Conditionally identically distributed species sampling sequences

Federico Bassetti*, Irene Crimaldi†, Fabrizio Leisen‡

12 October, 2009§

Abstract

In this paper the theory of species sampling sequences is linked to the theory of conditionally identically distributed sequences in order to enlarge the set of species sampling sequences which are mathematically tractable.

The Conditional identity in distribution (Berti, Pratelli and Rigo (2004)) is a new type of dependence for random variables, which generalizes the well-known notion of exchangeability. In this paper a class of random sequences, called *Generalized Species Sampling Sequences*, is defined and a condition to have conditional identity in distribution is given. Moreover, two types of generalized species sampling sequences that are conditionally identically distributed are introduced and studied: the *generalized Poisson-Dirichlet sequences* and the *generalized Ottawa sequences*. Some examples are discussed.

2000 MSC: 60F05, 60G57, 60B10

Key-words: species sampling sequence, conditional identity in distribution, stable convergence, almost sure conditional convergence, randomly reinforced urns, Poisson-Dirichlet sequences, random partitions, random probability measures.

1 Introduction

A sequence $(X_n)_{n\geq 1}$ of random variables defined on a probability space $(\Omega, \mathcal{A}, P)$ taking values in a Polish space, is a *species sampling sequence* if (a version) of the regular conditional distribution of X_{n+1} given $X(n) := (X_1, \ldots, X_n)$ is the transition kernel

$$K_{n+1}(\omega, \cdot) := \sum_{k=1}^{n} \tilde{p}_{n,k}(\omega)\delta_{X_k(\omega)}(\cdot) + \tilde{r}_n(\omega)\mu(\cdot) \tag{1}$$

where $\tilde{p}_{n,k}(\cdot)$ and $\tilde{r}_n(\cdot)$ are real–valued measurable functions of $X(n)$ and μ is a probability measure. See Pitman (1996).

As explained in Hansen and Pitman (2000), a species sampling sequence $(X_n)_{n\geq 1}$ can be interpreted as the sequential random sampling of individuals' species from a possibly infinite population of individuals belonging to several species. If, for the sake of simplicity, we assume that μ is diffuse, then the interpretation is the following. The species of the first individual to be observed is assigned a random tag X_1, distributed according to μ. Given the tags $X_1, \ldots X_n$ of the first n individuals observed, the species of the $(n+1)$-th individual is a new species with probability $\tilde{r}_n$ and it is equal to the observed species X_k with probability $\sum_{j=1}^{n} \tilde{p}_{n,j} I_{\{X_j = X_k\}}$.

The concept of species sampling sequence is naturally related to that of random partition induced by a sequence of observations (see Pitman (2006)). Given a random vector $X(n) = (X_1, \ldots, X_n)$, we denote by L_n the (random) number of distinct values of $X(n)$ and by $X^*(n) = (X_1^*, \ldots, X_{L_n}^*)$ the

*Department of Mathematics, University of Pavia, via Ferrata 1, 27100 Pavia, Italy. federico.bassetti@unipv.it

†Department of Mathematics, University of Bologna, Piazza di Porta San Donato 5, 40126 Bologna, Italy. crimaldi@dm.unibo.it

‡Faculty of Economics, University of Navarra, Campus Universitario, edificio de biblioteca (entrada este), 31008, Pamplona, Spain. fabrizio.leisen@unimore.it

§First version: 17 June, 2008. http://arxiv.org/abs/0806.2724

random vector of the distinct values of $X(n)$ in the order in which they appear. The *random partition induced by* $X(n)$ is the random partition of the set $\{1,\dots,n\}$ given by $\pi^{(n)} = [\pi_1^{(n)},\dots,\pi_{L_n}^{(n)}]$ where

$$i \in \pi_k^{(n)} \Leftrightarrow X_i = X_k^*.$$

Two distinct indices i and j clearly belong to the same block $\pi_k^{(n)}$ for a suitable k if and only if $X_i = X_j$. It follows that the *prediction rule* (1) can be rewritten as

$$K_{n+1}(\omega,\cdot) = \sum_{k=1}^{L_n(\omega)} \tilde{p}_{n,k}^*(\omega)\delta_{X_k^*(\omega)}(\cdot) + \tilde{r}_n(\omega)\mu(\cdot) \tag{2}$$

where

$$\tilde{p}_{n,k}^* := \sum_{j\in\pi_k^{(n)}} \tilde{p}_{n,j}.$$

In Hansen and Pitman (2000) it is proved that, if μ is diffuse and $(X_n)_{n\geq 1}$ is an exchangeable sequence, the coefficients $\tilde{p}_{n,k}^*$ are almost surely equal to some function of $\pi^{(n)}$ and they must satisfy a suitable recurrence relation. Although there are only a few explicit prediction rules which give rise to exchangeable sequences, this kind of prediction rules are appealing for many reasons. Indeed, exchangeability is a very natural assumption in many statistical problems, in particular from the Bayesian viewpoint, as well for many stochastic models. Moreover, remarkable results are known for exchangeable sequences: among others, such sequences satisfy a strong law of large numbers and they can be completely characterized by the well–known de Finetti representation theorem. See, e.g., Aldous (1985). Further, for an exchangeable sequence the empirical mean $\sum_{k=1}^n f(X_k)/n$ and the predictive mean, i.e. $E[f(X_{n+1})|X_1,\dots,X_n]$, converge to the same limit as the number of observations goes to infinity. This fact can be invoked to justify the use of the empirical mean in the place of the predictive mean, which is usually harder to compute. Nevertheless, in some situations the assumption of exchangeability can be too restrictive. For instance, instead of a classical Pólya urn scheme, it may be useful to deal with the so called randomly reinforced urn schemes. See, for example, Aletti, May and Secchi (2009), Bay and Hu (2005), Berti, Pratelli and Rigo (2004), Berti, Crimaldi, Pratelli and Rigo (2009), Crimaldi (2009), Crimaldi and Leisen (2008), Flournoy and May (2009), Janson (2005), May, Paganoni and Secchi (2005), Pemantle (2007) and the references therein. Such processes fail to be exchangeable. Our purpose is to introduce and study a class of *generalized species sampling sequences*, which are generally not exchangeable but which still have interesting mathematical properties.

We thus need to recall the notion of *conditional identity in distribution*, introduced and studied in Berti, Pratelli and Rigo (2004). Such form of dependence generalizes the notion of exchangeability preserving some of its nice predictive properties. One says that a sequence $(X_n)_{n\geq 1}$, defined on $(\Omega,\mathcal{A},P)$ and taking values in a measurable space $(E,\mathcal{E})$, is *conditionally identically distributed* with respect to a filtration $\mathcal{G} = (\mathcal{G}_n)_{n\geq 0}$ (in the sequel, $\mathcal{G}$-CID for short), whenever $(X_n)_{n\geq 1}$ is $\mathcal{G}$–adapted and, for each $n \geq 0$, $j \geq 1$ and every bounded measurable real–valued function f on E,

$$E[f(X_{n+j}) \,|\, \mathcal{G}_n] = E[f(X_{n+1}) \,|\, \mathcal{G}_n].$$

This means that, for each $n \geq 0$, all the random variables X_{n+j}, with $j \geq 1$, are identically distributed conditionally on $\mathcal{G}_n$. It is clear that every exchangeable sequence is a CID sequence with respect to its natural filtration but a CID sequence is not necessarily exchangeable. Moreover, it is possible to show that a $\mathcal{G}$–adapted sequence $(X_n)_{n\geq 1}$ is $\mathcal{G}$-CID if and only if, for each bounded measurable real–valued function f on E,

$$V_n^f := \mathrm{E}[f(X_{n+1}) \,|\, \mathcal{G}_n]$$

is a $\mathcal{G}$–martingale, see Berti, Pratelli and Rigo (2004). Hence, the sequence $(V_n^f)_{n\geq 0}$ converges almost surely to a random variable V_f. One of the most important features of CID sequences is the fact that this random variable V_f is also the almost sure limit of the empirical means. More precisely, CID sequences satisfy the following strong law of large numbers: for each bounded measurable real–valued function f on E, the sequence $(M_n^f)_{n\geq 1}$, defined by

$$M_n^f := \frac{1}{n}\sum_{k=1}^n f(X_k), \tag{3}$$

converges almost surely to V_f. It follows that also the predictive mean $\mathrm{E}[f(X_{n+1})|X_1,\dots,X_n]$ converges almost surely to V_f. In other words, CID sequences share with exchangeable sequences the remarkable fact that the predictive mean and the empirical mean merge when the number of observations diverges. Unfortunately, while, for an exchangeable sequence, we have $V_f = E[f(X_1)|\mathcal{T}] =$

$\int f(x)m(\omega, \mathrm{d}x)$, where $\mathcal{T}$ is the tail–σ–field and m is the random directing measure of the sequence, it is difficult to characterize explicitly the limit random variable V_f for a CID sequence. Indeed no representation theorems are available for CID sequences. See, e.g., Aletti, May and Secchi (2007).

The paper is organized as follows. In Section 2 we state our definition of generalized species sampling sequence and we give a condition under which a generalized species sampling sequence is CID with respect to a suitable filtration $\mathcal{G}$. After recalling the notion of stable convergence in Section 3, we introduce and analyze two types of generalized species sampling sequences which are CID: the *generalized Poisson-Dirichlet sequences* (see Section 4) and the *generalized Ottawa sequences* (see Section 5). We give some convergence results and we discuss some examples. The paper closes by a section devoted to proofs.

2 Generalized species sampling sequences

The Blackwell–MacQueen urn scheme provides the most famous example of exchangeable prediction rule, that is

$$P\{X_{n+1} \in \cdot \,|X_1, \ldots, X_n\} = \sum_{i=1}^n \frac{1}{\theta + n}\delta_{X_i}(\cdot) + \frac{\theta}{\theta + n}\mu(\cdot)$$

where θ is a strictly positive parameter and μ is a probability measure, see, e.g., Blackwell and MacQueen (1973) and Pitman (1996). This prediction rule determines an exchangeable sequence $(X_n)_{n\geq 1}$ whose directing random measure is a Dirichlet process with parameter $\theta\mu(\cdot)$, see Ferguson (1973). According to this prediction rule, if μ is diffuse, a new species is observed with probability $\theta/(\theta + n)$ and an old species X_j^* is observed with probability proportional to the cardinality of $\pi_j^{(n)}$, a sort of *preferential attachment principle*. In term of random partitions this rule corresponds to the so–called *Chinese restaurant process*, see Pitman (2006) and the references therein.

A *randomly reinforced prediction rule* of the same kind could work as follows:

$$P\{X_{n+1} \in \cdot \,|X_1, \ldots, X_n, Y_1, \ldots, Y_n\} = \sum_{i=1}^n \frac{Y_i}{\theta + \sum_{j=1}^n Y_j}\delta_{X_i}(\cdot) + \frac{\theta}{\theta + \sum_{j=1}^n Y_j}\mu(\cdot) \tag{4}$$

where μ is a probability measure and $(Y_n)_{n\geq 1}$ is a sequence of independent positive random variables. If μ is diffuse, then we have the following interpretation: each individual has a random positive weight Y_i and, given the first n tags $X(n) = (X_1, \ldots, X_n)$ together with the weights $Y(n) = (Y_1, \ldots, Y_n)$, it is supposed that the species of the next individual is a new species with probability $\theta/(\theta + \sum_{j=1}^n Y_j)$ and one of the species observed so far, say X_l^*, with probability $\sum_{i\in\pi_l^{(n)}} Y_i/(\theta + \sum_{j=1}^n Y_j)$. Again a preferential attachment principle. Note that, in this case, instead of describing the law of $(X_n)_{n\geq 1}$ with the sequence of the conditional distributions of X_{n+1} given $X(n)$, we have a latent process $(Y_n)_{n\geq 1}$ and we characterize $(X_n)_{n\geq 1}$ with the sequence of the conditional distributions of X_{n+1} given $(X(n), Y(n))$.

Now that we have given an idea, let us formalize what we mean by *generalized species sampling sequence*. Let $(\Omega, \mathcal{A}, P)$ be a probability space and E and S be two Polish spaces, endowed with their Borel σ-fields $\mathcal{E}$ and $\mathcal{S}$, respectively. In the sequel, $\mathcal{F}^Z = (\mathcal{F}_n^Z)_{n\geq 0}$ will stand for the natural filtration associated with any sequence of random variables $(Z_n)_{n\geq 1}$ on $(\Omega, \mathcal{A}, P)$ and we set $\mathcal{F}_\infty^Z = \vee_{n\geq 0}\mathcal{F}_n^Z$. Finally, $\mathcal{P}_n$ will denote the set of all partitions of $\{1, \ldots, n\}$.

We shall say that a sequence $(X_n)_{n\geq 1}$ of random variables on $(\Omega, \mathcal{A}, P)$, with values in E, is a generalized species sampling sequence if:

- (h_1) X_1 has distribution μ.
- (h_2) There exists a sequence $(Y_n)_{n\geq 1}$ of random variables with values in $(S, \mathcal{S})$ such that, for each $n \geq 1$, a version of the regular conditional distribution of X_{n+1} given

$$\mathcal{F}_n := \mathcal{F}_n^X \vee \mathcal{F}_n^Y$$

is

$$K_{n+1}(\omega, \cdot) = \sum_{i=1}^n p_{n,i}(\pi^{(n)}(\omega), Y(n)(\omega))\delta_{X_i(\omega)}(\cdot) + r_n(\pi^{(n)}(\omega), Y(n)(\omega))\mu(\cdot) \tag{5}$$

with $p_{n,i}(\cdot,\cdot)$ and $r_n(\cdot,\cdot)$ suitable measurable functions defined on $\mathcal{P}_n \times S^n$ with values in $[0, 1]$.
- (h_3) X_{n+1} and $(Y_{n+j})_{j\geq 1}$ are conditionally independent given $\mathcal{F}_n$.

Example 2.1. Let μ be a probability measure on E, $(\nu_n)_{n\geq 1}$ be a sequence of probability measures on S, $(r_n)_{n\geq 1}$ and $(p_{n,i})_{n\geq 1,\, 1\leq i\leq n}$ be measurable functions such that

$$r_n : \mathcal{P}_n \times S^n \to [0,1], \qquad p_{n,i} : \mathcal{P}_n \times Z^n \to [0,1]$$

and

$$\textstyle\sum_{i=1}^n p_{n,i}(q_n, y_1, \ldots, y_n) + r_n(q_n, y_1, \ldots, y_n) = 1 \tag{6}$$

for each $n \geq 1$ and each $(q_n, y_1, \ldots, y_n)$ in $\mathcal{P}_n \times S^n$. By the Ionescu Tulcea Theorem, there are two sequences of random variables $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$, defined on a suitable probability space $(\Omega, \mathcal{A}, P)$, taking values in E and S respectively, such that conditions (h_1), (h_2) and the following condition are satisfied:

- Y_{n+1} has distribution ν_{n+1} and it is independent of the σ-field

$$\mathcal{F}_n \vee \sigma(X_{n+1}) = \mathcal{F}_{n+1}^X \vee \mathcal{F}_n^Y.$$

This last condition implies that, for each n, $(Y_{n+j})_{j\geq 1}$ is independent of $\mathcal{F}_{n+1}^X \vee \mathcal{F}_n^Y$. It follows, in particular, that $(Y_n)_{n\geq 1}$ is a sequence of independent random variables. Therefore, also (h_3) holds true. Indeed, for each real–valued bounded $\mathcal{F}_n$-measurable random variable V, each bounded Borel function f on E, each $j \geq 1$ and each bounded Borel function h on S^j, we have

$$\begin{aligned}
\mathrm{E}[V f(X_{n+1}) h(Y_{n+1}, \ldots, Y_{n+j})] &= \mathrm{E}\big[\, V f(X_{n+1}) \mathrm{E}[h(Y_{n+1}, \ldots, Y_{n+j}) \,|\, \mathcal{F}_n \vee \sigma(X_{n+1})]\,\big] \\
&= \mathrm{E}[\, V f(X_{n+1}) \textstyle\int h(y_{n+1}, \ldots, y_{n+j})\, \nu_{n+1}(\mathrm{d}y_{n+1}) \ldots (\mathrm{d}y_{n+1})\,] \\
&= \mathrm{E}\big[\, V \mathrm{E}[f(X_{n+1}) \,|\, \mathcal{F}_n] \textstyle\int h(y_{n+1}, \ldots, y_{n+j})\, \nu_{n+1}(\mathrm{d}y_{n+1}) \ldots (\mathrm{d}y_{n+1})\,\big].
\end{aligned}$$

On the other hand, we have

$$\mathrm{E}[h(Y_{n+1}, \ldots, Y_{n+j}) \,|\, \mathcal{F}_n] = \textstyle\int h(y_{n+1}, \ldots, y_{n+j})\, \nu_{n+1}(\mathrm{d}y_{n+1}) \ldots (\mathrm{d}y_{n+1})$$

hence

$$\mathrm{E}[f(X_{n+1}) h(Y_{n+1}, \ldots, Y_{n+j}) \,|\, \mathcal{F}_n] = \mathrm{E}[f(X_{n+1}) \,|\, \mathcal{F}_n] \mathrm{E}[h(Y_{n+1}, \ldots, Y_{n+j}) \,|\, \mathcal{F}_n].$$

This fact is sufficient in order to conclude that also assumption (h_3) is verified. $\Diamond$

In order to state our first result concerning generalized species sampling sequences, we need some further notation. Set

$$p_{n,j}^*(\pi^{(n)}) = p_{n,j}^*\big(\pi^{(n)}, Y(n)\big) := \textstyle\sum_{i \in \pi_j^{(n)}} p_{n,i}\big(\pi^{(n)}, Y(n)\big) \quad \text{for } j = 1, \ldots, L_n$$

and

$$r_n := r_n\big(\pi^{(n)}, Y(n)\big).$$

Given a partition $\pi^{(n)}$, denote by $[\pi^{(n)}]_{j+}$ the partition of $\{1, \ldots, n+1\}$ obtained by adding the element $(n+1)$ to the j-th block of $\pi^{(n)}$. Finally, denote by $[\pi^{(n)}; (n+1)]$ the partition obtained by adding a block containing $(n+1)$ to $\pi^{(n)}$. For instance, if $\pi^{(3)} = [(1,3); (2)]$, then $[\pi^{(3)}]_{2+} = [(1,3); (2,4)]$ and $[\pi^{(3)}; (4)] = [(1,3); (2); (4)]$.

Theorem 2.2. *A generalized species sampling sequence $(X_n)_{n\geq 1}$ with μ diffuse is a CID sequence with respect to the filtration $\mathcal{G} = (\mathcal{G}_n)_{n\geq 0}$ with $\mathcal{G}_n := \mathcal{F}_n^X \vee \mathcal{F}_\infty^Y$ if and only if, for each n, the following condition holds P-almost surely:*

$$p_{n,j}^*(\pi^{(n)}) = r_n p_{n+1,j}^*([\pi^{(n)}; \{n+1\}]) + \textstyle\sum_{l=1}^{L_n} p_{n+1,j}^*([\pi^{(n)}]_{l+}) p_{n,l}^*(\pi^{(n)}) \tag{7}$$

for $1 \leq j \leq L_n$.

In the following sections, we shall introduce and study two types of generalized species sampling sequences that are CID.

We conclude this section with some remarks on the length L_n of the random partition induced by a generalized species sampling sequence at time n, i.e. the random number of distinct values of a generalized species sampling sequence until time n.

Let $A_0 := E$ and $A_n(\omega) := E \setminus \{X_1(\omega), \ldots, X_n(\omega)\} = \{y \in E : \ y \notin \{X_1(\omega), \ldots, X_n(\omega)\}\}$ for $n \geq 1$ and set $s_0 := 1$ and $s_n := r_n\big(\pi^{(n)}, Y(n)\big)\mu(A_n) = r_n \mu(A_n)$ for each $n \geq 1$. (If the probability measure μ is diffuse, then $s_n = r_n$.) Reconsidering the species interpretation, given

$X(n) = (X_1, \dots X_n)$ and $Y(n) = (Y_1, \dots, Y_n)$, the species of the $(n+1)$-th individual is a new species with probability s_n and one of the species observed so far with probability $1 - s_n$, that is

$$P[L_{n+1} = L_n + 1 \,|\, \mathcal{F}_n] = s_n = r_n \mu(A_n).$$

Moreover, setting $B_n = \{L_n = L_{n-1} + 1\} \in \mathcal{F}_n$ for each $n \geq 1$ (with $L_0 = 0$), we have

$$L_n = \sum_{k=1}^{n} I_{B_k}, \qquad \text{and} \qquad \sum_{k \geq 1} P[B_k \,|\, \mathcal{F}_{k-1}] = \sum_{k \geq 1} s_{k-1}.$$

Then, by Lévy's extension of Borel-Cantelli lemmas (see, for instance Williams (1991), sec. 12.15), we can obtain the following simple, but useful, result.

Proposition 2.3. *Let $(X_n)_{n \geq 1}$ be a generalized species sampling sequence. Then*
(i) $\sum_{k \geq 0} s_k < +\infty$ implies $L_n \xrightarrow{a.s.} L$, where L is a random variable with $P\{L < +\infty\} = 1$.
(ii) $\sum_{k \geq 0} s_k = +\infty$ implies $\frac{L_n}{\sum_{k=1}^{n} s_{k-1}} \xrightarrow{a.s.} 1$.
In particular, in case (ii), if there exists a sequence $(h_n)_{n \geq 1}$ of positive numbers and a random variable L such that

$$h_n \uparrow +\infty \qquad \text{and} \qquad \frac{1}{h_n} \sum_{k=1}^{n} s_{k-1} \xrightarrow{a.s.} L,$$

then $L_n / h_n \xrightarrow{a.s.} L$.

3 Stable convergence

Since in the sequel we shall deal with stable convergence, we briefly recall here this form of convergence.

Stable convergence has been introduced by Rényi (1963) and subsequently studied by various authors, see, for example, Aldous and Eagleson (1978), Jacod and Memin (1981), Hall and Heyde (1980). A detailed treatment, including some strengthened forms of stable convergence, can be found in Crimaldi, Letta and Pratelli (2007).

Given a probability space $(\Omega, \mathcal{A}, P)$ and a Polish space E (endowed with its Borel σ-field $\mathcal{E}$), recall that a kernel K on E is a family $K = (K(\omega, \cdot))_{\omega \in \Omega}$ of probability measure on E such that, for each bounded Borel function g on E, the function

$$K(g)(\omega) = \int g(x) K(\omega, \mathrm{d}x)$$

is measurable with respect to $\mathcal{A}$. Given a sub-σ-field $\mathcal{H}$ of $\mathcal{A}$, we say that the kernel K is $\mathcal{H}$-measurable if, for each bounded Borel function g on E, the random variable $K(g)$ is measurable with respect to $\mathcal{H}$. In the following, the symbol $\mathcal{N}$ will denote the sub-σ-field generated by the P-negligible events of $\mathcal{A}$. Given a sub-σ-field $\mathcal{H}$ of $\mathcal{A}$ and a $\mathcal{H} \vee \mathcal{N}$-measurable kernel K on E, a sequence $(Z_n)_{n \geq 1}$ of random variables on $(\Omega, \mathcal{A}, P)$ with values in E converges *$\mathcal{H}$-stably* to K if, for each bounded continuous function g on E and for each $\mathcal{H}$–measurable real–valued bounded random variable W

$$\mathrm{E}[g(Z_n)\, W] \longrightarrow \mathrm{E}[K(g)\, W].$$

If $(Z_n)_{n \geq 1}$ converges $\mathcal{H}$-stably to K then, for each $A \in \mathcal{H}$ with $P(A) \neq 0$, the sequence $(Z_n)_{n \geq 1}$ converges in distribution under the probability measure $P_A = P(\cdot|A)$ to the probability measure $P_A K$ on E given by

$$P_A K(B) = P(A)^{-1} \mathrm{E}[I_A K(\cdot, B)] = \int K(\omega, B)\, P_A(\mathrm{d}\omega) \qquad \text{for each } B \in \mathcal{E}. \tag{8}$$

In particular, if $(Z_n)_{n \geq 1}$ converges $\mathcal{H}$-stably to K, then $(Z_n)_{n \geq 1}$ converges in distribution to the probability measure PK on E given by

$$PK(B) = \mathrm{E}[K(\cdot, B)] = \int K(\omega, B)\, P(\mathrm{d}\omega) \qquad \text{for each } B \in \mathcal{E}. \tag{9}$$

Moreover, if all the random variables Z_n are $\mathcal{H}$-measurable, then the $\mathcal{H}$-stable convergence obviously implies the $\mathcal{A}$-stable convergence.

Throughout the paper, if U is a positive random variable, we shall call the Gaussian kernel associated with U the family

$$\mathcal{N}(0, U) = \big(\mathcal{N}(0, U(\omega))\big)_{\omega \in \Omega}$$

of Gaussian distributions with zero mean and variance equal to $U(\omega)$ (with $\mathcal{N}(0, 0) := \delta_0$). Note that, in this case, the probability measures defined in (8) and (9) are mixtures of Gaussian distributions.

4 Generalized Poisson Dirichlet sequences

Let $\alpha \geq 0$ and $\theta > -\alpha$. Moreover, let μ be a probability measure on E, ν_1 be a probability measure on $(\alpha, +\infty)$ and $(\nu_n)_{n\geq 2}$ be a sequence of probability measures on $[\alpha, +\infty)$. Consider the following sequence of functions

$$p_{n,i}(q_n, y(n)) := \frac{y_i - \alpha/C_i(q_n)}{\theta + \sum_{j=1}^n y_j}$$
$$r_n(q_n, y(n)) := \frac{\theta + \alpha L(q_n)}{\theta + \sum_{j=1}^n y_j}$$

where $y(n) = (y_1, \ldots, y_n) \in (\alpha, +\infty) \times [\alpha, +\infty)^{n-1}$, $q_n \in \mathcal{P}_n$, $C_i(q_n)$ is the cardinality of the block in q_n which contains i and $L(q_n)$ is the number of blocks of q_n. It is easy to see that such functions satisfy (6). Hence, by Example 2.1, there exists a generalized species sampling sequence $(X_n)_{n\geq 1}$ for which

$$P\{X_{n+1} \in \cdot \,|X(n), Y(n)\} = \sum_{l=1}^{L_n} \frac{\left(\sum_{i\in\pi_l^{(n)}} Y_i\right) - \alpha}{\theta + S_n} \delta_{X_l^*}(\cdot) + \frac{\theta + \alpha L_n}{\theta + S_n}\mu(\cdot), \tag{10}$$

where $(Y_n)_{n\geq 1}$ is a sequence of independent random variables such that each Y_n has law ν_n and $S_n = \sum_{j=1}^n Y_j$ (with $S_0 = 0$). If μ is *diffuse*, one can easily check that (7) of Theorem 2.2 holds and so $(X_n)_{n\geq 1}$ is a CID sequence with respect to $\mathcal{G} = (\mathcal{F}_n^X \vee \mathcal{F}_\infty^Y)_{n\geq 1}$.

It is worthwhile noting that if μ is diffuse, $Y_n = 1$ for every $n \geq 1$, $\alpha \in [0, 1)$ and $\theta > -\alpha$, then we get an exchangeable sequence directed by the well-known two parameter Poisson-Dirichlet process: i.e. an exchangeable sequence described by the prediction rule

$$P\{X_{n+1} \in \cdot \,|X_1, \ldots, X_n\} = \sum_{l=1}^{L_n} \frac{\mathrm{card}(\pi_l^{(n)}) - \alpha}{\theta + n} \delta_{X_l^*}(\cdot) + \frac{\theta + \alpha L_n}{\theta + n}\mu(\cdot). \tag{11}$$

See, e.g., Pitman and Yor (1997) and Pitman (2006).

The case $\alpha = 0$ have been deeply studied by many authors (see, for instance, Aletti, May and Secchi (2009), Bay and Hu (2005), Berti, Crimaldi, Pratelli and Rigo (2009), Crimaldi (2009), Flournoy and May (2009), Janson (2005), May, Paganoni and Secchi (2005), Pemantle (2007) and the references therein). The case when μ is discrete and $\alpha > 0$ has been treated in Berti, Crimaldi, Pratelli and Rigo (2009). Here, we present some results for the case when μ is *diffuse and* $\alpha > 0$.

Proposition 4.1. *If* $\sup_n \mathrm{E}[Y_n^2] < +\infty$ *and* $\lim_n \mathrm{E}[Y_n] = m$, *then*

$$r_n = \frac{\theta + \alpha L_n}{\theta + S_n} \xrightarrow{a.s.} R \qquad \text{and} \qquad \frac{L_n}{n} \xrightarrow{a.s.} R,$$

where R is a random variable such that $P\{0 \leq R \leq 1\} = 1$.
In particular, if $m > \alpha$, we have $P\{R = 0\} = 1$.

Later on we shall see some examples in which $P\{R > 0\} > 0$.

Let us take $A \in \mathcal{E}$ and set $V_n^A := P[X_{n+1} \in A \,|\, \mathcal{F}_n]$. Since, $(X_n)_{n\geq 1}$ is CID, we have

$$V_n^A \xrightarrow{a.s.} V_A \qquad \text{and} \qquad M_n^A := \frac{1}{n}\sum_{k=1}^n I_A(X_k) \xrightarrow{a.s.} V_A.$$

We shall prove the following central limit theorem.

Theorem 4.2. *Let us assume the following conditions:*
(i) $\sup_n \mathrm{E}[Y_n^u] < +\infty$ *for some* $u > 2$
(ii) $m = \lim_n \mathrm{E}[Y_n]$, $\quad q = \lim_n \mathrm{E}[Y_n^2]$.
Then

$$\left(\sqrt{n}(M_n^A - V_n^A),\, \sqrt{n}(V_n^A - V_A)\right) \xrightarrow{\mathcal{A}-stably} \mathcal{N}(0, U_A) \times \mathcal{N}(0, \Sigma_A),$$

where

$$U_A = \left(\frac{q}{m^2} - 1\right) V_A(1 - V_A) + \frac{\alpha^2}{m^2} R\mu(A)[1 - \mu(A)].$$
$$\Sigma_A = \frac{q}{m^2} V_A(1 - V_A) + \frac{\alpha}{m}\left(\frac{\alpha}{m} - 2\right) R\mu(A)[1 - \mu(A)].$$

In particular, $\sqrt{n}(M_n^A - V_A) \overset{\mathcal{A}-stably}{\longrightarrow} \mathcal{N}(0, U_A + \Sigma_A)$. *Moreover,*

$$\mathrm{E}[g\big(\sqrt{n}(V_n^A - V_A)\big) \,|\, \mathcal{G}_n] \xrightarrow{a.s.} \mathcal{N}(0, \Sigma_f))(g)$$

for each $g \in \mathcal{C}_b(\mathbb{R})$.

4.1 Case $m > \alpha$

By Proposition 4.1, we have $P\{R = 0\} = 1$ and so

$$U_A = \Big(\frac{q}{m^2} - 1\Big) V_A(1 - V_A). \qquad \Sigma_A = \frac{q}{m^2} V_A(1 - V_A).$$

Taking into account the analogy with randomly reinforced Pólya urns, it is natural to think that the random variable V_A is generally not degenerate (see Aletti, May and Secchi (2009)). This fact implies that Σ_A is not degenerate, while U_A is degenerate if and only if Y_n converges in L^2 to the constant m. This happens, for example, in the classical case (see (11)) studied by Pitman and Yor (1997) and Pitman (2006).

4.2 Case $m = \alpha$

If $m = \alpha$ and $q = \alpha^2$ (i.e. $Y_n \overset{L^2}{\to} \alpha$), then

$$U_A = R\mu(A)[1 - \mu(A)], \qquad \Sigma_A = V_A(1 - V_A) - R\mu(A)[1 - \mu(A)] \qquad \text{and} \qquad U_A + \Sigma_A = V_A(1 - V_A).$$

The following examples show that, if $m = \alpha$, we can have $P\{R > 0\} > 0$.

Example 4.3. Let us take $\alpha > 0$ and $-\alpha < \theta \leq 0$. Setting

$$W_n = \frac{\alpha L_n}{\alpha + \theta + S_{n-1}},$$

we have (see the following Lemma 6.2)

$$\Delta_n = \mathrm{E}[W_{n+1} - W_n \,|\, \mathcal{F}_n] = \frac{(\alpha - Y_n)W_n}{\theta + S_n} + \frac{\alpha\theta}{(\theta + S_n)(\alpha + \theta + S_n)} \geq \frac{(\alpha - Y_n)\alpha n}{(\theta + \alpha n)^2} + \frac{\alpha\theta}{(\theta + \alpha n)(\alpha + \theta + \alpha n)}.$$

Therefore, we have

$$\mathrm{E}[W_{n+1} \,|\, \mathcal{F}_n] - W_1 = \sum_{k=1}^{n} \mathrm{E}[W_{k+1} - W_k \,|\, \mathcal{F}_k] = \sum_{k=1}^{n} \Delta_k \geq \alpha \sum_{k=1}^{n} \frac{(\alpha - Y_k)k}{(\theta + \alpha k)^2} + \alpha\theta \sum_{k=1}^{n} \frac{1}{(\theta + \alpha k)(\alpha + \theta + \alpha k)}.$$

and so

$$\mathrm{E}[W_{n+1}] \geq \frac{\alpha}{\alpha + \theta} + \alpha \sum_{k=1}^{n} \frac{\mathrm{E}[\alpha - Y_k]k}{(\theta + \alpha k)^2} + \alpha\theta \sum_{k=1}^{n} \frac{1}{(\theta + \alpha k)(\alpha + \theta + \alpha k)}.$$

Letting $n \to +\infty$, we obtain

$$\mathrm{E}[R] \geq \frac{\alpha}{\alpha + \theta} + \alpha \sum_{k \geq 1} \frac{\mathrm{E}[\alpha - Y_k]k}{(\theta + \alpha k)^2} + \alpha\theta \sum_{k \geq 1} \frac{1}{(\theta + \alpha k)(\alpha + \theta + \alpha k)}.$$

Therefore, if

$$\alpha \sum_{k \geq 1} \frac{\mathrm{E}[Y_k - \alpha]k}{(\theta + \alpha k)^2} - \alpha\theta \sum_{k \geq 1} \frac{1}{(\theta + \alpha k)(\alpha + \theta + \alpha k)} < \frac{\alpha}{\alpha + \theta}, \tag{12}$$

then $\mathrm{E}[R] > 0$ and so $P\{R > 0\} > 0$. Note that, in order to have (12), it must be

$$\sum_{k \geq 1} \frac{\mathrm{E}[Y_k - \alpha]}{k} < +\infty.$$

$\Diamond$

Example 4.4. Let us take $\alpha > 0$, $\theta = 0$, $Y_1 > \alpha$ and $Y_n = \alpha$ for each $n \geq 2$. Then, using the same notation as the one in the previous example, we have $\Delta_n = 0$ for each $n \geq 2$ and so $\mathrm{E}[R] = \mathrm{E}[W_2]$. On the other hand, we have

$$0 < \mathrm{E}[W_2] \leq \mathrm{E}\left[\frac{2\alpha}{\alpha + Y_1}\right] < 1.$$

Then we get $\min\big[P\{R > 0\}, P\{R < 1\}\big] > 0$. Moreover, since it must be $P\{\Sigma_A \geq 0\} = 1$, we obtain that, if $0 < \mu(A) < 1$, then $P\{V_A = 0, R > 0\} = 0$ and $P\{V_A = 1, R > 0\} = 0$. $\Diamond$

5 Generalized Ottawa sequences

We shall say that a generalized species sampling sequence $(X_n)_{n\geq 1}$ is a *generalized Ottawa sequence* or, more briefly, a GOS, if the following conditions are satisfied for every $n \geq 1$:

- The functions r_n and $p_{n,i}$ (for $i = 1, \ldots, n$) do not depend on the partition, hence
$$K_{n+1}(\omega, \cdot) = \textstyle\sum_{i=1}^{n} p_{n,i}(Y(n)(\omega))\delta_{X_i(\omega)}(\cdot) + r_n(Y(n)(\omega))\mu(\cdot). \tag{13}$$
- The functions r_n are strictly positive and
$$r_n(Y_1, \ldots, Y_n) \geq r_{n+1}(Y_1, \ldots, Y_n, Y_{n+1}) \tag{14}$$
almost surely.
- The functions $p_{n,i}$ satisfy
$$\begin{aligned} p_{n,i} &:= \frac{r_n}{r_{n-1}}\, p_{n-1,i} \qquad \text{for } i = 1, \ldots, n-1 \\ p_{n,n} &:= 1 - \frac{r_n}{r_{n-1}} \end{aligned} \tag{15}$$
with $r_0 = 1$.

For simplicity, from now on, we shall denote by r_n and $p_{n,i}$ the $\mathcal{F}_n^Y$-measurable random variables $r_n(Y(n))$ and $p_{n,i}(Y(n))$, that is $r_n := r_n(Y(n))$ and $p_{n,i} := p_{n,i}(Y(n))$.

First of all let us stress that any GOS is a CID sequence with respect to the filtration $\mathcal{G} = (\mathcal{F}_n^X \vee \mathcal{F}_\infty^Y)_{n\geq 0}$. Indeed, since $\mathcal{G}_n = \mathcal{F}_n \vee \sigma(Y_{n+j} : j \geq 1)$, condition $(h3)$ implies that
$$\mathrm{E}[f(X_{n+1}) \,|\, \mathcal{G}_n] = \mathrm{E}[f(X_{n+1}) \,|\, \mathcal{F}_n] \tag{16}$$
for each bounded Borel real–valued function f on E and hence, by $(h2)$, one gets
$$V_n^f := \mathrm{E}[f(X_{n+1}) \,|\, \mathcal{G}_n] = \textstyle\sum_{i=1}^{n} p_{n,i} f(X_i) + r_n \mathrm{E}[f(X_1)].$$
Since the random variables $p_{n+1,i}$ are $\mathcal{G}_n$-measurable it follows that
$$\begin{aligned} \mathrm{E}[V_{n+1}^f \,|\, \mathcal{G}_n] &= \textstyle\sum_{i=1}^{n} p_{n+1,i} f(X_i) + p_{n+1,n+1}\mathrm{E}[f(X_{n+1}) \,|\, \mathcal{G}_n] + r_{n+1}\mathrm{E}[f(X_1)] \\ &= \frac{r_{n+1}}{r_n} \textstyle\sum_{i=1}^{n} p_{n,i} f(X_i) + V_n^f - \frac{r_{n+1}}{r_n} V_n^f + r_{n+1}\mathrm{E}[f(X_1)] \\ &= \frac{r_{n+1}}{r_n} V_n^f - r_{n+1}\mathrm{E}[f(X_1)] + V_n^f - \frac{r_{n+1}}{r_n} V_n^f + r_{n+1}\mathrm{E}[f(X_1)] = V_n^f. \end{aligned}$$

Some examples follow.

Example 5.1. Consider a GOS for which $Y_n = a_n$, where $(a_n)_{n\geq 0}$ is a *decreasing numerical* sequence with $a_0 = 1$, $a_n > 0$ and $r_n(y_1, \ldots, y_n) = y_n$.

If μ is diffuse, by Proposition 2.3, we can say that L_n converges almost surely to an integrable random variable if and only if $\sum_k a_k < +\infty$. $\Diamond$

Example 5.2. Consider a GOS for which $(Y_n)_{n\geq 1}$ is a Markov chain taking values in $(0, 1]$, with $Y_1 = 1$ and transition probability kernel given by
$$P\{Y_{n+1} \leq x | Y_n\} = \frac{x}{Y_n} I_{(0,Y_n)}(x) + I_{[Y_n,+\infty)}(x) \quad n \geq 1$$
and $r_n(y_1, \ldots, y_n) = y_n$.

If μ is diffuse, we have $\mathrm{E}[s_n] = \mathrm{E}[Y_n] = (1/2)^{n-1}$ and so $\sum_{k\geq 0} \mathrm{E}[s_k] < +\infty$. Therefore, by Proposition 2.3, we can say that L_n converges almost surely to an integrable random variable. $\Diamond$

Example 5.3. Consider a GOS for which $(Y_n)_{n\geq 1}$ is a sequence of random variable taking values on $(0, 1)$ and
$$r_n(y_1, \ldots, y_n) = \prod_{i=1}^{n} y_i.$$
Note that in this case
$$P\{X_{n+1} \in \cdot \,|X(n), Y(n)\} = \sum_{j=1}^{n} \left[(1 - Y_j) \prod_{i=j+1}^{n} Y_i \right] \delta_{X_j}(\cdot) + \left[\prod_{i=1}^{n} Y_i \right] \mu(\cdot).$$

Assume that μ is diffuse and that $(Y_j)_{j\geq 1}$ is a sequence of independent random variables distributed according to a Beta distribution of parameter $(j, 1-\alpha)$ with α in $[0,1)$. That is each Y_j has density (with respect to the Lebesgue measure) on $[0,1]$ given by

$$x \mapsto \frac{\Gamma(j+1-\alpha)}{\Gamma(j)\Gamma(1-\alpha)} \frac{x^{j-1}}{(1-x)^{\alpha}},$$

where $\Gamma(z) = \int_0^{+\infty} x^{z-1}e^{-x}dx$. Set $m_{1,n} := \mathrm{E}[L_n]$ and $m_{2,n} := \mathrm{E}[L_n^2]$. Note that

$$m_{1,n+1} = m_{1,n} + \mathrm{E}[r_n] = \sum_{j=0}^{n} \mathrm{E}[r_j]. \tag{17}$$

and

$$m_{2,n+1} = 3m_{1,n+1} - 2 + 2\sum_{j=2}^{n}\sum_{i=1}^{j-1} \mathrm{E}[r_i r_j]. \tag{18}$$

If $\alpha = 0$, (17) gives

$$m_{1,n+1} = 1 + \sum_{j=1}^{n}\prod_{i=1}^{j} \frac{i}{1+i} = \sum_{j=0}^{n} \frac{1}{1+j}$$

and, after some computations, from (18), one gets also

$$m_{2,n+1} = 1 + 3\sum_{j=1}^{n} \frac{1}{1+j} + 2\sum_{j=2}^{n}\sum_{i=1}^{j-1}\prod_{h=1}^{i} \frac{h}{h+2}\prod_{k=i+1}^{j} \frac{k}{k+1} = 1 + 3\sum_{j=1}^{n} \frac{1}{1+j} + 4\sum_{j=3}^{n+1} \frac{1}{j}\sum_{i=3}^{j} \frac{1}{i}.$$

Now recall that

$$\lim_{n\to+\infty} \frac{1}{\log n}\sum_{j=1}^{n} \frac{1}{j} = 1 \tag{19}$$

and, moreover, observe that

$$\lim_{n\to+\infty} \frac{1}{\log^2(n)}\sum_{j=1}^{n} \frac{1}{j}\sum_{i=1}^{j} \frac{1}{i} = \frac{1}{2}.$$

This shows that the mean of L_n diverges as $\log n$ and the second moment diverges as $\log^2(n)$. More precisely,

$$\lim_{n\to+\infty} \frac{m_{1,n+1}}{\log n} = \lim_{n\to+\infty} \frac{m_{2,n+1}}{2\log^2(n)} = 1.$$

If $\alpha \neq 0$, in the same way, one gets

$$m_{1,n+1} = 1 + \Gamma(2-\alpha)\sum_{j=1}^{n} \frac{\Gamma(j+1)}{\Gamma(j+2-\alpha)}.$$

Now recall that

$$\frac{\Gamma(j+1)}{\Gamma(j+2-\alpha)} = \frac{1}{j^{1-\alpha}}\big(1 + O(1/j)\big)$$

for $j \to +\infty$ and that

$$\lim_{n\to+\infty} \frac{1}{n^{\alpha}}\textstyle\sum_{j=1}^{n} \frac{1}{j^{1-\alpha}} = \frac{1}{\alpha} \qquad \text{for } \alpha \in (0,1). \tag{20}$$

Hence, when $\alpha \neq 0$, we have

$$\lim_{n\to+\infty} \frac{m_{1,n+1}}{n^{\alpha}} = \frac{\Gamma(2-\alpha)}{\alpha}.$$

$\Diamond$

Example 5.4. Consider a GOS for which $(Y_n)_{n\geq 1}$ is a sequence of random variable taking values on $\mathbb{R}_+$ and

$$r_n(y_1, \ldots, y_n) = \frac{\theta}{\theta + \sum_{j=1}^{n} y_j}$$

with $\theta > 0$. Note that the randomly reinforced Blackwell–McQueen urn scheme (described by (4)) gives rise to a GOS. This example will be reconsidered later on. $\Diamond$

For the length L_n of the random partition induced by a GOS, we shall prove the following central limit theorem.

Theorem 5.5. *Let $(X_n)_{n\geq 1}$ be a GOS with μ diffuse and suppose there exists a sequence $(h_n)_{n\geq 1}$ of positive numbers and a positive random variable σ^2 such that the following properties hold:*

$$h_n \uparrow +\infty, \quad \textit{and} \quad \sigma_n^2 := \frac{\sum_{j=1}^n r_{j-1}(1-r_{j-1})}{h_n} \xrightarrow{a.s.} \sigma^2.$$

Then, setting $R_n := \sum_{j=1}^n r_{j-1}$, we have

$$T_n := \frac{L_n - R_n}{\sqrt{h_n}} \xrightarrow{\mathcal{A}\text{-}stably} \mathcal{N}(0, \sigma^2).$$

Corollary 5.6. *Under the same assumptions of Theorem 5.5, if $P(\sigma^2 > 0) = 1$, then we have*

$$\frac{T_n}{\sigma_n} = \frac{(L_n - R_n)}{\sqrt{\sum_{j=1}^n r_{j-1}(1-r_{j-1})}} \xrightarrow{\mathcal{A}\text{-}stably} \mathcal{N}(0,1).$$

Example 5.7. Let us consider Example 5.1 with μ diffuse and

$$a_n = \frac{\theta}{\theta + n^{1-\alpha}}$$

with $\theta > 0$ and $0 < \alpha < 1$. We have $s_n = r_n = a_n$ and, setting $h_n = n^\alpha$ and $L = \theta/\alpha$, from (20) we get

$$\frac{1}{n^\alpha} R_n = \frac{1}{n^\alpha} \sum_{j=0}^{n-1} \frac{\theta}{\theta + j^{1-\alpha}} \longrightarrow \frac{\theta}{\alpha}.$$

Thus, by Proposition 2.3, we obtain that $L_n/n^\alpha \xrightarrow{a.s.} \theta/\alpha$. Further, since

$$\frac{1}{h_n} \sum_{j=1}^n a_j b_j \to b, \tag{21}$$

provided that $a_j \geq 0$, $\sum_{j=1}^n a_j/h_n \to 1$ and $b_n \to b$ as $n \to +\infty$, it is easy to see that

$$\sigma_n^2 = \frac{\sum_{j=0}^{n-1} r_j(1-r_j)}{n^\alpha} = \frac{\theta}{n^\alpha} \sum_{j=1}^{n-1} \frac{j^{1-\alpha}}{(\theta + j^{1-\alpha})^2} = \frac{\theta}{n^\alpha} \sum_{j=1}^{n-1} \left(\frac{j^{1-\alpha}}{\theta + j^{1-\alpha}} \right)^2 \frac{1}{j^{1-\alpha}} \to \theta/\alpha.$$

Therefore, by Theorem 5.5, we obtain

$$T_n = \frac{L_n - R_n}{n^{\alpha/2}} \xrightarrow{\mathcal{A}\text{-stably}} \mathcal{N}(0, \theta).$$

$\Diamond$

Example 5.8. Let us consider a GOS with μ diffuse and

$$r_n = \frac{\theta}{\theta + \sum_{i=1}^n Y_i}.$$

where $\theta > 0$ and the random variables Y_n are independent positive random variable such that $\sum_n \mathrm{E}[Y_n^2]/n^2 < +\infty$ and $\lim_n \mathrm{E}[Y_n] = m > 0$. Then $s_n = r_n$ and (see, for instance, Lemma 3 in Berti, Crimaldi, Pratelli and Rigo (2009)) we have

$$\left(\frac{\theta}{j} + \frac{1}{j} \sum_{i=1}^j Y_i \right)^{-1} \xrightarrow{a.s.} 1/m.$$

Setting $h_n = \log n$ and $L = c/m$, by (19) and (21), we obtain

$$\frac{1}{\log n} R_n = \frac{1}{\log n} + \frac{\theta}{\log n} \sum_{j=1}^{n-1} \frac{1}{\theta + \sum_{i=1}^j Y_i} \sim \frac{\theta}{\log n} \sum_{j=1}^{n-1} \frac{1}{j} \left(\frac{\theta}{j} + \frac{1}{j} \sum_{i=1}^j Y_i \right)^{-1} \xrightarrow{a.s.} \frac{\theta}{m}$$

and so, by Proposition 2.3, we can conclude that $L_n/\log n \xrightarrow{a.s.} \theta/m$. Moreover, by (21), we have

$$\begin{aligned} \sigma_n^2 &= \frac{\sum_{j=0}^{n-1} r_j(1-r_j)}{\log n} = \frac{\theta}{\log n} \sum_{j=1}^{n-1} \frac{\sum_{i=1}^j Y_i}{(\theta + \sum_{i=1}^j Y_i)^2} \\ &= \frac{\theta}{\log n} \sum_{j=1}^{n-1} \left(\frac{\sum_{i=1}^j Y_i/j}{\theta/j + \sum_{i=1}^j Y_i/j} \right)^2 \frac{j}{\sum_{i=1}^j Y_i} \frac{1}{j} \to \theta/m. \end{aligned}$$

Therefore, by Theorem 5.5, we obtain

$$T_n = \frac{L_n - R_n}{\sqrt{\log n}} \xrightarrow{\mathcal{A}\text{–stably}} \mathcal{N}(0, \theta/m)$$

and so

$$\frac{L_n - R_n}{\sqrt{\frac{\theta}{m}\log n}} \xrightarrow{\mathcal{A}\text{–stably}} \mathcal{N}(0,1).$$

If we take $Y_i = 1$ for all i, we find the well known results for the asymptotic distribution of the length of the random partition obtained for the Blackwell–McQueen urn scheme. Indeed, since $\sum_{j=1}^{n} j^{-1} - \log n = \gamma + O(\frac{1}{n})$, one gets

$$\frac{L_n - \theta \log n}{\sqrt{\theta \log n}} \xrightarrow{\mathcal{A}\text{–stably}} \mathcal{N}(0,1).$$

See, for instance, pages 68-69 in Pitman (2006). $\Diamond$

We recall that, since a GOS $(X_n)_{n\geq 1}$ is CID, then, for each bounded Borel real–valued function f on E, we have

$$V_n^f = \mathrm{E}[f(X_{n+1}) \,|\, \mathcal{F}_n] \xrightarrow{a.s.} V_f \qquad \text{and} \qquad M_n^f = \frac{1}{n}\sum_{k=1}^{n} f(X_k) \xrightarrow{a.s.} V_f.$$

Inspired by Theorem 3.3 in Berti, Pratelli and Rigo (2004) and the results in Crimaldi (2009), we conclude this section with the statements of some central limit theorems for a GOS.

Theorem 5.9. *Let $(X_n)_{n\geq 1}$ be a GOS. For each bounded Borel real–valued function f and each $n \geq 1$, let us set*

$$C_n^f = \sqrt{n}(M_n^f - V_n^f)$$

and, for $1 \leq j \leq n$,

$$Z_{n,j}^f = \frac{1}{\sqrt{n}}\big[f(X_j) - jV_j^f + (j-1)V_{j-1}^f\big] = \frac{1}{\sqrt{n}}(1 + jp_{j,j})\big[f(X_j) - V_{j-1}^f\big].$$

Suppose that:

(a) $U_n^f := \sum_{j=1}^{n}(Z_{n,j}^f)^2 \xrightarrow{P} U_f$.

(b) $(Z_n^f)^* := \sup_{1\leq j\leq n} |Z_{n,j}^f| \xrightarrow{L^1} 0$.

Then the sequence $(C_n^f)_{n\geq 1}$ converges $\mathcal{A}$-stably to the Gaussian kernel $\mathcal{N}(0, U_f)$.

In particular, condition (a) and (b) are satisfied if the following conditions hold:

(a1) $U_n^f \xrightarrow{a.s.} U_f$.

(b1) $\sup_{n\geq 1} \mathrm{E}[(C_n^f)^2] < +\infty$.

Theorem 5.10. *Let $(X_n)_{n\geq 1}$ be a GOS and f be a bounded Borel real–valued function. Using the previous notation, for $n \geq 0$ set*

$$Q_n := p_{n+1,n+1} = 1 - \frac{r_{n+1}}{r_n} \qquad \text{and} \qquad D_n^f := \sqrt{n}(V_n^f - V_f).$$

Suppose that the following conditions are satisfied:

(i) $n\sum_{k\geq n} Q_k^2 \xrightarrow{a.s.} H$, where H is a positive real random variable.

(ii) $\sum_{k\geq 0} k^2 \,\mathrm{E}[Q_k^4] < \infty$.

Then

$$\mathrm{E}\big[g(D_n^f) \,|\, \mathcal{F}_n\big] \xrightarrow{a.s.} \mathcal{N}\big(0, H(V_{f^2} - V_f^2)\big)(g)$$

for each $g \in \mathcal{C}_b(\mathbb{R})$. In particular, we have $D_n^f \xrightarrow{\mathcal{A}\text{–stably}} \mathcal{N}\big(0, H(V_{f^2} - V_f^2)\big)$.

Corollary 5.11. *Using the notation of Theorem 5.10, let us set for* $k \geq 0$

$$\rho_k = \frac{1}{r_{k+1}} - \frac{1}{r_k}$$

and assume the following conditions:

(a) $r_k \leq c_k$ *a.s. with* $\sum_{k\geq 0} k^2 c_{k+1}^4 < \infty$ *and* $k r_k \stackrel{a.s.}{\to} \alpha$*, where* c_k*,* α *are strictly positive constants.*

(b) The random variable ρ_k *are independent and identically distributed with* $\mathrm{E}[\rho_k^4] < \infty$*.*

Finally, let us set $\beta := \mathrm{E}[\rho_k^2]$ *and* $h := \alpha^2 \beta$*.*

Then, the conclusion of Theorem 5.10 holds true with H *equal to the constant* h*.*

Furthermore, if the assumptions of both Theorems 5.9 and 5.10 hold true, then, by Lemma 1 in Berti, Crimaldi, Pratelli and Rigo (2009), we get

$$\big(C_n^f, D_n^f\big) \stackrel{\mathcal{A}-stably}{\longrightarrow} \mathcal{N}(0, U_f) \times \mathcal{N}\big(0, H(V_{f^2} - V_f^2)\big).$$

In particular, $\sqrt{n}(M_n^f - V_f) = C_n^f + D_n^f \stackrel{\mathcal{A}-stably}{\longrightarrow} \mathcal{N}\big(0, U_f + H(V_{f^2} - V_f^2)\big)$.

Since the proofs of these results are essentially the same as those in Berti, Pratelli and Rigo (2004), in Crimaldi (2009) and Berti, Crimaldi, Pratelli and Rigo (2009), we shall skip them. The interested reader can find all the details and some simple examples in the first version of this paper Bassetti, Crimaldi and Leisen (2008).

6 Proofs

This section contains all the proofs of the paper. Recall that

$$\mathcal{F}_n = \mathcal{F}_n^X \vee \mathcal{F}_n^Y \qquad \text{and} \qquad \mathcal{G}_n = \mathcal{F}_n^X \vee \mathcal{F}_\infty^Y = \mathcal{F}_n \vee \sigma(Y_{n+j} : j \geq 1)$$

and so condition $(h3)$ of the definition of generalized species sampling sequence implies that

$$V_n^g := \mathrm{E}[g(X_{n+1}) \,|\, \mathcal{G}_n] = \mathrm{E}[g(X_{n+1}) \,|\, \mathcal{F}_n]$$

for each bounded Borel real–valued function g on E.

6.1 Proof of Theorem 2.2

We start with a useful lemma.

Lemma 6.1. *If* $(X_n)_{n\geq 1}$ *is a generalized species sampling sequence, then we have*

$$P[n+1 \in \pi_l^{(n+1)} \,|\, \mathcal{G}_n] = P[X_{n+1} = X_l^* \,|\, \mathcal{F}_n] = \textstyle\sum_{j \in \pi_l^{(n)}} p_{n,j}(\pi^{(n)}, Y(n)) + r_n(\pi^{(n)}, Y(n))\mu(\{X_l^*\})$$

for each $l = 1, \ldots, L_n$*. Moreover, for each bounded Borel real–valued function* f *on* E*,*

$$\mathrm{E}[I_{\{L_{n+1}=L_n+1\}} f(X_{n+1}) \,|\, \mathcal{G}_n] = \mathrm{E}[I_{\{L_{n+1}=L_n+1\}} f(X_{n+1}) \,|\, \mathcal{F}_n] = r_n(\pi^{(n)}, Y(n)) \textstyle\int_{A_n} f(y)\, \mu(\mathrm{d}y).$$

holds true with $A_0 := E$ *and* A_n *the random "set" defined by*

$$A_n(\omega) := E \setminus \{X_1(\omega), \ldots, X_n(\omega)\} = \{y \in E : \ y \notin \{X_1(\omega), \ldots, X_n(\omega)\}\} \quad \text{for } n \geq 1.$$

In particular, we have

$$P[L_{n+1} = L_n + 1 \,|\, \mathcal{G}_n] = P[L_{n+1} = L_n + 1 \,|\, \mathcal{F}_n] = r_n(\pi^{(n)}, Y(n))\mu(A_n) := s_n(\pi^{(n)}, Y(n)).$$

If μ *is diffuse, we have*

$$P[n+1 \in \pi_l^{(n+1)} \,|\, \mathcal{G}_n] = \mathrm{P}[X_{n+1} = X_l^* \,|\, \mathcal{F}_n] = \textstyle\sum_{j \in \pi_l^{(n)}} p_{n,j}(\pi^{(n)}, Y(n))$$

for each $l = 1, \ldots, L_n$ *and*

$$\mathrm{E}[I_{\{L_{n+1}=L_n+1\}} f(X_{n+1}) \,|\, \mathcal{G}_n] = \mathrm{E}[I_{\{L_{n+1}=L_n+1\}} f(X_{n+1}) \,|\, \mathcal{F}_n] = r_n(\pi^{(n)}, Y(n))\mathrm{E}[f(X_1)]$$

and

$$P[L_{n+1} = L_n + 1 \,|\, \mathcal{G}_n] = P[L_{n+1} = L_n + 1 \,|\, \mathcal{F}_n] = r_n(\pi^{(n)}, Y(n)).$$

Proof. Since $\mathcal{G}_n = \mathcal{F}_n \vee \sigma(Y_{n+j} : j \geq 1)$, condition (h_3) implies that

$$P[n+1 \in \pi_l^{(n+1)} \,|\, \mathcal{G}_n] = P[X_{n+1} = X_l^* \,|\, \mathcal{G}_n] P[X_{n+1} = X_l^* \,|\, \mathcal{F}_n].$$

Hence, by assumption $(h2)$, we have

$$\begin{aligned} P[X_{n+1} = X_l^* \,|\, \mathcal{F}_n] &= \textstyle\sum_{i=1}^n p_{n,i}(\pi^{(n)}, Y(n)) \delta_{X_i}(X_l^*) + r_n(\pi^{(n)}, Y(n)) \mu(\{X_l^*\}) \\ &= \textstyle\sum_{j \in \pi_l^{(n)}} p_{n,j}(\pi^{(n)}, Y(n)) + r_n(\pi^{(n)}, Y(n)) \mu(\{X_l^*\}). \end{aligned}$$

for each $l = 1, \ldots, L_n$. If μ is diffuse, we obtain

$$P[X_{n+1} = X_l^* \,|\, \mathcal{F}_n] = \textstyle\sum_{j \in \pi_l^{(n)}} p_{n,j}(\pi^{(n)}, Y(n))$$

for each $l = 1, \ldots, L_n$.

Now, we observe that

$$I_{\{L_{n+1} = L_n + 1\}} = I_{B_{n+1}}(X_1, \ldots, X_n, X_{n+1})$$

where $B_{n+1} = \{(x_1, \ldots, x_{n+1}) : \; x_{n+1} \notin \{x_1, \ldots, x_n\}\}$. Thus, by (h_3) and (h_2), we have

$$\begin{aligned} \mathrm{E}[I_{\{L_{n+1} = L_n+1\}} f(X_{n+1}) \,|\, \mathcal{G}_n] &= \mathrm{E}[I_{\{L_{n+1} = L_n+1\}} f(X_{n+1}) \,|\, \mathcal{F}_n] \\ &= \textstyle\int I_{B_{n+1}}(X_1, \ldots, X_n, y) f(y) K_{n+1}(\cdot, \mathrm{d}y) \\ &= \textstyle\sum_{i=1}^n p_{n,i}(\pi^{(n)}, Y(n)) \int_{A_n} f(y) \delta_{X_i}(\mathrm{d}y) + r_n(\pi^{(n)}, Y(n)) \int_{A_n} f(y) \mu(\mathrm{d}y) \\ &= r_n(\pi^{(n)}, Y(n)) \textstyle\int_{A_n} f(y) \mu(\mathrm{d}y). \end{aligned}$$

If we take $f = 1$, we get

$$P[U_{n+1} = 1 \,|\, \mathcal{G}_n] = P[U_{n+1} = 1 \,|\, \mathcal{F}_n] = r_n(\pi^{(n)}, Y(n)) \mu(A_n).$$

Finally, if μ is diffuse, then $\mu(A_n(\omega)) = 1$ for each ω and so we have

$$\textstyle\int_{A_n} f(y) \mu(\mathrm{d}y) = \mathrm{E}[f(X_1)].$$

□

Proof of Theorem 2.2. Let us fix a bounded Borel real–valued function f on E. Using the given prediction rule, we have

$$\begin{aligned} V_n^f &= \textstyle\sum_{i=1}^n p_{n,i}(\pi^{(n)}, Y(n)) f(X_i) + r_n(\pi^{(n)}, Y(n)) \mathrm{E}[f(X_1)] \\ &= \textstyle\sum_{j=1}^{L_n} p_{n,j}^*(\pi^{(n)}) f(X_j^*) + r_n \mathrm{E}[f(X_1)]. \end{aligned}$$

The sequence (X_n) is $\mathcal{G}$-cid if and only if for each bounded Borel real–valued function f on E, the sequence $(V_n^f)_{n \geq 0}$ is a $\mathcal{G}$-martingale. We observe that we have (for the sake of simplicity we skip the dependence on $(Y_n)_{n \geq 1}$)

$$\begin{aligned} \mathrm{E}[V_{n+1}^f \,|\, \mathcal{G}_n] &= \textstyle\sum_{i=1}^n f(X_i) E_i + \mathrm{E}[p_{n+1,n+1}(\pi^{(n+1)}) f(X_{n+1}) \,|\, \mathcal{G}_n] + \mathrm{E}[r_{n+1} \,|\, \mathcal{G}_n] \bar{f} \\ &= \textstyle\sum_{j=1}^{L_n} f(X_j^*) \sum_{i \in \pi_j^{(n)}} E_i + \mathrm{E}[p_{n+1,n+1}(\pi^{(n+1)}) f(X_{n+1}) \,|\, \mathcal{G}_n] + \mathrm{E}[r_{n+1} \,|\, \mathcal{G}_n] \bar{f} \end{aligned}$$

where $E_i = \mathrm{E}[p_{n+1,i}(\pi^{(n+1)}) \,|\, \mathcal{G}_n]$ and $\bar{f} = \mathrm{E}[f(X_1)]$.
Now we are going to compute the various conditional expectations which appear in the second member of above equality. Since μ is diffuse, using Lemma 6.1, we have

$$\begin{aligned} E_i &= \mathrm{E}[p_{n+1,i}(\pi^{(n+1)}) \,|\, \mathcal{G}_n] \\ &= \textstyle\sum_{l=1}^{L_n} \mathrm{E}[I_{\{n+1 \in \pi_l^{(n+1)}\}} p_{n+1,i}(\pi^{(n+1)}) \,|\, \mathcal{G}_n] + \mathrm{E}[I_{\{L_{n+1} = L_n+1\}} p_{n+1,i}(\pi^{(n+1)}) \,|\, \mathcal{G}_n] \\ &= \textstyle\sum_{l=1}^{L_n} p_{n+1,i}([\pi^{(n)}]_{l+}) \mathrm{E}[I_{\{n+1 \in \pi_l^{(n+1)}\}} \,|\, \mathcal{G}_n] + \mathrm{E}[I_{\{L_{n+1} = L_n+1\}} \,|\, \mathcal{G}_n] p_{n+1,i}([\pi^{(n)}; n+1]) \\ &= \textstyle\sum_{l=1}^{L_n} p_{n+1,i}([\pi^{(n)}]_{l+}) \sum_{j \in \pi_l^{(n)}} p_{n,j}(\pi^{(n)}) + r_n p_{n+1,i}([\pi^{(n)}; n+1]) \\ &= \textstyle\sum_{l=1}^{L_n} p_{n+1,i}([\pi^{(n)}]_{l+}) p_{n,l}^*(\pi^{(n)}) + r_n p_{n+1,i}([\pi^{(n)}; n+1]) \end{aligned}$$

and so

$$\sum_{i\in\pi_j^{(n)}} E_i = \sum_{l=1,l\neq j}^{L_n} p^*_{n+1,j}([\pi^{(n)}]_{l^+})p^*_{n,l}(\pi^{(n)}) + \sum_{i\in\pi_j^{(n)}} p_{n+1,i}([\pi^{(n)}]_{j^+})p^*_{n,j}(\pi^{(n)}) + r_n p^*_{n+1,j}([\pi^{(n)};n+1])$$
$$= \textstyle\sum_{l=1}^{L_n} p^*_{n+1,j}([\pi^{(n)}]_{l^+})p^*_{n,l}(\pi^{(n)}) - p_{n+1,n+1}([\pi^{(n)}]_{j^+})p^*_{n+1,j}(\pi^{(n)}) + r_n p^*_{n+1,j}([\pi^{(n)};n+1])$$

Moreover, using Lemma 6.1 again, we have

$$\mathrm{E}[p_{n+1,n+1}(\pi^{(n+1)})f(X_{n+1})\,|\,\mathcal{G}_n] =$$
$$\sum_{l=1}^{L_n} \mathrm{E}[I_{\{n+1\in\pi_l^{(n+1)}\}}p_{n+1,n+1}(\pi^{(n+1)})f(X_{n+1})\,|\,\mathcal{G}_n] + \mathrm{E}[I_{\{L_{n+1}=L_n+1\}}p_{n+1,n+1}(\pi^{(n+1)})f(X_{n+1})\,|\,\mathcal{G}_n] =$$
$$\sum_{l=1}^{L_n} \mathrm{E}[I_{\{n+1\in\pi_l^{(n+1)}\}}\,|\,\mathcal{G}_n]p_{n+1,n+1}([\pi^{(n)}]_{l^+})f(X^*_l) + \mathrm{E}[I_{\{L_{n+1}=L_n+1\}}f(X_{n+1})\,|\,\mathcal{G}_n]p_{n+1,n+1}([\pi^{(n)}];n+1) =$$
$$\textstyle\sum_{l=1}^{L_n}\left(\sum_{k\in\pi_l^{(n)}} p_{n,k}(\pi^{(n)})\right)p_{n+1,n+1}([\pi^{(n)}]_{l^+})f(X^*_l) + r_n p_{n+1,n+1}([\pi^{(n)}];n+1)\bar{f} =$$
$$\textstyle\sum_{l=1}^{L_n} p^*_{n,l}(\pi^{(n)})p_{n+1,n+1}([\pi^{(n)}]_{l^+})f(X^*_l) + r_n p_{n+1,n+1}([\pi^{(n)}];n+1)\bar{f}.$$

Finally we have

$$\begin{aligned}\mathrm{E}[r_{n+1}\,|\,\mathcal{G}_n] &= 1 - \textstyle\sum_{i=1}^{n+1}\mathrm{E}[p_{n+1,i}(\pi^{(n+1)})\,|\,\mathcal{G}_n]\\ &= 1 - \textstyle\sum_{i=1}^{n} E_i - E_{n+1}\\ &= 1 - \textstyle\sum_{i=1}^{n} E_i - \sum_{l=1}^{L_n} p^*_{n,l}(\pi^{(n)})p_{n+1,n+1}([\pi^{(n)}]_{l^+}) - r_n p_{n+1,n+1}([\pi^{(n)}];n+1)\end{aligned}$$

Thus we get

$$\mathrm{E}[V^f_{n+1}\,|\,\mathcal{G}_n] = \textstyle\sum_{j=1}^{L_n} c_{n,j}f(X^*_j) + (1-\sum_{j=1}^{L_n} c_{n,j})\bar{f}$$

where

$$\begin{aligned}c_{n,j} &= \textstyle\sum_{i\in\pi_j^{(n)}} E_i + p_{n+1,n+1}([\pi^{(n)}]_{j^+})p^*_{n,j}(\pi^{(n)})\\ &= r_n p^*_{n+1,j}([\pi^{(n)};n+1]) + \textstyle\sum_{l=1}^{L_n} p^*_{n+1,j}([\pi^{(n)}]_{l^+})p^*_{n,l}(\pi^{(n)})\end{aligned}$$

We can conclude that $(X_n)_{n\geq 1}$ is $\mathcal{G}$-cid if and only if we have, for each bounded Borel function f on E and each n

$$\textstyle\sum_{j=1}^{L_n} p^*_{n,j}f(X^*_j) + r_n\bar{f} = \sum_{j=1}^{L_n} c_{n,j}f(X^*_j) + (1-\sum_{j=1}^{L_n} c_{n,j})\bar{f} \qquad P\text{-almost surely.}$$

Since E is a Polish space, we may affirm that $(X_n)_{n\geq 1}$ is $\mathcal{G}$-cid if and only if, for each n, we have P-almost surely

$$\textstyle\sum_{j=1}^{L_n} p^*_{n,j}\delta_{X^*_k}(\cdot) + r_n\mu(\cdot) = \sum_{j=1}^{L_n} c_{n,j}\delta_{X^*_k}(\cdot) + (1-\sum_{j=1}^{L_n} c_{n,j})\mu(\cdot)$$

But this last equality holds if and only if, for each n, we have P-almost surely

$$p^*_{n,j} = c_{n,j} \qquad \text{for } 1\leq j\leq L_n\ ;$$

that is

$$p^*_{n,j}(\pi^{(n)}) = r_n p^*_{n+1,j}([\pi^{(n)};\{n+1\}]) + \textstyle\sum_{l=1}^{L_n} p^*_{n+1,j}([\pi^{(n)}]_{l+},)p^*_{n,l}(\pi^{(n)})$$

This is exactly the condition in the statement of the Theorem 2.2. □

6.2 Proofs of section 4

We need the following preliminary lemma.

Lemma 6.2. *Let us set $S_n = \sum_{j=1}^n Y_j$. Then*

$$W_n = \frac{\alpha L_n}{\alpha+\theta+S_{n-1}} \overset{a.s./L^1}{\longrightarrow} R.$$

where R is a random variable such that $P\{0\leq R\leq 1\} = 1$.

Proof. We have

$$\mathrm{E}[L_{n+1} \,|\, \mathcal{F}_n] = L_n + r_n = \frac{(\alpha+\theta+S_n)L_n + \theta}{\theta + S_n}.$$

Hence, we get

$$\begin{aligned}
\Delta_n = \mathrm{E}[W_{n+1} \,|\, \mathcal{F}_n] - W_n &= \frac{\alpha L_n}{\theta+S_n} - \frac{\alpha L_n}{\alpha+\theta+S_{n-1}} + \frac{\alpha\theta}{(\theta+S_n)(\alpha+\theta+S_n)} \\
&= \frac{(\alpha - Y_n)}{(\theta+S_n)} \frac{\alpha L_n}{(\alpha+\theta+S_{n-1})} + \frac{\alpha\theta}{(\theta+S_n)(\alpha+\theta+S_n)} \\
&= \frac{(\alpha-Y_n)W_n}{(\theta+S_n)} + \frac{\alpha\theta}{(\theta+S_n)(\alpha+\theta+S_n)}.
\end{aligned}$$

If $-\alpha < \theta \leq 0$, then Δ_n is negative for each n and so $(W_n)_n$ is a positive supermartingale. Therefore it converges almost surely to a random variable R.

If $\theta > 0$, let us set $Z_n = W_n + \frac{\theta}{(\theta+S_{n-1})}$. For each n, we have

$$\begin{aligned}
\mathrm{E}[Z_{n+1} \,|\, \mathcal{F}_n] - Z_n &= \Delta_n - \frac{\theta Y_n}{(\theta+S_{n-1})(\theta+S_n)} \\
&= \frac{(\alpha-Y_n)W_n}{(\theta+S_n)} + \frac{\alpha\theta}{(\theta+S_n)}\left[\frac{1}{\alpha+\theta+S_n} - \frac{Y_n}{\alpha(\theta+S_{n-1})}\right] \\
&= \frac{(\alpha-Y_n)W_n}{(\theta+S_n)} + \frac{\theta(\alpha-Y_n)}{(\theta+S_{n-1})(\theta+S_n)} - \frac{\alpha\theta(Y_n+\alpha)}{(\theta+S_{n-1})(\theta+S_n)(\alpha+\theta+S_n)} \leq 0
\end{aligned}$$

Therefore, the sequence $(Z_n)_n$ is a positive $\mathcal{F}$-supermartingale and so it converges almost surely to a random variable R. Since S_n goes to $+\infty$, we get

$$W_n = Z_n - \frac{\theta}{\theta+S_{n-1}} \overset{a.s./L^1}{\longrightarrow} R.$$

Finally, we observe that $0 \leq W_n \leq \frac{\alpha n}{\theta+\alpha n} \to 1$. □

Proof of Proposition 4.1. It is easy to verify that $\frac{S_n}{n} \overset{a.s.}{\to} m$ (see, for instance, Lemma 3 in Berti, Crimaldi, Pratelli and Rigo (2009)) and so, by Lemma 6.2, we get

$$r_n = \frac{\theta+\alpha L_n}{\theta+S_n} = \frac{\theta}{\theta+S_n} + W_n \frac{\alpha+\theta+S_{n-1}}{\theta+S_n} \overset{a.s.}{\to} R.$$

Moreover, we have

$$\sum_{k\geq 1} r_{k-1} \geq 1 + (\theta+\alpha)\sum_{k\geq 1}\frac{1}{\theta+S_k} \overset{a.s.}{\sim} \frac{\alpha+\theta}{m}\sum_{k\geq 1}\frac{1}{k} = \infty.$$

Then, by Proposition 2.3, we find

$$\frac{L_n}{\sum_{k=1}^n r_{k-1}} \overset{a.s.}{\longrightarrow} 1.$$

Since Cesaro's lemma implies

$$\frac{1}{n}\sum_{k=1}^n r_{k-1} \overset{a.s.}{\longrightarrow} R,$$

we get $L_n/n \overset{a.s.}{\to} R$. On the other hand, we have

$$\frac{L_n}{n} \overset{a.s.}{\sim} \frac{m}{\alpha} r_n \overset{a.s.}{\longrightarrow} \frac{m}{\alpha} R.$$

Therefore, we have $\frac{m}{\alpha} R \overset{a.s.}{\sim} R$ and so, if $m \neq \alpha$, it must be $P(R=0) = 1$. □

Proof of Theorem 4.2. As we have already observed, assumption *(ii)* implies $\frac{S_n}{n} \overset{a.s.}{\to} m$ and $r_n \overset{a.s.}{\to} R$. After some calculations, we find

$$V_{n+1}^A - V_n^A = \frac{Y_{n+1}[I_A(X_{n+1}) - V_n^A]}{\theta+S_{n+1}} + \alpha\frac{[\mu(A) - I_A(X_{n+1})]}{\theta+S_{n+1}} I_{\{L_{n+1}=L_n+1\}}.$$

We want to apply Lemma 1, Theorem 2 and Remark 4 of Berti, Crimaldi, Pratelli and Rigo (2009). Therefore, we have to prove the following conditions:

(1) $\frac{1}{\sqrt{n}}\mathrm{E}[\max_{1\le k\le n} k|V_{k-1}^A - V_k^A|] \longrightarrow 0$

(2) $\mathrm{E}[\sup_{k\ge 1}\sqrt{k}|V_{k-1}^A - V_k^A|] < +\infty$

(3) $n\sum_{k\ge n}(V_{k-1}^A - V_k^A)^2 \xrightarrow{a.s.} \Sigma_A$.

(4) $\frac{1}{n}\sum_{k=1}^n\{I_A(X_k) - V_{k-1}^A + k(V_{k-1}^A - V_k^A)\}^2 \xrightarrow{P} U_A$

Conditions (1) and (2) hold true: we observe that

$$|V_{n+1}^A - V_n^A| \le \frac{Y_{n+1}+\alpha}{\theta+\alpha(n+1)}.$$

This inequality and assumption *(i)* imply

$$\frac{1}{n^{u/2}}\left(\mathrm{E}[\max_{1\le k\le n} k|V_{k-1}^A - V_k^A|]\right)^u \le \frac{1}{n^{u/2}}\sum_{k=1}^n k^u \frac{\mathrm{E}\big[(Y_k+\alpha)^u\big]}{(\theta+\alpha k)^u} \longrightarrow 0$$

and

$$\mathrm{E}\big[\,(\sup_{k\ge 1}\sqrt{k}|V_{k-1}^A - V_k^A|)^u\,\big] \le \sum_{k\ge 1} k^{u/2}\frac{\mathrm{E}\big[(Y_k+\alpha)^u\big]}{(\theta+\alpha k)^u} < +\infty.$$

Condition (3) holds true: we observe that

$$\begin{aligned}
n\sum_{k\ge n}(V_{k-1}^A - V_k^A)^2 &= n\sum_{k\ge n}\frac{Y_k^2[I_A(X_k) - V_{k-1}^A]^2}{(\theta+S_k)^2} + n\alpha^2\sum_{k\ge n}\frac{[\mu(A) - I_A(X_k)]^2}{(\theta+S_k)^2}I_{\{L_k=L_{k-1}+1\}} \\
&\quad + 2\alpha n\sum_{k\ge n}\frac{Y_k[I_A(X_k) - V_{k-1}^A][\mu(A) - I_A(X_k)]}{(\theta+S_k)^2}I_{\{L_k=L_{k-1}+1\}} \\
&\overset{a.s.}{\sim} \frac{n}{m^2}\sum_{k\ge n}\frac{Y_k^2[I_A(X_k) - V_{k-1}^A]^2}{k^2} + \frac{\alpha^2}{m^2}n\sum_{k\ge n}\frac{[\mu(A) - I_A(X_k)]^2}{k^2}I_{\{L_k=L_{k-1}+1\}} \\
&\quad + 2\frac{\alpha}{m^2}n\sum_{k\ge n}\frac{Y_k[I_A(X_k) - V_{k-1}^A][\mu(A) - I_A(X_k)]}{k^2}I_{\{L_k=L_{k-1}+1\}}.
\end{aligned}$$

We want to use Lemma 3 in Berti, Crimaldi, Pratelli and Rigo (2009). Therefore, we observe that

$$\mathrm{E}\{Y_k^2[I_A(X_k) - V_{k-1}^A]^2\,|\,\mathcal{F}_{k-1}\} = \mathrm{E}[Y_k^2]\mathrm{E}\{[I_A(X_k) - V_{k-1}^A]^2\,|\,\mathcal{F}_{k-1}\} \overset{a.s}{\to} qV_A(1-V_A)$$

and so the first term converges almost surely to

$$\frac{q}{m^2}V_A(1-V_A).$$

Moreover, we observe that we have

$$\mathrm{E}\{[\mu(A) - I_A(X_k)]^2 I_{\{L_k=L_{k-1}+1\}}\,|\,\mathcal{F}_{k-1}\} = r_{k-1}\mu(A)[1-\mu(A)] \overset{a.s.}{\to} R\mu(A)[1-\mu(A)]$$

and so the second term converges almost surely to

$$\frac{\alpha^2}{m^2}\,R\mu(A)[1-\mu(A)]. \tag{22}$$

Finally, we have

$$\begin{aligned}
&\mathrm{E}\{Y_k[I_A(X_k) - V_{k-1}^A][\mu(A) - I_A(X_k)]I_{\{L_k=L_{k-1}+1\}}\,|\,\mathcal{F}_{k-1}\} = \\
&-\mathrm{E}[Y_k]r_{k-1}\mu(A)[1-\mu(A)] \overset{a.s.}{\to} -mR\mu(A)[1-\mu(A)]
\end{aligned}$$

and so the third term converges almost surely to

$$-2\frac{\alpha}{m}\,R\mu(A)[1-\mu(A)].$$

Condition (4) holds true: we observe that

$$\frac{1}{n}\sum_{k=1}^{n}\{I_A(X_k)-V_{k-1}^A+k(V_{k-1}^A-V_k^A)\}^2=$$
$$\frac{1}{n}\sum_{k=1}^{n}\left\{[I_A(X_k)-V_{k-1}^A]\left[1-\frac{kY_k}{\theta+S_k}\right]+\frac{k\alpha[\mu(A)-I_A(X_k)]}{\theta+S_k}I_{\{L_k=L_{k-1}+1\}}\right\}^2\overset{a.s.}{\sim}$$
$$\frac{1}{n}\sum_{k=1}^{n}\left\{[I_A(X_k)-V_{k-1}^A]\left[1-\frac{Y_k}{m}\right]+\frac{\alpha[\mu(A)-I_A(X_k)]}{m}I_{\{L_k=L_{k-1}+1\}}\right\}^2=$$
$$\frac{1}{n}\sum_{k=1}^{n}[I_A(X_k)-V_{k-1}^A]^2\left[1-\frac{Y_k}{m}\right]^2+\frac{\alpha^2}{m^2n}\sum_{k=1}^{n}[\mu(A)-I_A(X_k)]^2I_{\{L_k=L_{k-1}+1\}}+$$
$$\frac{2\alpha}{m}\frac{1}{n}\sum_{k=1}^{n}[I_A(X_k)-V_{k-1}^A]\left[1-\frac{Y_k}{m}\right][\mu(A)-I_A(X_k)]I_{\{L_k=L_{k-1}+1\}}$$

We want to use Lemma 3 in Berti, Crimaldi, Pratelli and Rigo (2009) once again. Therefore, we observe the second term converges almost surely to (22). Moreover, we have

$$\mathrm{E}\left\{[I_A(X_k)-V_{k-1}^A]^2\left[1-\frac{Y_k}{m}\right]^2\mid\mathcal{F}_{k-1}\right\}=\mathrm{E}\left\{\left[1-\frac{Y_k}{m}\right]^2\right\}\mathrm{E}\left\{[I_A(X_k)-V_{k-1}^A]^2\mid\mathcal{F}_{k-1}\right\}$$

and so the first term converges almost surely to

$$\left(\frac{q}{m^2}-1\right)V_A(1-V_A).$$

Finally, we have

$$\mathrm{E}\left\{\left[1-\frac{Y_k}{m}\right][I_A(X_k)-V_{k-1}^A][\mu(A)-I_A(X_k)]I_{\{L_k=L_{k-1}+1\}}\mid\mathcal{F}_{k-1}\right\}=$$
$$\mathrm{E}\left[1-\frac{Y_k}{m}\right]\mathrm{E}\left\{[I_A(X_k)-V_{k-1}^A][\mu(A)-I_A(X_k)]I_{\{L_k=L_{k-1}+1\}}\mid\mathcal{F}_{k-1}\right\}$$

and so the third term converges almost surely to zero. □

6.3 Proof of Theorem 5.5

It will be useful to introduce the sequence of the increments

$$U_1:=L_1=1\qquad\text{and}\qquad U_n:=L_n-L_{n-1}\quad\text{for } n\geq 2.$$

We need a preliminary lemma.

Lemma 6.3. *If $(X_n)_{n\geq1}$ is a GOS with μ diffuse, then, for each fixed k, a version of the conditional distribution of $(U_j)_{j\geq k+1}$ given $\mathcal{G}_k$ is the kernel Q_k so defined:*

$$Q_k(\omega,\cdot):=\bigotimes_{j=k+1}^{\infty}\mathcal{B}\big(1,r_{j-1}(\omega)\big)$$

where $\mathcal{B}\big(1,r_{j-1}(\omega)\big)$ denotes the Bernoulli distribution with parameter $r_{j-1}(\omega)$.

Proof. It is enough to verify that, for each $n\geq1$, for each $\epsilon_{k+1},\dots,\epsilon_{k+n}\in\{0,1\}$ and for each $\mathcal{G}_k$-measurable real–valued bounded random variable Z, we have

$$\mathrm{E}\big[ZI_{\{U_{k+1}=\epsilon_{k+1},\dots,U_{k+n}=\epsilon_{k+n}\}}\big]=\mathrm{E}\big[Z\textstyle\prod_{j=k+1}^{k+n}r_{j-1}^{\epsilon_j}(1-r_{j-1})^{1-\epsilon_j}\big].\tag{23}$$

We go on with the proof by induction on n. For $n=1$, by Lemma 6.1, we have

$$\mathrm{E}\big[ZI_{\{U_{k+1}=\epsilon_{k+1}\}}\big]=\mathrm{E}\big[Z\mathrm{E}[I_{\{U_{k+1}=\epsilon_{k+1}\}}\mid\mathcal{G}_k]\big]=\mathrm{E}[Zr_k^{\epsilon_{k+1}}(1-r_k)^{1-\epsilon_{k+1}}].$$

Assume that (23) is true for $n-1$ and let us prove it for n. Let us fix an $\mathcal{G}_k$-measurable real–valued bounded random variable Z. By Lemma 6.1, we have

$$\mathrm{E}\big[ZI_{\{U_{k+1}=\epsilon_{k+1},\dots,U_{k+n}=\epsilon_{k+n}\}}\big]=\mathrm{E}\big[ZI_{\{U_{k+1}=\epsilon_{k+1},\dots,U_{k+n-1}=\epsilon_{k+n-1}\}}\mathrm{E}[U_{k+n}=\epsilon_{k+n}\mid\mathcal{G}_{k+n-1}]\big]$$
$$=\mathrm{E}\big[Zr_{k+n-1}^{\epsilon_{k+n}}(1-r_{k+n-1})^{1-\epsilon_{k+n}}I_{\{U_{k+1}=\epsilon_{k+1},\dots,U_{k+n-1}=\epsilon_{k+n-1}\}}\big].$$

We have done because also the random variable $Zr_{k+n-1}^{\epsilon_{k+n}}(1-r_{k+n-1})^{1-\epsilon_{k+n}}$ is $\mathcal{G}_k$-measurable and (23) is true for $n-1$. □

We need also the following known result.

Theorem 6.4. *Let $(Z_{n,i})_{n\geq 1,\, 1\leq i\leq k_n}$ be a triangular array of square integrable centered random variables on a probability space $(\Omega, \mathcal{A}, P)$. Suppose that, for each fixed n, $(Z_{n,i})_i$ is independent ("row-independence property"). Moreover, set*

$$\sigma_{n,i}^2 := \mathrm{E}[Z_{n,i}^2] = \mathrm{Var}[Z_{n,i}], \qquad \sigma_n^2 := \textstyle\sum_{i=1}^{k_n} \sigma_{n,i}^2,$$
$$V_n := \textstyle\sum_{i=1}^{k_n} Z_{n,i}^2, \qquad Z_n^* := \sup_{1\leq i\leq k_n} |Z_{n,i}|$$

and assume that $(V_n)_{n\geq 1}$ is uniformly integrable, $Z_n^* \xrightarrow{P} 0$ and $\sigma_n^2 \longrightarrow \sigma^2$.

Then $\sum_{i=1}^{k_n} Z_{n,i} \xrightarrow{\text{in law}} \mathcal{N}(0,\sigma^2)$.

Proof. In Hall and Heyde (1980) (see pp. 53–54) it is proved that, under the uniform integrability of (V_n), the convergence in probability to zero of $(Z_n^*)_{n\geq 1}$ is equivalent to the Lindeberg condition. Hence, it is possible to apply Corollary 3.1 (pp. 58-59) in Hall and Heyde (1980) with $\mathcal{F}_{n,i} = \sigma(Z_{n,1}, \ldots, Z_{n,i})$. □

Proof of Theorem 5.5. Without loss of generality, we can assume $h_n > 0$ for each n. In order to prove the desidered $\mathcal{A}$-stable convergence, it is enough to prove the $\mathcal{F}_\infty^X \vee \mathcal{F}_\infty^Y$-stable convergence of (T_n) to $\mathcal{N}(0,\sigma^2)$. But, in order to prove this last convergence, since we have $\mathcal{F}_\infty^X \vee \mathcal{F}_\infty^Y = \bigvee_k \mathcal{G}_k$, it suffices to prove that, for each k and A in $\mathcal{G}_k$ with $P(A) \neq 0$, the sequence (T_n) converges in distribution under P_A to the probability measure $P_A\mathcal{N}(0,\sigma^2)$. In other words, it is sufficient to fix k and to verify that $(T_{k+n})_n$ (and so $(T_n)_n$) converges $\mathcal{G}_k$-stably to $\mathcal{N}(0,\sigma^2)$. (Note that the kernel $\mathcal{N}(0,\sigma^2)$ is $\mathcal{G}_k \vee \mathcal{N}$-measurable for each fixed k.) To this end, we observe that we have

$$T_{k+n} = \frac{\sum_{j=1}^{k+n}(U_j - r_{j-1})}{\sqrt{h_{k+n}}} = \frac{\sum_{j=1}^{k}(U_j - r_{j-1})}{\sqrt{h_{k+n}}} + \frac{\sum_{j=k+1}^{k+n}(U_j - r_{j-1})}{\sqrt{h_{k+n}}}.$$

Obviously, for $n \to +\infty$, we have

$$\frac{\sum_{j=1}^{k}(U_j - r_{j-1})}{\sqrt{h_{k+n}}} \xrightarrow{a.s.} 0.$$

Therefore we have to prove

$$\frac{\sum_{j=k+1}^{k+n}(U_j - r_{j-1})}{\sqrt{h_{k+n}}} \xrightarrow{\mathcal{G}_k\text{–stably}} \mathcal{N}(0,\sigma^2). \tag{24}$$

From Lemma 6.3 we know that a version of the the conditional distribution of $(U_j)_{j\geq k+1}$ given $\mathcal{G}_k$ is the kernel Q_k so defined:

$$Q_k(\omega,\cdot) = \textstyle\bigotimes_{j=k+1}^{\infty} \mathcal{B}\big(1, r_{j-1}(\omega)\big).$$

On the canonical space $\mathbb{R}^{\mathbb{N}^*}$ let us consider the canonical projections $(\xi_j)_{j\geq k+1}$. Then, for each $n \geq 1$, a version of the conditional distribution of

$$\frac{\sum_{j=k+1}^{k+n}(U_j - r_{j-1})}{\sqrt{h_{k+n}}}$$

given $\mathcal{G}_k$ is the kernel N_{k+n} so characterized: for each ω, the probability measure $N_{k+n}(\omega,\cdot)$ is the distribution, under the probability measure $Q_k(\omega,\cdot)$, of the random variable (which is defined on the canonical space)

$$\frac{\sum_{j=k+1}^{k+n}\big(\xi_j - r_{j-1}(\omega)\big)}{\sqrt{h_{k+n}}}.$$

On the other hand, for almost every ω, under $Q_k(\omega,\cdot)$, the random variables

$$Z_{n,i} := \frac{\xi_{k+i} - r_{k+i-1}(\omega)}{\sqrt{h_{k+n}}} \qquad \text{for } n \geq 1,\ 1 \leq i \leq n$$

form a triangular array which satisfies the assumptions of Theorem 6.4. Indeed, we have the row-independence property and

$$\mathrm{E}^{Q_k(\omega,\cdot)}[Z_{n,i}] = 0, \qquad \mathrm{E}^{Q_k(\omega,\cdot)}[Z_{n,i}^2] = \frac{r_{k+i-1}(\omega)\big(1 - r_{k+i-1}(\omega)\big)}{h_{k+n}}.$$

Therefore, by assumption, for $n \to +\infty$, we have for almost every ω,

$$\sum_{i=1}^n \mathrm{E}^{Q_k(\omega,\cdot)}[Z_{n,i}^2] = \frac{\sum_{i=1}^n r_{k+i-1}(\omega)\big(1 - r_{k+i-1}(\omega)\big)}{h_{k+n}} = \sigma_{k+n}^2(\omega) - \frac{h_{k-1}\sigma_{k-1}^2(\omega)}{h_{k+n}} \longrightarrow \sigma^2(\omega).$$

Moreover, under $Q_k(\omega,\cdot)$, we have $Z_n^* := \sup_i Z_{n,i} \leq 2/\sqrt{h_{k+n}} \longrightarrow 0$. Finally, we observe that, setting $V_n := \sum_{i=1}^n Z_{n,i}^2$, we have

$$\mathrm{E}^{Q_k(\omega,\cdot)}[V_n^2] = \mathrm{Var}^{Q_k(\omega,\cdot)}[V_n] + \left(\sigma_{k+n}^2(\omega) - \frac{h_{k-1}\sigma_{k-1}^2(\omega)}{h_{k+n}}\right)^2$$

with

$$\begin{aligned}\mathrm{Var}^{Q_k(\omega,\cdot)}[V_n] = \sum_{i=1}^n \mathrm{Var}^{Q_k(\omega,\cdot)}[Z_{n,i}^2] &\leq \sum_{i=1}^n \mathrm{E}^{Q_k(\omega,\cdot)}[Z_{n,i}^4] \\ &\leq 4\left(\sigma_{k+n}^2(\omega) - \frac{h_{k-1}\sigma_{k-1}^2(\omega)}{h_{k+n}}\right)\frac{1}{h_{k+n}}.\end{aligned}$$

Since, for almost every ω, the sequence $(\sigma_n^2(\omega))_n$ is bounded and $h_n \uparrow +\infty$, it follows that, for almost every ω, the sequence $(V_n)_n$ is bounded in L^2 under $Q_k(\omega,\cdot)$ and so uniformly integrable. Theorem 6.4 assures that, for almost every ω, the sequence of probability measures

$$\big(N_{k+n}(\omega,\cdot)\big)_{n\geq 1}$$

weakly converges to the Gaussian distribution $\mathcal{N}\big(0,\sigma^2(\omega)\big)$. This fact implies that, for each bounded continuous function g, we have

$$\mathrm{E}\left[g\left(\frac{\sum_{j=k+1}^{k+n}(U_j - r_{j-1})}{\sqrt{h_{k+n}}}\right) \mid \mathcal{G}_k\right] \xrightarrow{a.s.} \mathcal{N}(0,\sigma^2)(g).$$

It obviously follows the $\mathcal{G}_k$-stable convergence (24). □

Acknowledgements

This research work is supported by funds of GNAMPA 2008/09. Irene Crimaldi would like to thank Luca Pratelli for useful discussions on the generalized Poisson-Dirichlet sequences.

References

Aldous D.J and Eagleson G.K. (1978). On mixing and stability of limit theorems. *Ann. Prob.* **6** 325–331.

Aldous D.J. (1985). Exchangeability and related topics. *Ecole d'Eté de Probabilités de Saint-Flour*, XIII—1983, 1–198, Lecture Notes in Math., 1117, Springer, Berlin.

Aletti G., May C., Secchi P. (2009). A central limit theorem, and related results, for a two-color randomly reinforced urn. Currently available at: http://arxiv.org/abs/0811.2097.

Aletti G., May C., Secchi P. (2007). On the distribution of the limit proportion for a two-color, randomly reinforced urn with equal reinforcement distributions. *Adv. in Appl. Probab.* **39** no. 3 690–707.

Bassetti F., Crimaldi I. and Leisen F. (2008). Conditionally identically distributed species sampling sequences. Currently available at: http://arxiv.org/abs/0806.2724

Bay Z.D. and Hu F. (2005). Asymptotics in randomized urn models. *A*nn. Appl. Probab. **15**, 914-940.

Berti P., Crimaldi I., Pratelli L. and Rigo P. (2009). A central limit theorem and its applications to multicolor randomly reinforced urns. Currently available at: http://arxiv.org/abs/0904.0932

Berti P., Pratelli L. and Rigo P. (2004). Limit Theorems for a Class of Identically Distributed Random Variables. *Ann. Probab.* **32** 2029–2052.

Blackwell D. and MacQueen J.B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1** 353–355.

Crimaldi I. and Leisen F. (2008). Asymptotic results for a generalized Pólya urn with "multi-updating" and applications to clinical trials. *Communications in Statistics - Theory and Methods*, **37**(17), 2777-2794.

Crimaldi I. (2009). An almost sure conditional convergence result and an application to a generalized Pólya urn. *International Mathematical Forum*, **23**(4), 1139-1156.

Crimaldi I., Letta G. and Pratelli L. (2007). A strong form of stable convergence. *Séminaire de Probabilités XL*, LNM **1899** 203–225.

Ferguson T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 1047–1054.

Flournoy N. and May C. (2009). Asymptotics in response-adaptive designs generated by a two-colors, randomly reinforced urn. *Ann. Statist.* **37**, 1058-1078.

Hall P. and Heyde C.C. (1980). Martingale limit theory and its application. Academic press.

Hansen B. and Pitman J. (2000). Prediction rules for exchangeable sequences related to species sampling. *Statist. Probab. Lett.* **46** 251–256.

Jacod J. and Memin J. (1981). Sur un type de convegence en intermédiaire entre la convergence en loi et la convergence en probabilité. *Sem. de Prob. XV*, LNM **850** 529–546.

Janson S. (2005). Limit theorems for triangular urn schemes. *P*robab. Theo. Rel. Fields. **134**, 417-452.

May C., Paganoni A. and Secchi P. (2005). On a two-color generalized Pólya urn. *Metron* **63** no. 1 115–134.

Pemantle R. (2007). A survey of random processes with reinforcement. *P*robab. Surveys **4**, 1-79.

Pitman J. (2006) *Combinatorial Stochastic Processes.* Ecole d'Eté de Probabilités de Saint-Flour XXXII, LNM 1875, Springer.

Pitman J. (1996). Some developments of the Blackwell-MacQueen urn scheme, In: Ferguson, T.S. et al. (Eds.), *Statistics, Probability and Game Theory; Papers in honor of David Blackwell,* Lecture Notes-Monograph Series, Institute of Mathematical Statistics, Hayward, California, **30** 245–267.

Pitman J. and Yor M. (1997). The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25** 855–900.

Rényi A. (1963). On Stable Sequences of Events, *Sankhya* Ser. A. **25** 293–302.

Williams D. (1991). *Probability with martingales.* Cambridge University Press.