Riv. Mat. Univ. Parma (4) 8 (1982), 243-250

CARLO SEMPI (*)

On the space of distribution functions (**)

1 - Introduction

The aim of this note is to study a few properties of the set of distribution functions. To this end a definition of distribution function is adopted which, while not new (see [**10**] and [**12**]) is not common in the literature on probability.

Definition 1. A function $F: \boldsymbol{R}^* \to [0,1]$ is said to be a distribution function (d.f.) if it fulfils the following conditions:

(i) F is non-decreasing, i.e. $x' < x''$ implies $F(x') \leqslant F(x'')$;
(ii) F is continuous on the right on $\boldsymbol{R}$: $F(x+0) = F(x)$;
(iii) $F(-\infty) = 0 \leqslant l'(F) := \lim\limits_{x \to -\infty} F(x)$;
(iv) $F(+\infty) = 1 \geqslant l''(F) := \lim\limits_{x \to +\infty} F(x)$.

The set of d.f.'s will be denoted by $\mathscr{F}$. Here $R^* = R \cup \{-\infty, +\infty\}$.

The notation $:=$ is used to define the l.h.s. by means of the r.h.s.

It should be noted that a function that is constant on the reals, i.e. $F(x) = c$ $(x \in \boldsymbol{R})$ is a d.f. provided $c \in [0,1]$, $F(-\infty) = 0$ and $F(+\infty) = 1$. If F is the d.f. of a random variable X then $P[X = -\infty] = l'(F)$ and $P[X = +\infty] = 1 - l''(F)$, whilst $F(x) = P[X \leqslant x]$ if $x \in \boldsymbol{R}$. The definition of d.f. just given allows to consider random variables that take the values

---

(*) Indirizzo: Dipartimento di Matematica, Università, 73100 Lecce, Italy.
(**) Lavoro eseguito nell'ambito del G.N.F.M. (C.N.R.). — Ricevuto: 11-XI-1980.

$-\infty$ and/or $+\infty$ with non-zero probability; these latter random variables are excluded by the usual definition that stipulates

(v) $$\lim_{x\to-\infty} F(x) = 0 \quad \text{and} \quad \lim_{x\to+\infty} F(x) = 1$$

instead of (iii) and (iv) respectively. We shall denote by $\mathscr{F}_0$ the subset of $\mathscr{F}$ composed by the d.f.'s that satisfy (v).

There are situations in which $P[|X| = +\infty] > 0$, viz. the probability of a random variable taking an infinite value is not zero. Such examples are encountered in renewal theory, in a theory of physical measurement ([**12**]) or in the theory of probabilistic metric spaces ([**11**]). Moreover d.f.'s that agree with Definition 1, except possibly for (ii), occur in the theory of finitely additive probabilities.

One is confronted with a variety of different ways of defining weak convergence for d.f.'s. With the usual definition of d.f. those definitions turn out to be equivalent (see, e.g., [**4**], § 9). That need not continue to be so, as will shortly be seen, if Definition 1 is adopted. One will therefore have to choose among the several possible definitions of weak convergence. We shall adopt the following

Definition 2. A sequence $\{F_n\}_{n\in N} \subset \mathscr{F}$ will be said to *converge weakly* to the d.f. $F_n \in \mathscr{F}$ if

(1) $$\lim_{n\to\infty} F_n(x) = F(x) \qquad \forall x \in C(F)\,,$$

where $C(F) \subset \boldsymbol{R}^*$ is the set of continuity points of F; viz. $x \in \boldsymbol{R}$ belongs to $C(F)$ iff $F(x-0) = F(x+0)$; $-\infty \in C(F)$ iff $l'(F) = 0$; and $+\infty \in C(F)$ iff $l''(F) = 1$. The weak convergence of F_n to F will be denoted by $F_n \rightharpoonup F$.

2 - Properties of weak convergence

Helly's second theorem still holds, with the same proof, e.g., as in [**5**] (pp. 282-283).

Theorem 1. *If* $F_n \rightharpoonup F(F_n, F \in \mathscr{F})$, *if* $a, b \in C(F) \cap \boldsymbol{R}$ *and if* φ: $[a, b] \to \boldsymbol{R}$ *is continuous, then*

$$\lim_{n\to\infty} \int_{[a,b]} \varphi \, dF_n = \int_{[a,b]} \varphi \, dF\,.$$

Theorem 1 admitted of an extension and a converse (see [**5**], p. 283; [**4**], p. 33) as follows: *$F_n \rightharpoonup F(F_n, F \in \mathscr{F}_0)$ if, and only if, for every bounded continuous function $\varphi\colon \boldsymbol{R} \to \boldsymbol{R}$ one has*

$$\lim_{n\to\infty} \int_{\boldsymbol{R}} \varphi \, \mathrm{d}F_n = \int_{\boldsymbol{R}} \varphi \, \mathrm{d}F . \tag{2}$$

This result ceases to be valid as is shown by the following

Example 1. Let $\{a_n\}$ and $\{b_n\}$ be two convergent sequences of reals such that

(i) $0 \leqslant a_n < b_n \leqslant 1 \quad (n \in \boldsymbol{N})$; (ii) $\lim\limits_{n\to\infty} a_n = a < b = \lim\limits_{n\to\infty} b_n$.

Let $F_n \in \mathscr{F}$ $(n \in \boldsymbol{N})$ be defined for $x \in \boldsymbol{R}$ by

$$F_n(x) = \begin{cases} a_n & x < -n \\ a_n + (b_n - a_n)(x+n)/2n & x \in [-n, n] \\ b_n & x \geqslant n . \end{cases} \tag{3}$$

Then if

$$F(x) = (a+b)/2 \qquad (x \in \boldsymbol{R}) \tag{4}$$

one has $F_n \rightharpoonup F$. However for $\varphi(1) = 1$ $(x \in \boldsymbol{R})$ one has

$$\int_{\boldsymbol{R}} \varphi \, \mathrm{d}F_n = \{(b_n - a_n)/2n\} \int_{-n}^{n} \mathrm{d}x = b_n - a_n \to b - a > 0 ,$$

whilst $\int_{\boldsymbol{R}} \varphi \, \mathrm{d}F = 0$. This shows that $F_n \rightharpoonup F$ does not imply (2). Conversely, that (2) does not imply $F_n \rightharpoonup F$ can be immediately seen by taking, e.g., $F_n(x) = a_n$ $(n \in \boldsymbol{N})$, $F(x) = b$ $(x \in \boldsymbol{R})$.

Theorem 2. *If $F_n \rightharpoonup F(F_n, F \in \mathscr{F})$, then*

$$\limsup_{n\to\infty} l'(F_n) \leqslant l'(F) \tag{5}$$

and

$$\liminf_{n\to\infty} l''(F_n) \geqslant l''(F) . \tag{6}$$

Proof. One need only prove (5) since the proof of (6) is analogous. By definition, for arbitrary $\varepsilon > 0$, there exist $x_0 \in C(F)$ and $n_0 \in \boldsymbol{N}$ such that $l'(F) > F(x_0) - \varepsilon/2$ and $|F_n(x_0) - F(x_0)| < \varepsilon/2$ whenever $n \geqslant n_0$. Thus, for $n \geqslant n_0$ and for every $x < x_0$, one has $l'(F) > F(x_0) - \varepsilon/2 > F_n(x_0) - \varepsilon > F_n(x) - \varepsilon$, whence $l'(F_n) < l'(F) + \varepsilon$; now (5) follows on account of the arbitrariness of ε.

Corollary 1. *Under the assumption of Theorem 2 if* $l'(F) = 0$, *then* $\lim_{n\to\infty} l'(F_n) = 0$, *while if* $l''(F) = 1$, *then* $\lim_{n\to\infty} l''(F_n) = 1$.

Example 2. Let F_n and F be defined by (3) and (4) respectively. Then $l'(F_n) = a_n$, $l''(F_n) = b_n$, $l'(F) = l''(F) = (a+b)/2$ and therefore $\lim_{n\to\infty} l'(F_n) = a < l'(F)$ and $\lim_{n\to\infty} l''(F_n) = b > l''(F)$, which shows that the inequalities in (5) and in (6) can be strict.

Helly's first theorem holds with the usual proof (see, e.g. [**1**] or [7]); it can be stated in the form

Theorem 3. *Any sequence of d. f.'s* $\{F_n\}_{n\in N} \subset F$ *contains (at least) one subsequence* $\{F_{n(k)}\}_{k\in N}$ *that converges weakly to a* $F \in F$.

3 - F as a metric space

It will be shown presently that a metric can be introduced in F. To this end, let $a, b \in Q$ (viz. a and b are rationals) with $a < b$; then define $\varphi_{ab}\colon R^* \to [0, 1]$ by

$$\varphi_{ab}(x) := \begin{cases} 1 & x < a \\ (b-a)^{-1}(b-x) & x \in [a, b] \\ 0 & x > b\,, \end{cases}$$

if $x \in \boldsymbol{R}$, $\varphi_{ab}(-\infty) = 1$, $\varphi_{ab}(+\infty) = 0$.

The set of functions $\{\varphi_{ab}\colon a, b \in Q,\ a < b\}$ is countable and can be enumerated as $\{\theta_1, \theta_2, \ldots\}$. It is now easy to prove

Theorem 4. *The function* $d_F\colon F \times F \to R^+$ *defined by*

$$d_F(F, G) := \sum_{r=1}^{\infty} 2^{-r} \Big| \int_{R^*} \theta_r \,\mathrm{d}F - \int_{R^*} \theta_r \,\mathrm{d}G \Big| \qquad (F, G \in \mathscr{F}) \tag{7}$$

is a metric on F.

The distance (7) is a slight alteration of that used in [**6**].

Theorem 5. *Convergence in the distance* (7) *is equivalent to weak convergence, that is* $F_n \rightharpoonup F$ *if and only if* $d_F(F_n, F) \to 0$.

Proof. For every $G \in \mathscr{F}$ one has

$$\int_{R^*} \varphi_{ab}\, dG = l'(G) + \int_R \varphi_{ab}\, dG = l'(G) + \{-\int_R G\, d\varphi_{ab} + [G(x)\varphi_{ab}(x)]_{-\infty}^{+\infty}\} \tag{8}$$

$$= l'(G) + (b-a)^{-1}\int_a^b G\, dx - l'(G) = (b-a)^{-1}\int_a^b G\, dx ,$$

where use has been made of the formula of integration by parts for Stieltjes integrals ([**9**] p. 118).

Now let $F_n, F \in \mathscr{F}$ $(n \in \boldsymbol{N})$ and set

$$\delta(r, n) := |\int_{R^*} \theta_r\, dF_n - \int_{R^*} \theta_r\, dF| \qquad (r, n \in \boldsymbol{N}) . \tag{9}$$

Assume $d_F(F_n, F) \to 0$. It follows from (7) that $0 \leqslant \delta(r, n) \leqslant 2^r d_F(F_n, F)$ $(r, n \in \boldsymbol{N})$, so that $\lim_{n\to\infty} \delta(r, n) = 0$ for every $r \in \boldsymbol{N}$. Because of (9) and (8) this means

$$\lim_{n\to\infty} \int_a^b F_n\, dx = \int_a^b F\, dx$$

and therefore $F_n \rightharpoonup F$ as in [**6**] (p. 317).

Conversely if $F_n \rightharpoonup F$ then $\lim_{n\to\infty} \delta(r, n) = 0$ for every $r \in \boldsymbol{N}$, in view of (8). Since $\delta(r, n) \leqslant 2$ $(r, n \in \boldsymbol{N})$ one has $\lim_{n\to\infty} d_F(F_n, F) = \lim_{n\to\infty} \sum_{r=1}^{\infty} 2^{-r}\delta(r, n) = 0$ by the dominated convergence theorem applied to the counting measure on $\{1, 2, ...\}$.

The proof of Theorem 5 is a modified version of that of theorem 12.2 in [**6**].

Convergence in the Lévy distance (see [**4**] or [**8**])

$$d_L(F, G) := \inf\{h > 0 : F(x-h) - h \leqslant G(x) \leqslant F(x+h) + h \qquad \forall x \in \boldsymbol{R}\}$$

is not equivalent to weak convergence in $\mathscr{F}$, although it can proved, in the usual way, that $d_L(F_n, F) \to 0$ implies $F_n \rightharpoonup F$. In the following two examples sequences in $\mathscr{F}$ are given that converge weakly but which do not converge in the Lévy distance.

Example 3. If F_n is defined as in (3) with $a_n = 0$ and $b_n = 1$ $(n \in \boldsymbol{N})$ and consequently $F(x) = \frac{1}{2}$ $(x \in \boldsymbol{R})$, then a simple calculation yields $d_L(F_n, F) = \frac{1}{2}$ so that $d_L(F_n, F)$ does not tend to zero as n goes to infinity although $F_n \rightharpoonup F$.

Example 4. If H_n is the right-continuous Heaviside function with jump at $x = n$, i.e. if

$$H_n(x) = \begin{cases} 0 & x < n \\ 1 & x \geqslant n \end{cases}$$

then $H_n \rightharpoonup N$, where $N \in \mathscr{F}$ is identically zero on $\boldsymbol{R}$. Now $d_L(H_n, N) = 1$, so that $\{H_n\}$ does not converge to N in the Lévy metric. However since $\int_{R^*} \varphi_{ab} \, \mathrm{d}H_n = \varphi_{ab}(n)$ one has, explicitly for (7)

$$d_F(H_n, N) = \sum_{r=1}^{\infty} 2^{-r} \left| \int_{R^*} \theta_r \, \mathrm{d}H_n \right| = \sum_{r=1}^{\infty} 2^{-r} \theta_r(n) \,.$$

Thus given $\varepsilon > 0$, let k be the smallest natural such that $2^{-k} < \varepsilon$; one can then choose n large enough, say $n \geqslant \nu$, to have $\theta_r(n) = 0$ for $r = 1, 2, \ldots, k$ and, as a consequence

$$d_F(H_n, N) = \sum_{r=k+1}^{\infty} 2^{-r} \theta_r(n) \leqslant \sum_{r=k+1}^{\infty} 2^{-r} = 2^{-k} < \varepsilon \,.$$

Theorem 6. *The metric space* (F, d_F) *is complete.*

Proof. This theorem is a consequence of Theorem 7 below, but it can be proved independently as follows. Let $\{F_n\}_{n \in N} \subset \mathscr{F}$ be a Cauchy sequence, i.e. for every $\varepsilon > 0$ there exists $n(\varepsilon) \in \boldsymbol{N}$ such that $d_F(F_n, F_m) < \varepsilon$ whenever $n, m \geqslant n(\varepsilon)$. By Theorem 3, $\{F_n\}$ has a subsequence $\{F_{n(k)}\}$ that converges weakly to $F \in \mathscr{F}$. It follows from the previous theorem that $\lim_{k \to \infty} d_F(F_{n(k)}, F) = 0$. The triangle inequality can now be used to show that $\lim_{n \to \infty} d_F(F_n, F) = 0$.

Contrary to what happens with the usual definition of d.f. (see, e.g., [**1**] p. 329, [**6**] p. 319, [**2**] p. 160) the space $(\mathscr{F}, d_F)$ is compact. ($\mathscr{F}_0$ is complete but not compact; see [**8**]).

Theorem 7. *The metric space* (F, d_F) *is compact.*

Proof. The space $(\mathscr{F}, d_F)$ is sequentially compact; but a metric space is sequentially compact if and only if it is compact (see, e.g., [**3**] (3.16.1)).

The metric (7) is different although equivalent to that proposed by Sibley ([**13**]) and modified by Schweizer ([**10**]). It is, however, worth noticing that the metric proposed here renders evident an important property of weak convergence, namely that it is compatible with the existence of a norm. To be sure, it would not be proper to speak of a norm on $\mathscr{F}$ since this space is not linear. But it is easy to regard $\mathscr{F}$ as a subset of the linear space $BV(\boldsymbol{R}^*)$ of the functions of bounded variation on $\boldsymbol{R}^*$. It is a simple task now to verify that the map $\| \ \|: BV(\boldsymbol{R}^*) \to \boldsymbol{R}^+$ defined by

$$\|F\| = \sum_{r=1}^{\infty} 2^{-r} \left| \int_{\boldsymbol{R}^*} \theta_r \, \mathrm{d}F \right| \qquad F \in BV(\boldsymbol{R}^*)$$

is a norm on $BV(\boldsymbol{R}^*)$ if the functions of $BV(\boldsymbol{R}^*)$ are supposed to be continuous on the right.

References

[1] R. B. Ash, *Real Analysis and Probability*, Academic Press, New York - London 1972.

[2] L. Breiman, *Probability*, Addison-Wesley, Reading (Mass.) 1968.

[3] J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York - London 1969.

[4] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading (Mass.) 1954.

[5] B. V. Gnedenko, *The Theory of Probability*, Chelsea, New York 1968.

[6] J. F. C. Kingman and S. J. Taylor, *Introduction to Measure and Probability*, Cambridge University Press, Cambridge 1966.

[7] M. Loève, *Probability Theory*, Springer-Verlag, New York - Heidelberg - Berlin 1977 (4th ed.).

[8] E. Lukacs, *Stochastic Convergence*, Academic Press, New York, London 1975.

[9] F. Riesz et B. Sz. Nagy, *Leçons d'Analyse Fonctionnelle*, Gauthier-Villars, Paris 1968 (5ième éd.).

[10] B. Schweizer: *Multiplication on the space of distribution functions*, Aequationes Math. **12** (1975) 156-183.

[11] B. Schweizer and A. Sklar, *Probabilistic metric spaces*, Elzevier North-Holland, New York 1982.

[12] C. Sempi, *The entropies with error*, Rend. Mat. (6) **9** (1976) 1-15.

[13] D. A. Sibley, *A metric for weak convergence of distribution functions*, Rocky Mountain J. Math. **1** (1971) 427-430.

Riassunto

Si studiano le proprietà della convergenza debole per le funzioni di ripartizione usando una definizione che, pur discostandosi da quella usuale, presenta il vantaggio di includere anche variabili aleatorie che assumono i valori $-\infty$ *e/o* $+\infty$ *con probabilità non nulla; essa risponde inoltre alle proprietà richieste alle funzioni di ripartizione nella teoria delle probabilità finitamente additive. La convergenza debole equivale alla convergenza in un'opportuna metrica e si mostra che, dotato di tale metrica, lo spazio delle funzioni di ripartizione risulta essere compatto e completo. Si può inoltre far discendere la metrica introdotta da una norma.*

* * *