Electron. Commun. Probab. **31** (2026), article no. 16, 1–15.
https://doi.org/10.1214/26-ECP765
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Weak convergence in probability of predictive distributions

Fabrizio Leisen* Luca Pratelli† Pietro Rigo‡

Abstract

Let (X_n) be a sequence of random variables with values in a standard Borel space S. We investigate the condition

$$E\big\{f(X_{n+1}) \mid X_1, \ldots, X_n\big\} \quad \text{converges in probability,} \qquad (*)$$
$$\text{as } n \to \infty, \text{ for each bounded Borel function } f : S \to \mathbb{R}.$$

Some consequences of (*) are highlighted and various sufficient conditions for it are obtained. In particular, (*) is characterized in terms of stable convergence. It is also shown that, under (*), there is a random probability measure α on S such that $E\big\{f(X_{n+1}) \mid X_1, \ldots, X_n\big\} \xrightarrow{P} \int f\, d\alpha$ for each bounded Borel f. Moreover, since (*) holds whenever (X_n) is conditionally identically distributed, three weak versions of the latter condition are investigated. For each version, our main goal is proving (or disproving) that (*) holds. Several counterexamples are given as well.

Keywords: asymptotic exchangeability; conditional identity in distribution; predictive distribution; random probability measure; stable convergence.
MSC2020 subject classifications: 60B10; 60G57; 60G09; 60F99.
Submitted to ECP on July 25, 2025, final version accepted on February 28, 2026.

1 Introduction

Throughout, $(S, \mathcal{B})$ is a standard Borel space, that is, S is a Borel subset of some Polish space and $\mathcal{B}$ the Borel σ-field on S. We denote by $M_b(S)$ the set of real bounded $\mathcal{B}$-measurable functions on S and we let $C_b(S) = \big\{f \in M_b(S) : f \text{ is continuous}\big\}$.

A *random probability measure* (r.p.m.) on S, defined on a probability space $(\Omega, \mathcal{A}, P)$, is a map γ on $\Omega \times \mathcal{B}$ such that:

$-$ $\gamma(\omega, \cdot)$ is a probability measure on $\mathcal{B}$ for fixed $\omega \in \Omega$;

$-$ the function $\omega \mapsto \gamma(\omega, B)$ is $\mathcal{A}$-measurable for fixed $B \in \mathcal{B}$.

If $f \in M_b(S)$, we denote by $\gamma(f)$ the real random variable $\omega \mapsto \gamma(\omega, f)$ where

$$\gamma(\omega, f) = \int f(x)\, \gamma(\omega, dx).$$

Moreover, if $f = 1_B$ for some $B \in \mathcal{B}$, we let $\gamma(B) = \gamma(1_B)$.

Let γ_n and γ be r.p.m.'s on S. In the sequel, we write

$$\gamma_n \Rightarrow \gamma \text{ a.s.} \quad \text{or} \quad \gamma_n \Rightarrow \gamma \text{ in probability}$$

*King's College London, UK. E-mail: fabrizio.leisen@gmail.com
†Accademia Navale di Livorno, Italy. E-mail: luca_pratelli@marina.difesa.it
‡Universitá di Bologna, Italy. E-mail: pietro.rigo@unibo.it

to denote that $\gamma_n(f) \xrightarrow{a.s.} \gamma(f)$ or $\gamma_n(f) \xrightarrow{P} \gamma(f)$, respectively, for each $f \in C_b(S)$. By a diagonalization argument, it can be shown that $\gamma_n \Rightarrow \gamma$ in probability if and only if, for each subsequence (n_j), there is a sub-subsequence (n_{j_k}) of (n_j) such that $\gamma_{n_{j_k}} \Rightarrow \gamma$ a.s. as $k \to \infty$; see [4].

1.1 Content of this paper

Let $X = (X_1, X_2, \ldots)$ be a sequence of S-valued random variables on the probability space $(\Omega, \mathcal{A}, P)$. Define

$$\mathcal{F}_0 = \big\{\emptyset, \Omega\big\}, \quad \mathcal{F}_n = \sigma(X_1, \ldots, X_n), \quad \text{and}$$
$$\alpha_n(\cdot) = P(X_{n+1} \in \cdot \mid \mathcal{F}_n) \qquad \text{for each } n \geq 0.$$

Each α_n is a r.p.m. on S such that

$$\alpha_n(f) = E\big\{f(X_{n+1}) \mid \mathcal{F}_n\big\} \qquad \text{a.s. for every } f \in M_b(S).$$

The r.p.m.'s α_n are usually called the *predictive distributions* of the sequence X.

In various frameworks, a (natural) question is whether α_n converges, in some sense, as $n \to \infty$. For instance, convergence of α_n is an important issue in Bayesian inference, empirical processes for dependent data, species sampling sequences, urn models and bootstrap; see e.g. [5, 6, 7, 8, 9, 12, 13, 15, 16, 19, 21, 22, 24] and references therein.

Two meaningful modes of convergence for α_n are

$$\alpha_n(f) \text{ converges in probability, as } n \to \infty, \text{ for each } f \in M_b(S) \tag{1.1}$$

and

$$\alpha_n(f) \text{ converges a.s., as } n \to \infty, \text{ for each } f \in M_b(S). \tag{1.2}$$

The latter is strictly stronger than the former, that is, (1.2) trivially implies (1.1) while the converse is not true; see Example 4.1. In addition, (1.2) is more popular than (1.1). In fact, (1.2) implies a.s. weak convergence of predictive distributions, which is the usual notion of convergence for r.p.m.'s.

Even if almost neglected, however, condition (1.1) is of interest as well. Some of its consequences are listed in Section 1.2. Here, we anticipate that (1.1) implies the weak law of large numbers and asymptotic exchangeability. Moreover, under (1.1), the limit of $\alpha_n(f)$ can be written as $\alpha(f)$ for a single r.p.m. α. In particular, (1.1) implies $\alpha_n \Rightarrow \alpha$ in probability for some r.p.m. α.

This paper focuses on condition (1.1). Our main results are a characterization of (1.1) in terms of stable convergence and various sufficient (but not necessary) conditions. Moreover, we prove the consequences of (1.1) mentioned above, namely, asymptotic exchangeability and the existence of a r.p.m. α such that $\alpha_n(f) \xrightarrow{P} \alpha(f)$ for every $f \in M_b(S)$. Another main contribution of this paper is the examples. A number of examples, highlighting the connections between (1.1) and other related conditions, are given.

Finally, this paper has also another target. A well known sufficient condition for (1.2) is that X is *conditionally identically distributed* (c.i.d.), that is

$$(X_1, \ldots, X_n, X_k) \text{ is distributed as } (X_1, \ldots, X_n, X_{n+1}) \quad \text{for all } k > n \geq 0;$$

see [3]. A natural question is whether (1.2) is still true if the c.i.d. condition is weakened. Hence, we introduce three weak versions of the c.i.d. condition. In the spirit of this paper, for each version, our main goal is proving (or disproving) that (1.1) or (1.2) hold. However, (1.1) and (1.2) are not the only reasons for studying these versions. The latter are in fact three possible generalizations of the c.i.d. condition, which could be of interest independently of (1.1) and (1.2).

1.2 Some consequences of condition (1.1)

To motivate condition (1.1), we briefly list some of its consequences.

(i) One of our main results (Theorem 2.1) is that, under (1.1), the limit of $\alpha_n(f)$ can be written as $\alpha(f)$ for a single r.p.m. α. Precisely, if (1.1) holds, there is a r.p.m. α on S such that $\alpha_n(f) \xrightarrow{P} \alpha(f)$ for each $f \in M_b(S)$. In particular, condition (1.1) yields $\alpha_n \Rightarrow \alpha$ in probability for some r.p.m. α.

(ii) Let μ_n be the *empirical measure*, i.e., $\mu_n = (1/n)\sum_{i=1}^n \delta_{X_i}$ where δ_x denotes the unit mass at the point $x \in S$. Say that X *satisfies the weak law of large numbers* if $\mu_n(f) = (1/n)\sum_{i=1}^n f(X_i)$ converges in probability for each $f \in M_b(S)$. By a (well known) martingale argument, one obtains

$$\mu_n(f) - \frac{1}{n}\sum_{i=1}^n \alpha_{i-1}(f) = \frac{1}{n}\sum_{i=1}^n \Big\{ f(X_i) - \alpha_{i-1}(f) \Big\} \xrightarrow{a.s.} 0 \quad \text{for each } f \in M_b(S).$$

Hence, under (1.1), X satisfies the weak law of large numbers. Precisely, $\mu_n(f) \xrightarrow{P} \alpha(f)$ for all $f \in M_b(S)$ where α is the r.p.m. involved in (i).

(iii) By (ii), condition (1.1) implies $\mu_n(f) - \alpha_n(f) \xrightarrow{P} 0$ for each $f \in M_b(S)$. Hence, for large n, the empirical measure μ_n is a good approximation of the predictive α_n. Using a statistical language, under (1.1), μ_n is a consistent estimator of α_n.

(iv) Condition (1.1) implies that X is *asymptotically exchangeable*, that is

$$(X_{n+1}, X_{n+2}, \ldots) \longrightarrow (Z_1, Z_2, \ldots) \quad \text{in distribution, as } n \to \infty,$$

where $(Z_1, Z_2, \ldots)$ is an exchangeable sequence. This is proved in forthcoming Theorem 2.2, which slightly improves a result by Aldous. More precisely, in Lemma (8.2) of [1], Aldous proves that X is asymptotically exchangeable provided $\alpha_n \Rightarrow \alpha$ a.s. for some r.p.m. α. Instead, in Theorem 2.2, asymptotic exchangeability is obtained under the weaker condition that $\alpha_n(f)$ converges in probability for each $f \in C_b(S)$.

2 An analysis of condition (1.1)

This section includes our main results (apart from the examples). We recall that, in such results, $(S, \mathcal{B})$ is a standard Borel space. This assumption is basic. Indeed, if $(S, \mathcal{B})$ is arbitrary, a *regular* conditional distribution for X_{n+1} given $\mathcal{F}_n$ may fail to exist. In addition, convergence of $\alpha_n(f)$ for each f in a given class of test functions does not necessarily imply $\alpha_n \to \alpha$, in some sense, for some r.p.m. α; see Example 2.1 of [4].

Given a sub-σ-field $\mathcal{G} \subset \mathcal{A}$, a r.p.m. α is said to be $\mathcal{G}$-measurable if the function $\omega \mapsto \alpha(\omega, B)$ is $\mathcal{G}$-measurable for fixed $B \in \mathcal{B}$.

Theorem 2.1. *Let $\mathcal{T} = \bigcap_n \sigma(X_n, X_{n+1}, \ldots)$ be the tail σ-field of X. If condition (1.1) holds, there is a r.p.m. α on S such that α is $\mathcal{T}$-measurable and $\alpha_n(f) \xrightarrow{P} \alpha(f)$ for each $f \in M_b(S)$.*

Proof. Suppose condition (1.1) holds. The proof can be split into three steps. The first is just an application of Corollary 2.4 of [4]. In fact, since $(S, \mathcal{B})$ is standard Borel and $\alpha_n(f)$ converges in probability whenever $f \in C_b(S)$, Corollary 2.4 of [4] yields $\alpha_n \Rightarrow \alpha$ in probability for some r.p.m. α on S. The second step is proving that α can be taken to be $\mathcal{T}$-measurable. The third step, based on a monotone class argument, lies in showing that $\alpha_n(f) \xrightarrow{P} \alpha(f)$ for each $f \in M_b(S)$ (and not only for each $f \in C_b(S)$).

We next turn to the second step, that is, we obtain a $\mathcal{T}$-measurable version of α. Let $\mu_n = (1/n)\sum_{i=1}^n \delta_{X_i}$ be the empirical measure. By point (ii) of Section 1.2, since $\alpha_n \Rightarrow \alpha$

in probability, one obtains $\mu_n \Rightarrow \alpha$ in probability. By a diagonalization argument, there is a subsequence (n_j) such that $\mu_{n_j} \Rightarrow \alpha$ a.s. as $j \to \infty$; see Section 1. Define

$$A = \big\{\omega \in \Omega : \mu_{n_j}(\omega, \cdot) \text{ converges weakly as } j \to \infty\big\}$$

and note that $A \in \mathcal{T}$ and $P(A) = 1$. Moreover, fix a probability measure β_0 on $\mathcal{B}$ and define $\beta(\omega, \cdot) = \beta_0$ for $\omega \notin A$ and $\beta(\omega, \cdot) = \lim_j \mu_{n_j}(\omega, \cdot)$ for $\omega \in A$, where $\lim_j$ is meant as the weak limit as $j \to \infty$. Then, β is $\mathcal{T}$-measurable and $\alpha(\omega, \cdot) = \beta(\omega, \cdot)$ for almost all $\omega \in A$. Therefore, β is a $\mathcal{T}$-measurable version of α.

Finally, we focus on the third step, i.e., we prove that $\alpha_n(f) \xrightarrow{P} \alpha(f)$ for all $f \in M_b(S)$. For each $f \in M_b(S)$, fix a random variable $\nu(f)$ satisfying $\alpha_n(f) \xrightarrow{P} \nu(f)$ and define $L = \big\{f \in M_b(S) : \alpha(f) = \nu(f) \text{ a.s.}\big\}$. Then, $L \supset C_b(S)$ and L is a linear space including the constants. Hence, it suffices to show that, for each $f \in M_b(S)$,

$$f \in L \text{ provided } f_k \to f \text{ pointwise for some monotone sequence } (f_k) \subset L.$$

Under this condition, in fact, the monotone class theorem yields $L = M_b(S)$.

We now prove such condition. Fix $f \in M_b(S)$ and a monotone sequence $(f_k) \subset L$ such that $f_k \to f$ pointwise. Since L is a linear space, (f_k) can be assumed to be increasing, i.e. $f_1 \le f_2 \le \dots$ In this case,

$$\alpha(f) = \sup_k \alpha(f_k) = \sup_k \nu(f_k) \le \nu(f) \qquad \text{a.s.}$$

In addition, since $P(X_n \in B) = E\big\{\alpha_{n-1}(B)\big\}$, condition (1.1) implies that $\lim_n P(X_n \in B)$ exists for every $B \in \mathcal{B}$. Define

$$\lambda(B) = \lim_n P(X_n \in B) \quad \text{and} \quad \lambda^*(B) = E\big\{\alpha(B)\big\} \quad \text{for all } B \in \mathcal{B}.$$

By the Vitali-Hahn-Saks theorem, λ is a probability measure on $\mathcal{B}$. Moreover,

$$\lambda(g) = \lim_n E\big\{\alpha_n(g)\big\} = E\big\{\nu(g)\big\} = E\big\{\alpha(g)\big\} = \lambda^*(g) \quad \text{for all } g \in C_b(S).$$

Therefore, $\lambda = \lambda^*$ and this in turn implies

$$E\big\{\nu(f) - \alpha(f)\big\} = E\big\{\nu(f)\big\} - E\big\{\alpha(f)\big\} \le \lim_n E\big\{\alpha_n(f)\big\} - \lambda(f) = \lambda(f) - \lambda(f) = 0,$$

where the inequality is due to Fatou's lemma. Since $\nu(f) - \alpha(f) \ge 0$ a.s., it follows that $\alpha(f) = \nu(f)$ a.s. Hence, $f \in L$ and this concludes the proof. □

An obvious weakening of condition (1.1) is

$$\alpha_n(f) \text{ converges in probability, as } n \to \infty, \text{ for each } f \in C_b(S). \tag{2.1}$$

Condition (2.1) is strictly weaker than (1.1); see Example 4.2. However, it arises quite frequently. For instance, condition (2.1) amounts to $\alpha_n \Rightarrow \alpha$ in probability for some r.p.m. α; see Corollary 2.4 of [4]. To make a further example, let us introduce the condition

$$\lim_n P(X_n \in B) \text{ exists for every } B \in \mathcal{B}. \tag{2.2}$$

Then, by the same argument of Theorem 2.1, it can be shown that

$$(1.1) \iff (2.1) \text{ and } (2.2).$$

Note also that condition (2.2) is trivially true if the sequence X is identically distributed. Finally, condition (2.1) implies asymptotic exchangeability.

Theorem 2.2. *Condition* (2.1) *implies asymptotic exchangeability of* X*.*

As discussed in point (iv), Theorem 2.2 slightly improves a result by Aldous [1].

The proof of Theorem 2.2 is postponed to the end of this section. Here, we note that the converse of Theorem 2.2 does not hold, i.e., asymptotic exchangeability does not imply condition (2.1); see Example 4.3. A further remark is that, since exchangeability amounts to stationarity and asymptotic exchangeability, Theorem 2.2 yields

$$X \text{ exchangeable} \quad \Leftrightarrow \quad X \text{ stationary and condition (2.1) holds.}$$

We now give some characterizations of condition (1.1). Let $\mathcal{A}^+ = \big\{A \in \mathcal{A} : P(A) > 0\big\}$. Recall that X_n *converges stably to* α, where α is a r.p.m. on S, if

$$P(X_n \in \cdot \mid H) \overset{weakly}{\longrightarrow} E\big\{\alpha(\cdot) \mid H\big\} \qquad \text{for all } H \in \mathcal{A}^+.$$

Equivalently, $E\big\{\alpha(f) \mid H\big\} = \lim_n E\big\{f(X_n) \mid H\big\}$ for all $f \in C_b(S)$ and $H \in \mathcal{A}^+$.

Theorem 2.3. *Condition* (1.1) *holds if and only if condition* (2.2) *holds,* X_n *converges stably to* α*, for some r.p.m.* α *on* S*, and*

$$E\big\{\alpha(f)^2\big\} = \lim_n E\big\{\alpha_n(f)^2\big\} \qquad \textit{for all } f \in C_b(S). \tag{2.3}$$

Proof. First note that, if $H \in \mathcal{F}_k$ for some k, then

$$E\big\{1_H\, f(X_n)\big\} = E\big\{1_H\, E(f(X_n) \mid \mathcal{F}_{n-1})\big\} = E\big\{1_H\, \alpha_{n-1}(f)\big\} \quad \text{a.s.} \tag{2.4}$$

for all $f \in M_b(S)$ and $n > k$.

Assume condition (1.1). By Theorem 2.1, there is a r.p.m. α on S such that α is $\mathcal{T}$-measurable and $\alpha_n(f) \xrightarrow{P} \alpha(f)$ for each $f \in M_b(S)$. Hence, condition (2.3) follows from the bounded convergence theorem. Next, fix $f \in C_b(S)$. If $H \in \bigcup_k \mathcal{F}_k$ and $P(H) > 0$, one obtains

$$E\big\{\alpha(f) \mid H\big\} = \lim_n E\big\{\alpha_n(f) \mid H\big\} = \lim_n E\big\{f(X_n) \mid H\big\}$$

where the second equality depends on (2.4). Since $\mathcal{T} \subset \sigma(X)$ and α is $\mathcal{T}$-measurable, by standard arguments, this implies

$$E\big\{\alpha(f) \mid H\big\} = \lim_n E\big\{f(X_n) \mid H\big\} \qquad \text{for all } H \in \mathcal{A}^+.$$

Hence, $X_n \to \alpha$ stably. Finally, condition (2.2) holds since (1.1) $\Leftrightarrow$ (2.1)-(2.2).

Conversely, assume conditions (2.2)-(2.3) and $X_n \to \alpha$ stably. Since (2.2) holds, to obtain (1.1), it suffices to prove condition (2.1). Fix $f \in C_b(S)$. Since $X_n \to \alpha$ stably, using equation (2.4) and arguing as above, one obtains

$$E\big\{\alpha(f) \mid H\big\} = \lim_n E\big\{\alpha_n(f) \mid H\big\} \qquad \text{for all } H \in \mathcal{A}^+. \tag{2.5}$$

Since $\alpha(f)$ is bounded and $\mathcal{A}$-measurable, $\alpha(f)$ is the uniform limit of a sequence (U_k) of $\mathcal{A}$-simple functions. Since $U_k \to \alpha(f)$ uniformly, as $k \to \infty$, equation (2.5) implies

$$\begin{aligned} E\big\{\alpha(f)^2\big\} &= \lim_k E\big\{U_k\, \alpha(f)\big\} = \lim_k \lim_n E\big\{U_k\, \alpha_n(f)\big\} \\ &= \lim_n \lim_k E\big\{U_k\, \alpha_n(f)\big\} = \lim_n E\big\{\alpha(f)\, \alpha_n(f)\big\}. \end{aligned}$$

Hence, condition (2.3) yields

$$E\Big\{\big(\alpha_n(f) - \alpha(f)\big)^2\Big\} = E\big\{\alpha(f)^2\big\} + E\big\{\alpha_n(f)^2\big\} - 2\, E\big\{\alpha_n(f)\, \alpha(f)\big\} \longrightarrow 0.$$

Therefore, $\alpha_n(f) \xrightarrow{P} \alpha(f)$ and this concludes the proof. □

The next result deals with a special case. Recall that $\mu_n = (1/n)\sum_{i=1}^n \delta_{X_i}$ denotes the empirical measure.

Theorem 2.4. *Suppose that, for each $f \in M_b(S)$,*

$$\mu_n(f) \quad \textit{converges in probability to a degenerate random variable.} \tag{2.6}$$

Then, condition (1.1) *is equivalent to condition* (2.2) *and*

$$\lim_n E\big\{\alpha_n(f)^2\big\} = \lim_n E\big\{f(X_n)\big\}^2 \quad \textit{for each } f \in M_b(S). \tag{2.7}$$

Proof. Assume condition (1.1). Then, condition (2.2) holds. Moreover, as noted in point (ii), there is a r.p.m. α such that $\alpha_n(f) \xrightarrow{P} \alpha(f)$ and $\mu_n(f) \xrightarrow{P} \alpha(f)$ for all $f \in M_b(S)$. Since $\mu_n(f) \xrightarrow{P} \alpha(f)$, condition (2.6) implies that $\alpha(f)$ is degenerate. Therefore,

$$\lim_n E\big\{\alpha_n(f)^2\big\} = E\big\{\alpha(f)^2\big\} = \Big(E\big\{\alpha(f)\big\}\Big)^2 = \Big(\lim_n E\big\{\alpha_{n-1}(f)\big\}\Big)^2 = \lim_n E\big\{f(X_n)\big\}^2$$

where the second equality is because $\alpha(f)$ is a degenerate random variable.

Conversely, assume conditions (2.2) and (2.7). By (2.2) and the Vitali-Hahn-Saks theorem, $\lambda(\cdot) := \lim_n P(X_n \in \cdot)$ is a probability measure on $\mathcal{B}$ such that $\lambda(f) = \lim_n E\big\{f(X_n)\big\}$ for each $f \in M_b(S)$. Hence, condition (2.7) yields

$$E\Big\{\big(\alpha_n(f) - \lambda(f)\big)^2\Big\} = E\big\{\alpha_n(f)^2\big\} + \lambda(f)^2 - 2\,\lambda(f)\,E\big\{f(X_{n+1})\big\} \longrightarrow 0. \qquad \square$$

Condition (2.6) is known to be true in various meaningful cases. For instance, it is true if X is stationary and ergodic, or pairwise independent and identically distributed, or an irreducible positive recurrent Markov chain. More generally, condition (2.6) holds if X satisfies the weak law of large numbers (see point (ii) of Section 1.2) and $P(H) \in \{0,1\}$ for each H in the σ-field generated by the random variables $\limsup_j \mu_{n_j}(f)$, where $f \in M_b(S)$ and (n_j) is any subsequence. We also note that, in the special case where X is pairwise independent, Theorem 2.4 is analogous to a characterization of convergence in probability of X_n proved in [14].

To get condition (1.1) via Theorem 2.3, one should first identify the r.p.m. α. We next provide a version of Theorem 2.3 not involving α explicitly.

Theorem 2.5. *Condition* (1.1) *holds if and only if condition* (2.2) *holds and*

$$\lim_n E\big\{f(X_n) \mid H\big\} \quad \textit{exists,} \quad \mu_n(f) \textit{ converges in probability,} \tag{2.8}$$
$$\textit{and} \quad \lim_n E\Big\{\alpha_n(f)^2 - \mu_n(f)^2\Big\} = 0 \quad \textit{for all } f \in C_b(S) \textit{ and } H \in \mathcal{A}^+.$$

Proof. The "only if" part follows trivially from Theorem 2.3 and point (ii). Conversely, assume conditions (2.2) and (2.8). Since $(S, \mathcal{B})$ is standard Borel and $\lim_n E\big\{f(X_n) \mid H\big\}$ exists for all $f \in C_b(S)$ and $H \in \mathcal{A}^+$, there is a r.p.m. α on S such that $X_n \to \alpha$ stably. By Corollary 2.4 of [4], since $\mu_n(f)$ converges in probability for all $f \in C_b(S)$, there is also a r.p.m. μ on S such that $\mu_n \Rightarrow \mu$ in probability. Moreover, by (2.8),

$$E\big\{\mu(f)^2\big\} = \lim_n E\big\{\mu_n(f)^2\big\} = \lim_n E\big\{\alpha_n(f)^2\big\} \qquad \text{for all } f \in C_b(S).$$

Hence, because of Theorem 2.3, it suffices to see that $\alpha(f) = \mu(f)$ a.s. for each $f \in C_b(S)$. Fix $f \in C_b(S)$ and recall that $\mu_n(f) - \frac{1}{n}\sum_{i=1}^n \alpha_{i-1}(f) \xrightarrow{a.s.} 0$; see point (ii) again. For all $H \in \mathcal{A}^+$, one obtains

$$E\big\{\alpha(f) \mid H\big\} = \lim_n E\big\{f(X_n) \mid H\big\} = \lim_n E\big\{\alpha_n(f) \mid H\big\}$$
$$= \lim_n \frac{1}{n}\sum_{i=1}^n E\big\{\alpha_{i-1}(f) \mid H\big\} = \lim_n E\big\{\mu_n(f) \mid H\big\} = E\big\{\mu(f) \mid H\big\}.$$

It follows that $\alpha(f) = \mu(f)$ a.s., and this concludes the proof. $\square$

In view of the next example (and of Section 3.2) we briefly recall the notion of *quasi-martingale*. Let (Y_n) be a sequence of real integrable random variables adapted to a filtration $(\mathcal{G}_n)$. Then, (Y_n) is a quasi-martingale with respect to $(\mathcal{G}_n)$ if

$$\sum_n E\Big|\, E(Y_{n+1} \mid \mathcal{G}_n) - Y_n \Big| < \infty.$$

For instance, a non-negative supermartingale is a quasi-martingale (but not necessarily a martingale). Importantly, Y_n converges a.s. provided it is a quasi-martingale such that $\sup_n E|Y_n| < \infty$.

Example 2.6. (Empirical based predictive distributions) For each $n \geq 1$, fix a constant $c_n \in (0,1)$, a $\mathcal{F}_n$-measurable r.p.m. β_n, and define

$$\alpha_n = c_n\, \beta_n + (1 - c_n)\, \mu_n = \mu_n + c_n\, (\beta_n - \mu_n).$$

Predictive distributions of this type play a role in various frameworks, including empirical processes for dependent data, species sampling models, Bayesian predictive inference and bootstrap; see e.g. [6, 8, 12, 15, 18, 24] and references therein. Anyway, because of Theorem 2.5, condition (1.1) holds whenever

$$c_n \to 0 \quad \text{and} \quad \sum_n \frac{c_n}{n+1} < \infty.$$

To check the conditions of Theorem 2.5, fix $f \in M_b(S)$. Since f is bounded and $c_n \to 0$,

$$E\big\{\alpha_n(f)^2 - \mu_n(f)^2\big\} = c_n^2\, E\big\{(\beta_n(f) - \mu_n(f))^2\big\} + 2\, c_n\, E\big\{\mu_n(f)\, (\beta_n(f) - \mu_n(f))\big\} \longrightarrow 0.$$

Moreover, since

$$\begin{aligned} E\big\{\mu_{n+1}(f) \mid \mathcal{F}_n\big\} &= \frac{n\mu_n(f) + E\big\{f(X_{n+1}) \mid \mathcal{F}_n\big\}}{n+1} = \frac{n\mu_n(f) + \alpha_n(f)}{n+1} \\ &= \mu_n(f) + \frac{c_n}{n+1}\, (\beta_n(f) - \mu_n(f)) \qquad \text{a.s.,} \end{aligned}$$

then

$$\sum_n E\Big|E\big\{\mu_{n+1}(f) \mid \mathcal{F}_n\big\} - \mu_n(f)\Big| \leq 2\, \sup|f| \sum_n \frac{c_n}{n+1} < \infty.$$

Therefore, $\mu_n(f)$ is a quasi-martingale, and this implies that $\mu_n(f)$ converges a.s. In particular, condition (2.2) holds. In fact, letting $f = 1_B$, one obtains

$$\lim_n P(X_n \in B) = \lim_n E\big\{\alpha_{n-1}(B)\big\} = \lim_n E\big\{\mu_{n-1}(B)\big\} \text{ exists for every } B \in \mathcal{B}.$$

It remains only to show that $\lim_n E\big\{f(X_n) \mid H\big\}$ exists for each $H \in \mathcal{A}^+$. To this end, by standard arguments, it can be assumed $H \in \mathcal{F}_k \cap \mathcal{A}^+$ for some $k \geq 1$. In this case, the existence of $\lim_n E\big\{f(X_n) \mid H\big\}$ follows from $c_n \to 0$, $\mu_n(f)$ converges a.s., and

$$E\big\{1_H\, f(X_{n+1})\big\} = E\big\{1_H\, \alpha_n(f)\big\} = E\big\{1_H\, \mu_n(f)\big\} + c_n\, E\big\{1_H\, (\beta_n(f) - \mu_n(f))\big\} \quad \text{for all } n \geq k.$$

To close this section, we provide a proof of Theorem 2.2.

Proof of Theorem 2.2. Suppose condition (2.1) holds. The proof can be split into four steps. The first coincides with the corresponding step of the proof of Theorem 2.1. Indeed, by [4, Cor. 2.4], there is a r.p.m. α on S such that $\alpha_n \Rightarrow \alpha$ in probability. The second step is proving that $X_n \to \alpha$ stably. The third step is showing that

$$\lim_n E\big\{f(X_{n+1})\, g(X_{n+2})\big\} = E\big\{\alpha(f)\, \alpha(g)\big\} \qquad \text{for all } f,\, g \in C_b(S). \tag{2.9}$$

The fourth step is to generalize equation (2.9) by induction.

We skip the second step, since $X_n \to \alpha$ stably can be proved exactly as in the proof of Theorem 2.3. Let us turn to the third step, i.e. let us prove (2.9). For all $f,\, g \in C_b(S)$,

$$E\big\{f(X_{n+1})\, g(X_{n+2})\big\} = E\big\{f(X_{n+1})\, E(g(X_{n+2}) \mid \mathcal{F}_{n+1})\big\}$$
$$= E\big\{f(X_{n+1})\, \alpha_{n+1}(g)\big\} = E\big\{f(X_{n+1})\, \alpha(g)\big\} + E\big\{f(X_{n+1})\, \big(\alpha_{n+1}(g) - \alpha(g)\big)\big\}.$$

Observe now that

$$\Big|E\big\{f(X_{n+1})\, \big(\alpha_{n+1}(g) - \alpha(g)\big)\big\}\Big| \le \sup|f|\, E\Big|\alpha_{n+1}(g) - \alpha(g)\Big| \longrightarrow 0.$$

Moreover, since $X_n \to \alpha$ stably and $\alpha(g)$ is bounded and $\mathcal{A}$-measurable,

$$\lim_n E\big\{f(X_{n+1})\, \alpha(g)\big\} = E\big\{\alpha(f)\, \alpha(g)\big\}.$$

Hence, equation (2.9) follows immediately from the above limit relations.

Finally, regarding the fourth step, we note that

$$\lim_n E\left\{\prod_{j=1}^{k} f_j(X_{n+j})\right\} = E\left\{\prod_{j=1}^{k} \alpha(f_j)\right\} \tag{2.10}$$

for all $k \ge 2$ and all $f_1, \ldots, f_k \in C_b(S)$. In fact, (2.10) reduces to (2.9) for $k = 2$, and it can be proved by induction in the general case. Equation (2.10) precisely means that $(X_{n+1}, X_{n+2}, \ldots)$ converges in distribution, as $n \to \infty$, to the exchangeable sequence $(Z_1, Z_2, \ldots)$ whose probability distribution is given by

$$P(Z_1 \in B_1, \ldots, Z_k \in B_k) = E\left\{\prod_{j=1}^{k} \alpha(B_j)\right\} \quad \text{for all } k \ge 1 \text{ and } B_1, \ldots, B_k \in \mathcal{B}. \qquad \square$$

3 Weak versions of conditional identity in distribution

Recall that X is c.i.d. if $(X_1, \ldots, X_n, X_k)$ is distributed as $(X_1, \ldots, X_n, X_{n+1})$ for all $k > n \ge 0$. C.i.d. sequences have been introduced in [3] and [20] and then investigated or used in various papers. See e.g. [2, 6, 7, 8, 10, 11, 15, 17] and references therein.

Condition (1.2) holds whenever X is c.i.d. In this case, in fact, $(\alpha_n(f) : n \ge 0)$ is a uniformly bounded martingale for fixed $f \in M_b(S)$. Hence, $\alpha_n(f)$ converges a.s.

In this section, we focus on three weak versions of the c.i.d. condition. In the spirit of this paper, our main goal is to investigate whether they are still sufficient for (1.1) or (1.2). However, such versions are of interest independently of (1.1) and (1.2). Indeed, they provide three generalizations of the c.i.d. condition which could be useful in various contexts.

3.1 Conditional identity in distribution of higher order

Let m be a non-negative integer. Say that X is m-c.i.d. if

$$E\big\{f(X_k) \mid \mathcal{F}_n\big\} = E\big\{f(X_{n+m+1}) \mid \mathcal{F}_n\big\} \quad \text{a.s.}$$

for all $k > n + m$, all $n \ge 0$, and all $f \in M_b(S)$. If $m = 0$, then X is c.i.d. if and only if it is 0-c.i.d. Moreover, if X is m-c.i.d., it is p-c.i.d. for all integers $p > m$.

The m-c.i.d. condition looks reasonable in various problems. As a remarkable special case, X is m-c.i.d. provided it is identically distributed and m-dependent. (Recall that X is m-dependent if $(X_1, \ldots, X_n)$ is independent of $(X_i : i > n + m)$ for every n).

The m-c.i.d. condition has been included in this section for it is a natural weakening of the c.i.d. condition, which looks potentially useful in applications. However, condition (1.1) may fail when X is m-c.i.d. As shown in Example 4.4, it may be that X is stationary, m-dependent but not asymptotically exchangeable. In this case, X is m-c.i.d. (for it is identically distributed and m-dependent) but condition (1.1) fails (for X is not asymptotically exchangeable).

3.2 Quasi-martingale predictive distributions

Quasi-martingales have been defined in Section 2, just before Example 2.6.

It is straightforward to check that X is c.i.d. if and only if the sequence $(\alpha_n(f) : n \geq 0)$ is a martingale, with respect to $(\mathcal{F}_n)$, for fixed $f \in M_b(S)$. Hence, trivially, a weakening of the c.i.d. condition which still implies (1.2) is

$$(\alpha_n(f) : n \geq 0) \text{ is a quasi-martingale, with respect to } (\mathcal{F}_n), \tag{3.1}$$

for each $f \in M_b(S)$. In fact, under (3.1), $\alpha_n(f)$ converges a.s. since $(\alpha_n(f) : n \geq 0)$ is a quasi-martingale such that $E|\alpha_n(f)| \leq \sup|f|$ for all n. Note also that

$$\alpha_n(f) = E\big\{f(X_{n+1}) \mid \mathcal{F}_n\big\} \quad \text{and} \quad E\big\{\alpha_{n+1}(f) \mid \mathcal{F}_n\big\} = E\big\{f(X_{n+2}) \mid \mathcal{F}_n\big\} \quad \text{a.s.}$$

Therefore, condition (3.1) reduces to

$$\sum_n E\Big| E\big\{f(X_{n+2}) - f(X_{n+1}) \mid \mathcal{F}_n\big\}\Big| < \infty \quad \text{for each } f \in M_b(S).$$

In real problems, if X is required to be c.i.d. (or even exchangeable), the predictives α_n are subjected to a quite strong constraint, and this rules out various meaningful choices of α_n. An example are kernel estimators. Some choices of α_n, which make X c.i.d. or exchangeable, do not give rise to realistic kernels; see e.g. [2, 8, 9, 18, 24]. On the contrary, condition (3.1) allows to deal with actual kernel estimators, in the sense of statistics and learning theory. In fact, non-c.i.d. sequences satisfying condition (3.1) arise quite frequently in applications. We close this section with a few examples.

Example 3.1. (Recursive predictive distributions) If $(\Omega, \mathcal{A}) = (S, \mathcal{B})$, a r.p.m. on S is said to be a *kernel* on $(S, \mathcal{B})$. For each $n \geq 0$, suppose

$$\alpha_{n+1}(\cdot) = q_n(X_1, \ldots, X_n)\, \alpha_n(\cdot) \,+\, (1 - q_n(X_1, \ldots, X_n))\, K_n(X_{n+1}, \cdot) \tag{3.2}$$

where $q_n : S^n \to (0, 1)$ is a Borel function and K_n a kernel on $(S, \mathcal{B})$. This type of predictive distributions have an intuitive interpretation and are quite popular in Bayesian predictive inference and species sampling; see e.g. [2, 8, 18, 24]. Since

$$\begin{aligned} E\Big| E\big\{f(X_{n+2}) - f(X_{n+1}) \mid \mathcal{F}_n\big\}\Big| &\leq E\big\{|\alpha_{n+1}(f) - \alpha_n(f)|\big\} \\ &\leq 2\, \sup|f|\, E\Big\{1 - q_n(X_1, \ldots, X_n)\Big\}, \end{aligned}$$

condition (3.1) trivially holds whenever

$$\sum_n \sup_{x \in S^n} \big(1 - q_n(x)\big) < \infty.$$

Moreover, since the K_n are still arbitrary, they can be chosen so that X is not c.i.d.

Example 3.2. (Convex combinations of kernels) For each $n \geq 0$, choose a constant $d_n > 0$ and a kernel K_n. Define $D_n = \sum_{i=0}^{n-1} d_i$ and

$$\alpha_n(\cdot) = (1/D_n) \sum_{i=0}^{n-1} d_i\, K_i(X_{i+1}, \cdot) \qquad \text{where } n \geq 1.$$

Since (3.2) holds with $q_n = D_n/D_{n+1}$, a sufficient condition for (3.1) is $\sum_n \frac{d_n}{D_{n+1}} < \infty$.

Example 3.3. (Generalized Pólya urns) An urn contains $b > 0$ black balls and $r > 0$ red balls. At each time $n \geq 1$, a ball is drawn and then replaced together with a random number of balls of the same color. Say that B_n black balls or R_n red balls are added to the urn according to whether $Y_n = 1$ or $Y_n = 0$, where Y_n is the indicator of the event $\{\text{black ball at time } n\}$. To model this urn scheme, it is quite natural to let

$$P(Y_{n+1} = 1 \mid B_1, R_1, Y_1, \ldots, B_n, R_n, Y_n) = \frac{b + \sum_{i=1}^n B_i Y_i}{b + r + \sum_{i=1}^n \big(B_i Y_i + R_i(1 - Y_i)\big)}.$$

Urns of this type have an history; see [2, 3, 5, 21, 23] and references therein. Now, define $X_n = (B_n, R_n, Y_n)$ and suppose:

- (B_{n+1}, R_{n+1}) is independent of $(X_1, \ldots, X_n, Y_{n+1})$;
- $0 \leq B_n,\ R_n \leq b$ for some $b > 0$ and $E(B_n) = E(R_n)$;
- $\lim_n E(B_n)$, $\lim_n E(B_n^2)$, $\lim_n E(R_n^2)$ exist and $\lim_n E(B_n) > 0$.

Then, as shown in [5, Sect. 4.3], $P(Y_{n+1} = 1 \mid \mathcal{F}_n)$ is a quasi-martingale, so that it converges a.s. Thus, since (B_{n+1}, R_{n+1}) is independent of $(X_1, \ldots, X_n, Y_{n+1})$, condition (1.1) holds whenever $E\big\{f(B_n, R_n)\big\}$ approaches a limit for every bounded Borel f. This trivially happens, for instance, if the sequence (B_n, R_n) is identically distributed. However, X is not necessarily c.i.d.

3.3 Non-adapted c.i.d. sequences

Let $(\mathcal{G}_n)$ be a filtration on $(\Omega, \mathcal{A}, P)$ satisfying $\mathcal{G}_n \subset \mathcal{F}_n$ for each $n \geq 0$. Define the r.p.m.'s

$$\beta_n(\cdot) = P(X_{n+1} \in \cdot \mid \mathcal{G}_n)$$

and note that $\beta_n(f) = E\big\{f(X_{n+1}) \mid \mathcal{G}_n\big\}$ a.s. for all $f \in M_b(S)$.

In this section, we assume that

$$E\big\{f(X_k) \mid \mathcal{G}_n\big\} = E\big\{f(X_{n+1}) \mid \mathcal{G}_n\big\} \quad \text{a.s.} \tag{3.3}$$

for all $k > n \geq 0$ and all $f \in M_b(S)$. Then, X is identically distributed (since $\mathcal{G}_0 = \{\emptyset, \Omega\}$) and X is c.i.d. whenever $\mathcal{G}_n = \mathcal{F}_n$ for all n. More importantly, for fixed $f \in M_b(S)$, the sequence $(\beta_n(f) : n \geq 0)$ is a uniformly bounded martingale with respect to $(\mathcal{G}_n)$. Hence, $\beta_n(f)$ converges a.s. In addition,

$$\beta_n(f) = E\big\{f(X_{n+1}) \mid \mathcal{G}_n\big\} = E\Big\{E\big\{f(X_{n+1}) \mid \mathcal{F}_n\big\} \mid \mathcal{G}_n\Big\} = E\big\{\alpha_n(f) \mid \mathcal{G}_n\big\} \text{ a.s.}$$

Therefore, a (natural) question is whether condition (1.1) holds whenever $\mathcal{G}_n$ is close to $\mathcal{F}_n$ for large n. The next result is a possible answer to this question.

Theorem 3.4. *Suppose condition* (3.3) *holds and define* $\mathcal{F}_\infty = \sigma\big(\cup_n \mathcal{F}_n\big)$ *and* $\mathcal{G}_\infty = \sigma\big(\cup_n \mathcal{G}_n\big)$. *Then, a sufficient condition for* (1.1) *is*

$$\lim_n E\big\{\alpha_n(f)^2\big\} = \lim_n E\big\{\beta_n(f)^2\big\} \qquad \textit{for each } f \in M_b(S). \tag{3.4}$$

If $\mathcal{G}_\infty = \mathcal{F}_\infty$, *condition* (3.4) *is necessary as well.*

Proof. Fix $f \in M_b(S)$. Under (3.4),

$$\begin{aligned} E\Big\{\big(\alpha_n(f) - \beta_n(f)\big)^2\Big\} &= E\big\{\alpha_n(f)^2\big\} + E\big\{\beta_n(f)^2\big\} - 2\,E\Big\{\beta_n(f)\,E\big\{\alpha_n(f) \mid \mathcal{G}_n\big\}\Big\} \\ &= E\big\{\alpha_n(f)^2\big\} - E\big\{\beta_n(f)^2\big\} \longrightarrow 0. \end{aligned}$$

Hence, $\alpha_n(f)$ converges in probability since $\alpha_n(f) - \beta_n(f) \xrightarrow{P} 0$ and $\beta_n(f)$ converges a.s. (because of (3.3)). Next, suppose $\mathcal{G}_\infty = \mathcal{F}_\infty$ and condition (1.1) holds. By Theorem 2.1, $\alpha_n(f) \xrightarrow{P} \alpha(f)$ where the r.p.m. α is $\mathcal{T}$-measurable. By the martingale convergence theorem, $\beta_n(f) = E\big\{\alpha_n(f) \mid \mathcal{G}_n\big\} \xrightarrow{P} E\big\{\alpha(f) \mid \mathcal{G}_\infty\big\}$. Since $\mathcal{T} \subset \mathcal{F}_\infty = \mathcal{G}_\infty$ and α is $\mathcal{T}$-measurable, then $E\big\{\alpha(f) \mid \mathcal{G}_\infty\big\} = \alpha(f)$ a.s. Hence, both $\alpha_n(f)$ and $\beta_n(f)$ converge in probability to $\alpha(f)$, so that $\lim_n E\big\{\alpha_n(f)^2\big\} = E\big\{\alpha(f)^2\big\} = \lim_n E\big\{\beta_n(f)^2\big\}$. □

In general, under (3.3), condition (1.1) could be expected to be true if $\mathcal{G}_n$ is close to $\mathcal{F}_n$ for large n. However, $\mathcal{G}_\infty = \mathcal{F}_\infty$ is not enough. Indeed, as shown by Example 4.5, condition (1.1) may fail even if condition (3.3) holds and $\mathcal{G}_\infty = \mathcal{F}_\infty$.

Finally, apart from (1.1), condition (3.3) helps to get the weak law of large numbers.

Theorem 3.5. *Let $g : [0,\infty) \to [0,\infty)$ be an increasing function such that $g(x) \le x$ for all $x \ge 0$ and $\lim_n n^{-2} \int_1^n g(x)\,dx = 0$. Assume condition* (3.3) *and*

$$\mathcal{F}_{n-[g(n)]} \subset \mathcal{G}_n \qquad \text{for each } n \ge 0$$

where $[\cdot]$ is the integer part. Then, $\mu_n(f)$ converges in probability for each $f \in M_b(S)$.

Proof. Let $f \in M_b(S)$. Since $\beta_n(f)$ converges a.s., it suffices to show that

$$\lim_n \frac{1}{n^2} E\Big\{\Big(\sum_{i=1}^n \big(f(X_i) - \beta_{i-1}(f)\big)\Big)^2\Big\} = 0$$

or equivalently

$$\lim_n \frac{1}{n^2} \sum_{1 \le i < j \le n} E\Big\{\big(f(X_i) - \beta_{i-1}(f)\big)\big(f(X_j) - \beta_{j-1}(f)\big)\Big\} = 0.$$

For fixed j, since $\mathcal{F}_{j-1-[g(j-1)]} \subset \mathcal{G}_{j-1}$, one obtains

$$\begin{aligned} &E\Big\{\big(f(X_i) - \beta_{i-1}(f)\big)\big(f(X_j) - \beta_{j-1}(f)\big)\Big\} \\ &= E\Big\{\big(f(X_i) - \beta_{i-1}(f)\big)\, E\Big(f(X_j) - \beta_{j-1}(f) \mid \mathcal{G}_{j-1}\Big)\Big\} = 0 \end{aligned}$$

whenever $i < j - [g(j-1)]$. Therefore, it suffices noting that

$$\begin{aligned} \frac{1}{n^2} \sum_{j=2}^n \sum_{i=j-[g(j-1)]}^{j-1} E\Big|\big(f(X_i) - \beta_{i-1}(f)\big)\big(f(X_j) - \beta_{j-1}(f)\big)\Big| &\le \frac{4\,\big(\sup|f|\big)^2}{n^2} \sum_{j=2}^n g(j-1) \\ &\le \frac{4\,\big(\sup|f|\big)^2}{n^2} \int_1^n g(x)\,dx \longrightarrow 0. \end{aligned}$$

□

4 Counterexamples

All the counterexamples are collected in this final section.

Example 4.1. $\big($**Condition** (1.1) $\not\Rightarrow$ **condition** (1.2)$\big)$ In the following example, not only (1.1) holds and (1.2) fails, but X is also pairwise independent. Simpler examples are available if pairwise independence is not required.

Let $(\mathcal{X}, \mathcal{U}, Q)$ be a non-atomic probability space. Then, given $d \in (0, 1/2)$, there are three events A, B, $C \in \mathcal{U}$ such that

- A and B are independent and $C = (A \cap B) \cup G$ with $G \subset A^c \cap B^c$;
- $Q(A) = Q(B) = d$ and $Q(G) = d - d^2$.

In fact, since $(\mathcal{X},\mathcal{U},Q)$ is non-atomic, there is $A\in\mathcal{U}$ with $Q(A)=d$. Since $d<1/2$, then $Q(A^c)=1-d>2\,(d-d^2)$. Hence, using non-atomicity again, there are $A_0,\,B_0,\,G\in\mathcal{U}$ with $A_0\subset A$, $B_0\cup G\subset A^c$, $B_0\cap G=\emptyset$, $Q(A_0)=d^2$ and $Q(B_0)=Q(G)=d-d^2$. To achieve the desired conditions, it suffices to let $B=A_0\cup B_0$ and $C=A_0\cup G$. Note also that A, B, C are pairwise independent (even if not independent).

Next, define the countable product space $(\Omega,\mathcal{A},P)=(\mathcal{X}^\infty,\mathcal{U}^\infty,Q^\infty)$ and fix $d_n\in(0,1/2)$ for each $n\geq 1$. Then, $(\Omega,\mathcal{A},P)$ supports the events $A_n,B_n,C_n\in\mathcal{A}$ such that:

- The sequence of triples $\big\{(A_n,B_n,C_n):n\geq 1\big\}$ is independent;
- A_n and B_n are independent and $C_n=(A_n\cap B_n)\cup G_n$ with $G_n\subset A_n^c\cap B_n^c$;
- $P(A_n)=P(B_n)=d_n$ and $P(G_n)=d_n-d_n^2$.

Finally, for each $n\geq 1$, define $X_{3n-2}=1_{A_n}$, $X_{3n-1}=1_{B_n}$ and $X_{3n}=1_{C_n}$. Since each triple (A_n,B_n,C_n) is pairwise independent, X is pairwise independent. Since X_{3n-2} is independent of $\mathcal{F}_{3n-3}$ and X_{3n-1} is independent of $\mathcal{F}_{3n-2}$, one obtains

$$\alpha_{3n-3}(f)=E\big\{f(X_{3n-2})\mid\mathcal{F}_{3n-3}\big\}=E\big\{f(X_{3n-2})\big\}=E\big\{f(1_{A_n})\big\}=f(0)+(f(1)-f(0))\,d_n$$
$$\text{and}\quad \alpha_{3n-2}(f)=E\big\{f(X_{3n-1})\mid\mathcal{F}_{3n-2}\big\}=E\big\{f(1_{B_n})\big\}=f(0)+(f(1)-f(0))\,d_n\qquad\text{a.s.}$$

On noting that $P\big(A_n^c\cap B_n^c\setminus G_n\big)=(1-d_n)(1-2d_n)$, one also obtains

$$\frac{E\big\{f(1_{C_n})1_{A_n^c}1_{B_n^c}\big\}}{P(A_n^c\cap B_n^c)}=\frac{f(1)P(G_n)+f(0)P\big(A_n^c\cap B_n^c\setminus G_n\big)}{P(A_n^c)P(B_n^c)}=\frac{f(1)d_n+f(0)(1-2d_n)}{(1-d_n)}.$$

Therefore,

$$\begin{aligned}\alpha_{3n-1}(f)&=E\big\{f(X_{3n})\mid\mathcal{F}_{3n-1}\big\}=E\big\{f(1_{C_n})\mid 1_{A_n},1_{B_n}\big\}\\&=f(1)\,1_{A_n}1_{B_n}+f(0)\,\big\{1_{A_n^c}1_{B_n}+1_{A_n}1_{B_n^c}\big\}+\frac{E\big\{f(1_{C_n})1_{A_n^c}1_{B_n^c}\big\}}{P(A_n^c\cap B_n^c)}\;1_{A_n^c}1_{B_n^c}\\&=f(0)+(f(1)-f(0))\,\Big\{1_{A_n}1_{B_n}+\frac{d_n}{1-d_n}\,1_{A_n^c}1_{B_n^c}\Big\}\quad\text{a.s.}\end{aligned}$$

Thus, (1.1) holds and (1.2) fails whenever $d_n\to 0$ and $\sum_n d_n^2=\infty$. In fact, if $d_n\to 0$, then $\alpha_{3n-3}(f)\xrightarrow{P}f(0)$, $\alpha_{3n-2}(f)\xrightarrow{P}f(0)$, and

$$\begin{aligned}E\Big|\alpha_{3n-1}(f)-f(0)\Big|&\leq|f(1)-f(0)|\,\Big\{P(A_n\cap B_n)\;+\;\frac{d_n}{1-d_n}\Big\}\\&=|f(1)-f(0)|\,\Big\{d_n^2\;+\;\frac{d_n}{1-d_n}\Big\}\longrightarrow 0.\end{aligned}$$

However, if $f(0)\neq f(1)$ and $\sum_n d_n^2=\infty$, then $\alpha_{3n-1}(f)$ does not converge to $f(0)$ a.s., since the Borel-Cantelli lemma yields $P(A_n\cap B_n$ occurs for infinitely many $n)=1$.

Example 4.2. $\big($**Condition** (2.1) $\not\Rightarrow$ **condition** (1.1)$\big)$ First note that, by Theorem 2.1, condition (2.1) holds and condition (1.1) fails whenever $\alpha_n\Rightarrow\alpha$ in probability, for some r.p.m. α, but $\alpha_n(f)$ does not converge in probability to $\alpha(f)$ for some $f\in M_b(S)$. This situation can be realized in various ways. For instance, suppose (X_n) is independent, each X_n has a discrete distribution ν_n, and $\nu_n\overset{weakly}{\longrightarrow}\nu$ where ν is diffuse (i.e., ν is a probability measure on $\mathcal{B}$ such that $\nu(\{x\})=0$ for each $x\in S$). By independence, $\alpha_n=\nu_{n+1}\Rightarrow\nu$ a.s. Moreover, since the ν_n are discrete, there is a countable set $B\subset S$ such that $\nu_n(B)=1$ for all n. However, $\nu(B)=0$ since B is countable and ν is diffuse. As another (trivial) example, fix a random variable X_1 with values in $\big\{1,2,\ldots\big\}$ and define $X_n=\frac{1}{n}\,X_1$. Then, $\alpha_n=\delta_{X_1/(n+1)}\Rightarrow\delta_0$ a.s. However, letting $B=\big\{\frac{j}{n}:j,\,n\geq 1\big\}$, one obtains $\delta_0(B)=0$ and $\alpha_n(B)=1$ a.s. for all n.

Example 4.3. $\big($**Asymptotic exchangeability** $\not\Rightarrow$ **condition** (2.1)$\big)$ Let (Y_n, Z_n) be an independent sequence of bivariate random variables. Suppose that (Y_n, Z_n) has a density f_n, with respect to Lebesgue measure on $(0,1)^2$, where

$$f_n(y,z) = 1 + \sin(2\,\pi\,n\,y)\,\cos(2\,\pi\,z) \quad \text{for all } (y,z) \in (0,1)^2.$$

Define $X_1 = 0$, $X_{2n} = Y_n$ and $X_{2n+1} = Z_n$ for all $n \geq 1$. Since the pairs (Y_n, Z_n) are independent and Y_n is uniformly distributed on $(0,1)$, letting $f(z) = \cos(2\,\pi\,z)$, one obtains

$$\alpha_{2n-1}(f) = E\big\{f(Y_n) \mid Y_1, Z_1, \ldots, Y_{n-1}, Z_{n-1}\big\} = E\big\{f(Y_n)\big\} = \int_0^1 \cos(2\,\pi\,z)\,dz = 0$$

a.s. However,

$$\begin{aligned}\alpha_{2n}(f) &= E\big\{f(Z_n) \mid Y_1, Z_1, \ldots, Y_{n-1}, Z_{n-1}, Y_n\big\} \\ &= E\big\{f(Z_n) \mid Y_n\big\} = \int_0^1 \cos(2\,\pi\,z)\,f_n(Y_n, z)\,dz = \frac{\sin(2\,\pi\,n\,Y_n)}{2} \quad \text{a.s.}\end{aligned}$$

Hence, since $f \in C_b(S)$ and $\alpha_n(f)$ does not converge in probability, condition (2.1) fails. It remains to show that X is asymptotically exchangeable. First note that

$$\lim_n E\big\{g(Y_n)\,h(Z_n)\big\} = \lim_n \int_0^1 \int_0^1 g(y)\,h(z)\,f_n(y,z)\,dy\,dz = \int_0^1 g(y)\,dy \int_0^1 h(z)\,dz$$

for all g, $h \in C_b(S)$. Based on this fact, for all $k \geq 1$ and $g_1, \ldots, g_k \in C_b(S)$, one obtains

$$\lim_n E\left\{\prod_{i=1}^k g_i(X_{n+i})\right\} = \prod_{i=1}^k \int_0^1 g_i(t)\,dt.$$

This means that $(X_{n+1}, X_{n+2}, \ldots) \to (Z_1, Z_2, \ldots)$ in distribution, as $n \to \infty$, where $(Z_1, Z_2, \ldots)$ is i.i.d. with Z_1 uniformly distributed on $(0,1)$. Therefore, X is asymptotically exchangeable (it is even asymptotically i.i.d.).

Example 4.4. $\big($**Stationary and** m**-dependent** $\not\Rightarrow$ **asymptotically exchangeable**$\big)$ Let $X_n = Y_n - Y_{n+m}$ where $m \geq 1$ and (Y_n) is an i.i.d. sequence of real non-degenerate random variables. Then, X is stationary and m-dependent. Let $\mathcal{T} = \bigcap_n \sigma(X_n, X_{n+1}, \ldots)$. Since $\mathcal{T} \subset \bigcap_n \sigma(Y_n, Y_{n+1}, \ldots)$ and (Y_n) is i.i.d., then $P(A) \in \{0,1\}$ for each $A \in \mathcal{T}$. Toward a contradiction, suppose now that X is asymptotically exchangeable. Since X is stationary, this would imply exchangeability of X. In turn, exchangeability of X and $P(A) \in \{0,1\}$ for each $A \in \mathcal{T}$, would imply that X is i.i.d. But X is not i.i.d. Hence, X is not asymptotically exchangeable.

Example 4.5. $\big($**Condition** (3.3) **and** $\mathcal{G}_\infty = \mathcal{F}_\infty \not\Rightarrow$ **condition** (1.1)$\big)$ Take X as in Example 4.4 and define $\mathcal{G}_n = \big\{\emptyset, \Omega\big\}$ for $n \leq m$ and $\mathcal{G}_n = \mathcal{F}_{n-m}$ for $n > m$. Then, $\mathcal{G}_\infty = \mathcal{F}_\infty$ and condition (1.1) fails (due to X is not asymptotically exchangeable). Moreover, since X is m-dependent, $E\big\{f(X_k) \mid \mathcal{G}_n\big\} = E\big\{f(X_1)\big\}$ a.s. for all $k > n \geq 0$ and $f \in M_b(S)$. Thus, condition (3.3) holds.

To introduce our last example, recall that $E\big\{f(X_{n+2}) \mid \mathcal{F}_n\big\} = E\big\{\alpha_{n+1}(f) \mid \mathcal{F}_n\big\}$ a.s. Hence, a necessary condition for (1.1) is

$$E\big\{f(X_{n+2}) - f(X_{n+1}) \mid \mathcal{F}_n\big\} \xrightarrow{P} 0 \qquad \text{for each } f \in M_b(S). \tag{4.1}$$

However, (4.1) does not suffice for (1.1) even if various other conditions are satisfied. Interestingly, this may happen even in the classical CLT.

Example 4.6. **(Condition (1.1) may fail even if condition (4.1) holds, X_n converges stably, and X is asymptotically exchangeable)** Let $W, Z_1, Z_2, \ldots$ be real random variables on $(\Omega, \mathcal{A}, P)$. Suppose W has a standard normal distribution and the Z_n are i.i.d. with $E(Z_1) = 0$ and $E(Z_1^2) = 1$. Then, it is well known that

$$X_n := n^{-1/2} \sum_{i=1}^{n} Z_i \longrightarrow N(0,1) \qquad \text{stably.}$$

Next, for any $k \geq 1$, the vector $(X_{n+1}, \ldots, X_{n+k})$ can be written as

$$\left(X_{n+1},\, \sqrt{\frac{n+1}{n+2}}\, X_{n+1}, \ldots, \sqrt{\frac{n+1}{n+k}} X_{n+1} \right) + \left(0,\, \frac{Z_{n+2}}{\sqrt{n+2}}, \ldots, \frac{\sum_{i=n+2}^{n+k} Z_i}{\sqrt{n+k}} \right).$$

Hence, $(X_{n+1}, \ldots, X_{n+k}) \to (W, \ldots, W)$ in distribution for every $k \geq 1$, or equivalently $(X_{n+1}, X_{n+2}, \ldots) \longrightarrow (W, W, \ldots)$ in distribution. Thus, X is asymptotically exchangeable. To prove (4.1), we also assume that the distribution of Z_1 is absolutely continuous with respect to Lebesgue measure, so that $\sup_{A \in \mathcal{B}} |P(X_n \in A) - P(W \in A)| \to 0$. We also note that $X_{n+2} - X_{n+1} \xrightarrow{a.s.} 0$, and thus $g(X_{n+2}) - g(X_{n+1}) \xrightarrow{a.s.} 0$ for each Lipschitz function g. Fix $f \in M_b(S)$ and $\epsilon > 0$. Since X_n converges in total variation, there is a bounded Lipschitz function g such that $E|f(X_n) - g(X_n)| < \epsilon$ for large n. Therefore,

$$\begin{aligned} E|f(X_{n+2}) - f(X_{n+1})| \leq E|f(X_{n+2}) - g(X_{n+2})| + E|g(X_{n+2}) - g(X_{n+1})| + \\ + E|g(X_{n+1}) - f(X_{n+1})| < 3\,\epsilon \qquad \text{for large } n. \end{aligned}$$

This proves $\lim_n E|f(X_{n+2}) - f(X_{n+1})| = 0$, which in turn implies condition (4.1). Finally, we prove that condition (1.1) fails. Since X_n does not converge in probability, there is a bounded Lipschitz function g such that $g(X_n)$ does not converge in probability. Let ν denote the probability distribution of Z_1. Since

$$\alpha_n(g) = E\big\{ g(X_{n+1}) \mid \mathcal{F}_n \big\} = \int g\left(\frac{\sqrt{n}}{\sqrt{n+1}}\, X_n + \frac{x}{\sqrt{n+1}} \right) \nu(dx) \quad \text{a.s.,}$$

one obtains

$$\Big| \alpha_n(g) - g(X_n) \Big| \leq \int \Big| g\left(\frac{\sqrt{n}}{\sqrt{n+1}}\, X_n + \frac{x}{\sqrt{n+1}} \right) - g(X_n) \Big| \, \nu(dx) \leq c\, \frac{|X_n| + E|Z_1|}{\sqrt{n+1}} \xrightarrow{P} 0$$

where c is the Lipschitz constant of g. Hence, condition (1.1) fails. In fact, $\alpha_n(g)$ does not converge in probability since $g(X_n)$ does not converge in probability.

References

[1] Aldous, D.J.: Exchangeability and related topics. *Ecole d'eté de Probabilités de Saint- Flour XIII, Lecture Notes in Math.* **1117**, (1985), Springer, Berlin. MR0883646

[2] Bassetti, F., Crimaldi, I. and Leisen, F.: Conditionally identically distributed species sampling sequences. *Adv. Appl. Probab.* **42**, (2010), 433–459. MR2675111

[3] Berti, P., Pratelli, L. and Rigo, P.: Limit theorems for a class of identically distributed random variables. *Ann. Probab.* **32**, (2004), 2029–2052. MR2073184

[4] Berti, P., Pratelli, L. and Rigo, P.: Almost sure weak convergence of random probability measures. *Stochastics* **78**, (2006), 91–97. MR2236634

[5] Berti, P., Crimaldi, I., Pratelli, L. and Rigo, P.: A central limit theorem and its applications to multicolor randomly reinforced urns. *J. Appl. Probab.* **48**, (2011), 527–546. MR2840314

[6] Berti, P., Pratelli, L. and Rigo, P.: Limit theorems for empirical processes based on dependent data. *Electr. J. Probab.* **17**, (2012), 1–18. MR2878788

[7] Berti, P., Pratelli, L. and Rigo, P.: Exchangeable sequences driven by an absolutely continuous random measure. *Ann. Probab* **41**, (2013), 2090–2102. MR3098068

[8] Berti, P., Dreassi, E., Pratelli, L. and Rigo, P.: A class of models for Bayesian predictive inference. *Bernoulli* **27**, (2021), 702–726. MR4177386

[9] Berti, P., Dreassi, E., Leisen, F., Pratelli, L. and Rigo, P.: Kernel based Dirichlet sequences. *Bernoulli* **29**, (2023), 1321–1342. MR4550225

[10] Bissiri, P.G. and Walker, S.G.: Bayesian analysis with conditionally identically distributed sequences. *Electr. J. Statistics* **19**, (2025), 1609–1632. MR4888176

[11] Cassese, A., Zhu, W., Guindani, M. and Vannucci, M.: A Bayesian nonparametric spiked process prior for dynamic model selection. *Bayesian Anal.* **14**, (2019), 553–572. MR3934097

[12] Cui, F. and Walker, S.G.: A bayesian bootstrap for mixture models. *Bayesian Anal.*, (2024), to appear. MR1212168

[13] Dawid, A.P. and Vovk, V.G.: Prequential probability: principles and properties. *Bernoulli* **5**, (1999), 125–162. MR1673572

[14] Etemadi, N. and Lenzhen, A.: Convergence of sequences of pairwise independent random variables. *Proc. Amer. Math. Soc.* **132**, (2004), 1201–1202. MR2045438

[15] Fong, E., Holmes, C. and Walker, S.G.: Martingale posterior distributions (with discussion). *J. Royal Stat. Soc.* B **85**, (2023), 1357–1391. MR4726953

[16] Fong, E. and Yiu, A.: Asymptotics for parametric martingale posteriors. *Biometrika*, (2026), to appear.

[17] Fortini, S., Petrone, S. and Sporysheva, P.: On a notion of partially conditionally identically distributed sequences. *Stoch. Proc. Appl.* **128**, (2018), 819–846. MR3758339

[18] Gnedin, A.: A species sampling model with finitely many types. *Electron. Commun. Probab.* **15**, (2010), 79–88. MR2606505

[19] Hahn, P.R., Martin, R. and Walker, S.G.: On recursive Bayesian predictive distributions. *J.A.S.A.* **113**, (2018), 1085–1093. MR3862341

[20] Kallenberg, O.: Spreading and predictable sampling in exchangeable sequences and processes. *Ann. Probab.* **16**, (1988), 508–534. MR0929061

[21] May, C. and Flournoy, N.: Asymptotics in response-adaptive designs generated by a two-color, randomly reinforced urn. *Ann. Statist.* **37**, (2009), 1058–1078. MR2502661

[22] Moya, B. and Walker, S.G.: Martingale posterior distributions for time series models. *Statist. Sci.* **40**, (2025), 68–80. MR4859118

[23] Pemantle, R.: A survey of random processes with reinforcement. *Prob. Surveys* **4**, (2007), 1–79. MR2282181

[24] Pitman, J.: Some developments of the Blackwell-MacQueen urn scheme. *Statistics, Probability and Game Theory, IMS Lect. Notes Mon. Series* **30**, (1996), 245–267. MR1481784

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4], ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]LOCKSS: Lots of Copies Keep Stuff Safe `http://www.lockss.org/`
[2]EJMS: Electronic Journal Management System: `https://vtex.lt/services/ejms-peer-review/`
[3]IMS: Institute of Mathematical Statistics `http://www.imstat.org/`
[4]BS: Bernoulli Society `http://www.bernoulli-society.org/`
[5]Project Euclid: `https://projecteuclid.org/`
[6]IMS Open Access Fund: `https://imstat.org/shop/donation/`