MODES OF CONVERGENCE IN THE COHERENT FRAMEWORK

PATRIZIA BERTI, EUGENIO REGAZZINI, AND PIETRO RIGO

Abstract. Convergence in distribution is investigated in a finitely additive setting. Let X_n be maps, from any set Ω into a metric space S, and P a finitely additive probability (f.a.p.) on the field $\mathcal{F} = \bigcup_n \sigma(X_1, \ldots, X_n)$. Fix $H \subset \Omega$ and $X : \Omega \to S$. Conditions for $Q(H) = 1$ and $X_n \overset{d}{\to} X$ under Q, for some f.a.p. Q extending P, are provided. In particular, one can let $H = \{\omega \in \Omega : X_n(\omega) \text{ converges}\}$ and $X = \lim_n X_n$ on H. Connections between convergence in probability and in distribution are also exploited. A general criterion for weak convergence of a sequence (μ_n) of f.a.p.'s is given. Such a criterion grants a σ-additive limit provided each μ_n is σ-additive. Some extension results are proved as well. As an example, let X and Y be maps on Ω. Necessary and sufficient conditions for the existence of a f.a.p. on $\sigma(X, Y)$, which makes X and Y independent with assigned distributions, are given. As a consequence, a question posed by de Finetti in 1930 is answered.

1. Introduction and motivations

At pages 11-12 of [6], de Finetti raises the following question:

Let $X = \{X(t) : t \in [0,1]\}$ be a real process with continuous paths and $S_n = \frac{1}{n}\sum_{j=1}^n X(\frac{j}{n})$. Under some assumptions (such as X has independent and stationary increments), $P(S_n \leq t) \to F(t)$ for each continuity point t of F, where F is some distribution function. Also, continuity of the X-paths yields $S_n \to \int_0^1 X(t)dt$ pathwise. *Does it follow that* $\int_0^1 X(t)dt$ *has distribution function* F ?

The answer is clearly yes if probability measures are requested to be σ-additive. But at the end of the twenty's, while investigating processes with independent increments, de Finetti started to have some criticism on σ-additivity as a general axiom. This criticism led him to the notion of coherence, stated in [7] and [9]; see also [8] about his correspondence with Frechet. In any case, de Finetti's question is placed in a finitely additive framework.

Then, the answer is no. Indeed, in a finitely additive setting, $\int_0^1 X(t)dt$ can be given *any* distribution *independently* of the behaviour of X on finite dimensional sets. It may be, for instance, that X has the finite dimensional distributions of standard Brownian motion, $\int_0^1 X(t)dt$ has an *arbitrary* distribution, and $\int_0^1 X(t)dt$ is independent of $(X(t_1), \ldots, X(t_n))$ for each (finite) choice of $t_1, \ldots, t_n$.

Example 1.1. Let Ω be the set of real continuous functions on $[0,1]$, $X(t,\omega) = \omega(t)$ the coordinate process, and $\mathcal{A}$ the field of finite dimensional sets of the form $A = \{\omega \in \Omega : (\omega(t_1), \ldots, \omega(t_n)) \in D\}$, where $n \geq 1$, $t_1, \ldots, t_n \in [0,1]$ and $D \in \mathcal{B}(\mathbb{R}^n)$.

2000 *Mathematics Subject Classification.* 60A05, 60A10, 60B10.

Key words and phrases. Coherence – Convergence in distribution – Extension – Finitely additive probability.

Say that a real functional Y on Ω is *elastic* in case: For each finite collection of distinct points $t_1, \ldots, t_n \in [0,1]$ and each $x, x_1, \ldots, x_n \in \mathbb{R}$, there is $\omega \in \Omega$ satisfying $Y(\omega) = x$ and $\omega(t_i) = x_i$ for $i = 1, \ldots, n$. There are actually a number of elastic functionals on Ω, and one of these is the integral $Y(\omega) = \int_0^1 \omega(t)dt$. Fix now an elastic functional Y, a *finitely additive probability* (f.a.p.) α on $\mathcal{A}$ and a f.a.p. β on $\mathcal{B}(\mathbb{R})$. Here, α and β should be regarded as the finite dimensional distributions of X and the distribution of Y, respectively. Then, in view of the forthcoming Corollary 3.2, there is a f.a.p. P on $\mathcal{P}(\Omega)$ such that

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) = \alpha(A)\beta(B)$$

whenever $A \in \mathcal{A}$ and $B \in \mathcal{B}(\mathbb{R})$.

Among other things, Example 1.1 is one more validation of a known fact, dating to [6]: in a finitely additive setting, almost sure convergence (to some limit) does not imply convergence in distribution (to the same limit). In fact, almost sure convergence basically depends on the choice of some extensions. To see this, suppose (X_n) is a sequence of functions, from any set Ω into a metric space S, and P is a f.a.p. on the field

$$\mathcal{F} = \bigcup_n \sigma(X_1, \ldots, X_n)$$

of finite dimensional sets for (X_n). The set $C = \{\omega \in \Omega : X_n(\omega) \text{ converges}\}$ is typically not in $\mathcal{F}$. For deciding whether X_n converges a.s., thus, some extension Q of P to $\mathcal{F} \cup \{C\}$ is to be selected. In the *coherent framework*, all probability assessments (both conditional and unconditional) are only asked to meet de Finetti's coherence principle (Subsection 2.1). Then, $Q(C)$ can be taken to be any value between the inner measure $P_*(C)$ and the outer measure $P^*(C)$.

Suppose $P^*(C) = 1$. This includes in particular the case where X_n converges pointwise ($C = \Omega$). Denote Q the extension of P satisfying $Q(C) = 1$ and define $X = \lim_n X_n$ on C and $X = x_0$ on C^c, for some $x_0 \in S$. Then, $X_n \to X$ a.s. under Q, but X_n can fail to converge in distribution as well (even if $C = \Omega$). A trivial example is: $\Omega = \mathbb{N}$, $S = \{0,1\}$, $X_n = I_{\{1,\ldots,n\}}$ or $X_n = 1 - I_{\{n\}}$ according to n is even or odd, and P such that $P\{n\} = 0$ for all n. Furthermore, even when X_n converges in distribution, X need not be $\mathcal{F}$-measurable. Hence, to decide whether $X_n \xrightarrow{d} X$, a further (coherent) extension Q_0 of Q to $\mathcal{F} \cup \{C\} \cup \sigma(X)$ is to be chosen, and again it may be that $X_n \xrightarrow{d} X$ fails to be true under Q_0.

To sum up, if P is assessed only on $\mathcal{F}$, as it is reasonable in real problems, convergence issues in the coherent framework deal with existence of extensions. For instance, one question is: Under what conditions, is there an extension Q of P such that $Q(C) = 1$ and $X_n \xrightarrow{d} X$ under Q ?

We point out that the previous remarks do not concern the *strategic* approach, introduced by Dubins and Savage and developed by Purves and Sudderth in [10], [11], [16]. In that case, in fact, P is strategic on $\mathcal{F}$ and Q *must* be its strategic extension to $\sigma(\mathcal{F})$.

This paper contains two types of results, both concerning coherence. Related references are [4] and [19].

The first type (Section 3) are pure coherence results: existence of (coherent) extensions satisfying certain requirements. As an example, given two maps X and

Y on Ω, suppose we would like to model X and Y as independent, with assigned distributions. Is this possible ? That is, is there a f.a.p. P on $\sigma(X,Y)$ which makes X and Y independent with the assigned distributions ? Our main result (Theorem 3.1) provides necessary and sufficient conditions for such a P to exist.

The second type of results (Section 4) deals with convergence in distribution. Let $X:\Omega\rightarrow S$ be a map and $H\subset\Omega$ a subset. For instance, one could let $H=C$ and $X=\lim_n X_n$ on H. Conditions for $Q(H)=1$ and $X_n\overset{d}{\rightarrow}X$ under Q, for some extension Q of P, are given. Next, the connections between convergence in probability and convergence in distribution are investigated. It turns out that the former implies the latter provided S is Polish and each X_n has a tight distribution, but not in general. Next, let (μ_n) be a sequence of f.a.p.'s on $\mathcal{B}(S)$. A necessary and sufficient condition for $\mu_n\rightarrow\mu$ weakly, for some f.a.p. μ on $\mathcal{B}(S)$, is provided (Theorem 4.2). This condition is potentially useful in the standard setting as well. In fact, μ can be taken to be σ-additive whenever each μ_n is σ-additive. Finally, various examples are given.

2. Basic definitions

In the sequel, Ω is a set, $\mathcal{F}$ a field of subsets of Ω and P a f.a.p. on $\mathcal{F}$. We let $\mathcal{P}(\Omega)$ denote the power set of Ω and P^* and P_* the outer and inner measures induced by P, i.e.

$$P^*(H)=\inf\{P(A):H\subset A\in\mathcal{F}\},\quad P_*(H)=1-P^*(H^c),\quad H\subset\Omega.$$

2.1. **Coherence.** Let $\mathcal{D}$ be any class of bounded functions defined on (possibly different) subsets of Ω. The members of $\mathcal{D}$ are thus of the form $X|H$, where $X:\Omega\rightarrow\mathbb{R}$ is bounded and $H\subset\Omega$, $H\neq\emptyset$. In case $H=\Omega$, we write X instead of $X|\Omega$. According to de Finetti's coherence principle, a real function E on $\mathcal{D}$ is *coherent* if, for each $n\geq 1$, $X_1|H_1,\ldots,X_n|H_n\in\mathcal{D}$ and $c_1,\ldots,c_n\in\mathbb{R}$, one has

$$\inf G|H\leq 0\leq\sup G|H$$

where G and H are defined as

$$G=\sum_{i=1}^n c_iI_{H_i}\big(X_i-E(X_i\mid H_i)\big)\quad\text{and}\quad H=\bigcup_{i=1}^n H_i.$$

Heuristically, suppose E describes the opinions of a bookie on the elements of $\mathcal{D}$. If E is not coherent, a gambler can select a finite combination of bets (on $X_1|H_1,\ldots,X_n|H_n$ with stakes $c_1,\ldots,c_n$) which makes the bookie a sure loser. Instead, no Dutch book against a coherent bookie is possible.

A conditional expectation is a coherent function E and a conditional probability is the restriction of a conditional expectation to indicators, i.e.,

$$Prob(A\mid H)=E(I_A\mid H)\quad A\subset\Omega, I_A|H\in\mathcal{D}.$$

We refer to [17] for more on coherence. Here, we just note that a coherent function E can be extended, preserving coherence, to the class of *all* functions $X|H$ with $X:\Omega\rightarrow\mathbb{R}$ bounded and H nonempty subset of Ω. Moreover, if $\mathcal{G}$ is a field of subsets of Ω and $\mathcal{D}=\{I_A:A\in\mathcal{G}\}$, then E is coherent if and only if $A\mapsto E(I_A)$ is a f.a.p. on $\mathcal{G}$.

2.2. **Convergence in distribution.** Let S be a metric space. We write $\mathcal{B}(S)$ for the Borel σ-field of S and $C_b(S)$ for the set of real bounded continuous functions on S. Also, we let $\nu(f) = \int f d\nu$ whenever ν is a f.a.p. on $\mathcal{B}(S)$ and f a real bounded Borel function on S. Given f.a.p.'s μ_n and μ on $\mathcal{B}(S)$, say that μ_n *converges to* μ *weakly* in case $\mu_n(f) \to \mu(f)$ for all $f \in C_b(S)$.

A sufficient condition for $\mu_n \to \mu$ weakly is

$$\limsup_n \mu_n(F) \leq \mu(F) \quad \text{for all closed } F \subset S.$$

Such a condition is necessary as well provided the limit μ is *regular on open sets*, i.e.

$$\mu(U) = \sup\{\mu(F) : F \text{ closed}, F \subset U\} \quad \text{for all open } U \subset S.$$

This can be proved by the same argument as in the standard setting; see [13] and page 316 of [12]. We also recall that, if ϕ is a linear positive functional on $C_b(S)$ satisfying $\phi(1) = 1$, then $\phi(f) = \mu(f)$, $f \in C_b(S)$, for some f.a.p. μ on $\mathcal{B}(S)$ regular on open sets. Therefore, if $\mu_n \to \nu$ weakly for some ν, one also obtains $\mu_n \to \mu$ weakly for some μ regular on open sets.

We refer to [13] for more on weak convergence of f.a.p.'s.

Convergence in distribution of *measurable* maps is defined via weak convergence of their laws. For any function $Z : \Omega \to S$, denote

$$\sigma(Z) = \{\{Z \in B\} : B \in \mathcal{B}(S)\}.$$

Let $X_n, X : \Omega \to S$ be maps and μ a f.a.p. on $\mathcal{B}(S)$. Say that X_n *converges to* μ *in distribution* in case $\sigma(X_n) \subset \mathcal{F}$ for all n and $P \circ X_n^{-1} \to \mu$ weakly. Also, say that X_n *converges to* X *in distribution*, written $X_n \xrightarrow{d} X$, in case $\sigma(X_n) \subset \mathcal{F}$ for all n, $\sigma(X) \subset \mathcal{F}$ and $P \circ X_n^{-1} \to P \circ X^{-1}$ weakly.

Finally, we mention a result of Karandikar [14]. For our purposes, it is useful to extend it to S-valued maps (the original result is stated for real valued functions). As usual, a probability space is a triplet $(\Omega_0, \mathcal{F}_0, P_0)$ where P_0 is a σ-additive f.a.p. on the σ-field $\mathcal{F}_0$ of subsets of Ω_0.

Theorem 2.1. *Suppose P is a f.a.p. on*

$$\mathcal{F} = \bigcup_n \sigma(X_1, \ldots, X_n)$$

where each $X_n : \Omega \to S$ is tight (i.e., for all $\epsilon > 0$ there is a compact $K \subset S$ such that $P(X_n \notin K) < \epsilon$). Then, there are a probability space $(\Omega_0, \mathcal{F}_0, P_0)$ and a sequence (Z_n) of S-valued random variables on $(\Omega_0, \mathcal{F}_0, P_0)$ satisfying

$$E_P f(X_1, \ldots, X_n) = E_{P_0} f(Z_1, \ldots, Z_n)$$

for all n and $f \in C_b(S^n)$, and

$$P\big((X_1, \ldots, X_n) \in B\big) = P_0\big((Z_1, \ldots, Z_n) \in B\big)$$

for all n and $B \in \mathcal{B}(S^n)$ such that $P_0\big((Z_1, \ldots, Z_n) \in \partial B\big) = 0$.

Proof. We just give a sketch of the proof, since it coincides essentially with Karandikar's original one. First note that $E_P f(X_1, \ldots, X_n) = \int f(X_1, \ldots, X_n) dP$ is well defined for all n and $f \in C_b(S^n)$, since f is bounded and $\sigma(X_1, \ldots, X_n) \subset \mathcal{F}$. Fix n and define $\phi_n(f) = E_P f(X_1, \ldots, X_n)$ for all $f \in C_b(S^n)$. Then, ϕ_n is a linear positive functional on $C_b(S^n)$. Since X_i is tight for each $i \leq n$, given $\epsilon > 0$ one has

$P\big((X_1,\ldots,X_n)\notin K\big)<\epsilon$ for some compact $K\subset S^n$. Using this fact, it is straightforward to verify that $\lim_j \phi_n(f_j)=0$ whenever $(f_j)\subset C_b(S^n)$ is a non-increasing sequence converging to 0 pointwise. By Daniell theorem, there is a σ-additive f.a.p. λ_n on $\mathcal{B}(S^n)$ such that $\phi_n(f)=\lambda_n(f)$ for all $f\in C_b(S^n)$. Let $\Omega_0=S^\infty$ and $\mathcal{F}_0=\sigma(Z_1,Z_2,\ldots)$ where Z_n is the n-th coordinate projection on S^∞. Since $\lambda_{n+1}(B\times S)=\lambda_n(B)$ for all n and $B\in\mathcal{B}(S^n)$, one can define a f.a.p. $\mathbb{P}$ on the field $\mathbb{F}=\bigcup_n \sigma(Z_1,\ldots,Z_n)$ as

$$\mathbb{P}\big((Z_1,\ldots,Z_n)\in B\big)=\lambda_n(B).$$

It remains to prove that $\mathbb{P}$ is σ-additive on $\mathbb{F}$. Each Z_n has a σ-additive distribution (under $\mathbb{P}$) due to λ_n is σ-additive. Thus, by Theorem 6 of [20], for $\mathbb{P}$ to be σ-additive it is enough that Z_n is tight for all n. Fix n and $\epsilon>0$, and take a compact $K\subset S$ such that $P(X_n\notin K)<\epsilon$. Since Z_n has a σ-additive distribution, some closed set $F\subset K^c$ meets $\mathbb{P}(Z_n\notin K)<\epsilon+\mathbb{P}(Z_n\in F)$. Let $h\in C_b(S)$ be such that $I_F\le h\le I_{K^c}$. Then,

$$\begin{aligned}\mathbb{P}(Z_n\notin K)&<\epsilon+\mathbb{P}(Z_n\in F)\le\epsilon+E_{\mathbb{P}}h(Z_n)\\&=\epsilon+E_P h(X_n)\le\epsilon+P(X_n\notin K)<2\epsilon.\end{aligned}$$

□

3. Coherence results

Theorem 3.1. *Let $\mathcal{F}$ be the field generated by $\mathcal{F}_1\cup\mathcal{F}_2$, where $\mathcal{F}_1$ and $\mathcal{F}_2$ are fields of subsets of Ω, and let P_1 and P_2 be f.a.p.'s on $\mathcal{F}_1$ and $\mathcal{F}_2$. In order to*

$$P(A\cap B)=P_1(A)P_2(B)\quad\text{for all } A\in\mathcal{F}_1 \text{ and } B\in\mathcal{F}_2, \tag{1}$$

for some f.a.p. P on $\mathcal{F}$, it is necessary and sufficient that

$$P_1(A)P_2(B)=P_1(A')P_2(B') \tag{2}$$

whenever $A\cap B=A'\cap B'$ with $A,A'\in\mathcal{F}_1$ and $B,B'\in\mathcal{F}_2$.

Proof. Under (1), if $A\cap B=A'\cap B'$ for some $A,A'\in\mathcal{F}_1$ and $B,B'\in\mathcal{F}_2$, then

$$P_1(A)P_2(B)=P(A\cap B)=P(A'\cap B')=P_1(A')P_2(B').$$

Conversely, suppose (2) holds and define $\mathbb{P}(A\cap B)=P_1(A)P_2(B)$ for $A\in\mathcal{F}_1$ and $B\in\mathcal{F}_2$. By (2), $\mathbb{P}$ is well defined on $\mathcal{S}:=\{A\cap B: A\in\mathcal{F}_1, B\in\mathcal{F}_2\}$. Since $\mathcal{F}$ is the collection of finite disjoint unions of elements of $\mathcal{S}$, it is enough proving that $\mathbb{P}$ is finitely additive on $\mathcal{S}$. In fact, fix $H,F\in\mathcal{F}$ with $H=F$, and suppose $H=\cup_i H_i$ and $F=\cup_j F_j$, where the H_i and the F_j are disjoint members of $\mathcal{S}$. If $\mathbb{P}$ is finitely additive on $\mathcal{S}$, one obtains

$$\sum_i \mathbb{P}(H_i)=\sum_i \mathbb{P}\big(\cup_j(H_i\cap F_j)\big)=\sum_i\sum_j \mathbb{P}(H_i\cap F_j)=\sum_j \mathbb{P}(F_j).$$

Thus, it is possible to define $P(\cup_i H_i)=\sum_i \mathbb{P}(H_i)$, where $\cup_i H_i$ is a finite disjoint union of elements of $\mathcal{S}$, and such a P is a f.a.p. on $\mathcal{F}$ which meets condition (1).

It remains to show that $\mathbb{P}$ is finitely additive on $\mathcal{S}$. Let $A\cap B=\bigcup_{i=1}^m A_i\cap B_i$, where $A,A_1,\ldots,A_m\in\mathcal{F}_1$, $B,B_1,\ldots,B_m\in\mathcal{F}_2$ and $A_i\cap B_i\cap A_j\cap B_j=\emptyset$ for $i\ne j$. We have to prove that

$$P_1(A)P_2(B)=\sum_{i=1}^m P_1(A_i)P_2(B_i). \tag{3}$$

For $m = 1$, condition (3) reduces to (2). By induction, suppose that (3) holds for $m = n - 1$ where $n \geq 2$. Let $A \cap B = \bigcup_{i=1}^n A_i \cap B_i$, where $A, A_1, \ldots, A_n \in \mathcal{F}_1$, $B, B_1, \ldots, B_n \in \mathcal{F}_2$ and $A_i \cap B_i \cap A_j \cap B_j = \emptyset$ for $i \neq j$. On noting that

$$(A - A_n) \cap B = \big(\bigcup_{i=1}^n A_i \cap B_i\big) \cap A_n^c = \bigcup_{i=1}^{n-1} (A_i - A_n) \cap B_i,$$

the inductive assumption implies

$$P_1(A - A_n)P_2(B) = \sum_{i=1}^{n-1} P_1(A_i - A_n)P_2(B_i).$$

Accordingly,

$$\begin{aligned} P_1(A)P_2(B) &= P_1(A - A_n)P_2(B) + P_1(A \cap A_n)P_2(B) \\ &= \sum_{i=1}^{n-1} P_1(A_i - A_n)P_2(B_i) + P_1(A \cap A_n)P_2(B) \\ &= \sum_{i=1}^{n-1} P_1(A_i)P_2(B_i) + P_1(A \cap A_n)P_2(B) - \sum_{i=1}^{n-1} P_1(A_i \cap A_n)P_2(B_i). \end{aligned}$$

Observe now that, if $U \cap V = \emptyset$ with $U \in \mathcal{F}_1$ and $V \in \mathcal{F}_2$, then (2) yields $P_1(U)P_2(V) = 0$. Since $(A_n - A) \cap B_n = \emptyset$ and $A_n \cap (B_n - B) = \emptyset$, it follows that $P_1(A_n - A)P_2(B_n) = 0 = P_1(A_n)P_2(B_n - B)$. Hence,

$$\begin{aligned} P_1(A \cap A_n)P_2(B) &= P_1(A \cap A_n)P_2(B \cap B_n) + P_1(A \cap A_n)P_2(B - B_n) \\ &= P_1(A_n)P_2(B_n) + P_1(A \cap A_n)P_2(B - B_n) \end{aligned}$$

which implies

$$P_1(A)P_2(B) = \sum_{i=1}^{n} P_1(A_i)P_2(B_i) + P_1(A \cap A_n)P_2(B - B_n) - \sum_{i=1}^{n-1} P_1(A_i \cap A_n)P_2(B_i).$$

Finally, since $A_i \cap B_i \cap A_n \cap B_n = \emptyset$ for each $i < n$, one obtains

$$\bigcup_{i=1}^{n-1} (A_i \cap A_n) \cap B_i = \bigcup_{i=1}^{n-1} (A_i \cap A_n) \cap (B_i - B_n) = (A \cap A_n) \cap (B - B_n).$$

Therefore, the inductive assumption again implies

$$P_1(A \cap A_n)P_2(B - B_n) = \sum_{i=1}^{n-1} P_1(A_i \cap A_n)P_2(B_i),$$

and this concludes the proof. □

Theorem 3.1 is our main result. One of its consequences is next Corollary 3.2, already used in Example 1.1. See also [19] and Theorem (2.1.4) of [18].

Corollary 3.2. *Let* Ω *be the set of real continuous functions on* $[0, 1]$,

$$X(t, \omega) = \omega(t), \quad (t, \omega) \in [0, 1] \times \Omega$$

the coordinate process, and $\mathcal{A}$ *the union of* $\sigma(X(t_1), \ldots, X(t_n))$ *for all* $n \geq 1$ *and* $t_1, \ldots, t_n \in [0, 1]$. *Fix an elastic functional* Y *on* Ω, *a f.a.p.* α *on* $\mathcal{A}$ *and a f.a.p.*

β on $\mathcal{B}(\mathbb{R})$. (The definition of elastic functional has been given in Example 1.1). Then, there is a f.a.p. P on $\mathcal{P}(\Omega)$ satisfying

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) = \alpha(A)\beta(B)$$

whenever $A \in \mathcal{A}$ and $B \in \mathcal{B}(\mathbb{R})$.

Proof. Let $\mathcal{F}_1 = \mathcal{A}$, $P_1 = \alpha$ and $\mathcal{F}_2 = \sigma(Y)$. Fix $I, J \in \mathcal{B}(\mathbb{R})$. Since $Y(\Omega) = \mathbb{R}$ (by elasticity), $\{Y \in I\} = \{Y \in J\}$ implies $I = J$ and $\{Y \in I\} \cap \{Y \in J\} = \emptyset$ implies $I \cap J = \emptyset$. Thus, a f.a.p. P_2 on $\mathcal{F}_2$ can be defined as $P_2(Y \in I) = \beta(I)$, $I \in \mathcal{B}(\mathbb{R})$. By Theorem 3.1, it suffices to prove condition (2). In obvious notation, let $A \cap B = A' \cap B'$ where $A = \{\omega : (\omega(t_1), \ldots, \omega(t_n)) \in D\}$, $A' = \{\omega : (\omega(s_1), \ldots, \omega(s_m)) \in D'\}$, $B = \{\omega : Y(\omega) \in I\}$ and $B' = \{\omega : Y(\omega) \in I'\}$. It can be assumed $m = n$, $s_i = t_i$ for each i, and $t_1, \ldots, t_n$ all distinct. If one of D, D', I, I' is empty, say $D = \emptyset$, then $A' \cap B' = A \cap B = \emptyset$, and elasticity of Y yields $D' = \emptyset$ or $I' = \emptyset$. Hence, $P_1(A)P_2(B) = 0 = P_1(A')P_2(B')$. Suppose now that $D, D', I, I' \neq \emptyset$. If $D \neq D'$, say $D - D' \neq \emptyset$, elasticity of Y implies $(\omega(t_1), \ldots, \omega(t_n)) \in D - D'$ and $Y(\omega) \in I$ for some $\omega \in \Omega$, which is a contradiction. Assuming $I \neq I'$ yields a similar contradiction. Therefore, $D = D'$ and $I = I'$, that is, $A = A'$ and $B = B'$. □

We close this section with a result needed in Section 4. It improves Proposition 8 of [3].

Theorem 3.3. *Let P be a f.a.p. on a field $\mathcal{F}$ on Ω and μ a f.a.p. on a field $\mathcal{G}$ on T, where T is any set. Fix $H \subset \Omega$ and a map $X : \Omega \to T$. Then, there is a f.a.p. Q on $\mathcal{P}(\Omega)$ such that*

$$Q = P \text{ on } \mathcal{F}, \quad Q(H) = 1, \quad Q(X \in B) = \mu(B) \text{ for all } B \in \mathcal{G}$$

if and only if

$$\mu(B) \leq P^*\big(\{X \in B\} \cap H\big) \text{ for all } B \in \mathcal{G}. \tag{4}$$

Proof. Suppose (4) holds. Given $\epsilon > 0$, one has $\mu^*(X(H)) + \epsilon > \mu(B)$ for some $B \in \mathcal{G}$ with $B \supset X(H)$. Since $\{X \in B\} \supset H$, condition (4) yields

$$\begin{aligned}\mu^*(X(H)) + \epsilon > \mu(B) &= 1 - \mu(B^c) \geq 1 - P^*\big(\{X \in B^c\} \cap H\big) \\ &= P_*\big(\{X \in B\} \cup H^c\big) \geq P_*(H \cup H^c) = P(\Omega) = 1.\end{aligned}$$

Hence $\mu^*(X(H)) = 1$, so that μ can be extended to a f.a.p. λ on $\mathcal{P}(T)$ such that $\lambda(X(H)) = 1$. If $\{X \in B_1\} \cap H = \{X \in B_2\} \cap H$ for some $B_1, B_2 \in \mathcal{G}$, then $B_1 \cap X(H) = B_2 \cap X(H)$ and thus $\mu(B_1) = \lambda(B_1 \cap X(H)) = \lambda(B_2 \cap X(H)) = \mu(B_2)$. Therefore, one can define $P_0(\{X \in B\} \cap H) = \mu(B)$ for all $B \in \mathcal{G}$, and P_0 is a f.a.p. on the field $\mathcal{G}_0 = \{\{X \in B\} \cap H : B \in \mathcal{G}\}$ of subsets of H. Define further $\mathcal{G}_1 = \{C \subset \Omega : C \cap H \in \mathcal{G}_0\}$ and $P_1(C) = P_0(C \cap H)$ for $C \in \mathcal{G}_1$. Fix $C \in \mathcal{G}_1$ and $A \in \mathcal{F}$ with $C \subset A$. Since $P_1(H) = 1$ and $C \cap H = \{X \in B\} \cap H$ for some $B \in \mathcal{G}$, condition (4) implies

$$P_1(C) = P_1(C \cap H) = \mu(B) \leq P^*\big(\{X \in B\} \cap H\big) \leq P(A).$$

By Theorem 3.6.1 of [5], there exists a f.a.p. Q on $\mathcal{P}(\Omega)$ extending both P and P_1. This concludes the proof of the "if" part while the "only if" part is trivial. □

4. Convergence results

In this section, S is a metric space, X_n and X are S-valued maps on Ω, $n = 1, 2, \ldots$, and P is a f.a.p. on the field $\mathcal{F} = \bigcup_n \sigma(X_1, \ldots, X_n)$ of finite dimensional sets for the sequence (X_n).

Motivations for the next result have been given in Section 1.

Theorem 4.1. *Suppose X_n converges in distribution (to some f.a.p. on $\mathcal{B}(S)$) and $H \subset \Omega$ is any subset. In order that $Q(H) = 1$ and $X_n \overset{d}{\to} X$ under Q, for some f.a.p. Q on $\mathcal{P}(\Omega)$ extending P, it is enough that*

$$\lim_n P(X_n \in B) \leq P^*\big(\{X \in B\} \cap H\big)$$
$$\textit{whenever } B \in \mathcal{U} \textit{ and the limit exists}$$

where $\mathcal{U}$ denotes the field generated by the open subsets of S.

Proof. Since X_n converges in distribution, there is a f.a.p. μ on $\mathcal{B}(S)$, regular on open sets, such that X_n converges to μ in distribution (cf. Subsection 2.2). Let $\mathcal{G} = \{B \in \mathcal{U} : \mu(\partial B) = 0\}$. Then,

$$\mu(B) = \lim_n P(X_n \in B) \leq P^*\big(\{X \in B\} \cap H\big) \quad \text{for all } B \in \mathcal{G}$$

where the equality depends on μ being regular on open sets. By Theorem 3.3 (applied to $T = S$ and $\mu|\mathcal{G}$), some extension Q of P meets $Q(H) = 1$ and $Q(X \in B) = \mu(B)$ for all $B \in \mathcal{G}$. It follows that

$$E_Q f(X) = \mu(f) = \lim_n E_P f(X_n) \quad \text{for all } f \in C_b(S),$$

where the first equality depends on $Q \circ X^{-1} = \mu$ on $\mathcal{G}$ while the second is due to X_n converges to μ in distribution. □

As an example, suppose $H = \Omega$ and $X_n \to X$ pointwise. Then, roughly speaking, X is "logically independent" of $(X_1, \ldots, X_n)$ for every *fixed* n. Thus, since $\mathcal{F}$ contains finite dimensional sets for (X_n) only, it is not unusual that $P^*(X \in B) = 1$ for all $B \in \mathcal{U}$ with $B \neq \emptyset$. In a sense, this happens in Example 1.1 of Section 1. In all these situations, Theorem 4.1 applies as far as X_n converges in distribution. Analogous considerations can be repeated when $H = \{\omega : X_n(\omega) \text{ converges}\}$ and $X = \lim_n X_n$ on H, provided $P^*(H) = 1$.

In Theorem 4.1, X_n is assumed to converge in distribution. At least in principle, this can be checked via the following general criterion, whose first part is inspired by Proposition 7 of [3].

Theorem 4.2. *Let (μ_n) be a sequence of f.a.p.'s on $\mathcal{B}(S)$. Then $\mu_n \to \mu$ weakly, for some f.a.p. μ on $\mathcal{B}(S)$, if and only if, for each finite collection $F_1, \ldots, F_k \subset S$ of closed sets, there is a f.a.p. γ on $\mathcal{B}(S)$ satisfying*

$$\limsup_n \mu_n(F_i) \leq \gamma(F_i) \quad \textit{for } i = 1, \ldots, k. \tag{5}$$

Moreover, under (5), μ can be taken to be σ-additive provided infinitely many μ_n are σ-additive.

Proof. If $\mu_n \to \nu$ weakly for some ν, one also has $\mu_n \to \mu$ weakly for some μ regular on open sets, and thus (5) holds with $\gamma = \mu$ (cf. Subsection 2.2). Conversely, suppose (5) holds. Let $\mathcal{C}$ be the class of closed subsets of S and

$$\mathcal{R}_F = \{\gamma : \gamma \text{ is a f.a.p. on } \mathcal{B}(S) \text{ and } \limsup_n \mu_n(F) \leq \gamma(F)\}, \quad F \in \mathcal{C}.$$

It is enough proving that $\bigcap_{F\in\mathcal{C}}\mathcal{R}_F \neq \emptyset$. Let M be the space of $[0,1]$-valued functions on $\mathcal{B}(S)$, equipped with product topology. Then, M is compact, $\mathcal{R}_F$ is closed in M, and, by assumption, $\{\mathcal{R}_F : F \in \mathcal{C}\}$ has the finite intersection property. Hence, $\bigcap_{F\in\mathcal{C}}\mathcal{R}_F \neq \emptyset$. Finally, suppose (5) holds and there is a subsequence (n_j) such that μ_{n_j} is σ-additive for each j. By (5), $\mu_n \to \nu$ weakly for some ν, so that $\lim_j \mu_{n_j}(f)$ exists for each $f \in C_b(S)$. Since each μ_{n_j} is σ-additive, Alexandrov's theorem implies $\mu_{n_j} \to \mu$ weakly for some σ-additive f.a.p. μ on $\mathcal{B}(S)$; see [1] and [15]. Therefore, $\lim_n \mu_n(f) = \lim_j \mu_{n_j}(f) = \mu(f)$ for all $f \in C_b(S)$. □

The classical portmanteau theorem states that, if μ is regular on open sets, then $\mu_n \to \mu$ weakly if and only if $\limsup_n \mu_n(F) \leq \mu(F)$ for all closed $F \subset S$. Instead, Theorem 4.2 does not require μ to be specified in advance. Furthermore, it is enough considering *finite* collections of closed sets rather than *all* closed sets (in fact, γ may depend on $F_1, \ldots, F_k$).

Incidentally, Theorem 4.2 can be useful in the standard setting as well, since it provides a criterion for deciding whether a sequence of σ-additive f.a.p.'s has a σ-additive weak limit. We do not know whether such a criterion is already known.

We next investigate the connections between convergence in probability and convergence in distribution in the coherent framework. Say that X_n *converges to* X *in probability*, written $X_n \overset{P}{\to} X$, in case $\lim_n P^*\big(d(X_n, X) > \epsilon\big) = 0$ for all $\epsilon > 0$, where d denotes the distance on S.

Theorem 4.3. *Suppose S is a Polish space and each X_n is tight. If $d(X_n, X_m) \overset{P}{\to} 0$ as $n, m \to \infty$, then X_n converges to μ in distribution for some σ-additive f.a.p. μ on $\mathcal{B}(S)$. Moreover, if $X_n \overset{P}{\to} X$, then $X_n \overset{d}{\to} X$ under Q, for any f.a.p. Q on $\mathcal{P}(\Omega)$ extending P.*

Proof. On some probability space $(\Omega_0, \mathcal{F}_0, P_0)$, there is a sequence (Z_n) of S-valued random variables satisfying the conclusions of Theorem 2.1. Let

$$I = \{\epsilon > 0 : P_0\big(d(Z_n, Z_m) = \epsilon\big) > 0 \text{ for some } n, m\}.$$

By Theorem 2.1, $P_0\big(d(Z_n, Z_m) > \epsilon\big) = P\big(d(X_n, X_m) > \epsilon\big)$ for all n, m and all $\epsilon > 0$ such that $\epsilon \notin I$. Since I is countable and $d(X_n, X_m) \overset{P}{\to} 0$, it follows that $d(Z_n, Z_m) \overset{P_0}{\to} 0$. Since S is Polish, $Z_n \overset{P_0}{\to} Z$ for some S-valued random variable Z on $(\Omega_0, \mathcal{F}_0, P_0)$. Define $\mu(B) = P_0(Z \in B)$ for all $B \in \mathcal{B}(S)$. Then, μ is σ-additive and Theorem 2.1 yields

$$E_P f(X_n) = E_{P_0} f(Z_n) \to \mu(f) \quad \text{for all } f \in C_b(S).$$

Finally, suppose $X_n \overset{P}{\to} X$ and fix a closed set $F \subset S$. Since μ is regular on open sets (it is in fact σ-additive), given $\epsilon > 0$ one obtains

$$\begin{aligned} P^*(X \in F) &\leq \limsup_n P^*\big(X \in F, d(X_n, X) \leq \epsilon\big) \\ &\leq \limsup_n P\big(d(X_n, F) \leq \epsilon\big) \leq \mu\{x \in S : d(x, F) \leq \epsilon\}. \end{aligned}$$

By σ-additivity of μ, one also obtains $P^*(X \in F) \leq \mu(F)$. Thus, for each $B \in \mathcal{B}(S)$, taking complements yields

$$\mu(B^0) \leq P_*(X \in B) \leq P^*(X \in B) \leq \mu(\overline{B}) \tag{6}$$

where B^0 and $\overline{B}$ stand for the interior and the closure of B. Let Q be a f.a.p. on $\mathcal{P}(\Omega)$ extending P. By (6), $Q(X \in B) = \mu(B)$ whenever $B \in \mathcal{B}(S)$ and $\mu(\partial B) = 0$, and this routinely implies $E_Q f(X) = \mu(f)$ for all $f \in C_b(S)$. □

Example 4.4. (Empirical measures) Let (ξ_n) be a sequence of real valued functions on Ω, $\widetilde{P}$ a f.a.p. on $\widetilde{\mathcal{F}} = \bigcup_n \sigma(\xi_1, \ldots, \xi_n)$ and

$$X_n(B) = \frac{1}{n} \sum_{i=1}^n I_B(\xi_i), \quad B \in \mathcal{B}(\mathbb{R})$$

the n-th empirical measure. Define S to be the set of σ-additive f.a.p.'s on $\mathcal{B}(\mathbb{R})$, equipped with the topology of weak convergence, and suppose

$$\lim_{a\to\infty} \widetilde{P}(|\xi_n| > a) = 0 \quad \text{for all } n, \tag{7}$$

$$\widetilde{P}(\xi_i = \xi_j) = 0 \quad \text{for all } i \neq j, \tag{8}$$

$$\widetilde{P}\big(\xi_{j_1} < \xi_{j_2} < \ldots < \xi_{j_n}\big) = \widetilde{P}\big(\xi_1 < \xi_2 < \ldots < \xi_n\big) \tag{9}$$

for all n and all permutations $(j_1, \ldots, j_n)$ of $(1, \ldots, n)$.

Then, S is Polish and, by (7), every X_n is tight. By Theorem 4.4 of [4] (see also Theorem 4.2 of [4]), conditions (8)-(9) yield

$$\sup_t \Big| X_n(-\infty, t] - X_m(-\infty, t] \Big| \overset{P}{\to} 0$$

where $P = \widetilde{P}|\mathcal{F}$. Thus $d(X_n, X_m) \overset{P}{\to} 0$, where d is Levy distance on S, and Theorem 4.3 implies that X_n converges in distribution to a σ-additive f.a.p. μ on $\mathcal{B}(S)$.

Assuming X_n tight for each n is crucial in Theorem 4.3. Otherwise, X_n can fail to converge in distribution, even though $X_n \overset{P}{\to} X$.

Example 4.5. Let $X : \Omega \to \mathbb{R}$ be a map and P a f.a.p. on $\sigma(X)$ such that $P(X \in \mathbb{N}) = P(X > n) = 1$ for all n. Define $S = \mathbb{R}$ and $X_n = X$ or $X_n = X + \frac{1}{X}$ according to n is even or odd. Then, $X_n \overset{P}{\to} X$ but X_n fails to converge in distribution. In fact, since $\{n + \frac{1}{n} : n \geq 2\}$ and $\{2, 3, \ldots\}$ are closed and disjoint, some $f \in C_b(\mathbb{R})$ meets $f(n + \frac{1}{n}) = 1$ and $f(n) = 0$ for all $n \geq 2$, so that $E_P f(X_n)$ does not converge to a limit.

In the spirit of Theorems 4.1 and 4.3, one can ask whether $X_n \overset{Q}{\to} X$ for some extension Q of P, provided $X_n \to X$ pointwise, $d(X_n, X_m) \overset{P}{\to} 0$ and each X_n is tight. The answer is generally no, and this does not depend only on the possible non completeness of the space $L_1(P)$.

Example 4.6. Let $S = \mathbb{R}$, $\Omega = \mathbb{N}$, $X_n = I_{\{1,\ldots,n\}}$, $X = 1$ and P a f.a.p. on $\mathcal{F}$ satisfying $P\{n\} = 0$ for all n. Then, $X_n \to X$ pointwise, each X_n is tight, and $\mathcal{F} = \bigcup_n \sigma(X_1, \ldots, X_n)$ is the field of finite co-finite subsets of $\mathbb{N}$. Since P is 0-1-valued, $L_1(P)$ is complete. However, $X_n \overset{P}{\to} 0$, and thus $X_n - X_m \overset{P}{\to} 0$ but $X_n \overset{Q}{\to} 1$ fails to be true for every extension Q of P.

We close with a brief remark. In the coherent framework, the limit in distribution (provided it exists) is not unique, even if f.a.p.'s are restricted to the field generated by the open subsets of S. Two types of limit should be mentioned. One is a f.a.p. regular on open sets, while the other is

$$\mu_\pi(\cdot) = \int P(X_n \in \cdot)\pi(dn)$$

where π is a f.a.p. on $\mathcal{P}(\mathbb{N})$ satisfying $\pi\{n\} = 0$ for all n. Indeed, μ_π has the nice property that $\mu_\pi(B) = \lim_n P(X_n \in B)$ whenever $B \in \mathcal{B}(S)$ and the limit exists. Sometimes, this fact is useful for modelling real situations. A well known example is $S = \mathbb{Q}$ (the set of rational numbers) and μ_π such that $\mu_\pi\big([0,t] \cap \mathbb{Q}\big) = t$ for all $0 \le t \le 1$. Another example is in [2]: For a suitable choice of S and X_n, under μ_π, a pathwise (Lebesgue-Stieltjes) integral has the same distribution of an Ito integral (at least on a certain class of events).

References

[1] Alexandrov A.D. (1940) Additive set functions in abstract spaces, *Mat. Sb. I (Chap. 1)*, 8, 307-348; (1941) *II (Chap. 2-3)*, 9, 563-628; (1943) *III (Chap. 4-6)*, 13, 169-238.

[2] Berti P., Regazzini E., Rigo P. (2003) Integration with respect to Brownian-like processes, *Sankhya*, 65, 715-732.

[3] Berti P., Rigo P. (2004) Convergence in distribution of non measurable random elements, *Ann. Probab.*, 32, 365-379.

[4] Berti P., Rigo P. (2006) Finitely additive uniform limit theorems, *Sankhya*, 68, 24-44.

[5] Bhaskara Rao K.P.S., Bhaskara Rao M. (1983) *Theory of charges*, Academic Press.

[6] de Finetti B. (1930) Sui passaggi al limite nel calcolo delle probabilita', *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere*, vol. LXIII, 1-12.

[7] de Finetti B. (1930) Problemi determinati e indeterminati nel calcolo delle probabilita', *Rendiconti della R. Accademia Nazionale dei Lincei*, vol. XII, 327-333.

[8] de Finetti B. (1930) A proposito dell'estensione del teorema delle probabilita' totali alle classi numerabili, *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere*, vol. LXIII, 1-5.

[9] de Finetti B. (1931) Sul significato soggettivo della probabilita', *Fund. Math.*, 17, 298-329.

[10] Dubins L.E., Savage L.J. (1965) *How to gamble if you must: Inequalities for stochastic processes,* McGraw Hill, New York.

[11] Dubins L.E. (1974) On Lebesgue-like extensions of finitely additive measures, *Ann. Probab.*, 2, 456-463.

[12] Dunford N., Schwartz J.T. (1958) *Linear Operators, Part I: General Theory*, Wiley-Interscience, New York.

[13] Girotto B., Holzer S. (1993) Weak convergence of masses on normal topological spaces, *Sankhya*, 55, 188-201.

[14] Karandikar R.L. (1988) A general principle for limit theorems in finitely additive probability: the dependent case, *J. Multivariate Anal.*, 24, 189-206.

[15] Letta G. (1982) Sur le suites simplement convergentes de mesures de Daniell, *Bollettino* U.M.I., (6) 1-B, 1105-1115.

[16] Purves R.A., Sudderth W.D. (1976) Some finitely additive probability, *Ann. Probab.*, 4, 259-276.

[17] Regazzini E. (1985) Finitely additive conditional probabilities, *Rend. Sem. Mat. Fis. Milano*, 55, 69-89. Corrections in: *Rend. Sem. Mat. Fis. Milano* (1987) 57, 599.

[18] Regazzini E. (1991) Coherence, exchangeability and statistical models (de Finetti's stance "revisited"), In: *Sviluppi metodologici nei diversi approcci all'inferenza statistica*, Pitagora Editrice Bologna, 1-37.

[19] Regazzini E., Zanoni G. (1988) Stochastic convergence in a finitely additive setting, *Quaderno* n. 29/1988, Dip. Matematica "F. Enriques", Univ. Milano.

[20] Sazonov V.V. (1965) On perfect measures, *Amer. Math. Soc. Transl.*, (2), 48, 229-254.

Patrizia Berti, Dipartimento di Matematica Pura ed Applicata "G. Vitali", Universita' di Modena e Reggio-Emilia, via Campi 213/B, 41100 Modena, Italy
E-mail address: berti.patrizia@unimore.it

Eugenio Regazzini, Dipartimento di Matematica "F. Casorati", Università di Pavia, via Ferrata 1, 27100 Pavia, Italy
E-mail address: eugenio.regazzini@unipv.it

Pietro Rigo (corresponding author), Dipartimento di Economia Politica e Metodi Quantitativi, Universita' di Pavia, via S. Felice 5, 27100 Pavia, Italy
E-mail address: prigo@eco.unipv.it