The Annals of Probability
2013, Vol. 41, No. 3B, 2090–2102
DOI: 10.1214/12-AOP786

EXCHANGEABLE SEQUENCES DRIVEN BY AN ABSOLUTELY CONTINUOUS RANDOM MEASURE

BY PATRIZIA BERTI, LUCA PRATELLI AND PIETRO RIGO

Universita' di Modena e Reggio-Emilia, Accademia Navale di Livorno and Universita' di Pavia

Let S be a Polish space and $(X_n : n \geq 1)$ an exchangeable sequence of S-valued random variables. Let $\alpha_n(\cdot) = P(X_{n+1} \in \cdot \mid X_1, \dots, X_n)$ be the predictive measure and α a random probability measure on S such that $\alpha_n \xrightarrow{\text{weak}} \alpha$ a.s. Two (related) problems are addressed. One is to give conditions for $\alpha \ll \lambda$ a.s., where λ is a (nonrandom) σ-finite Borel measure on S. Such conditions should concern the finite dimensional distributions $\mathcal{L}(X_1, \dots, X_n)$, $n \geq 1$, only. The other problem is to investigate whether $\|\alpha_n - \alpha\| \xrightarrow{\text{a.s.}} 0$, where $\|\cdot\|$ is total variation norm. Various results are obtained. Some of them do not require exchangeability, but hold under the weaker assumption that (X_n) is conditionally identically distributed, in the sense of [*Ann. Probab.* **32** (2004) 2029–2052].

1. Two related problems. Throughout, S is a Polish space and

$$X = (X_1, X_2, \dots)$$

a sequence of S-valued random variables on the probability space $(\Omega, \mathcal{A}, P)$. We let $\mathcal{B}$ denote the Borel σ-field on S and $\mathbb{S}$ the set of probability measures on $\mathcal{B}$. A random probability measure on S is a map $\alpha : \Omega \to \mathbb{S}$ such that $\sigma(\alpha) \subset \mathcal{A}$, where $\sigma(\alpha)$ is the σ-field on Ω generated by $\omega \mapsto \alpha(\omega)(B)$ for all $B \in \mathcal{B}$.

For each $n \geq 1$, let α_n be the nth *predictive measure*. Thus, α_n is a random probability measure on S, and $\alpha_n(\cdot)(B)$ is a version of $P(X_{n+1} \in B \mid X_1, \dots, X_n)$ for all $B \in \mathcal{B}$. Define also $\alpha_0(\cdot) = P(X_1 \in \cdot)$.

If X is *exchangeable*, as assumed in this section, there is a random probability measure α on S such that

$$\alpha_n(\omega) \xrightarrow{\text{weak}} \alpha(\omega) \qquad \text{for almost all } \omega \in \Omega.$$

Such an α can also be viewed as

$$\mu_n(\omega) \xrightarrow{\text{weak}} \alpha(\omega) \qquad \text{for almost all } \omega \in \Omega,$$

Received March 2011; revised April 2012.
MSC2010 subject classifications. 60G09, 60G42, 60G57, 62F15.
Key words and phrases. Conditional identity in distribution, exchangeability, predictive measure, random probability measure.

where $\mu_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ is the empirical measure. Further, α grants the usual representation

$$P(X \in B) = \int \alpha(\omega)^\infty(B) P(d\omega) \qquad \text{for every Borel set } B \subset S^\infty,$$

where $\alpha(\omega)^\infty = \alpha(\omega) \times \alpha(\omega) \times \cdots$.

Let λ be a σ-finite measure on $\mathcal{B}$. Our *first problem* is to give conditions for

$$\alpha(\omega) \ll \lambda \qquad \text{for almost all } \omega \in \Omega. \tag{1}$$

The conditions should concern the finite dimensional distributions $\mathcal{L}(X_1, \dots, X_n)$, $n \geq 1$, only.

While investigating (1), one meets another problem, of possible independent interest. Let $\|\cdot\|$ denote total variation norm on $(S, \mathcal{B})$. Our *second problem* is to give conditions for

$$\|\alpha_n - \alpha\| \xrightarrow{\text{a.s.}} 0.$$

2. Motivations. Again, let $X = (X_1, X_2, \dots)$ be exchangeable.

Reasonable conditions for (1) look of theoretical interest. They are of practical interest as well thanks to Bayesian nonparametrics. In this framework, the starting point is a prior π on $\mathbb{S}$. Since $\pi = P \circ \alpha^{-1}$, condition (1) is equivalent to

$$\pi\{\nu \in \mathbb{S} : \nu \ll \lambda\} = 1.$$

This is a basic information for the subsequent statistical analysis. Roughly speaking, it means that the "underlying statistical model" consists of absolutely continuous laws.

Notwithstanding the significance of (1), however, there is a growing literature which gets around the first problem of this paper. Indeed, in a plenty of Bayesian nonparametric problems, condition (1) is just a crude *assumption* and the prior π is directly assessed on a set of densities (with respect to λ). See, for example, [11] and references therein. Instead, it seems reasonable to get (1) as a consequence of explicit assumptions on the finite dimensional distributions $\mathcal{L}(X_1, \dots, X_n)$, $n \geq 1$. From a foundational point of view, in fact, only assumptions on observable facts make sense. This attitude is strongly supported by de Finetti, among others. When dealing with the sequence X, the observable facts are events of the type $\{(X_1, \dots, X_n) \in B\}$ for some $n \geq 1$ and $B \in \mathcal{B}^n$. This is why, in this paper, the conditions for (1) are requested to concern $\mathcal{L}(X_1, \dots, X_n)$, $n \geq 1$, only.

Some references related to the above remarks are [3] and [5–10]. In particular, in [6] and [7], Diaconis and Freedman have an exchangeable sequence of indicators and give conditions for the mixing measure (i.e., the prior π) to be absolutely continuous with respect to Lebesgue measure. The present paper is much in the spirit of [6] and [7]. The main difference is that we give conditions for the mixands $\{\alpha(\omega) : \omega \in \Omega\}$, and not for the mixing measure π, to be absolutely continuous.

Next, a necessary condition for (1) is

$$\mathcal{L}(X_1, \dots, X_n) \ll \lambda^n \qquad \text{for all } n \geq 1, \tag{2}$$

where $\lambda^n = \lambda \times \cdots \times \lambda$. Condition (2) clearly involves the finite dimensional distributions only. Thus, a (natural) question is whether (2) suffices for (1) as well.

The answer is yes provided α can be approximated by the predictive measures α_n in some stronger sense. In fact, condition (2) can be written as

$$\alpha_n(\omega) \ll \lambda \qquad \text{for all } n \geq 0 \text{ and almost all } \omega \in \Omega.$$

Hence, if (2) holds and $\|\alpha_n - \alpha\| \xrightarrow{\text{a.s.}} 0$, the set

$$A = \{\|\alpha_n - \alpha\| \to 0\} \cap \{\alpha_n \ll \lambda \text{ for all } n \geq 0\}$$

has probability 1. And, for each $\omega \in A$, one obtains

$$\alpha(\omega)(B) = \lim_n \alpha_n(\omega)(B) = 0 \qquad \text{whenever } B \in \mathcal{B} \text{ and } \lambda(B) = 0.$$

Therefore, (1) follows from (2) and $\|\alpha_n - \alpha\| \xrightarrow{\text{a.s.}} 0$. In addition, a martingale argument implies the converse implication, that is,

$$\alpha \ll \lambda \text{ a.s.} \quad \Longleftrightarrow \quad \|\alpha_n - \alpha\| \xrightarrow{\text{a.s.}} 0 \quad \text{and} \quad \mathcal{L}(X_1, \dots, X_n) \ll \lambda^n \qquad \text{for all } n;$$

see Theorem 1. Thus, our first problem turns into the second one.

The question of whether $\|\alpha_n - \alpha\| \xrightarrow{\text{a.s.}} 0$ is of independent interest. Among other things, it is connected to Bayesian consistency. Surprisingly, however, this question seems not answered so far. To the best of our knowledge, $\|\alpha_n - \alpha\| \xrightarrow{\text{a.s.}} 0$ in every example known so far. And in fact, for some time, we conjectured that $\|\alpha_n - \alpha\| \xrightarrow{\text{a.s.}} 0$ under condition (2). But this is not true. As shown in Example 5, when $S = \mathbb{R}$ and $\lambda =$ Lebesgue measure, it may be that $\mathcal{L}(X_1, \dots, X_n) \ll \lambda^n$ for all n, and yet α is singular continuous a.s. Indeed, the (topological) support of $\alpha(\omega)$ has Hausdorff dimension 0 for almost all $\omega \in \Omega$.

Thus, (2) does not suffice for (1). To get (1), in addition to (2), one needs some growth conditions on the conditional densities. We refer to forthcoming Theorem 4 for such conditions. Here, we mention a result on the second problem. Actually, for $\|\alpha_n - \alpha\| \xrightarrow{\text{a.s.}} 0$, it suffices that

$$P\{\omega : \alpha_c(\omega) \ll \lambda\} = 1,$$

where $\alpha_c(\omega)$ denotes the continuous part of $\alpha(\omega)$; see Theorem 2.

Finally, most results mentioned above do not need exchangeability of X, but the weaker assumption

$$(X_1, \dots, X_n, X_{n+2}) \sim (X_1, \dots, X_n, X_{n+1}) \qquad \text{for all } n \geq 0.$$

Those sequences X satisfying the above condition, investigated in [2], are called *conditionally identically distributed* (c.i.d.).

3. Mixtures of i.i.d. absolutely continuous sequences. In this section, $\mathcal{G}_0 = \{\emptyset, \Omega\}$, $\mathcal{G}_n = \sigma(X_1, \dots, X_n)$ for $n \geq 1$ and $\mathcal{G}_\infty = \sigma(\bigcup_n \mathcal{G}_n)$. If μ is a random probability measure on S, we write $\mu(B)$ to denote the real random variable $\mu(\cdot)(B)$, $B \in \mathcal{B}$. Similarly, if $h : S \to \mathbb{R}$ is a Borel function, integrable with respect to $\mu(\omega)$ for almost all $\omega \in \Omega$, we write $\mu(h)$ to denote $\int h(x)\mu(\cdot)(dx)$.

3.1. *Preliminaries.* Let $X = (X_1, X_2, \dots)$ be c.i.d., as defined in Section 2. Since X needs not be exchangeable, the representation $P(X \in \cdot) = \int \alpha(\omega)^\infty(\cdot)P(d\omega)$ can fail for any α. However, there is a random probability measure α on S such that

$$\sigma(\alpha) \subset \mathcal{G}_\infty \quad \text{and} \quad \alpha_n(B) = E\{\alpha(B) \mid \mathcal{G}_n\} \qquad \text{a.s.} \tag{3}$$

for all $B \in \mathcal{B}$. In particular, $\alpha_n \xrightarrow{\text{weak}} \alpha$ a.s. Also, letting

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$$

be the empirical measure, one obtains $\mu_n \xrightarrow{\text{weak}} \alpha$ a.s. Such an α is of interest for one more reason. There is an exchangeable sequence $Y = (Y_1, Y_2, \dots)$ of S-valued random variables on $(\Omega, \mathcal{A}, P)$ such that

$$(X_n, X_{n+1}, \dots) \xrightarrow{d} Y \quad \text{and} \quad P(Y \in \cdot) = \int \alpha(\omega)^\infty(\cdot)P(d\omega).$$

See [2] for details.

We next recall some known facts about vector-valued martingales; see [14]. Let $(\mathcal{Z}, \|\cdot\|_*)$ be a separable Banach space. Also, let $\mathcal{F} = (\mathcal{F}_n)$ be a filtration and (Z_n) a sequence of $\mathcal{Z}$-valued random variables on $(\Omega, \mathcal{A}, P)$ such that $E\|Z_n\|_* < \infty$ for all n. Then, (Z_n) is an $\mathcal{F}$-martingale in case $(\phi(Z_n))$ is an $\mathcal{F}$-martingale for each linear continuous functional $\phi : \mathcal{Z} \to \mathbb{R}$. If (Z_n) is an $\mathcal{F}$-martingale, $(\|Z_n\|_*)$ is a real-valued $\mathcal{F}$-submartingale. So, Doob's maximal inequality yields

$$E\Big\{\sup_n \|Z_n\|_*^p\Big\} \leq \left(\frac{p}{p-1}\right)^p \sup_n E\{\|Z_n\|_*^p\} \qquad \text{for all } p > 1.$$

The following martingale convergence theorem is available as well. Let $Z : \Omega \to \mathcal{Z}$ be $\mathcal{F}_\infty$-measurable and such that $E\|Z\|_* < \infty$, where $\mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$. Then, $Z_n \xrightarrow{\text{a.s.}} Z$ provided $\phi(Z_n) = E\{\phi(Z) \mid \mathcal{F}_n\}$ a.s. for all n and all linear continuous functionals $\phi : \mathcal{Z} \to \mathbb{R}$.

3.2. *Results.* In the sequel, λ is a σ-finite measure on $\mathcal{B}$. When $S = \mathbb{R}$, it may be natural to think of λ as the Lebesgue measure, but this is only a particular case. Indeed, λ could be singular continuous or concentrated on any Borel subset. In addition, X is c.i.d. (in particular, exchangeable), and α is a random probability measure on S such that $\alpha_n \xrightarrow{\text{weak}} \alpha$ a.s. Equivalently, α can be obtained as $\mu_n \xrightarrow{\text{weak}} \alpha$ a.s. It can (and will) be assumed $\sigma(\alpha) \subset \mathcal{G}_\infty$.

THEOREM 1. *Suppose $X = (X_1, X_2, \dots)$ is c.i.d. Then, $\alpha \ll \lambda$ a.s. if and only if $\|\alpha_n - \alpha\| \xrightarrow{a.s.} 0$ and $\mathcal{L}(X_1, \dots, X_n) \ll \lambda^n$ for all n.*

PROOF. The "if" part can be proved exactly as in Section 2. Conversely, suppose $\alpha \ll \lambda$ a.s. It can be assumed $\alpha(\omega) \ll \lambda$ for all $\omega \in \Omega$. We let $L_p = L_p(S, \mathcal{B}, \lambda)$ for each $1 \le p \le \infty$.

Let $f : \Omega \times S \to [0, \infty)$ be such that $\alpha(\omega)(dx) = f(\omega, x)\lambda(dx)$ for all $\omega \in \Omega$. Since $\mathcal{B}$ is countably generated, f can be taken to be $\mathcal{A} \otimes \mathcal{B}$-measurable (see [4], V.5.58, page 52) so that

$$1 = \int 1\, dP = \int\int f(\omega, x)\lambda(dx)P(d\omega) = \int E\{f(\cdot, x)\}\lambda(dx).$$

Thus, given $n \ge 0$, $E\{f(\cdot, x) \mid \mathcal{G}_n\}$ is well defined for λ-almost all $x \in S$. Since X is c.i.d., condition (3) also implies

$$\begin{aligned}\int_B E\{f(\cdot, x) \mid \mathcal{G}_n\}\lambda(dx) &= E\left\{\int_B f(\cdot, x)\lambda(dx) \mid \mathcal{G}_n\right\}\\ &= E\{\alpha(B) \mid \mathcal{G}_n\} = \alpha_n(B) \qquad \text{a.s. for fixed } B \in \mathcal{B}.\end{aligned}$$

Since $\mathcal{B}$ is countably generated, the previous equality yields

$$\alpha_n(\omega)(dx) = E\{f(\cdot, x) \mid \mathcal{G}_n\}(\omega)\lambda(dx) \qquad \text{for almost all } \omega \in \Omega.$$

This proves that $\mathcal{L}(X_1, \dots, X_n) \ll \lambda^n$ for all n. In particular, up to modifying α_n on a P-null set, it can be assumed $\alpha_n(\omega)(dx) = f_n(\omega, x)\lambda(dx)$ for all $n \ge 0$, all $\omega \in \Omega$, and suitable functions $f_n : \Omega \times S \to [0, \infty)$.

Regard $f, f_n : \Omega \to L_1$ as L_1-valued random variables. Then, $f : \Omega \to L_1$ is $\mathcal{G}_\infty$-measurable for $\int h(x) f(\cdot, x)\lambda(dx) = \alpha(h)$ is $\mathcal{G}_\infty$-measurable for all $h \in L_\infty$. Clearly, $\|f(\omega, \cdot)\|_{L_1} = \|f_n(\omega, \cdot)\|_{L_1} = 1$ for all n and ω. Finally, X c.i.d. implies

$$\begin{aligned}E\left\{\int h(x) f(\cdot, x)\lambda(dx) \mid \mathcal{G}_n\right\} &= E\{\alpha(h) \mid \mathcal{G}_n\} = \alpha_n(h)\\ &= \int h(x) f_n(\cdot, x)\lambda(dx) \qquad \text{a.s. for all } h \in L_\infty.\end{aligned}$$

By the martingale convergence theorem (see Section 3.1) $f_n \xrightarrow{\text{a.s.}} f$ in the space L_1, that is,

$$\|\alpha_n(\omega) - \alpha(\omega)\| = \frac{1}{2}\int |f_n(\omega, x) - f(\omega, x)|\lambda(dx) \longrightarrow 0$$

for almost all $\omega \in \Omega$.

□

In the exchangeable case, the argument of the previous proof yields a little bit more. Indeed, if X is exchangeable and $\alpha \ll \lambda$ a.s., then

$$\sup_{B \in \mathcal{B}^k} |P\{(X_{n+1}, \dots, X_{n+k}) \in B \mid \mathcal{G}_n\} - \alpha^k(B)| \xrightarrow{\text{a.s.}} 0,$$

where $k \geq 1$ is any integer and $\alpha^k = \alpha \times \cdots \times \alpha$.

The next result deals with the second problem of Section 1. For each $\nu \in \mathbb{S}$, let ν_c and ν_d denote the continuous and discrete parts of ν, that is, $\nu_d(B) = \sum_{x \in B} \nu\{x\}$ for all $B \in \mathcal{B}$ and $\nu_c = \nu - \nu_d$.

THEOREM 2. *Suppose $X = (X_1, X_2, \dots)$ is c.i.d. and $P\{\omega : \alpha_c(\omega) \ll \lambda\} = 1$. Then, $\|\alpha_n - \alpha\| \xrightarrow{a.s.} 0$ if and only if*

$$\text{there is a set } A_0 \in \mathcal{A} \text{ such that } P(A_0) = 1 \text{ and} \\ \alpha_n(\omega)\{x\} \longrightarrow \alpha(\omega)\{x\} \text{ for all } x \in S \text{ and } \omega \in A_0. \tag{4}$$

(*Recall that $\mathcal{A}$ denotes the basic σ-field on Ω*). *Moreover, condition* (4) *is automatically true if X is exchangeable, so that $\|\alpha_n - \alpha\| \xrightarrow{a.s.} 0$ provided X is exchangeable and $\alpha_c \ll \lambda$ a.s.*

PROOF. The "only if" part is trivial. Suppose condition (4) holds. For each $n \geq 0$, take functions β_n and γ_n on Ω such that $\beta_n(\omega)$ and $\gamma_n(\omega)$ are measures on $\mathcal{B}$ for all $\omega \in \Omega$ and

$$\beta_n(B) = E\{\alpha_c(B) \mid \mathcal{G}_n\}, \qquad \gamma_n(B) = E\{\alpha_d(B) \mid \mathcal{G}_n\} \qquad \text{a.s.}$$

for all $B \in \mathcal{B}$. Since X is c.i.d., condition (3) yields $\alpha_n = \beta_n + \gamma_n$ a.s.

We first prove $\|\beta_n - \alpha_c\| \xrightarrow{\text{a.s.}} 0$. It can be assumed $\alpha_c(\omega) \ll \lambda$ for all $\omega \in \Omega$, so that $\alpha_c(\omega)(dx) = f(\omega, x)\lambda(dx)$ for all $\omega \in \Omega$ and some function $f : \Omega \times S \to [0, \infty)$. For fixed $B \in \mathcal{B}$, arguing as in the proof of Theorem 1, one has

$$\beta_n(B) = E\left\{\int_B f(\cdot, x)\lambda(dx) \mid \mathcal{G}_n\right\} = \int_B E\big(f(\cdot, x) \mid \mathcal{G}_n\big)\lambda(dx) \qquad \text{a.s.}$$

By standard arguments, it follows that $\beta_n \ll \lambda$ a.s. Again, it can be assumed $\beta_n(\omega)(dx) = f_n(\omega, x)\lambda(dx)$ for all $\omega \in \Omega$ and some function $f_n : \Omega \times S \to [0, \infty)$. Define $L_1 = L_1(S, \mathcal{B}, \lambda)$ and regard $f_n, f : \Omega \to L_1$ as L_1-valued random variables. By the same martingale argument used for Theorem 1, one obtains $f_n \xrightarrow{\text{a.s.}} f$ in the space L_1. That is, $\|\beta_n - \alpha_c\| \xrightarrow{\text{a.s.}} 0$.

We next prove $\|\gamma_n - \alpha_d\| \xrightarrow{\text{a.s.}} 0$. Take A_0 as in condition (4), and define

$$A_1 = \Big\{\lim_n \|f_n - f\|_{L_1} = 0 \text{ and } \alpha_n = \beta_n + \gamma_n \text{ for all } n \geq 0\Big\}.$$

Then, $P(A_0 \cap A_1) = 1$ and

$$\begin{aligned} \alpha_d(\omega)\{x\} &= \alpha(\omega)\{x\} - \alpha_c(\omega)\{x\} = \alpha(\omega)\{x\} - f(\omega, x)\lambda\{x\} \\ &= \lim_n\big(\alpha_n(\omega)\{x\} - f_n(\omega, x)\lambda\{x\}\big) = \lim_n\big(\alpha_n(\omega)\{x\} - \beta_n(\omega)\{x\}\big) \\ &= \lim_n \gamma_n(\omega)\{x\} \end{aligned}$$

for all $\omega \in A_0 \cap A_1$ and $x \in S$. Define also

$$A = A_0 \cap A_1 \cap \{\gamma_n(S) \longrightarrow \alpha_d(S)\}.$$

Since $\gamma_n(S) = 1 - \beta_n(S) \xrightarrow{\text{a.s.}} 1 - \alpha_c(S) = \alpha_d(S)$, then $P(A) = 1$. Fix $\omega \in A$ and let $D_\omega = \{x \in S : \alpha(\omega)\{x\} > 0\}$. Then

$$\alpha_d(\omega)(D_\omega) \le \liminf_n \gamma_n(\omega)(D_\omega)$$

since D_ω is countable and $\alpha_d(\omega)\{x\} = \lim_n \gamma_n(\omega)\{x\}$ for all $x \in D_\omega$. Further,

$$\limsup_n \gamma_n(\omega)(D_\omega) \le \limsup_n \gamma_n(\omega)(S) = \alpha_d(\omega)(S) = \alpha_d(\omega)(D_\omega).$$

Therefore, $\lim_n \|\gamma_n(\omega) - \alpha_d(\omega)\| = 0$ is an immediate consequence of

$$\gamma_n(\omega)\{x\} \longrightarrow \alpha_d(\omega)\{x\} \qquad \text{for each } x \in D_\omega,$$
$$\alpha_d(\omega)(D_\omega) = \lim_n \gamma_n(\omega)(D_\omega), \qquad \alpha_d(\omega)\bigl(D_\omega^c\bigr) = \lim_n \gamma_n(\omega)\bigl(D_\omega^c\bigr) = 0.$$

Finally, suppose X is exchangeable. We have to prove condition (4). If S is countable, condition (4) is trivial for $\alpha_n(B) \xrightarrow{\text{a.s.}} \alpha(B)$ for fixed $B \in \mathcal{B}$. If $S = \mathbb{R}$, the Glivenko–Cantelli theorem yields $\sup_x |\mu_n(I_x) - \alpha(I_x)| \xrightarrow{\text{a.s.}} 0$, where $I_x = (-\infty, x]$ and $\mu_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ is the empirical measure. Hence, (4) follows from

$$\sup_x \bigl|\alpha_n(I_x) - \mu_n(I_x)\bigr| \xrightarrow{\text{a.s.}} 0;$$

see Corollary 3.2 of [1]. If S is any uncountable Polish space, take a Borel isomorphism $\psi : S \to \mathbb{R}$. (Thus ψ is bijective with ψ and ψ^{-1} Borel measurable). Then $(\psi(X_n))$ is an exchangeable sequence of real random variables, and condition (4) is a straightforward consequence of

$$\begin{aligned} P\bigl\{\psi(X_{n+1}) \in B \mid \psi(X_1), \dots, \psi(X_n)\bigr\} &= P\bigl\{\psi(X_{n+1}) \in B \mid \mathcal{G}_n\bigr\} \\ &= \alpha_n\bigl(\psi^{-1}B\bigr) \qquad \text{a.s.} \end{aligned}$$

for each Borel set $B \subset \mathbb{R}$. This concludes the proof. $\square$

When X is c.i.d. (but not exchangeable) $\|\alpha_n - \alpha\| \xrightarrow{\text{a.s.}} 0$ needs not be true even if $\alpha_c \ll \lambda$ a.s.

EXAMPLE 3. Let (Z_n) and (U_n) be independent sequences of independent real random variables such that $Z_n \sim \mathcal{N}(0, b_n - b_{n-1})$ and $U_n \sim \mathcal{N}(0, 1 - b_n)$, where $0 = b_0 < b_1 < b_2 < \cdots < 1$ and $\sum_n (1 - b_n) < \infty$. As shown in Example 1.2 of [2],

$$X_n = \sum_{i=1}^n Z_i + U_n$$

is c.i.d. and $X_n \xrightarrow{\text{a.s.}} V$ for some real random variable V. Since $\mu_n \xrightarrow{\text{weak}} \delta_V$ a.s., then $\alpha = \delta_V$ and $\alpha_c \ll \lambda$ a.s. (in fact, $\alpha_c = 0$ a.s.). However, condition (4) fails. In fact, $\mathcal{L}(X_1, \dots, X_n) \ll \lambda^n$ for all n, where λ is Lebesgue measure. Hence, $\alpha_n(\omega)\{V(\omega)\} = 0$ while $\alpha(\omega)\{V(\omega)\} = 1$ for all n and almost all $\omega \in \Omega$.

We now turn to the first problem of Section 1. Recall that condition (2) amounts to $\alpha_n \ll \lambda$ a.s. for all $n \geq 0$. Therefore, up to modifying α_n on a P-null set, under condition (2) one can write

$$\alpha_n(\omega)(dx) = f_n(\omega, x)\lambda(dx)$$

for each $\omega \in \Omega$, each $n \geq 0$ and some function $f_n : \Omega \times S \to [0, \infty)$. We also let

$$\mathcal{K} = \{K : K \text{ compact subset of } S \text{ and } \lambda(K) < \infty\} \quad \text{and}$$
$$\lambda_B(\cdot) = \lambda(\cdot \cap B) \qquad \text{for all } B \in \mathcal{B}.$$

THEOREM 4. *Suppose $X = (X_1, X_2, \dots)$ is c.i.d. and $\mathcal{L}(X_1, \dots, X_n) \ll \lambda^n$ for all n. Then $\alpha \ll \lambda$ a.s. if and only if, for each $K \in \mathcal{K}$,*

(5) *the sequence $(f_n(\omega, \cdot) : n \geq 1)$ is uniformly integrable, in the space $(S, \mathcal{B}, \lambda_K)$, for almost all $\omega \in \Omega$.*

In particular, $\alpha \ll \lambda$ a.s. provided, for each $K \in \mathcal{K}$, there is $p > 1$ such that

$$\sup_n \int_K f_n(\omega, x)^p \lambda(dx) < \infty \qquad \text{for almost all } \omega \in \Omega. \tag{6}$$

Moreover, for condition (6) to be true, it suffices that

$$\sup_n E\left\{\int_K f_n^p \, d\lambda\right\} < \infty.$$

PROOF. If $\alpha \ll \lambda$ a.s., Theorem 1 yields $\|\alpha_n - \alpha\| \xrightarrow{\text{a.s.}} 0$. Thus, $f_n(\omega, \cdot)$ converges in $L_1(S, \mathcal{B}, \lambda)$, for almost all $\omega \in \Omega$, and this implies condition (5). Conversely, we now prove that $\alpha \ll \lambda$ a.s. under condition (5).

Fix a nondecreasing sequence $B_1 \subset B_2 \subset \cdots$ such that $B_n \in \mathcal{B}$, $\lambda(B_n) < \infty$, and $\bigcup_n B_n = S$. Since $\lambda(B_1) < \infty$ and S is Polish, there is $K_1 \in \mathcal{K}$ satisfying $K_1 \subset B_1$ and $\lambda(B_1 \cap K_1^c) < 1$. By induction, for each $n \geq 2$, there is $K_n \in \mathcal{K}$ such that $K_{n-1} \subset K_n \subset B_n$ and $\lambda(B_n \cap K_n^c) < 1/n$. Since X is c.i.d., condition (3) implies

$$\alpha(K_m) = \lim_n E\{\alpha(K_m) \mid \mathcal{G}_n\} = \lim_n \alpha_n(K_m) \qquad \text{a.s. for all } m \geq 1.$$

Define $H = \bigcup_m K_m$ and $A_H = \{\alpha(H) = 1\}$. If $\omega \in A_H$, then

$$\alpha(\omega)(B) = \alpha(\omega)(B \cap H) = \sup_m \alpha(\omega)(B \cap K_m) \qquad \text{for all } B \in \mathcal{B}.$$

Moreover, $P(A_H) = 1$. In fact, $\lambda(H^c) = 0$ and $\alpha_n \ll \lambda$ a.s. for all n, so that

$$\alpha(H) = \lim_n E\{\alpha(H) \mid \mathcal{G}_n\} = \lim_n \alpha_n(H) = 1 \qquad \text{a.s.}$$

Thus, to prove $\alpha \ll \lambda$ a.s., it suffices to see that $\alpha(\cdot \cap K_m) \ll \lambda$ a.s. for all m.

Suppose (5) holds. Fix $m \geq 1$, define $K = K_m$ and take a set $A \in \mathcal{A}$ such that $P(A) = 1$ and, for each $\omega \in A$,

$$\alpha(\omega)(K) = \lim_n \alpha_n(\omega)(K), \qquad \alpha_n(\omega) \xrightarrow{\text{weak}} \alpha(\omega),$$

$$\big(f_n(\omega, \cdot) : n \geq 1\big) \qquad \text{is uniformly integrable in } (S, \mathcal{B}, \lambda_K).$$

Let $\omega \in A$. Since $\lambda_K(S) = \lambda(K) < \infty$ and $(f_n(\omega, \cdot) : n \geq 1)$ is uniformly integrable under λ_K, there is a subsequence (n_j) and a function $\psi_\omega \in L_1(S, \mathcal{B}, \lambda_K)$ such that $f_{n_j}(\omega, \cdot) \longrightarrow \psi_\omega$ in the weak-topology of $L_1(S, \mathcal{B}, \lambda_K)$. This means that

$$\int_{B \cap K} \psi_\omega(x)\lambda(dx) = \lim_j \int_{B \cap K} f_{n_j}(\omega, x)\lambda(dx) = \lim_j \alpha_{n_j}(\omega)(B \cap K) \qquad \text{for all } B \in \mathcal{B}.$$

Therefore,

$$\int_K \psi_\omega(x)\lambda(dx) = \lim_j \alpha_{n_j}(\omega)(K) = \alpha(\omega)(K) \quad \text{and}$$

$$\int_{F \cap K} \psi_\omega(x)\lambda(dx) = \lim_j \alpha_{n_j}(\omega)(F \cap K) \leq \alpha(\omega)(F \cap K) \qquad \text{for each closed } F \subset S.$$

By standard arguments, the previous two relations yield

$$\alpha(\omega)(B \cap K) = \int_{B \cap K} \psi_\omega(x)\lambda(dx) \qquad \text{for all } B \in \mathcal{B}.$$

Thus, $\alpha(\omega)(\cdot \cap K) \ll \lambda$. This proves that condition (5) implies $\alpha \ll \lambda$ a.s.

Next, since $p > 1$, it is obvious that (6) $\Longrightarrow$ (5). Hence, it remains only to see that condition (6) follows from $\sup_n E\{\int_K f_n^p \, d\lambda\} < \infty$.

Fix $B \in \mathcal{B}$, $p > 1$, and suppose $\sup_n E\{\int_B f_n^p \, d\lambda\} < \infty$. Let $L_r = L_r(S, \mathcal{B}, \lambda_B)$ for all r. It can be assumed $\int_B f_n(\omega, x)^p \lambda(dx) < \infty$ for all $\omega \in \Omega$ and $n \geq 1$. Thus, each $f_n : \Omega \to L_p$ can be seen as an L_p-valued random variable such that

$$E\{\|f_n\|_{L_p}\} = E\left\{\left(\int_B f_n^p \, d\lambda\right)^{1/p}\right\} \leq \left(E\left\{\int_B f_n^p \, d\lambda\right\}\right)^{1/p} < \infty.$$

Further, $\int f_n(\cdot, x)h(x)\lambda_B(dx) = \alpha_n(I_B h)$ is $\mathcal{G}_n$-measurable for all $h \in L_q$, where $q = p/(p-1)$. Since X is c.i.d., condition (3) also implies

$$\begin{aligned}
E&\left\{\int f_{n+1}(\cdot, x)h(x)\lambda_B(dx) \mid \mathcal{G}_n\right\} \\
&= E\{\alpha_{n+1}(I_B h) \mid \mathcal{G}_n\} \\
&= E\{E(\alpha(I_B h) \mid \mathcal{G}_{n+1}) \mid \mathcal{G}_n\} \\
&= E\{\alpha(I_B h) \mid \mathcal{G}_n\} = \alpha_n(I_B h) \\
&= \int f_n(\cdot, x)h(x)\lambda_B(dx) \qquad \text{a.s. for all } h \in L_q.
\end{aligned}$$

Hence, (f_n) is a $(\mathcal{G}_n)$-martingale. By Doob's maximal inequality,

$$E\left\{\sup_n \int_B f_n^p \, d\lambda\right\} = E\left\{\sup_n \|f_n\|_{L_p}^p\right\}$$
$$\le q^p \sup_n E\{\|f_n\|_{L_p}^p\} = q^p \sup_n E\left\{\int_B f_n^p \, d\lambda\right\} < \infty.$$

In particular, $\sup_n \int_B f_n^p \, d\lambda < \infty$ a.s., and this completes the proof. □

Some remarks on Theorem 4 are in order.

First, for $S = [0, 1]$ and a particular class of exchangeable sequences, results similar to Theorem 4 are in [12] and [13].

Second,

$$f_n(\omega, \cdot) = \frac{g_{n+1}(X_1(\omega), \dots, X_n(\omega), \cdot)}{g_n(X_1(\omega), \dots, X_n(\omega))} \qquad \text{for almost all } \omega \in \Omega,$$

where each $g_n : S^n \to [0, \infty)$ is a density of $\mathcal{L}(X_1, \dots, X_n)$ with respect to λ^n. Thus, more concretely, one obtains

$$\int_K f_n^p \, d\lambda = \frac{\int_K g_{n+1}(X_1, \dots, X_n, x)^p \lambda(dx)}{g_n(X_1, \dots, X_n)^p} \qquad \text{a.s.}$$

Third, suppose X exchangeable, and fix *any* random probability measure γ on S such that $P(X \in \cdot) = \int \gamma(\omega)^\infty(\cdot) P(d\omega)$. Then $\gamma \ll \lambda$ a.s. under the assumptions of Theorem 4. In fact, α and γ have the same probability distribution, when regarded as $\mathbb{S}$-valued random variables.

A last (and important) remark deals with condition (2). Indeed, even if X is exchangeable, condition (2) is not enough for $\alpha \ll \lambda$ a.s. We close the paper showing this fact.

EXAMPLE 5. Let $S = \mathbb{R}$ and $\lambda =$ Lebesgue measure. All random variables are defined on the probability space $(\Omega, \mathcal{A}, P)$. We now exhibit an exchangeable sequence X such that $\mathcal{L}(X_1, \dots, X_n) \ll \lambda^n$ for all $n \ge 1$ and yet $P(\alpha \ll \lambda) = 0$. In fact, the support of $\alpha(\omega)$ has Hausdorff dimension 0 for almost all $\omega \in \Omega$.

Two known facts are to be recalled. First, if T and Z are independent $\mathbb{R}^n$-valued random variables, then

$$P(T + Z \in B) = \int P(T + z \in B) P_Z(dz),$$

where $B \in \mathcal{B}^n$ and P_Z is the distribution of Z. Hence, $\mathcal{L}(T + Z) \ll \lambda^n$ provided $\mathcal{L}(T) \ll \lambda^n$. The second fact is the following:

THEOREM 6 (Pratsiovytyi and Feshchenko). *Let $Z_1, Z_2, \dots$ be i.i.d. random variables with $P(Z_1 = 0) = P(Z_1 = 1) = 1/2$ and $b_1 > b_2 > \cdots > 0$ real numbers such that $\sum_m b_m < \infty$. Then the support of $\mathcal{L}(\sum_m b_m Z_m)$ has Hausdorff dimension 0 whenever $\lim_m (\sum_{j>m} b_j)^{-1} b_m = \infty$.*

Theorem 6 is a consequence of Theorem 8 of [15] (which is actually much more general).

Next, let U_m and $Y_{m,n}$ be independent real random variables such that:

- U_m is uniformly distributed on $(\frac{1}{m+1}, \frac{1}{m})$ for each $m \geq 1$;
- $P(Y_{m,n} = 0) = P(Y_{m,n} = 1) = \frac{1}{2}$ for all $m, n \geq 1$.

Define $V_m = U_m^m$ and

$$X_n = \sum_{m=1}^{\infty} U_m^m Y_{m,n} = \sum_{m=1}^{\infty} V_m Y_{m,n}.$$

Then, $X = (X_1, X_2, \dots)$ is conditionally i.i.d. given $\mathcal{V} = \sigma(V_1, V_2, \dots)$. Precisely, for $\omega \in \Omega$ and $B \in \mathcal{B}$, define

$$\alpha(\omega)(B) = P\left\{u \in \Omega : \sum_m V_m(\omega) Y_{m,1}(u) \in B\right\}.$$

Then, $\alpha(B)$ is a version of $P(X_1 \in B \mid \mathcal{V})$ and $P(X \in \cdot) = \int \alpha(\omega)^\infty(\cdot) P(d\omega)$. In particular, X is exchangeable. Moreover, $\mu_n \xrightarrow{\text{weak}} \alpha$ a.s. for

$$P(\mu_n \xrightarrow{\text{weak}} \alpha \mid \mathcal{V}) = 1 \qquad \text{a.s.}$$

The (topological) support of $\alpha(\omega)$ has Hausdorff dimension 0 for almost all $\omega \in \Omega$. Define in fact $b_m = V_m(\omega)$ and $Z_m = Y_{m,1}$. By Theorem 6, it suffices to verify that

$$\lim_m \frac{V_m(\omega)}{\sum_{j>m} V_j(\omega)} = \infty \qquad \text{for almost all } \omega \in \Omega. \tag{7}$$

And condition (7) follows immediately from

$$(j+1)^{-j} < V_j < j^{-j} \quad \text{and}$$

$$\sum_{j>m} V_j \leq \sum_{j>m} j^{-j} \leq \sum_{j>m} (m+1)^{-j} = \frac{(m+1)^{-m}}{m} \qquad \text{a.s.}$$

We finally prove that $\mathcal{L}(X_1, \dots, X_n) \ll \lambda^n$ for all $n \geq 1$. Given the array $y = (y_{m,n} : m, n \geq 1)$, with $y_{m,n} \in \{0, 1\}$ for all m, n, define

$$X_{n,y} = \sum_m V_m y_{m,n}.$$

Fix $n \geq 1$ and denote I_n the $n \times n$ identity matrix. If y satisfies

$$\begin{pmatrix} y_{m+1,1} & \cdots & y_{m+1,n} \\ \cdots & \cdots & \cdots \\ y_{m+n,1} & \cdots & y_{m+n,n} \end{pmatrix} = I_n \qquad \text{for some } m \geq 0, \tag{8}$$

then

$$(X_{1,y}, \dots, X_{n,y}) = (V_{m+1}, \dots, V_{m+n}) + (R_1, \dots, R_n)$$
$$\text{with } (R_1, \dots, R_n) \text{ independent of } (V_{m+1}, \dots, V_{m+n}).$$

In this case, since $\mathcal{L}(V_{m+1}, \dots, V_{m+n}) \ll \lambda^n$, then $\mathcal{L}(X_{1,y}, \dots, X_{n,y}) \ll \lambda^n$. Hence, letting $Y = (Y_{m,n} : m, n \geq 1)$, the conditional distribution of $(X_1, \dots, X_n)$ given $Y = y$ is absolutely continuous with respect to λ^n as far as y satisfies (8). To complete the proof, it suffices to note that

$$P\big(Y = y \text{ for some } y \text{ satisfying (8)}\big) = 1.$$

REFERENCES

[1] BERTI, P., MATTEI, A. and RIGO, P. (2002). Uniform convergence of empirical and predictive measures. *Atti Semin. Mat. Fis. Univ. Modena* **50** 465–477. MR1958292

[2] BERTI, P., PRATELLI, L. and RIGO, P. (2004). Limit theorems for a class of identically distributed random variables. *Ann. Probab.* **32** 2029–2052. MR2073184

[3] CIFARELLI, D. M. and REGAZZINI, E. (1996). De Finetti's contribution to probability and statistics. *Statist. Sci.* **11** 253–282. MR1445983

[4] DELLACHERIE, C. and MEYER, P. A. (1982). *Probabilities and Potential. B: Theory of Martingales. North-Holland Mathematics Studies* **72**. North-Holland, Amsterdam. MR0745449

[5] DIACONIS, P. and FREEDMAN, D. (1984). Partial exchangeability and sufficiency. In *Statistics: Applications and New Directions* (*Calcutta*, 1981) (J. K. Ghosh and J. Roy, eds.) 205–236. Indian Statist. Inst., Calcutta. MR0786142

[6] DIACONIS, P. and FREEDMAN, D. (2004). The Markov moment problem and de Finetti's theorem. I. *Math. Z.* **247** 183–199. MR2054525

[7] DIACONIS, P. and FREEDMAN, D. (2004). The Markov moment problem and de Finetti's theorem. II. *Math. Z.* **247** 201–212. MR2054526

[8] DIACONIS, P. and FREEDMAN, D. A. (1988). Conditional limit theorems for exponential families and finite versions of de Finetti's theorem. *J. Theoret. Probab.* **1** 381–410. MR0958245

[9] DIACONIS, P. and FREEDMAN, D. A. (1990). Cauchy's equation and de Finetti's theorem. *Scand. J. Stat.* **17** 235–249. MR1092946

[10] FORTINI, S., LADELLI, L. and REGAZZINI, E. (2000). Exchangeability, predictive distributions and parametric models. *Sankhyā Ser. A* **62** 86–109. MR1769738

[11] GHOSAL, S. and VAN DER VAART, A. W. (2013). *Fundamentals of Nonparametric Bayesian Inference*. Cambridge Univ. Press. To appear.

[12] KRAFT, C. H. (1964). A class of distribution function processes which have derivatives. *J. Appl. Probab.* **1** 385–388. MR0171296

[13] MÉTIVIER, M. (1971). Sur la construction de mesures aléatoires presque sûrement absolument continues par rapport à une mesure donnée. *Z. Wahrsch. Verw. Gebiete* **20** 332–344. MR0310931

[14] NEVEU, J. (1975). *Discrete-parameter Martingales*, revised ed. North-Holland, Amsterdam. MR0402915

[15] PRATSIOVYTYI, M. V. and FESHCHENKO, O. Y. (2007). Topological, metric and fractal properties of probability distributions on the set of incomplete sums of positive series. *Theory Stoch. Process.* **13** 205–224. MR2343824

P. BERTI
DIPARTIMENTO DI MATEMATICA PURA
ED APPLICATA "G. VITALI"
UNIVERSITA' DI MODENA E REGGIO-EMILIA
VIA CAMPI 213/B
41100 MODENA
ITALY
E-MAIL: patrizia.berti@unimore.it

L. PRATELLI
ACCADEMIA NAVALE DI LIVORNO
VIALE ITALIA 72
57100 LIVORNO
ITALY
E-MAIL: pratel@mail.dm.unipi.it

P. RIGO
DIPARTIMENTO DI MATEMATICA "F. CASORATI"
UNIVERSITA' DI PAVIA
VIA FERRATA 1
27100 PAVIA
ITALY
E-MAIL: pietro.rigo@unipv.it