

# WEAK CONVERGENCE ON NON-SEPARABLE METRIC SPACES

DAVID POLLARD

(Received 2 January; revised 27 February 1979)

Communicated by R. Tweedie

## Abstract

For certain problems in the weak convergence theory of probability measures on non-separable metric spaces it is convenient to consider measures defined on a  $\sigma$ -field  $\mathcal{B}_0$  smaller than the Borel  $\sigma$ -field. A suitable theory has been developed by Dudley and Wichura. In this paper it is shown that some of the key results in that theory can be deduced directly from the better known weak convergence theory for Borel measures. This is achieved by a remetrization of the underlying space to make it separable. The  $\sigma$ -field  $\mathcal{B}_0$  contains the new Borel  $\sigma$ -field and weak convergence in Dudley's sense becomes equivalent to convergence of restrictions of the measures to the latter  $\sigma$ -field.

*1980 Mathematics Subject Classification (Amer. Math. Soc.):* primary 60 B 10, 60 B 05.

*Keywords and phrases:* weak convergence of non-Borel measures, non-separable metric spaces,  $\sigma$ -field generated by the balls.

## 1. Introduction

A theory of weak convergence of (non-Borel) measures on (non-separable) metric spaces was introduced by Dudley (1966) in order to cope with measurability problems arising in the study of empirical processes. The difficulties occur if a (suitably normed) empirical distribution function  $F_n$  is to be regarded as a random element of a space such as  $D[0, 1]$ . Under the sup norm topology on  $D[0, 1]$  the map  $F_n$  need not be Borel measurable, essentially because of the non-separability of this topology (see Billingsley (1968), Section 18). Dudley's solution was to equip  $D[0, 1]$  with a weaker  $\sigma$ -field: the  $\sigma$ -field  $\mathcal{B}_0$  generated by the class of all balls.

---

Supported by a fellowship of the Alexander von Humboldt Foundation while visiting the Ruhr-Universität Bochum.

Since the  $F_n$ 's are  $\mathcal{B}_0$  measurable, they induce a sequence of probability measures on that  $\sigma$ -field for which the usual limit theorems can be proved once the weak convergence concept has been suitably defined. Dudley also showed that the same approach can be used to prove the corresponding weak convergence results for multidimensional empirical processes.

A better known solution to the problem lies in equipping  $D[0,1]$  with the weaker (separable) Skorokhod  $J_1$ -topology, under which the maps  $F_n$  become Borel measurable. The weak convergence result for the sequence of induced Borel measures in this setting can be proved equivalent to the one proved by Dudley.

What is more, it can also be shown that the Borel  $\sigma$ -field of the Skorokhod topology coincides with  $\mathcal{B}_0$ , since both are generated by the finite dimensional projection maps (see Theorem 14.5 of Billingsley (1968) and Proposition 1.1 of Dudley (1978)).

The advantage of using Dudley's approach derives from the simpler structure of the uniform topology; the disadvantage is that it is apparently necessary to develop a separate theory of weak convergence for measures defined only on the  $\sigma$ -field  $\mathcal{B}_0$ . In this paper it will be shown that, by means of a simple device, the main results for the non-separable theory can in fact be deduced from the corresponding theory for weak convergence of Borel measures. Perhaps this will encourage a greater interest in the non-separable theory; in view of recent applications by Dudley (1978), to cases where there is no obvious analogue of the  $J_1$ -metric available, that theory clearly has something to offer.

## 2. Weak convergence of non-Borel probability measures

Let  $(S, d)$  be a (possibly non-separable) metric space. The  $\sigma$ -field on  $S$  generated by the family of all balls will be denoted by  $\mathcal{B}_0$ . This is the smallest  $\sigma$ -field with respect to which all of the functions  $d(\cdot, x)$ , for each fixed  $x \in S$ , are measurable. Thus  $\mathcal{B}_0$  forms a sub  $\sigma$ -field of the Borel  $\sigma$ -field, with equality at least in the case where the topology is separable. (Talagrand (1978) has constructed an example of a non-separable space for which  $\mathcal{B}_0$  coincides with the Borel  $\sigma$ -field.) Write  $C(S)$  for the class of all bounded continuous real functions on  $S$ .

A sequence  $\{P_n\}$  of probability measures on  $\mathcal{B}_0$  will be said to converge weakly (in Dudley's sense) to a probability  $P$  on  $\mathcal{B}_0$  if

- (i)  $P$  has separable support,
- (ii)  $\lim_n \int^* f dP_n = \lim_n \int_* f dP_n = \int f dP$  for every  $f \in C(S)$ .

Here  $\int^*$  and  $\int_*$  denote upper and lower integrals. (For example,

$$\int^* f dP_n = \inf \left\{ \int g dP_n : \text{measurable } g \geq f \right\}.$$

The stars are not needed for the last integral since (i) implies that  $f$  is  $P$  integrable, that is  $\int^* f dP = \int_* f dP$ . This form of weak convergence will be denoted by  $P_n \rightarrow P$ . Some equivalent forms for (ii), due to Dudley (1967) and Wichura (1968), will arise naturally out of the results to be proved in the next section.

Three of the more important results for such a notation of weak convergence run as follows. The set of all points within a distance  $\delta$  of a set  $K$  is denoted by  $K^\delta$ .

**THEOREM A** (compare with Proposition 1 of Dudley (1966)). *Suppose  $P_n \rightarrow P$  and that  $F$  is a bounded real function on  $S$  which is  $P$  almost surely continuous. Then*

$$\lim_n \int^* F dP_n = \lim_n \int_* F dP_n = \int F dP.$$

**THEOREM B** (compare with Theorem 1 of Dudley (1966)). *Let  $\{P_n\}$  be a sequence of probability measures on  $\mathcal{B}_0$  having the property that for each  $\varepsilon > 0$  there exists a compact set  $K(\varepsilon) \subseteq S$  such that for every  $\delta > 0$*

$$\liminf_n P_n(K^\delta(\varepsilon)) > 1 - \varepsilon.$$

*Then there exists a weakly convergent subsequence (in Dudley's sense).*

**THEOREM C** (compare with Theorem 2 of Wichura (1970)). *Suppose  $P_n \rightarrow P$ . Then there exists a probability space  $(\Omega, \mathcal{F}, \mu)$  and  $\mathcal{F}$ - $\mathcal{B}_0$ -measurable random maps  $X_n, X$  from  $\Omega$  into  $S$  such that*

$$\mu X_n^{-1} = P_n, \quad \mu X^{-1} = P$$

*and*

$$X_n \rightarrow X, \quad \mu \text{ almost surely.}$$

(Recall that a map  $X$  from  $\Omega$  into  $S$  is said to be  $\mathcal{F}$ - $\mathcal{B}_0$ -measurable if  $X^{-1} B$  belongs to  $\mathcal{F}$  for every  $B$  in  $\mathcal{B}_0$ .)

Once again the absence of stars on the integral  $\int F dP$  in Theorem A indicates that the function is  $P$  integrable in the sense that the upper and lower integrals agree. Similarly in Theorem B, the set  $K^\delta(\varepsilon)$  of all points within distance  $\delta$  of  $K(\varepsilon)$  is actually  $\mathcal{B}_0$ -measurable, since it can be represented as a countable union of open balls whose centres form a dense subset of the compact set  $K(\varepsilon)$ .

The corresponding result for sequences of Borel measures might be more familiar to some readers if  $K^\delta(\varepsilon)$  were to be replaced by  $K(\varepsilon)$ . However, this slightly weaker condition is to be preferred since it can easily be shown to be necessary for the weak convergence of  $\{P_n\}$  to a tight measure (use the fact that  $K^\delta(\varepsilon)$  is open). The form of Theorem C given here differs from that proved by Wichura in that he considered nets of probabilities whose domains of definition may be larger than  $\mathcal{B}_0$ ; Wichura also proved convergence in the almost uniform sense, which is stronger than the almost sure convergence in Theorem C. The less

general form suffices for most applications though (Pyke (1969, 1970)). The analogous result in the separable case was proved by Dudley (1968), generalizing the original theorem of Skorokhod (1956) for complete, separable metric spaces.

In the next two sections a method for deducing these theorems from the corresponding results for separable metric spaces will be described.

### 3. The weaker separable topology on $S$

As before  $(S, d)$  is a possibly non-separable metric space. Let  $S_0$  be a closed separable subset of  $S$ , chosen to support the limit measure. As in the classical method for generating a separable metric topology, a countable collection of continuous real functions on  $S$  will be used to define a pseudometric. In our case the functions should be  $\mathcal{B}_0$ -measurable, in order that the Borel  $\sigma$ -field be weaker than  $\mathcal{B}_0$ , and be such that the neighbourhood base for the topology at points of  $S_0$  agrees with that of the original  $d$ -topology; the analogy with the behaviour of the Skorokhod topology at points of  $C[0, 1]$  will become clear (compare with Billingsley (1968), page 112).

Let  $T = \{x_1, x_2, \dots\}$  be a countable dense subset of  $S_0$ . Since the distance function  $d(\cdot, S_0)$  can be represented as  $\inf \{d(\cdot, x_n) : x_n \in T\}$  which is  $\mathcal{B}_0$ -measurable, it follows that  $S_0$  itself belongs to  $\mathcal{B}_0$ . Now consider the uniformity on  $S$  generated by the countable family of real continuous functions  $\{d(\cdot, S_0), d(\cdot, x_n) : x_n \in T\}$ . One possible pseudometric for this uniformity is

$$p(x, y) := |d(x, S_0) - d(y, S_0)| + \sum_{n=1}^{\infty} 2^{-n} \min \{1, |d(x, x_n) - d(y, x_n)|\}.$$

(The connection between uniformities and pseudometrics is described by Ash (1972), page 404.) In general the uniformity need not be Hausdorff, but that creates no great problems. Since the generating family of functions is countable, the  $p$ -topology is separable: separability of  $(S, p)$  is equivalent to separability of the associated metric space of equivalence classes  $[x] := \{y \in S : p(x, y) = 0\}$ ; the map  $[x] \rightarrow (d(x, x_n) : x_n \in T)$  is a homeomorphism between this metric space and a subset of the countable product of real lines  $\mathbf{R}^T$ . Because each of the functions in the definition of  $p$  is  $d$ -uniformity continuous, the  $p$ -topology and uniformity are weaker than the  $d$ -topology and uniformity. Denote the generated Borel  $\sigma$ -field by  $\mathcal{B}(p)$ . The key properties of the  $p$ -topology can be summarized in six points.

#### THEOREM 1.

- (i) *The  $p$ -topology is separable.*
- (ii) *The  $p$ -topology (uniformity) is weaker than the  $d$ -topology (uniformity).*
- (iii)  *$S_0$  is closed in the  $p$ -topology.*

- (iv) If  $p(y_\alpha, y) \rightarrow 0$  and  $y \in S_0$  then  $d(y_\alpha, y) \rightarrow 0$ .
- (v) The  $p$ - and  $d$ -topologies agree on  $S_0$ .
- (vi)  $\mathcal{B}_0 \supseteq \mathcal{B}(p)$ .

PROOF. (i) and (ii) have already been noted. Assertion (iii) follows from the continuity of  $d(\cdot, S_0)$  under the  $p$ -topology. For part (iv) first choose an  $x_m \in T$  for which  $d(y, x_m) < \varepsilon$ , then notice that  $d(y_\alpha, x_m) \rightarrow d(y, x_m)$ . It follows that

$$\limsup d(y_\alpha, y) < 2\varepsilon.$$

Part (v) is a direct consequence of (ii) and (iv). Finally, from (i) it suffices to show that each of the functions  $p(\cdot, x)$ , for  $x \in S$ , is  $\mathcal{B}_0$ -measurable;  $\mathcal{B}_0$ -measurability of the functions  $d(\cdot, x_n)$  takes care of this.

Notice that if  $\mathcal{B}_0$  is not countably generated then we cannot hope to strengthen (vi) to equality. However, it can be arranged that  $\mathcal{B}(p)$  contains any fixed countably generated sub  $\sigma$ -field of  $\mathcal{B}_0$ . First observe that such a sub  $\sigma$ -field must be generated by some countable collection of balls; if the countable set  $\{x_1, x_2, \dots\}$  is expanded to include the centres of these balls then the balls themselves become closed sets in the (new)  $p$ -topology. This will apply to both of the spaces  $D_0(\mathcal{C}, P)$  and  $D(\mathcal{C}, P)$ , defined by Dudley (1978), in the case when  $\mathcal{C}$  possesses a countable subset  $\mathcal{D}$  with the property (Dudley (1978), page 902) that for each  $C \in \mathcal{C}$  there exists a sequence  $\{D(n)\} \subseteq \mathcal{D}$  such that  $1_{D(n)}(x) \rightarrow 1_C(x)$  for all  $x \in X$ . This is analogous with the classical  $D[0, 1]$  situation.

The most important aspect of the new topology is that the notion of weak convergence is preserved. For a measure  $P$  defined on  $\mathcal{B}_0$ , write  $P'$  for its restriction to  $\mathcal{B}(p)$ . If  $P$  happens to concentrate on  $S_0$  this entails no loss of information. Use the usual symbol  $\Rightarrow$  to denote weak convergence of measures on the Borel  $\sigma$ -field  $\mathcal{B}(p)$ .

**THEOREM 2.** *Let  $\{P_n\}$ ,  $P$  be probabilities on  $\mathcal{B}_0$ , with  $P$  concentrated on the closed separable set  $S_0$ . Then  $P_n \rightarrow P$  if and only if  $P'_n \Rightarrow P'$ .*

PROOF. Suppose that  $P_n \rightarrow P$  and that  $f$  is bounded and  $p$ -continuous. Then from parts (ii) and (vi) of Theorem 1 it follows that  $f$  is both  $d$ -continuous and  $\mathcal{B}_0$ -measurable. Hence

$$\begin{aligned} \int f dP'_n &= \int^* f dP_n \quad \left( = \int_* f dP_n \right) \\ &\rightarrow \int f dP \\ &= \int f dP'; \end{aligned}$$

that is,  $P'_n \Rightarrow P'$ .

On the other hand, suppose that  $P'_n \Rightarrow P'$ . Somewhat more than  $P_n \rightarrow P$  will be proved. For  $F$  bounded and  $P$  almost surely  $d$ -continuous, let  $\bar{F}$  denote the smallest function greater than  $F$  which is upper semicontinuous with respect to the  $p$ -topology, and  $\tilde{F}$  the largest  $p$ -lower semicontinuous function less than  $F$ . Then (Topsøe (1970), page XI) the set  $\{\bar{F} = F = \tilde{F}\}$  consists of all those points at which  $F$  is  $p$ -continuous. By assumption  $F$  is  $d$ -continuous  $P$  almost surely; since  $P(S_0) = 1$ , part (iv) of Theorem 1 implies that  $\bar{F} = F = \tilde{F}$  on a set of  $P'$  measure one. Hence the equalities  $\int \bar{F} dP' \geq \int^* F dP \geq \int_* F dP \geq \int \tilde{F} dP'$  become equalities throughout. The convergence  $P_n \rightarrow P$  now follows by the standard argument (Topsøe (1970), page 41):

$$\begin{aligned} \limsup \int^* F dP_n &\leq \limsup \int \bar{F} dP'_n \\ &\leq \int \bar{F} dP' \\ &= \int F dP \\ &= \int \tilde{F} dP' \\ &\leq \liminf \int \tilde{F} dP'_n \\ &\leq \liminf \int_* F dP_n. \end{aligned}$$

**COROLLARY 1.** *The second part of the proof was given in greater generality than needed there (only  $d$ -continuous function  $F$  need have been considered) because we have thereby obtained the proof of Theorem A.*

**COROLLARY 2** (Dudley (1967) and Theorem 1.2 of Wichura (1968)). *With  $\{P_n\}$  and  $P$  as above, the convergence  $P_n \rightarrow P$  is equivalent to  $\int f dP_n \rightarrow \int f dP$  for all bounded uniformly  $d$ -continuous  $\mathcal{B}_0$ -measurable functions  $f$ .*

This follows immediately from Theorem 1 part (ii) and Theorem 2, since the class of all such  $f$ 's contains all bounded  $p$ -uniformly continuous real functions. The latter determine the weak convergence  $P'_n \Rightarrow P'$ .

#### 4. Theorems B and C

For the first of these theorems the aim is to transform the problem to the corresponding one for Borel measures, where known results may be applied. Specifically, we can use a variant of the well-known Prohorov theorem for separable metric spaces.

LEMMA (see Lemma 4 of Štěpán (1970), or Theorem 6.1 of Billingsley (1968)).  
*Let  $\{P_n\}$  be a sequence of Borel probability measures on a separable pseudometric space. If, for every  $\varepsilon > 0$ , there exists a compact set  $K(\varepsilon)$  such that*

$$\liminf P_n(G) > 1 - \varepsilon$$

*for every open  $G \supseteq K(\varepsilon)$ , then there exists a subsequence converging weakly to a tight measure.*

PROOF. Using Štěpán's technique we obtain a tight measure  $P$  as a cluster point of  $\{P_n\}$ . Apply the Cantor diagonalization procedure to find a subsequence  $\{P_{n_i}\}$  for which  $\liminf P_{n_i}(G_i) \geq P(G_i)$  for each  $G_i$ , where  $\{G_1, G_2, \dots\}$  forms a countable base for the topology which is closed under the formation of finite unions. It follows that  $P_{n_i} \Rightarrow P$ . (See, for example, Theorem II.5 of Varadarajan (1965).)

The appropriate separable set to choose will be the closure of  $\bigcup_{m=1}^{\infty} K(m^{-1})$ . Using this as  $S_0$ , construct the pseudometric  $p$  as in Section 3. By virtue of the  $d$ -compactness of  $K(m^{-1})$  and parts (iii) and (v) of Theorem 1, these sets will also be closed (since the relative topology on  $S_0$  is Hausdorff) and compact in the  $p$ -topology. Also for each  $p$ -open set  $G$  containing  $K(m^{-1})$  there exists, by Theorem 1 (iv), a  $\delta > 0$  depending on the set  $G$  for which  $G \supseteq K^\delta(m^{-1})$ . Hence

$$\liminf P_n(G) > 1 - \varepsilon$$

for every such  $p$ -open  $G$ . The above Lemma can be applied to the Borel measures  $P'_n$  obtained by restricting the  $P_n$ 's to  $\mathcal{B}(p)$ , to obtain a subsequence  $P'_{n_i} \Rightarrow P'$ . Since  $P'$  concentrates on the  $p$ -closed set  $S_0$ , and since by Theorem 1 part (v) the  $\sigma$ -fields  $\mathcal{B}_0$  and  $\mathcal{B}(p)$  agree on  $S_0$ , the probability  $P'$  can be regarded as a measure on  $\mathcal{B}_0$  having support in  $S_0$ . Application of Theorem 2 then shows that  $P_{n_i} \rightarrow P'$ . This completes the proof of Theorem B.

Theorem C can be deduced directly from Theorem 3 of Dudley (1968), provided two trivial modifications are made to his proof. First notice that Dudley's arguments carry over without change to the pseudometric case. Also, observe that if  $P_n$  is defined on a  $\sigma$ -field  $\mathcal{B}_n$  bigger than the Borel  $\sigma$ -field, the random map  $X_n$  can be chosen as  $\mathcal{B}_n$ -measurable; only changes in interpretation of the wording are needed for this. Using this modified form of Dudley's result with  $\mathcal{B}_n = \mathcal{B}_0$  for all  $n$ , and with the separable topology induced by the  $p$ -pseudometric, we obtain

$\mathcal{B}_0$ -measurable random maps  $X_n$  and  $X$ , inducing the desired measures on  $\mathcal{B}_0$ , for which  $p(X_n, X)$  converges to zero almost surely. Since  $X$  concentrates on  $S_0$ , part (iv) of Theorem 1 implies that the almost sure convergence can be strengthened to almost sure convergence with respect to the  $d$  metric. This is the desired form of the result.

For the special case of  $D[0, 1]$  under its uniform metric, the above type of argument has already been used by Pyke and Shorack (1968) and Pyke (1969).

### Acknowledgment

I wish to thank Peter Gänsler not only for some helpful comments on this paper, but also for his general hospitality and assistance during my stay in Bochum.

### References

- R. B. Ash (1972), *Real analysis and probability* (Academic Press, New York).  
 P. Billingsley (1968), *Convergence of probability measures* (Wiley, New York).  
 R. M. Dudley (1966), 'Weak convergence of probabilities on non-separable metric spaces and empirical measures on Euclidean spaces', *Ill. J. Math.* **10**, 109–126.  
 R. M. Dudley (1967), 'Measures on non-separable metric spaces', *Ill. J. Math.* **11**, 449–453.  
 R. M. Dudley (1968), 'Distances of probability measures and random variables', *Ann. Math. Statist.* **39**, 1563–1572.  
 R. M. Dudley (1978), 'Central limit theorems for empirical measures', *Ann. Probability* **6**, 899–929.  
 R. Pyke (1969), 'Applications of almost surely convergent constructions of weakly convergent processes', *Proc. Internat. Symp. on Prob. and Inform. Theory*, Lecture Notes in Mathematics 89, 187–200 (Springer-Verlag, Berlin).  
 R. Pyke (1970), 'Asymptotic results for rank statistics', *Nonparametric techniques in statistical inference*, 21–37, edited by M. L. Puri (Cambridge Univ. Press).  
 R. Pyke and G. R. Shorack (1968), 'Weak convergence of a two-sample empirical process and a new approach to Chernoff–Savage theorems', *Ann. Math. Statist.* **39**, 755–771.  
 A. V. Skorokhod (1956), 'Limit theorems for stochastic processes', *Theor. Prob. Appl.* **1**, 261–290.  
 J. Štěpán (1970), 'On the family of translations of a tight probability measure on a topological group', *Z. Wahrscheinlichkeitstheorie verw. Geb.* **15**, 131–138.  
 M. Talagrand (1978), 'Les boules peuvent elles engendrer la tribu borélienne d'un espace métrisable non séparable?' *Communication au séminaire Choquet* (Paris).  
 F. Topsøe (1970), *Topology and measure*, Lecture Notes in Mathematics 133 (Springer-Verlag, Berlin).  
 V. S. Varadarajan (1965), 'Measures on topological spaces', *Amer. Math. Soc. Transl.*, Series 2, **48**, 161–228.  
 M. J. Wichura (1968), *On the weak convergence of non-Borel probabilities on a metric space* (Ph.D. Dissertation, Columbia University).  
 M. J. Wichura (1970), 'On the construction of almost uniformly convergent random variables with given weakly convergent image laws', *Ann. Math. Statist.* **41**, 284–291.

Department of Statistics  
 Yale University  
 New Haven, CT 06520  
 USA