Riv. Mat. Univ. Parma (4) **16** (1989), 219-226

MARIUSZ STARTEK and DOMINIK SZYNAL (*)

On a metric defined on the space of probability measures (**)

1 - Introduction

Let S be a separable metric space and $\mathcal{S}$ the Borel σ-field of subsets of S. We denote by $\mathscr{P}$ the space of all probability measures on $(S, \mathcal{S})$ with defective measures, i.e. $P \in \mathscr{P}$ iff $P(S) \leq 1$. Examples of defective measures are encountered in renewal theory, in a theory of physical measurement $[4]_1$ or in the theory of probabilistic metric spaces [3].

It is known that weak convergence of a sequence $\{P_n, n \geq 1\}$ of probability measures is equivalent to the convergence in the Prokhorov distance. However, this statement is not still true for a sequence $\{P_n, n \geq 1\}$ of measures of the larger class $\mathscr{P}$. In $[4]_2$ it has been introduced a distance between distribution functions (not in the probabilistic sense) such that the convergence in that distance is equivalent to the weak convergence of a sequence $\{F_n, n \geq 1\}$ of distribution functions on $\mathbb{R}$. The aim of this note is to extend this result to measures on metric spaces. To do this we need to modify the distance introduced in $[4]_2$ in such way which is suitable to that extension.

2 - A distance between distribution functions

We shall see that the modified distance $d_{\mathscr{F}}$ of $[4]_2$ can be used to extend the results of 2 to measures defined on a metric space.

(*) Indirizzo degli AA.: M. STARTEK, Instytut Matematyki i Fizyki, Politechnika Rzeszowska, ul.W. Pola 2, PL-35-959 Rzeszów; D. SZYNAL, Instytut Matematyki UMCS, ul. Nowotki 10, PL-20-031 Lublin.
(**) Ricevuto: 2-VIII-1989.

Let $\mathscr{F}$ be a set of functions $F\colon \mathbb{R}^* \to [0, 1]$ such that:

(i) F is non-decreasing, i.e. $x < y$ implies $F(x) \leqslant F(y)$;

(ii) F is continuous on the right on $\mathbb{R}$: $F(x+0) = F(x)$;

(iii) $F(-\infty) = 0 \leqslant \lim_{x \to -\infty} F(x)$;

(iv) $F(+\infty) = 1 \geqslant \lim_{x \to +\infty} F(x)$.

Here $\mathbb{R}^* = \mathbb{R} \cup \{-\infty, +\infty\}$.

We say that a sequence $\{F_n, n \geqslant 1\} \subset \mathscr{F}$ *converges weakly* to the function $F \in \mathscr{F}$, $F_n \Rightarrow F$, if for every bounded and continuous function $f\colon \mathbb{R}^* \to \mathbb{R}$

$$\lim_{n \to \infty} \int_{\mathbb{R}^*} f(x)\, \mathrm{d}F_n(x) = \int_{\mathbb{R}^*} f(x)\, \mathrm{d}F(x).$$

Let D be a countable and dense subset of $\mathbb{R}$, $D^* = D \cup \{-\infty, +\infty\}$. We introduce the functions $\varphi_{abr}\colon \mathbb{R}^* \to [0, 1]$, $a, b \in D^*$, $a < b$, $r > 0$, $r \in \mathrm{Q}$, Q is the set of rational numbers, by

$$\varphi_{abr}(x) = \begin{array}{ll} (x-a+r)/r & x \in [a-r, a) \\ 1 & x \in [a, b] \\ (b+r-x)/r & x \in (b, b+r] \\ 0 & x \in \mathbb{R}^* \setminus [a-r, b+r] \end{array}$$

where $[-\infty, -\infty) = (+\infty, +\infty] := \emptyset$. The bounded continuous functions φ_{abr} are the set standard approximations of the indicators.

The set $\{\varphi_{abr}: a, b \in D^*, a < b, 0 < r < \infty, r \in \mathrm{Q}\}$ is countable and can be enumerated as $\{f_1^D, f_2^D, \ldots\}$.

Now we define a function $d_{\mathscr{F}}^D$ on $\mathscr{F} \times \mathscr{F}$ as follows

$$d_{\mathscr{F}}^D(F, G) = \sum_{k=1}^{\infty} 2^{-k} \Big| \int_{\mathbb{R}^*} f_k^D\, \mathrm{d}F - \int_{\mathbb{R}^*} f_k^D\, \mathrm{d}G \Big| \qquad F, G \in \mathscr{F}.$$

Lemma 1. *The function $d_{\mathscr{F}}^D\colon \mathscr{F} \times \mathscr{F} \to \mathbb{R}_+$ is a metric.*

Proof. By the definition $d_{\mathscr{F}}^D$ we have $d_{\mathscr{F}}^D(F, G) = 0$ iff $F = G$ and $d_{\mathscr{F}}^D(F, G) = d_{\mathscr{F}}^D(G, F)$ for all $F, G \in \mathscr{F}$.

Let now $F, G, H \in \mathscr{F}$. Then

$$d^D_{\mathscr{F}}(F, H) = \sum_{k=1}^{\infty} 2^{-k} \Big| \int_{R^*} f^D_k \, dF - \int_{R^*} f^D_k \, dH \Big|$$

$$= \sum_{k=1}^{\infty} 2^{-k} \Big| \int_{R^*} f^D_k \, dF - \int_{R^*} f^D_k \, dG + \int_{R^*} f^D_k \, dG - \int_{R^*} f^D_k \, dH \Big|$$

$$\leqslant \sum_{k=1}^{\infty} 2^{-k} \Big(\Big| \int_{R^*} f^D_k \, dF - \int_{R^*} f^D_k \, dG \Big| + \Big| \int_{R^*} f^D_k \, dG - \int_{R^*} f^D_k \, dH \Big| \Big)$$

$$= \sum_{k=1}^{\infty} 2^{-k} \Big| \int_{R^*} f^D_k \, dF - \int_{R^*} f^D_k \, dG \Big| + \sum_{k=1}^{\infty} 2^{-k} \Big| \int_{R^*} f^D_k \, dG - \int_{R^*} f^D_k \, dH \Big|$$

$$= d^D_{\mathscr{F}}(F, G) + d^D_{\mathscr{F}}(G, H)$$

which completes the proof.

Now we prove that the convergence in the distance $d^D_{\mathscr{F}}$ is independent of a choice of $D \subset \mathbb{R}$.

Proposition 1. *Let D, E be two any given dense and countable subsets of $\mathbb{R}$, $D^* = D \cup \{-\infty, +\infty\}$, $E^* = E \cup \{-\infty, +\infty\}$ and $\{F_n, n \geqslant 1\}$ be a sequence of functions of $\mathscr{F}$. The convergence F_n to F in the metric $d^D_{\mathscr{F}}$ is equivalent to the convergence in the metric $d^E_{\mathscr{F}}$.*

Proof. Let $d^D_{\mathscr{F}}(F_n, F) \to 0$, $n \to \infty$. Then

$$\Big| \int_{R^*} \varphi_{pqr} \, dF_n - \int_{R^*} \varphi_{pqr} \, dF \Big| \to 0 \qquad n \to \infty \tag{1}$$

for every $p, q \in D^*$, $r \in \mathbb{Q}$. Write

$$\delta(p, q, r, n) := \Big| \int_{R^*} \varphi_{pqr} \, dF_n - \int_{R^*} \varphi_{pqr} \, dF \Big|$$

and let $a, b \in E^*$, $\varepsilon > 0$. We note that there exist $c, d \in D^*$ such that

$$|\varphi_{abr}(x) - \varphi_{cdr}(x)| \leqslant \varepsilon/2 \qquad x \in \mathbb{R}^*.$$

Hence

$$\Big| \int_{R^*} \varphi_{abr} \, dF_n - \int_{R^*} \varphi_{cdr} \, dF_n \Big| \leqslant \varepsilon/2 \qquad \Big| \int_{R^*} \varphi_{abr} \, dF - \int_{R^*} \varphi_{cdr} \, dF \Big| \leqslant \varepsilon/2 .$$

Therefore, $|\delta(a, b, r, n)-\delta(c, d, r, n)| \leqslant \varepsilon$, so, by (1), $\delta(a, b, r, n) \to 0$, $n \to \infty$, for every $a, b \in E^*$, or

$$\Big|\int\limits_{R^*} f_k^E \, dF_n - \int\limits_{R^*} f_k^E \, dF\Big| \to 0 \qquad n \to \infty \qquad \text{for every } k \in \mathrm{N}.$$

Moreover, we have

$$\Big|\int\limits_{R^*} f_k^E \, dF_n - \int\limits_{R^*} f_k^E \, dF\Big| \leqslant 2 \qquad k \in \mathrm{N} \qquad n \in \mathrm{N}$$

which imply

$$\lim_{n\to\infty} d_{\mathscr{F}}^E(F_n, F) = \lim_{n\to\infty} \sum_{k=1}^{\infty} 2^{-k} \Big|\int\limits_{R^*} f_k^E \, dF_n - \int\limits_{R^*} f_k^E \, dF\Big| = 0$$

and, by symmetry of D and E, completes the proof of Proposition 1.

Theorem 1. *The weak convergence of a sequence* $\{F_n, n \geqslant 1\}$ *of functions of* $\mathscr{F}$ *is equivalent to the convergence in the distance* $d_{\mathscr{F}}^D$, *i.e.* $F_n \Rightarrow F$ *iff* $d_{\mathscr{F}}^D(F_n, F) \to 0$, $n \to \infty$, *for any countable and dense subset D of continuity points of F.*

Proof. Choose $a, b \in D^*$, and $q < (b-a)/2$. Let $\delta(\alpha, b, r, n) = \big|\int\limits_{R^*} \varphi_{abr} \, dF_n - \int\limits_{R^*} \varphi_{abr} \, dF\big|$. If $d_{\mathscr{F}}^D(F_n, F) \to 0$, $n \to \infty$, then $\delta(a, b, r, n) \to 0$, $n \to \infty$, for every a, b and $r < q$, so

$$\lim_{n\to\infty} \int\limits_{R^*} \varphi_{abr} \, dF_n = \int\limits_{R^*} \varphi_{abr} \, dF.$$

Suppose that $a \neq -\infty$ and $b \neq +\infty$. Then we have

$$\limsup_{n\to\infty} (F_n(b) - F_n(a-)) \leqslant \limsup_{n\to\infty} \int\limits_{R^*} \varphi_{abr} \, dF_n = \int\limits_{R^*} \varphi_{abr} \, dF \leqslant F(b+r) - F(a-r).$$

Letting $r \to 0$ we get

$$\limsup_{n\to\infty} (F_n(b) - F_n(a-)) \leqslant F(b) - F(a-). \tag{2}$$

Taking $p \in (r/2, r]$ such that $a+p \in D$ and $b-p \in D$ we have

$$\liminf_{n\to\infty} (F_n(b) - F_n(a-)) \geqslant \liminf_{n\to\infty} \int_{R^*} \varphi_{a+p, b-p, r/2}\, dF_n = \int_{R^*} \varphi_{a+p, b-p, r/2}\, dF$$

$$\geqslant F(b-p) - F(a+p) \geqslant F(b-r) - F(a+r).$$

Then letting $r \to 0$, we have

$$\liminf_{n\to\infty} (F_n(b) - F_n(a-)) \geqslant F(b-) - F(a+).$$

Taking into account that a, b are continuity points of F we get

$$\liminf_{n\to\infty} (F_n(b) - F_n(a-)) \geqslant F(b) - F(a-). \tag{3}$$

If $a = -\infty$ then $a+r = a = a-$, so (2) and (3) are also true. Similarly these inequalities are true if $b = +\infty$.

From (2) and (3) we have

$$\limsup_{n\to\infty} (F_n(b) - F_n(a-)) \leqslant F(b) - F(a-) \leqslant \liminf_{n\to\infty} (F_n(b) - F(a-)).$$

Thus

$$\lim_{n\to\infty} (F_n(b) - F_n(a-)) = F(b) - F(a-)$$

for every a, $b \in D^*$ and so $F_n \Rightarrow F$.

Let now $F_n \Rightarrow F$. Then $\delta(a, b, r, n) \to 0$, $n \to \infty$, for every a, $b \in D^*$ and $r \in Q$, so

$$\left| \int_{R^*} f_k^D\, dF_n - \int_{R^*} f_k^D\, dF \right| \to 0 \qquad n \to \infty \qquad k = 1, 2, \ldots .$$

Since $\left| \int_{R^*} f_k^D\, dF_n - \int_{R^*} f_k^D\, dF \right| \leqslant 2$, therefore,

$$\lim_{n\to\infty} d_{\mathscr{F}}^D(F_n, F) = \lim_{n\to\infty} \sum_{k=1}^{\infty} 2^{-k} \left| \int_{R^*} f_k^D\, dF_n - \int_{R^*} f_k^D\, dF \right| = 0$$

which completes the proof.

By virtue of the Proposition 1 we have the following

Corollary 1. *The weak convergence of functions of $\mathscr{F}$ is equivalent to the convergence in the distance $d^D_{\mathscr{F}}$ for any dense countable set $D \subset \mathbb{R}$.*

3 - A distance in the space $\mathscr{P}$

Let $\mathscr{K}$ be the family of the closed ball $K(s_i, r_j)$, where $\{s_1, s_2, \dots\}$ is a countable and dense subsets of S and $D = \{r_1, r_2, \dots\}$ is a dense sequence of positive numbers. Denote by $\mathscr{U}$ the family containing S and all the finite intersections of $K(s_i, r_j)$. $\mathscr{U}$ is the convergence determining class [1]. The family $\mathscr{U}$ is countable and can be enumerated as $\{U_1, U_2, \dots\}$. Let $U_i = \bigcap_{n=1}^{N} K(s_{in}, r_{in})$, $r \in \mathbb{R}_+$ and $r'_{in} = \max\{0, r_{in} - r\}$. Then we denote $U_{i_r} := \bigcap_{n=1}^{N} K(s_{in}, r'_{in})$ and $\mathscr{K}_{i_r} := \{K(s_{in}, r'_{in}): n = 1, 2, \dots, \mathrm{N}\}$. Now we define the functions $g_{ir}: S \to [0, 1]$ as follows

$$g_{ir}(x) = \begin{cases} 1 & \text{if } x \in U_i \\ 1 - \operatorname{dist}(x, U_i)/r & \text{if } \operatorname{dist}(x, U_i) \leqslant r \\ 0 & \text{otherwise} \end{cases}$$

for $i = 1, 2, \dots$ and $r \in \mathbb{Q}_+$.

The set $\{g_{ir}: i = 1, 2, \dots, r > 0, r \in \mathbb{Q}\}$ is countable and can be enumerated as $\{f_1^D, f_2^D, \dots\}$.

Define now a function d^D on $\mathscr{P} \times \mathscr{P}$ as follows

$$d^D(P, Q) = \sum_{k=1}^{\infty} 2^{-k} \left| \int f_k^D \, dP - \int f_k^D \, dQ \right|. \tag{4}$$

Lemma 2. *The function d^D: $\mathscr{P} \times \mathscr{P} \to \mathbb{R}_+$ defined by* (4) *is a metric.*

Proof. The proof is similar to that of Lemma 1.

Proposition 2. *Let D, E be two any given dense and countable subsets of $\mathbb{R}$ and $\{P_n, n \geqslant 1\}$ be a sequence of functions of $\mathscr{P}$. The convergence P_n to P in the metric d^D is equivalent to the convergence in the metric d^E.*

Proof. The proof is similar to that of Proposition 1.

Now we prove the following

Theorem 2. *The weak convergence of a sequence* $\{P_n,\ n \geq 1\}$ *of measures of* $\mathscr{P}$ *is equivalent to the convergence in the distance* (4), *i.e.* $P_n \Rightarrow P$ *iff* $d^D(P_n, P) \to 0$, $n \to \infty$, *for any countable and dense subset* D *of* R_+ *such that* $K(s_i, r)$, $r \in D$, $i \in \mathrm{N}$, *is a continuity set of* P.

Proof. Choose $U_i \in \mathscr{U}$ and q such that $U_{i_q} \neq \emptyset$.

Let $\gamma(i, r, n) := |\int g_{ir}\, \mathrm{d}P_n - \int g_{ir}\, \mathrm{d}P|$. If $d^D(P_n, P) \to 0$, $n \to \infty$, then $\gamma(i, r, n) \to 0$, $n \to \infty$, for every i and $r < q$, so $\lim_{n\to\infty} \int g_{ir}\, \mathrm{d}P_n = \int g_{ir}\, \mathrm{d}P$.

We have

$$\limsup_{n\to\infty} P_n(U_i) \leqslant \limsup_{n\to\infty} \int g_{ir}\, \mathrm{d}P_n = \int g_{ir}\, \mathrm{d}P \leqslant P(U_i^r)$$

where $U_i^r = \{x \in S\colon \mathrm{dist}(x, U_i) < r\}$. Letting $r \to 0$, we get

$$\limsup_{n\to\infty} P_n(U_i) \leqslant P(U_i).$$

Taking $p \in (r/2, r]$ such that $\mathscr{K}_{i_p} \subset \mathscr{U}$, we have

$$\liminf_{n\to\infty} P_n(U_i) \geqslant \liminf_{n\to\infty} \int g_{i_p r/2}\, \mathrm{d}P_n = \int g_{i_p r/2}\, \mathrm{d}P \geqslant P(U_{i_p}) \geqslant P(U_{i_r}).$$

Letting $r \to 0$, taking into account that U_i is a continuity set of P and by virtue of the equality $\lim_{r\to 0} U_{i_r} = U_i$ we get

$$\liminf_{n\to\infty} P_n(U_i) \geqslant P(U_i).$$

Thus $$\limsup_{n\to\infty} P_n(U_i) \leqslant P(U_i) \leqslant \liminf_{n\to\infty} P_n(U_i).$$

Hence $$\lim_{n\to\infty} P_n(U_i) = P(U_i) \quad \text{for every } i \quad \text{and so } P_n \Rightarrow P.$$

Let now $P_n \Rightarrow P$. Then $\gamma(i, r, n) \to 0$, $n \to \infty$, for every i and r, so

$$\left|\int f_k^D\, \mathrm{d}P_n - \int f_k^D\, \mathrm{d}P\right| \to 0 \qquad n \to \infty \qquad k = 1, 2, \ldots.$$

Since $|\int f_k^D \, dP_n - \int f_k^D \, dP| \leqslant 2$, therefore,

$$\lim_{n\to\infty} d^D(P_n, P) = \lim_{n\to\infty} \sum_{k=1}^{\infty} 2^{-k} |\int f_k^D \, dP_n - \int f_k^D \, dP| = 0$$

which completes the proof.

By virtue of the Proposition 2 we have the following

Corollary 2. *The weak convergence of measures of* $\mathscr{P}$ *is equivalent to the convergence in the distance* d^D *for any dense countable* $D \subset R_+$.

References

[1] P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968.

[2] K. R. Parthasarathy, *Probability measures on metric spaces*, Academic Press, 1967.

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

[4] C. Sempi: [•]$_1$ *The entropies with error*, Rend. Math. (6) 9 (1976), 1-15; [•]$_2$ *On the space of distribution functions*, Riv. Mat. Univ. Parma (4) 8 (1982), 243-250.

Summary

See Introduction.

***