti-type result. More precisely, we consider random selection processes under fairly general conditions, and we show that, in some cases, their associated indicator random variables are conditionally independent though not identically distributed.

In order to give a precise definition of a random selection process, let E be an arbitrary nonempty set. For simplicity, one can think of E as a denumerable set, but this is by no means necessary. We are interested in random selection procedures of elements of E that will be characterized by a family of random variables

$$N(x) = \begin{cases} 1 & \text{if } x \text{ is selected} \\ 0 & \text{otherwise,} \end{cases} \qquad (x \in E)$$

defined on a common probability space $(\Omega, \mathscr{F}, \mathrm{P})$. We will write $\boldsymbol{N} \equiv \{N(x)\}_{x\in E}$. For each finite subset $A \subseteq E$, we define the random variable

$$N(A) = \sum_{x\in A} N(x).$$

Let us also consider a *weight function* $m : E \to (0,\infty)$ which assigns to each element x of E a weight $m(x) > 0$. We will say that $\boldsymbol{N}$ is a *random selection process*, with respect to m, if the following conditions are satisfied for every finite and nonempty $A \subseteq E$:

(a) Conditional on $N(A)$, $\{N(x)\}_{x\in A}$ and $\{N(x)\}_{x\in A^c}$ are independent.

(b) $\mathrm{P}\{N(x) = 1 \mid N(A) = 1\} = \dfrac{m(x)}{m(A)}$ for all $x \in A$, where $m(A) = \sum_{x\in A} m(x)$.

The motivation of the above assumptions is clear. Assumption (a) states a Markovian property: once we know the number of elements selected in A, the choices inside A are independent of the choices outside A. Assumption (b) means, roughly speaking, that an element is chosen with a probability proportional to its weight.

In Section 2 we prove the existence of a family of such random selection processes for any set E and any weight function m. In terms of the $\{0,1\}$-valued stochastic process $\{N(x)\}_{x\in E}$, this means being able to specify a set of compatible finite-dimensional distributions satisfying the Assumptions (a) and (b). We will also analyze the conditional and marginal distributions of $\boldsymbol{N}$. This gives rise to interesting generalizations of the hypergeometric and binomial distributions, respectively, in which the weight function is involved.

In Section 3 we prove two De Finetti-type results on the conditional independence of the indicator variables $\{N(x)\}_{x\in E}$. The first one assumes the existence of a sequence of elements of E with the same weight (the corresponding $N(x)$ are then exchangeable) and the conditional independence property is extended to the remaining variables that are no longer identically distributed.

The second De Finetti-type result assumes the existence of a sequence of elements in E with weights in $\mathbb{N}$, and then proves the conditional independence of the whole family of random variables $\{N(x)\}_{x\in E}$, which are not identically distributed.

In Section 4 we explore the relations of our setting with the generalized Stirling triangles of the first kind, and we conclude in Section 5 with some general remarks.

2 Generalized hypergeometric and binomial distributions

Let us consider a random selection process $\boldsymbol{N}$ on a set E with positive weights $\{m(x)\}_{x\in E}$. In the sequel, for a finite subset A of E, we will denote by $\#A$ the cardinal of A.

Conditional distributions of $\boldsymbol{N}$

Our first result is a generalization of Assumption (b).

Proposition 2.1. *Let $A \subseteq E$ be finite and, for each $1 \le r \le \#A$, let $x_1, x_2, \dots, x_r$ be distinct elements of A. Then*

$$\mathrm{P}\{N(x_1) = 1, N(x_2) = 1, \dots, N(x_r) = 1 \mid N(A) = r\} = \frac{1}{M_r(A)} \prod_{j=1}^{r} m(x_j), \tag{2.1}$$

where $M_r(A)$ is the sum of the products $\prod_{j=1}^{r} m(x_j)$ over all the subsets $\{x_1, \dots, x_r\}$ of A with r elements.

Proof: We will make the proof by induction on r. For $r = 1$, (2.1) holds as a direct consequence of Assumption (b). Suppose now that (2.1) holds for $r-1$. Fix $A \subseteq E$ with $\#A \ge r$ and r distinct points $x_1, x_2, \dots, x_r$ in A. We have

$$\begin{aligned}
&\mathrm{P}\{N(x_1) = 1, \dots, N(x_r) = 1 \mid N(A) = r\} = \\
&\quad = \mathrm{P}\{N(x_1) = 1 \mid N(A) = r\}\, \mathrm{P}\{N(x_2) = 1, \dots, N(x_r) = 1 \mid N(A) = r, N(x_1) = 1\} \\
&\quad = \mathrm{P}\{N(x_1) = 1 \mid N(A) = r\}\, \mathrm{P}\{N(x_2) = 1, \dots, N(x_r) = 1 \mid N(A - \{x_1\}) = r - 1\} \\
&\quad = \mathrm{P}\{N(x_1) = 1 \mid N(A) = r\}\, \frac{1}{M_{r-1}(A - \{x_1\})} \prod_{j=2}^{r} m(x_j) \\
&\quad = \frac{\mathrm{P}\{N(x_1) = 1 \mid N(A) = r\}}{m(x_1) M_{r-1}(A - \{x_1\})} \prod_{j=1}^{r} m(x_j)
\end{aligned}$$

(where the second equality follows from Assumption (a) and the third equality from the induction hypothesis). Now, x_1 may be replaced with any other x_j and thus the left-hand term of the last expression must necessarily have the same value, say C, for any element of A. Therefore,

$$\mathrm{P}\{N(x_1) = 1, \dots, N(x_r) = 1 \mid N(A) = r\} = C \prod_{l=1}^{r} m(x_l).$$

Since the above expressions sum up to one (for all possible $x_1 \neq \dots \neq x_r$ in A), we get $C = 1/M_r(A)$, which completes the proof. ∎

Remark 2.2. *The values* $M_0(A)=1$*,* $M_1(A)=m(A)$*,* $M_2(A),\ldots,M_{\#A}(A)=\prod_{x\in A}m(x)$ *can be easily determined as the successive coefficients of the polynomial*

$$\prod_{x\in A}[1+m(x)z].$$

For instance, if $m(x)=1$ *for all* $x\in A$*, then* $M_r(A)$ *is the binomial coefficient* $M_r(A)=\binom{\#A}{r}$*.*

Proposition 2.3 below gives an explicit expression for the conditional distributions of $\boldsymbol{N}$.

Proposition 2.3. *Let* $B\subseteq A$ *be finite subsets of* E *and fix* $1\le r\le \#A$*. For each* s *such that* $0\le s\le \#B$

$$\mathrm{P}\{N(B)=s\,|\,N(A)=r\}=\begin{cases} M_s(B)\,M_{r-s}(A-B)/M_r(A) & \text{if } r+\#B-\#A\le s\le r,\\ 0 & \text{otherwise.}\end{cases} \tag{2.2}$$

Proof: The result easily follows by summing up (2.1) for all the r-uples in A with $x_1,\ldots,x_s\in B$ and $x_{s+1},\ldots,x_r\in A-B$. ■

It is worth noting that the distribution (2.2) is a generalization of the hypergeometric distribution, to which it reduces when $m(x)=1$ for all $x\in A$, and thus it could be named "hypergeometric distribution with weights". As a first conclusion, we have proved that the Assumptions (a) and (b) uniquely determine the conditional distribution of $N(B)$ given $N(A)$ (for all finite sets $B\subseteq A$), which coincides with the above mentioned "hypergeometric distribution".

Marginal distributions of $\boldsymbol{N}$

In case that E is a finite set, the distribution of $N(E)$ can be arbitrarily fixed (among the distributions on $\{0,1,\ldots,\#E\}$) and then the distribution of $N(B)$, for any subset B, can be deduced from (2.2). The set of all possible distributions for $N(B)$ will be the convex hull of the extreme "hypergeometric distributions" $\big\{\mathrm{P}\{N(B)=s\,|\,N(E)=r\}:s=0,1,\ldots,\#B\big\}$ with $r=0,1,\ldots,\#E$. Observe that each of these distributions depends on the weights of all the elements in E (as well on the weights in B as on those in B^c).

If E is not finite, the key issue is to determine a family of marginal distributions $\mathrm{P}\{N(B)=s\}$ that are compatible with the conditional distributions given in Proposition 2.3, that is, a family of marginal distributions satisfying the total probability rule

$$\mathrm{P}\{N(B)=s\}=\sum_{r=0}^{\#C}\mathrm{P}\{N(B\cup C)=s+r\}\,\frac{M_s(B)M_r(C)}{M_{s+r}(B\cup C)} \tag{2.3}$$

for every finite sets B and $C\subseteq B^c$, and every $0\le s\le \#B$. For future reference, we state this result, without proof, in our next lemma.

Lemma 2.4. *A family of probability distributions* $\{p_j(A)\}_{0\le j\le \#A}$*, for each finite* $A\subseteq E$*, are the marginal distributions of a random selection process* $\boldsymbol{N}$ *(that is,* $\mathrm{P}\{N(A)=j\}=p_j(A)$*) with respect to a weight function* m*, if and only if*

$$p_s(B)=\sum_{r=0}^{\#C}p_{s+r}(B\cup C)\,\frac{M_s(B)M_r(C)}{M_{s+r}(B\cup C)} \tag{2.4}$$

for every finite sets B and $C \subseteq B^c$ and every $0 \le s \le \#B$.

In this case, $\mathrm{P}\{N(A) = j\} = p_j(A)$ for each finite $A \subseteq E$ and $0 \le j \le \#A$, together with (2.2), gives the joint distribution of $(N(A_1), \dots, N(A_k))$ for any finite sets $A_1, \dots, A_k$ and, therefore, the finite-dimensional distributions of $\{N(x)\}_{x \in E}$.

Obviously, there are two degenerate solutions: $N(B) = 0,\ \forall\ B$, and $N(B) = \#B,\ \forall\ B$, in which nothing is random. For the existence of a (properly speaking) random selection process, we are interested in solutions such that $p_j(A) > 0$ for some finite A and some $0 < j < \#A$.

The following theorem gives a universal solution to the existence problem.

Theorem 2.5. *Let m be a weight function on an infinite set E, and fix an arbitrary $\xi \in [0,1]$. For any finite $B \subseteq E$ and $s = 0, 1, \dots, \#B$ define $p_s(B)$ as*

$$p_s(B) = \frac{M_s(B)\, \xi^s (1-\xi)^{\#B-s}}{\prod_{x \in B} [1 - \xi + \xi m(x)]}. \tag{2.5}$$

Then the $p_s(B)$ verify Lemma 2.4. Similarly, given any fixed distribution F on $[0,1]$, the family

$$p_s(B) = \int_0^1 \frac{M_s(B)\, \xi^s (1-\xi)^{\#B-s}}{\prod_{x \in B} [1 - \xi + \xi m(x)]}\, F(d\xi) \tag{2.6}$$

*also verifies Lemma 2.4, so that there exists a random selection process **N** with marginal distributions given by (2.6).*

Remark 2.6. *Note that the obvious equality*

$$\sum_{s=0}^{\#B} M_s(B) \xi^s (1-\xi)^{\#B-s} = \prod_{x \in B} [1 - \xi + \xi m(x)]$$

ensures that (2.5) is a probability distribution. The distribution (2.5) is a generalization of the ordinary binomial distribution, which is obtained when $m(x) = 1$ for all $x \in B$. It can be observed that (2.5) only depends on the weights of the elements in B (and is independent of those in B^c). The set of solutions given in (2.6) is the convex hull of the extreme "generalized binomial distributions" (2.5). The two degenerate solutions correspond, respectively, to $\xi = 0$ and $\xi = 1$, and for any $\xi \in (0,1)$ we get a non-degenerate solution.

Proof: In (2.4), we replace $p_s(B)$ and $p_{s+r}(B \cup C)$ with their values taken from (2.5), and it follows that (2.4) is equivalent to the identity

$$\prod_{x \in B \cup C} [1 - \xi + \xi m(x)] = \prod_{x \in B} [1 - \xi + \xi m(x)] \sum_{r=0}^{\#C} M_r(C) \xi^r (1-\xi)^{\#C-r}.$$

By averaging (2.4) with respect to an arbitrary distribution $F(d\xi)$, it follows that (2.4) also holds when the $p_s(B)$ are defined as in (2.6). ∎

Remark 2.7. *For any fixed $\xi \in [0,1]$, the distributions in (2.5) give, for all $x \in E$,*

$$\mathrm{P}\{N(x) = 0\} = \frac{1-\xi}{1 - \xi + \xi m(x)} \quad \textit{and} \quad \mathrm{P}\{N(x) = 1\} = \frac{\xi m(x)}{1 - \xi + \xi m(x)}, \tag{2.7}$$

while if $B = \{x_1, \ldots, x_n\}$, according to Proposition 2.1 we have

$$\begin{aligned}
\mathrm{P}\{N(x_1) = 1, \ldots, N(x_s) = 1, N(x_{s+1}) = 0, \ldots, N(x_n) = 0\} = \\
&= \mathrm{P}\{N(B) = s\}\ \mathrm{P}\{N(x_1) = 1, \ldots, N(x_s) = 1 \mid N(B) = s\} \\
&= \frac{M_s(B)\,\xi^s(1-\xi)^{\#B-s}}{\prod_{x\in B}[1-\xi+\xi m(x)]}\ \frac{\prod_{j=1}^{s} m(x_j)}{M_s(B)} \\
&= \prod_{j=1}^{s} \mathrm{P}\{N(x_j) = 1\} \prod_{j=s+1}^{n} \mathrm{P}\{N(x_j) = 0\}.
\end{aligned} \tag{2.8}$$

Hence, under (2.5), the random variables $\{N(x)\}_{x\in E}$ are independent but not identically distributed (indeed, the distributions (2.7) depend on $m(x)$). When ξ has a certain distribution F, as in (2.6), independence usually does not hold, and the average of (2.8) yields a wide class of finite-dimensional distributions for the process $\{N(x)\}_{x\in E}$. Now, the obvious question is whether there exist further classes of distributions satisfying Lemma 2.4 (as in the case of a finite E, where the $p_s(E)$ can be arbitrarily chosen, yielding marginal distributions $p_s(B)$ that depend on the weights in B^c).

The above Remark 2.7 bears in mind De Finetti's theorem for exchangeable sequences of random variables (see, e.g., [1, Section 7.3], [5, Chapter 27] or [8, Chapter 9]). In fact, if $E = \mathbb{N}$ and the weight function were a constant m, then $\{N(x)\}_{x\in\mathbb{N}}$ would be a sequence of exchangeable random variables and De Finetti's theorem would ensure that, necessarily, the distribution of $\{N(x)\}_{x\in E}$ corresponds to the choice of a value θ, with any distribution F on $[0,1]$, while, conditional on θ, the $N(x)$ are independent with common distribution: $\mathrm{P}\{N(x) = 1 \mid \theta\} = \theta = \xi m/(1-\xi+\xi m)$. A similar result will hold provided that we can prove that the distributions described in Theorem 2.5 are the only possible marginal distributions for a random selection process. This issue will be analyzed in Section 3.

3 Extensions of De Finetti's theorem

Thus, our goal is to give conditions on the weight function m ensuring that the marginal probability distributions of the random selection process are precisely of the form (2.6) for some distribution F on $[0,1]$.

But first we need some preliminary results. In the following we suppose that the $p_s(B)$, for finite $B \subseteq E$ and $s = 0, 1, \ldots, \#B$, verify the hypotheses of Lemma 2.4.

Lemma 3.1. *Let $B = \{x_1, \ldots, x_k\}$ and $B' = \{x'_1, \ldots, x'_k\}$ be finite subsets of E and suppose that there exists a permutation π of $\{1, \ldots, k\}$ such that $m(x_j) = m(x'_{\pi(j)})$ for each $1 \le j \le k$. Then $p_s(B) = p_s(B')$ for every $0 \le s \le k$. (In short, this lemma states that $p_s(B)$ is a symmetric function of the weights in B.)*

Proof: Let $C = B' - B$ and $C' = B - B'$. Then (2.4) gives

$$p_s(B) = \sum_{r=0}^{\#C} p_{s+r}(B \cup C)\,\frac{M_s(B)M_r(C)}{M_{s+r}(B\cup C)} = \sum_{r=0}^{\#C'} p_{s+r}(B' \cup C')\,\frac{M_s(B')M_r(C')}{M_{s+r}(B'\cup C')} = p_s(B')$$

since $B \cup C = B' \cup C'$ and the weights in B and C are the same as those in B' and C', respectively. This completes the proof. ■

As a consequence of Lemma 3.1, we may write $p_s(B) = p_s(m_1, \ldots, m_k)$ as a function of the weights in B.

Lemma 3.2. *Suppose that the families $\{p_s(B)\}$ and $\{\overline{p}_s(B)\}$ satisfy Lemma 2.4, and that $p_0(B) = \overline{p}_0(B)$ for every finite set $B \subseteq E$. Then $\{p_s(B)\} \equiv \{\overline{p}_s(B)\}$. (Loosely speaking, this lemma shows that $\{p_s(B)\}$ is uniquely determined by $\{p_0(B)\}$.)*

Proof: We will show that $p_s(B) = \overline{p}_s(B)$ for every finite set B by induction on s. The case $s = 0$ is precisely the hypothesis of this lemma. Suppose now that the stated property holds for some $s \geq 0$. Given a finite subset B of E and $x \in B$, we have

$$p_s(B - \{x\}) = p_s(B)\, \frac{M_s(B - \{x\})}{M_s(B)} + p_{s+1}(B)\, \frac{M_s(B - \{x\})m(x)}{M_{s+1}(B)}$$

and

$$p_s(B - \{x\}) = p_s(B)\, \frac{M_s(B - \{x\})}{M_s(B)} + \overline{p}_{s+1}(B)\, \frac{M_s(B - \{x\})m(x)}{M_{s+1}(B)},$$

which proves that $p_{s+1}(B) = \overline{p}_{s+1}(B)$. ■

The next result exploits the well known result for exchangeable sequences.

Theorem 3.3. *Consider a random selection process $\mathbf{N}$ with marginal distributions $p_s(B)$ satisfying Lemma 2.4, and let us assume that there exists a denumerable subset $E^\star = \{x_1^\star, x_2^\star, \ldots\}$ of E with $m(x_i^\star) = h$ for every $i \geq 1$ and some $h > 0$. Then, there exists a probability distribution F on $[0,1]$ such that*

$$p_s(B) = \int_0^1 \frac{M_s(B)\, \xi^s (1-\xi)^{\#B-s}}{\prod_{x \in B}[1 - \xi + \xi m(x)]}\, F(d\xi) \tag{3.1}$$

for every finite $B \subseteq E$ and $0 \leq s \leq \#B$. Hence the random variables $\{N(x)\}_{x \in E}$ are conditionally independent given ξ and satisfy (2.7).

Proof: Without any loss of generality we can assume that $h = 1$, since a change of the weights scale leaves the problem unaffected.

First of all, observe that the sequence of random variables $\{N(x^\star)\}_{x^\star \in E^\star}$ is exchangeable. Indeed, according to (2.1) and letting $E_n^\star = \{x_1^\star, \ldots, x_n^\star\}$, we have that $p_s(E_n^\star)/M_s(E_n^\star)$ is the probability that s arbitrary variables among $\{N(x^\star)\}_{x^\star \in E_n^\star}$ take the value 1 while the remaining $n - s$ take the value 0. Therefore (cf. [4, p. 228]), there exists a probability distribution F on $[0,1]$ such that

$$\frac{p_s(E_n^\star)}{M_s(E_n^\star)} = \int_0^1 \xi^s (1-\xi)^{n-s}\, F(d\xi). \tag{3.2}$$

Next, we will focus on the extreme points of (3.2), that is, we will assume that F is concentrated at some $\xi \in [0,1)$ (the case $\xi = 1$ will be analyzed later). We thus have

$$p_s(E_n^\star) = \binom{n}{s} \xi^s (1-\xi)^{n-s} \quad \text{for some } \xi \in [0,1),$$

so that the random variables $N(x^\star)$ for $x \in E^\star$ are i.i.d. with $\mathrm{P}\{N(x^\star)=1\}=\xi$, for $x^\star \in E^\star$.

Now, let B be a finite subset of $E-E^\star$. The strong law of large numbers ensures that

$$\frac{N(B\cup E_n^\star)}{n} \xrightarrow{\text{a.s.}} \xi \qquad \text{as } n\to\infty. \tag{3.3}$$

On the other hand, for each $y\in[0,1]$ and $n\geq 1$, we define

$$h_n(y)=\left[1+\sum_{s=1}^{\#B\wedge yn} M_s(B)\,\frac{\Gamma(yn+1)\Gamma(n-yn+1)}{\Gamma(yn-s+1)\Gamma(n-yn+s+1)}\right]^{-1}. \tag{3.4}$$

The h_n are continuous functions bounded by 1. The Stirling approximation easily gives that, whenever $y_n\to y\in[0,1)$, the coefficient of $M_s(B)$ in (3.4) converges to $y^s/(1-y)^s$ and, therefore,

$$h_n(y_n)\longrightarrow\left[1+\sum_{s=1}^{\#B}M_s(B)\frac{y^s}{(1-y)^s}\right]^{-1}=\left[\prod_{x\in B}\left[1+m(x)\frac{y}{1-y}\right]\right]^{-1}=\frac{(1-y)^{\#B}}{\prod_{x\in B}[1-y+m(x)y]}.$$

Now, according to (2.4), the value of $p_0(B)$ is given, for every n, by

$$p_0(B)=\sum_{r=0}^{n}p_r(B\cup E_n^\star)\,\frac{M_r(E_n^\star)}{M_r(B\cup E_n^\star)}=\sum_{r=0}^{n}p_r(B\cup E_n^\star)h_n(r/n)$$

since $M_r(B\cup E_n^\star)=\sum_{s=0}^{\#B\wedge r}M_s(B)\binom{n}{r-s}$. Therefore, we have

$$p_0(B)=\mathrm{E}\left[h_n\left(\frac{N(B\cup E_n^\star)}{n}\right)\mathbb{1}_{\{N(B\cup E_n^\star)\leq n\}}\right]$$

for every n. Then, by dominated convergence and (3.3),

$$p_0(B)=\frac{(1-\xi)^{\#B}}{\prod_{x\in B}[1-\xi+m(x)\xi]}\,,$$

as we wanted to prove. In case that $\xi=1$, $p_n(E_n^\star)=1$ and we easily obtain $p_0(B)=0$. The stated result is now derived from Lemma 3.2.

To conclude the proof, note that the extreme points of the convex set of solutions $p_s(B)$ are given by the integrand of (3.1), and then (3.1) is the general solution. ■

In summary, using the classical result about conditional independence of the random variables $N(x^\star)$, we have established the additional conditional independence of the remaining variables $N(x)$ that are no longer identically distributed. Now, our purpose is to show that the same result may be established without any reference to exchangeability properties, though this will require further efforts.

For finite $B\subseteq E$ and $0\leq s\leq \#B$, we define $g_s(B)=p_s(B)/M_s(B)$, for which the equation (2.4) gives

$$g_s(B)=\sum_{r=0}^{\#C}g_{s+r}(B\cup C)M_r(C). \tag{3.5}$$

By Lemma 3.1, g_s is a symmetric function of the weights of the elements in B. Thus, we will also write $g_s(m_1,\dots,m_n)$ when $m_1,\dots,m_n$ are the weights of the elements in B.

Lemma 3.4. *Suppose that $E' \subseteq E$ is such that $m(x) \neq m(x')$ for $x \neq x' \in E'$. Then the functions $g_s(B)$, for finite $B \subseteq E'$, are determined by the function $g_1(m)$ by means of the recursive formulas*

$$g_{k+1}(m_1, m_2, \dots, m_{n+1}) = \frac{g_k(m_1, \dots, m_n) - g_k(m_2, \dots, m_{n+1})}{m_{n+1} - m_1} \tag{3.6}$$

$$g_k(m_1, m_2, \dots, m_{n+1}) = \frac{m_{n+1} g_k(m_2, \dots, m_{n+1}) - m_1 g_k(m_1, \dots, m_n)}{m_{n+1} - m_1}. \tag{3.7}$$

Proof: The equations (3.6) and (3.7) are directly obtained from the particular cases of (3.5):

$$g_k(m_1, \dots, m_n) = g_k(m_1, \dots, m_{n+1}) + g_{k+1}(m_1, \dots, m_{n+1})\, m_{n+1}$$
$$g_k(m_2, \dots, m_{n+1}) = g_k(m_1, \dots, m_{n+1}) + g_{k+1}(m_1, \dots, m_{n+1})\, m_1. \quad \blacksquare$$

By means of the preceding lemma, we will now examine the case $E = \mathbb{N}$ and $m(i) = i$ for $i \geq 1$.

Proposition 3.5. *If $\{N(i)\}_{i \in \mathbb{N}}$ is a random selection process on $\mathbb{N}$ with $m(i) = i$ and marginal distributions $p_s(B)$, there exists a probability distribution F on $[0,1]$ such that*

$$p_s(B) = \int_0^1 \frac{M_s(B)\, \xi^s (1-\xi)^{\#B-s}}{\prod_{x \in B}[1 - \xi + \xi m(x)]}\, F(d\xi). \tag{3.8}$$

Proof: Consider the sequence $a_i = g_1(i+1)$, for $i \geq 0$, and note that, as a consequence of Lemma 3.4, it satisfies

$$g_2(i+1, i+2) = -\Delta a_i, \quad g_3(i+1, i+2, i+3) = \frac{1}{2}\, \Delta^2 a_i$$

and, in general,

$$g_k(i+1, i+2, \dots, i+k) = \frac{(-1)^{k-1}}{(k-1)!}\, \Delta^{k-1} a_i \quad \text{for } i \geq 1.$$

Since the left hand side of the above equation is nonnegative, it follows that $\{a_i\}_{i \geq 0}$ is a completely monotone sequence that can be expressed (see [4, Section VII.3]) as

$$a_i = a_0 \int_0^1 \theta^i\, \boldsymbol{A}(d\theta)$$

for a unique probability distribution $\boldsymbol{A}$ on $[0,1]$; also, $\boldsymbol{A}$ is the weak limit, as $n \to \infty$, of the discrete distributions giving probability

$$p_i^{(n)} = a_0^{-1} \binom{n}{i} (-1)^{n-i} \Delta^{n-i} a_i \tag{3.9}$$

to each point i/n, for $0 \leq i \leq n$. Note that

$$p_k(i+1, \dots, i+k) = \frac{(i+k)!}{i!(k-1)!}\, (-1)^{k-1} \Delta^{k-1} a_i = a_0\, p_i^{(i+k-1)} (i+k)$$

is less than or equal to 1, so that

$$p_i^{(n)} = \frac{p_{n+1-i}(i+1, \dots, n+1)}{a_0(n+1)} \leq \frac{1}{a_0(n+1)}.$$

Then $\boldsymbol{A}$ is absolutely continuous with respect to the Lebesgue measure and its density, denoted by $\boldsymbol{a}$, is bounded above by a_0^{-1} (see [2, Theorem 6] for a detailed proof). Moreover, since $p_{n-i}(i+2,\ldots,n+1) \geq p_{n-i+1}(i+1,\ldots,n+1)$, the terms $p_i^{(n)}$ are increasing in i and the density $\boldsymbol{a}$ is an increasing function (see [2, Theorem 10] and [2, Proposition 2]).

We have thus proved that

$$g_1(i+1) = a_0 \int_0^1 \theta^i \, \boldsymbol{a}(\theta) \, d\theta$$

and

$$(i+1)g_1(i+1) = a_0 \boldsymbol{a}(1) - a_0 \int_0^1 \theta^{i+1} \boldsymbol{a}(d\theta), \tag{3.10}$$

where (3.10) is obtained after integration by parts.

Now, the sequence $b_i = g_1(i+1,i+2) = (i+2)g_1(i+2) - (i+1)g_1(i+1)$, defined for $i \geq 0$, satisfies (recall (3.6))

$$g_{k-1}(i+1,\ldots,i+k) = \frac{(-1)^{k-2}}{(k-1)!} \Delta^{k-2} b_i \geq 0 \quad \text{for every } k \geq 2.$$

Thus the completely monotone sequence $\{b_i\}_{i\geq 0}$ has the form

$$b_i = b_0 \int_0^1 \theta^i \, \boldsymbol{B}(d\theta)$$

where $\boldsymbol{B}$ is the distribution function on $[0,1]$ obtained as the weak limit of (3.9) with the a_i replaced with the b_i, that is,

$$p_i^{(n)} = b_0^{-1} \binom{n}{i} (-1)^{n-i} \Delta^{n-i} b_i \quad \text{for } 0 \leq i \leq n \text{ and } n \geq 1.$$

We also have

$$\begin{aligned}
1 \geq p_{k-1}(i+1,\ldots,i+k) &= M_{k-1}(i+1,\ldots,i+k) \frac{(-1)^{k-2}}{(k-1)!} \Delta^{k-2} b_i \\
&= \frac{(i+k)!}{i!(k-1)!} \left[\frac{1}{i+1} + \cdots + \frac{1}{i+k} \right] \frac{b_0 \, p_i^{(i+k-2)}}{\binom{i+k-2}{i}} \\
&= \left[\frac{1}{i+1} + \cdots + \frac{1}{i+k} \right] b_0 \, p_i^{(i+k-2)} \frac{(i+k)(i+k-1)}{k-1} \\
&\geq b_0 \, p_i^{(i+k-2)} (i+k-1)
\end{aligned}$$

and, therefore, $p_i^{(n)} \leq b_0^{-1}/(n+1)$. Again, $\boldsymbol{B}$ has a density $\boldsymbol{b}$ which is bounded above by b_0^{-1}. The definition of b_i yields

$$(i+2)g_1(i+2) - (i+1)g_1(i+1) = b_0 \int_0^1 \theta^i \, \boldsymbol{b}(\theta) \, d\theta \tag{3.11}$$

and, summing these equalities from some i onwards, we obtain

$$(i+1)g_1(i+1) = p_1(\infty) - b_0 \int_0^1 \theta^i \frac{\boldsymbol{b}(\theta)}{1-\theta}\, d\theta \tag{3.12}$$

where $p_1(\infty)$ is the limit of the increasing function p_1. Comparing this expression with (3.10), we conclude that $a_0\boldsymbol{a}(1) = p_1(\infty)$ and that $\boldsymbol{a}'(\theta)$ exists and verifies

$$a_0\, \theta \boldsymbol{a}'(\theta) = b_0 \frac{\boldsymbol{b}(\theta)}{1-\theta}.$$

Our next step is to show that $\theta\boldsymbol{a}'(\theta)$, or $\boldsymbol{b}(\theta)/(1-\theta)$, is an increasing function. To this end, let $\{\overline{b}_k\}_{k\geq 0}$ be the moments of $\boldsymbol{b}(\theta)/(1-\theta)$ and define

$$\overline{p}_i^{(n+1)} = \frac{p_i^{(n)}}{1-i/(n+1)} = b_0^{-1} \frac{(n+1)!}{i!} g_{n-i+1}(i+1,\ldots,n+2),$$

which verifies

$$\overline{p}_0^{(n+1)} \leq \overline{p}_1^{(n+1)} \leq \ldots \leq \overline{p}_n^{(n+1)}. \tag{3.13}$$

On the other hand,

$$\overline{p}_i^{(n+1)} = \frac{1}{b_0}\binom{n+1}{i}(-1)^{n-i+1}\Delta^{n-i+1}\overline{b}_i \quad \text{for } 0 \leq i \leq n+1.$$

Therefore, if the inequalities in (3.13) were also true for $\overline{p}_{n+1}^{(n+1)}$, by [2, Theorem 10], the density of $\boldsymbol{b}(\theta)/(1-\theta)$, would be increasing. In fact, this can be overcome (and we skip the details) just by defining $\overline{p}_{n+1}^{(n+1)}$ as $\overline{p}_n^{(n+1)}$, so that the whole sequence (3.13) is now increasing, and then by suitably rescaling the $\overline{p}_i^{(n+1)}$. Thus, we conclude that $\boldsymbol{b}(\theta)/(1-\theta)$ increases with θ.

Once we know that $\theta\boldsymbol{a}'(\theta)$ is an increasing function, we perform another integration by parts in (3.10) in order to get

$$(i+1)^2 g_1(i+1) = (i+1)p_1(\infty) - a_0\boldsymbol{a}'(1) + a_0 \int_0^1 \theta^{i+1}\, d\big(\theta\boldsymbol{a}'(\theta)\big). \tag{3.14}$$

Define now $c_i = 2g_1(i+1,i+2,i+3) = (i+3)^2 g_1(i+3) - 2(i+2)^2 g_1(i+2) + (i+1)^2 g_1(i+1)$ for $i \geq 0$. This sequence verifies

$$0 \leq g_{k+1}(i+1,\ldots,i+k+3) = \frac{1}{(k+2)!}(-1)^k \Delta^k c_i \quad \text{for } k \geq 0,$$

so that $\{c_i\}$ is a completely monotone sequence, and it can be written

$$c_i = c_0 \int_0^1 \theta^i \boldsymbol{C}(d\theta), \quad \text{for } i \geq 0,$$

where $\boldsymbol{C}$ is a probability distribution on $[0,1]$ which is the weak limit of (3.9) with a_i replaced with c_i, that is

$$p_i^{(n)} = c_0^{-1}\binom{n}{i}(-1)^{n-i}\Delta^{n-i}c_i.$$

Since

$$\begin{aligned}
p_{k-2}(i+1,\ldots,i+k) &= \frac{M_{k-2}(i+1,\ldots,i+k)}{(k-1)!}\,\frac{c_0\,p_i^{(i+k-3)}}{\binom{i+k-3}{i}} \\
&= \frac{(i+k)!}{i!(k-1)!}\sum_{j>l=1}^{k}\frac{1}{(i+j)(i+l)}\,\frac{c_0\,p_i^{(i+k-3)}\,i!(k-3)!}{(i+k-3)!} \\
&\geq \frac{c_0}{2}\,p_i^{(i+k-3)}(i+k-2),
\end{aligned}$$

we have $p_i^{(n)} \leq 2c_0^{-1}/(n+1)$ and $\boldsymbol{C}$ has a density $\boldsymbol{c}$ bounded above by $2c_0^{-1}$. Hence,

$$(i+3)^2 g_1(i+3) - 2(i+2)^2 g_1(i+2) + (i+1)^2 g_1(i+1) = c_0\int_0^1 \theta^i \boldsymbol{c}(\theta)\,d\theta,$$

and summing these inequalities from i onwards, yields

$$(i+2)^2 g_1(i+2) - (i+1)^2 g_1(i+1) = p_1(\infty) - c_0\int_0^1 \theta^i\,\frac{\boldsymbol{c}(\theta)}{1-\theta}\,d\theta.$$

Combining this equation with (3.14), we derive that

$$p_1(\infty) - c_0\int_0^1 \theta^i\,\frac{\boldsymbol{c}(\theta)}{1-\theta}\,d\theta = p_1(\infty) - a_0\int_0^1 \theta^{i+1}(1-\theta)\,d\big(\theta\boldsymbol{a}'(\theta)\big).$$

This proves that $\boldsymbol{a}''(\theta)$ exists and verifies

$$a_0\,\theta\,[\boldsymbol{a}'(\theta)+\theta\boldsymbol{a}''(\theta)] = c_0\,\frac{\boldsymbol{c}(\theta)}{(1-\theta)^2}.$$

Since

$$\frac{p_i^{(n)}}{(1-i/(n+1))(1-i/(n+2))} = c_0^{-1}\frac{(n+2)!}{i!}g_{n+1-i}(i+1,\ldots,n+3)$$

increases with $0 \leq i \leq n$, we can prove as we did before that $\boldsymbol{c}(\theta)/(1-\theta)^2$ or $\theta\boldsymbol{a}'(\theta)+\theta^2\boldsymbol{a}''(\theta)$ are increasing functions. The reasoning can then be pursued iteratively.

As a first conclusion, we have that $\boldsymbol{a}(\theta)$ possesses derivatives of all orders, whose properties are more easily analyzed if we transform (3.10) by means of the change of variable $\theta = e^{-\lambda}$. We thus have

$$g_1(i+1) = a_0\int_0^\infty e^{-\lambda(i+1)}\tilde{\boldsymbol{a}}(\lambda)\,d\lambda$$

where $\tilde{\boldsymbol{a}}(\lambda) = \boldsymbol{a}(e^{-\lambda})$ is a decreasing function on $[0,\infty)$ with first derivative

$$\tilde{\boldsymbol{a}}'(\lambda) = -e^{-\lambda}\boldsymbol{a}'(e^{-\lambda}) = -\theta\boldsymbol{a}'(\theta) \leq 0,$$

which is an increasing function of λ. In turn,

$$\tilde{\boldsymbol{a}}''(\lambda) = e^{-\lambda}\boldsymbol{a}'(e^{-\lambda}) + e^{-2\lambda}\boldsymbol{a}''(e^{-\lambda}) = \theta\boldsymbol{a}'(\theta)+\theta^2\boldsymbol{a}''(\theta) \geq 0$$

is a decreasing function of λ; therefore $\tilde{\boldsymbol{a}}'''(\lambda) \leq 0$, etc. We conclude that $\tilde{\boldsymbol{a}}(\lambda)$ is a completely monotone function that can be therefore expressed as

$$\tilde{\boldsymbol{a}}(\lambda) = \int_0^\infty e^{-\lambda\rho} H(d\rho)$$

where H is a measure on $[0,\infty)$ with $H(\infty) = \tilde{\boldsymbol{a}}(0) = \boldsymbol{a}(1)$. Recalling that $a_0\boldsymbol{a}(1) = p_1(\infty)$ and letting $\bar{H}(\rho) = H(\rho)/H(\infty)$, we have

$$g_1(i) = p_1(\infty) \int_0^\infty e^{-\lambda i} \int_0^\infty e^{-\lambda\rho}\, \bar{H}(d\rho)\, d\lambda = p_1(\infty) \int_0^\infty \frac{1}{i+\rho}\, \bar{H}(d\rho).$$

If $\bar{F}$ is the probability distribution function on $(0,1]$ associated to $\bar{H}$ by means of the change of variable $\xi = (1+\rho)^{-1}$, we have

$$g_1(i) = p_1(\infty) \int_{(0,1]} \frac{\xi}{1-\xi+\xi i}\, \bar{F}(d\xi) = \int_0^1 \frac{\xi}{1-\xi+\xi i}\, F(d\xi)$$

where F gives probability $1 - p_1(\infty)$ to $\xi = 0$ and coincides with $p_1(\infty)\bar{F}$ on $(0,1]$. According to Lemma 3.4, this identifies the solution (3.8). ■

Finally, we extend the result to the case of a set E including a sequence with natural weights.

Theorem 3.6. *Consider a random selection process $\boldsymbol{N}$ with marginal distributions $p_s(B)$ satisfying Lemma 2.4, and let us assume that there exists a denumerable subset $E^\star = \{x_1^\star, x_2^\star, \ldots\}$ of E with $m(x_i^\star) = ih$ for every $i \geq 1$ and some $h > 0$. Then, there exists a probability distribution F on $[0,1]$ such that*

$$p_s(B) = \int_0^1 \frac{M_s(B)\, \xi^s (1-\xi)^{\#B-s}}{\prod_{x\in B}[1-\xi+\xi m(x)]}\, F(d\xi) \tag{3.15}$$

for every finite $B \subseteq E$ and $0 \leq s \leq \#B$. Hence the random variables $\{N(x)\}_{x\in E}$ are conditionally independent given ξ and they satisfy (2.7).

Proof: We can assume that $h = 1$. According to Proposition 3.5, there exists a random variable Ξ in $[0,1]$, with distribution F, such that, conditional on $\Xi = \xi$, the random variables $\{N(x_j^\star)\}_{x_j^\star \in E^\star}$ are independent with

$$\mathrm{P}\{N(x_j^\star) = 1\} = \frac{\xi j}{1-\xi+\xi j}.$$

Let $E_n^\star = \{x_1^\star, \ldots, x_n^\star\}$ and assume $\xi \in (0,1)$. Since $\sigma^2[N(x_j^\star)] = \xi(1-\xi)j/(1-\xi+\xi j)^2$ and

$$\sum_{j=2}^\infty \frac{\sigma^2[N(x_j^\star)]}{(\log j)^2} < \infty$$

the strong law of large numbers, applied to the sequence $1 - N(x_j^\star)$, gives

$$\frac{\bar{N}(E_n^\star)}{\log n} - \frac{1}{\log n} \sum_{j=1}^n \frac{1-\xi}{1-\xi+\xi j} \xrightarrow{\text{a.s.}} 0 \qquad \text{as } n \to \infty$$

where $\bar{N}(E_n^\star) = n - N(E_n^\star)$ is the number of elements not selected in $E_n^\star$. In other words

$$\frac{\bar{N}(E_n^\star)}{\log n} \xrightarrow{\text{a.s.}} \frac{1-\xi}{\xi} \qquad \text{and thus} \qquad \frac{N(E_n^\star)}{n} \xrightarrow{\text{a.s.}} 1 \qquad \text{as } n \to \infty. \tag{3.16}$$

We now need to show that, if r varies with n in such a way that $y_n = r_n/\log n \longrightarrow y \in (0,\infty)$, it is

$$\frac{M_{n-r-s}(E_n^\star)}{M_{n-r}(E_n^\star)} \longrightarrow y^{-s}.$$

This will be proved in Lemma 3.7 below and gives, for any finite set $B \subset E - E^\star$,

$$h_n(y_n) = \left[1 + \sum_{s=1}^{\#B\wedge(n-r_n)} M_s(B)\,\frac{M_{n-r_n-s}(E_n^\star)}{M_{n-r_n}(E_n^\star)}\right]^{-1} \longrightarrow \frac{y^{\#B}}{\prod_{x\in B}[y+m(x)]}.$$

But, according to (2.4), for all n, it is

$$p_0(B) = \sum_{r=0}^{n} p_{n-r}(B\cup E_n^\star)\,h_n\Big(\frac{r}{\log n}\Big) = \mathrm{E}\left[h_n\left(\frac{\bar{N}(B\cup E_n^\star)}{\log n}\right)\mathbb{1}_{\bar{N}(B\cup E_n^\star)\le n}\right]$$

hence, since (3.16) shows that $\bar{N}(B\cup E_n^\star)/\log n \xrightarrow{\text{a.s.}} (1-\xi)/\xi$, we get

$$p_0(B) = \frac{(1-\xi)^{\#B}}{\prod_{x\in B}[1-\xi+\xi m(x)]}.$$

In case that $\xi = 0$, it is $\mathrm{P}\{N(x_j^\star)=0\} = 1$ or $p_0(j) = 1$ for all $j \ge 1$. But $p_0(m)$ is a non increasing function of m, since (2.4) gives

$$p_0(m_1) - p_0(m_2) = p_1(m_1, m_2)\,\frac{m_2 - m_1}{m_1 + m_2}.$$

Thus $p_0(m) \equiv 1$ and $p_0(B) = 1$. Similarly, when $\xi = 1$, it is $\mathrm{P}\{N(x_j^\star) = 1\} = 1$, thus $p_0(m) \equiv 0$ and $p_0(B) = 0$. ■

Lemma 3.7. *Under the hypotheses of Theorem 3.6, if s is fixed and r varies with n in such a way that $r_n/\log n \to y \in (0,\infty)$, then*

$$\frac{M_{n-r-s}(E_n^\star)}{M_{n-r}(E_n^\star)} \longrightarrow y^{-s}.$$

Proof: Let be

$$m_n(r) = \frac{M_{n-r}(E_n^\star)}{n!} = \sum_{1\le j_1<\cdots<j_r\le n} \frac{1}{j_1 j_2 \cdots j_r}$$

and observe that

$$\sum_{1\le j_1<\cdots<j_r\le n} \frac{1}{j_1 j_2\cdots j_r}\left[\sum_{j=1}^{n}\frac{1}{j} - \sum_{i=1}^{r}\frac{1}{j_i}\right] = (r+1)m_n(r+1)$$

We thus have

$$m_n(r)\sum_{j=r+1}^{n}\frac{1}{j}\le (r+1)m_n(r+1)\le m_n(r)\sum_{j=1}^{n-r}\frac{1}{j}$$

or as well

$$\frac{r+1}{\sum_{j=1}^{n-r}1/j}\le\frac{m_n(r)}{m_n(r+1)}\le\frac{r+1}{\sum_{j=r+1}^{n}1/j}.$$

But $\sum_{j=1}^{n-r}1/j\sim\log(n-r)$ and $\sum_{j=r+1}^{n}1/j\sim\log n-\log r$; and, if $r/\log n\to y$, both extremes of the inequality converge to y. Therefore

$$\frac{M_{n-r-s}(E_n^\star)}{M_{n-r}(E_n^\star)}=\frac{m_n(r+s)}{m_n(r)}\longrightarrow y^{-s}.\quad\blacksquare$$

In summary, Theorem 3.6 states a De Finetti-type theorem for a class of non exchangeable and non identically distributed $\{0,1\}$-valued random variables.

Example 3.8. *If E is an interval in $\mathbb{R}$ and the weight function satisfies either*

- $m(x)$ *takes some value infinitely often or*
- $m(x)$ *is continuous and unbounded*

then either Theorem 3.3 or Theorem 3.6 hold. That is to say that real numbers can only be "randomly selected", according to their weights $m(x)$, if each one is independently chosen with probability

$$\frac{\xi m(x)}{1-\xi+\xi m(x)}$$

for some fixed or random value of ξ.

4 Relations with generalized Stirling triangles of the first kind

An anonymous referee has brought to our attention the fact that our setting is intimately related to the generalized Stirling triangles of the first kind, as defined in [9, Section 1.4]. We now summarize this point of view, with a terminology that emphasizes the relation with our model.

Let $E^\star=\{x_1^\star,x_2^\star,\ldots\}$ be any fixed sequence in E. A generalized Stirling triangle of the first kind is a branching graph, with rows corresponding to successive $x_n^\star$, in which each vertex $(x_{n-1}^\star,k)$ $(0\le k\le n-1)$ is connected to $(x_n^\star,k)$ and $(x_n^\star,k+1)$, and these edges have attributed multiplicities $a(x_n^\star)$ and $b(x_n^\star)$, respectively. A path from the initial vertex $(0,0)$ to $(x_n^\star,k)$ may be interpreted as the successive selections or rejections of the elements $(x_1^\star,\ldots,x_n^\star)$. More precisely, such a path $(x_i^\star,k_i)$ selects those $x_i^\star$ with $k_i=k_{i-1}+1$ and rejects $x_i^\star$ if $k_i=k_{i-1}$. Thus k_i gives the number of selected elements among $x_1^\star,\ldots,x_i^\star$.

Kerov [9] associates to each path a weight given by the product of the multiplicities of its edges and defines the *dimension* of the vertex $(x_n^\star,k)$ as the sum of the weights of all paths from $(0,0)$ to $(x_n^\star,k)$.

Since there is a bijective correspondence between these paths and the subsets B of $E_n^\star = \{x_1^\star, \ldots, x_n^\star\}$ with $\#B = k$, in our terms:

$$w(B) = \prod_{x \in B} b(x) \prod_{x \in E_n^\star - B} a(x) = \prod_{x \in E_n^\star} a(x) \prod_{x \in B} m(x) \quad \text{with} \quad m(x) = b(x)/a(x)$$

gives the weight of each path and the dimension of the vertex $(x_n^\star, k)$ is

$$\dim(n,k) = M_k(E_n^\star) \prod_{x \in E_n^\star} a(x).$$

Following [6, Section 3], we can consider the Markovian process $\{K_n\}$, with $0 \le K_n \le n$, such that

$$\mathrm{P}\{K_1 = k_1, K_2 = k_2, \ldots, K_{n-1} = k_{n-1} \mid K_n = k\} = \frac{\prod_{x \in B} m(x)}{M_k(E_n^\star)}$$

where $B = \{x_j^\star \in E_n^\star : k_j - k_{j-1} = 1\}$, which is the weight of the path associated with B divided by the dimension of the last vertex. Thus

$$\mathrm{P}\{K_m = j \mid K_n = k\} = \frac{M_j(E_m^\star) M_{k-j}(E_n^\star - E_m^\star)}{M_k(E_n^\star)}$$

for $m \le n$ and with $M_i(A) = 0$ if $i \notin \{1, \ldots \#A\}$. Obviously the last two formulas are closely related to (2.1) and (2.2), respectively. The distribution of the random walk $\{K_n\}$ depends on a family of marginal distributions, $\mathrm{P}\{K_n = k\}$, which have a meaning similar to that of our marginal distributions $p_s(B)$ and analogous requirements of compatibility:

$$V_{n,k} = a(x_{n+1}^\star) V_{n+1,k} + b(x_{n+1}^\star) V_{n+1,k+1} \quad \text{where} \quad V_{n,k} = \frac{\mathrm{P}\{K_n = k\}}{\dim(n,k)}.$$

The functions $V = (V_{n,k})$ satisfying the preceding backward recurrence are called *harmonic* (for the generalized Stirling triangle) and they constitute a convex set $\mathscr{V}$. Its extremes may be characterized in terms of the Martin boundary of the chain $\{K_n\}$ (cf. [6, Lemma 5]). Specifically, each extreme of $\mathscr{V}$ can be obtained as the limit

$$V_{n,k} = \lim_{\nu \to \infty} \frac{M_{k_\nu - k}(E_\nu^\star - E_n^\star)}{M_{k_\nu}(E_\nu^\star)}$$

along a suitable sequence k_ν. Thus, the analysis in [6, Section 3] is mainly concerned with the Martin boundary of the Markov chain $\{K_n\}$ and its properties. In this context, a proof of Proposition 3.5 is given in [6, Section 4.3] (together with similar results for the case where $m(x_n^\star) = n - \alpha k$ ($\alpha < 1$) depends also on k). More generally, Kerov conjectures in [9, Section 1.4] that, under the hypothesis $\sum_{i=1}^\infty 1/m(x_i^\star) = \infty$, the extremes of $\mathscr{V}$ are given by

$$V_{n,k}(\theta) = \frac{\theta^{n-k}}{\prod_{i=1}^n [a(x_i)\theta + b(x_i)]} \quad \text{for each} \quad \theta \in [0, \infty].$$

With the change of variables $\xi = 1/(\theta + 1)$, this yields the marginal distributions of the random walk

$$\mathrm{P}_\xi\{K_n = k\} = \frac{M_k(E_n^\star)\, \xi^k (1-\xi)^{n-k}}{\prod_{i=1}^n [1 - \xi + \xi m(x_i)]} \quad \text{for each} \quad \xi \in [0,1]$$

in total concordance with (2.5). In this sense, the Kerov conjecture asserts that a De Finetti-type result holds for any random selection process on (E, m) if E is countable and $\sum_{i=1}^\infty 1/m(i) = \infty$.

5 Conclusions

Before knowing the connection of our framework with the generalized Stirling triangles of the first kind, our own conjecture was that a De Finetti-like theorem should hold "for a large variety of weights functions". Kerov asserts that, for the case $E = \mathbb{N}$, the divergence of $\sum_{i\in E} 1/m(i)$ is a necessary condition and shows, in fact, that, when this series converges, the result does not hold (cf. [9, Theorem 5 p. 71]). However, our recent research suggests that the assumption $\sum_{i\in E} 1/m(i) = \infty$ is somehow inappropriate:

- As the problem arises from combinatorial problems, Kerov wishes to control the "rate of growth of the sequence $m(i)$". For us, $m(i)$ may be as well increasing or decreasing and, in this last case, the appropriate condition is $\sum_{i\in E} m(i) = \infty$. In fact, there is an obvious duality between these two cases, that can be observed from the symmetry of the generalized Stirling triangle: the change of selections/rejections and the interchange of $a(x)$ and $b(x)$ (in the description of the last section) leaves the problem invariant; while, if $a(x)$ and $b(x)$ remain unchanged, $m(x) = b(x)/a(x)$ is replaced with $a(x)/b(x)$.
- For the general case (E uncountable), the divergence of $\sum_k 1/m(x_k)$ along some subsequence x_k would suffice to establish the desired result. If E is countable there is no difference between both assumptions. But, if E is uncountable, it is easy to see that this second assumption is always satisfied.

It seems that Kerov's conjecture, stated in 1989, has been verified in a few number of cases. Kerov himself (cf. [9, Section 1.4]) refers to the case where all weights $m(i)$ appear infinitely often. Our Theorem 3.3 shows that only a single weight needs to be infinitely repeated. As just said, the case $E = \mathbb{N}$, $m(i) = i$ was proved in [6, Section 4.3] (and now, independently of the results therein, in Proposition 3.5).

An interesting contribution of this paper is that the weights of a suitable subsequence $\{x_i^\star\}_{i\geq 1}$ in E (in case that either $m(x_i^\star) = h$ or $m(x_i^\star) = ih$ for all $i \geq 1$ and some $h > 0$) suffices to determine the extreme probability distributions of the whole random selection process $\boldsymbol{N}$. In particular, Theorem 3.6 shows that the Kerov conjecture is also true for $E = \mathbb{N}$, $m(i) = i^{1/q}$ or $m(i) = \log_q i$ (with $q \in \mathbb{N}, q \geq 2$ fixed) since $m(i^q) = i$ and $m(q^i) = i$, respectively.

When dealing with an arbitrary weight function, it seems that the Martin boundary approach has not been fruitful. An approach through the techniques used in Proposition 3.5, seemingly requires more general tools than the characterization of completely monotone sequences given in [2; 3] and [4]. More specifically, it would be desirable to dispose of analogous results for which difference ratios

$$\Delta_{h_i} a(x) = [a(x_i + h_i) - a(x_i)]/h_i,$$

with different spans h_i, are allowed. We think that a new approach will be, therefore, needed because the methods used in [1; 5; 8] for exchangeable random variables seem definitely adapted to this case and useless in the present context, and the reference [10] neither seems to be useful for the proposed problem. Thus, our current research is focused on a generalization of the results obtained in the present paper, which in particular would establish the Kerov conjecture.

On the other hand, in the same way that the classical De Finetti theorem was first stated for $\{0, 1\}$-valued random variables and then vastly extended in [7], more general cases are currently being studied. Results on a De Finetti theorem for families of $\{0, 1, \ldots, \ell\}$-valued random variables are given in [11]. These results can be seen in the context of a multidimensional Stirling "triangle", with several edges in different directions which have multiplicities depending of the vertex and the direction. We also hope that this would be a fruitful line of investigation.

References

[1] Chow, Y.S., Teicher, H (1988). *Probability Theory. Independence, Interchangeability, Martingales*. Springer-Verlag, New York. MR0953964

[2] Diaconis, P., Freedman, D. (2004). The Markov moment problem and De Finetti's theorem, I. *Math. Z.*, **247**, No. 1, pp. 183–199. Available at *www.stat.berkeley.edu/∼census/631.pdf* MR2054525

[3] Diaconis, P., Freedman, D. (2004). The Markov moment problem and De Finetti's theorem, II. *Math. Z.*, **247**, No. 1, pp. 201–212. Available at *www.stat.berkeley.edu/∼census/631.pdf* MR2054526

[4] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*. John Wiley, New York. MR0270403

[5] Fristedt, B., Gray, L. (1997). *A Modern Approach to Probability Theory*. Birkhäuser, Boston. MR1422917

[6] Gnedin, A., Pitman, J. (2006). Exchangeable Gibbs partitions and Stirling triangles. *J. Math. Sci.*, 138, No. 3, pp. 5674–5685. Available at *www.pdmi.ras.ru/znsl/2005/v325.html* MR2160320

[7] Hewitt, E., Savage, L.J. (1955). Symmetric measures on cartesian products. *Trans. Amer. Math. Soc.*, **80**, No. 2, pp. 470–501. MR0076206

[8] Kallenberg, O. (1997). *Foundations of Modern Probability*. Springer-Verlag, New York. MR1464694

[9] Kerov, S.V. (2003). *Asymptotic representation theory of the symmetric group and its applications in analysis*. Translations of Mathematical Monographs 219. American Mathematical Society, Providence. MR1984868

[10] Ressel, P. (1985). De Finetti-type theorems: an analytical approach. *Ann. Probab.* **13**, No. 3, pp. 898–922. MR0799427

[11] Vélez Ibarrola, R., Prieto-Rumeau, T. (2008). A De Finetti-type theorem for nonexchangeable finite-valued random variables. *J. Math. Anal. Appl.* **347**, pp. 407–415. MR2440337