MATH. SCAND. 25 (1969), 97–104

A CRITERION FOR WEAK CONVERGENCE OF MEASURES WITH AN APPLICATION TO CONVERGENCE OF MEASURES ON $D[0,1]$

FLEMMING TOPSØE

Introduction.

The usual way to establish weak convergence of a sequence of probability measures on $C[0,1]$ is to prove that all the finite dimensional distributions converge weakly to the "right" limits and that the sequence is tight; then a theorem, which is quite easy to establish, tells us that we have weak convergence. If we turn our attention to the Skorohod space $D[0,1]$ we find that the analogous theorem is much harder to obtain. Recently, P. Billingsley has obtained a suitable result published in [1, p. 124]. My aim has been to find a general theorem, valid in any Polish space, which implies the desired result in $D[0,1]$.

All measures below are supposed to be defined on the Borel σ-field.

A reasonable problem inspired by the concrete question about $D[0,1]$ is the following: *Let X be a Polish space (that is, separable and metrizable in such a way that it becomes complete), and let $\mathscr{A}$ be a field of Borel sets generating the entire Borel σ-field $\mathscr{B}(X)$. If P, $(P_n)_{n\geqq 1}$ are probability measures on X such that (P_n) is tight and such that $P_nA \to PA$ for all sets A in $\mathscr{A}$, is it then true, or under what additional assumptions is it true, that P_n converges weakly to P?*

Equivalently, we could ask if, under the just mentioned hypotheses on X and $\mathscr{A}$, the facts that $P_n \to Q$ for some probability measure Q and $P_nA \to PA$ for all A in $\mathscr{A}$, imply $P=Q$. Here, as below, the sign $\to$ is used to indicate weak convergence of probability measures as well as ordinary convergence of real numbers.

Since $P_n \to Q$ and $P_nA \to PA$ imply

$$Q\mathring{A} \leqq \liminf P_n\mathring{A} \leqq PA \leqq \limsup P_n\bar{A} \leqq Q\bar{A} ,$$

a problem related to the problem above arises: *Let again X be Polish and $\mathscr{A}$ a field generating $\mathscr{B}(X)$. If P and Q are two probability measures*

Received November 13, 1967; in revised form January 28, 1969.

such that $Q\mathring{A} \leqq PA \leqq Q\bar{A}$ *holds for all* A *in* $\mathscr{A}$, *is it then true, or under what additional assumptions is it true, that* $P = Q$?

Without additional assumptions the answers to both problems are negative. Indeed, O. Björnsson observed that there is a counterexample in which P and Q are as simple as possible, namely point masses.

1. A convergence criterion.

A class $\mathscr{A}$ of subsets of the topological space X is said to *separate points* T_2 if, for every pair of distinct points x and y in X there exists a set A in $\mathscr{A}$ such that $x \in \mathring{A}$ and $y \notin \bar{A}$.

THEOREM 1. *Let* $\mathscr{A}$ *be a lattice of subsets of the Polish space* X *and suppose that* $\mathscr{A}$ *separates points* T_2.

(i) *If* P *and* Q *are probability measures on* X *such that* $Q\mathring{A} \leqq P\bar{A}$ *for all* $A \in \mathscr{A}$, *then* $P = Q$.

(ii) *If* P, $(P_n)_{n \geqq 1}$ *are probability measures on* X *such that* $(P_n)_{n \geqq 1}$ *is tight and such that* $\limsup P_n\mathring{A} \leqq P\bar{A}$ *for all* $A \in \mathscr{A}$, *then* $P_n \to P$.

Part (ii) is the hoped for convergence criterion.

PROOF. (i) If K_1 and K_2 are disjoint compact subsets of X, we can find $A \in \mathscr{A}$ such that $K_1 \subset \mathring{A}$ and $\bar{A} \cap K_2 = \emptyset$. It follows that $QK_1 \leqq 1 - PK_2$. Employing the tightness of P and Q, one deduces from this that $P = Q$.

(ii) Let $(P_{n_k})_{k \geqq 1}$ be a convergent subsequence of $(P_n)_{n \geqq 1}$, say $P_{n_k} \to Q$. Then

$$Q\mathring{A} \leqq \liminf_{k \to \infty} P_{n_k}\mathring{A} \leqq \limsup_{n \to \infty} P_n\mathring{A} \leqq P\bar{A}$$

for all $A \in \mathscr{A}$. By (i), $Q = P$ follows.

Part (i) in the theorem can of course be considered as a special case of part (ii). It is easy to see, by means of a simple counterexample, that we can not drop the assumption that $\mathscr{A}$ be a lattice. We have been unable to decide whether one can relax this assumption assuming only that $\mathscr{A}$ is closed under finite intersections. The condition in Theorem 1 that $\mathscr{A}$ separates points T_2 is necessary, as is easily seen.

It is not difficult to extend the convergence criterion from measures on Polish spaces to tight measures (Radon measures) on arbitrary Hausdorff spaces; also, one may consider a lattice of functions in stead of a lattice of sets. It is our intention to publish elsewhere some results on weak convergence of tight measures on arbitrary Hausdorff spaces.

If $\mathscr{A}$ in the convergence criterion is a subclass of $\mathscr{B}(X)$, then $\mathscr{A}$ generates $\mathscr{B}(X)$; to see this, note that $\mathscr{A}$ contains a countable subclass separating points and apply Theorem 3.3 of [3]. This argument was pointed out to us by E. T. Kehlet.

2. An application to measures on the Skorohod space $D[0,1]$.

$D[0,1]$ consists of those real-valued functions on $[0,1]$ which are continuous from the right for $0 \leqq t < 1$ and have limits from the left for $0 < t \leqq 1$. The distance $d(x,y)$ between two functions in $D[0,1]$ is the infimum of those $\varepsilon \geqq 0$ for which there exists an increasing homeomorphism λ from $[0,1]$ onto itself such that $\|\lambda - i\| < \varepsilon$ and $\|x - y \circ \lambda\| < \varepsilon$; here i denotes the identity map and $\|\cdot\|$ the uniform norm. The metric space $D[0,1]$ is known to be Polish. For a finite (ordered) subset $\boldsymbol{t} = \{t_1, \ldots, t_k\}$ of $[0,1]$ we denote by $\pi_{\boldsymbol{t}}$ the projection from $D[0,1]$ onto R^k. The projections are all measurable. For a probability measure P on $D[0,1]$ we denote by T_P the set of $t \in [0,1]$ such that π_t is continuous a.e. P. There are at most countably many points in $[0,1] \setminus T_P$; these points are called fixed points of discontinuity for P.

The purpose of this section is to prove the following result:

THEOREM 2. *Let $(P_n)_{n \geqq 1}$ be a tight sequence of probability measures on $D[0,1]$ and let T be a dense subset of $[0,1]$ containing the point 1. Suppose that for each finite subset $\boldsymbol{t}$ of T there is a probability measure $P_{\boldsymbol{t}}$ on the proper Euclidean space such that $P_n \pi_{\boldsymbol{t}}^{-1}$ converges weakly to $P_{\boldsymbol{t}}$. Then the sequence $(P_n)_{n \geqq 1}$ converges weakly. Furthermore, the limit measure P can be identified by the formula*

$$P\pi_{\boldsymbol{t}}^{-1} = P_{\boldsymbol{t}+}, \tag{1}$$

which holds for any finite subset $\boldsymbol{t}$ of $[0,1]$.

The formula $P\pi_{\boldsymbol{t}}^{-1} = P_{\boldsymbol{t}+}$ means, first of all, that the limit from the right, in the sense of weak convergence, exists at $\boldsymbol{t}$ and, secondly, that this limit is the finite dimensional distribution of P at $\boldsymbol{t}$. In more detail, what we claim is the following: Let $\boldsymbol{t} = \{t_1, \ldots, t_k\}$ be any finite subset of $[0,1]$; let, for each $\nu = 1, 2, \ldots,$

$$\boldsymbol{t}_\nu = \{t_{\nu 1}, \ldots, t_{\nu k}\}$$

be a finite subset of T with $t_{\nu i} > t_i$ for each $i = 1, \ldots, k$ (unless $t_i = 1$ in which case we demand $t_{\nu i} = 1$); suppose further that

$$t_{\nu i} \downarrow t_i \quad \text{as } \nu \to \infty \quad \text{for each } i = 1, \ldots, k$$

(shortly: $\boldsymbol{t}_\nu \downarrow \boldsymbol{t}$). Then we claim that the measures $P_{\boldsymbol{t}_\nu}$ converge weakly in R^k to $P\pi_{\boldsymbol{t}}^{-1}$ as $\nu \to \infty$.

In Theorem 2, the limit measure is not supposed to be known in advance; thus the result can be used to construct various measures.

A special case of Theorem 2 (with known limit measure and special T) has been established by Billingsley (Theorem 15.1 of [1]). Our method is completely different from Billingsley's.

To prove that (P_n) converges weakly in Theorem 2 we need some simple lemmas.

LEMMA 1. *For any finite subset* $\boldsymbol{t} = \{t_1, \ldots, t_k\}$ *of* $[0,1]$ *and any subset* E *of* R^k *we have*

$$\pi_{\boldsymbol{t}}^{-1}(\bar{E}) \subset \overline{\pi_{\boldsymbol{t}}^{-1}(E)} \quad \textit{and} \quad \pi_{\boldsymbol{t}}^{-1}(\mathring{E}) \supset \widehat{\pi_{\boldsymbol{t}}^{-1}(E)}^{\circ} .$$

In other words, all projections are open mappings.

PROOF. Assume that $x \in \pi_{\boldsymbol{t}}^{-1}(\bar{E})$. Then $(x(t_1), \ldots, x(t_k))$ lies in $\bar{E}$. Thus, to any $\delta > 0$ we can find real numbers $r_1, \ldots, r_k$ with $|r_i| < \delta$ for all i, and such that $(x(t_1) + r_1, \ldots, x(t_k) + r_k)$ lies in E. Clearly then, there exists a function y in D with $y(t_i) = x(t_i) + r_i$ for all i, and such that the uniform distance from x to y is less than δ. Then the Skorohod distance from x to y is also less than δ. Since $y \in \pi_{\boldsymbol{t}}^{-1}(E)$ and δ is arbitrary the first inclusion follows. The second inclusion is a consequence of the first.

LEMMA 2. *Let s be a point in* $[0,1]$ *and E a subset of* R. *Denote by A the cylinder set* $\pi_s^{-1}(E) = \{x \in D : x(s) \in E\}$. *If* $0 < s < 1$, *then we have*

$$\bar{A} = \{x \in D : \ x(s-) \in \bar{E} \ \text{ or } \ x(s) \in \bar{E}\} ,$$
$$\mathring{A} = \{x \in D : \ x(s-) \in \mathring{E} \ \text{ and } \ x(s) \in \mathring{E}\} .$$

If s is either 0 *or* 1, *then*

$$\bar{A} = \{x \in D : \ x(s) \in \bar{E}\} ,$$
$$\mathring{A} = \{x \in D : \ x(s) \in \mathring{E}\} .$$

PROOF. The case $s = 0$ or 1 is easily treated. Now assume that $s \in (0,1)$. If $x \in \bar{A}$ then there exists a sequence (x_n) of functions in A and a sequence (λ_n) of increasing homeomorphisms of $[0,1]$ onto $[0,1]$ such that

$$\|\lambda_n - i\| \to 0 \quad \text{and} \quad \|x_n - x \circ \lambda_n\| \to 0 .$$

Put $s_n = \lambda_n(s)$. We may assume that either $s_n < s$ holds for all n or else $s_n \geqq s$ holds for all n. If the first alternative takes place, then $x(s_n) \to x(s-)$ and one finds that $x_n(s) \to x(s-)$ so that $x(s-) \in \bar{E}$. The second alternative leeds to $x(s) \in \bar{E}$.

To prove the reverse inclusion, assume first that $x(s) \in \bar{E}$; it follows from Lemma 2 that $x \in \bar{A}$. Now assume that $x(s-) \in \bar{E}$. By moving the function a little to the right and then adding a small constant function, one arrives at a function in A. Intuitively, it is thus clear that $x \in \bar{A}$. It is left to the reader to make this argument rigorous.

Lemma 3. *Let T be a dense subset of $[0,1]$ containing the point 1. Denote by $\mathscr{A}$ the class of cylinder sets based on time-points in T, that is, $\mathscr{A}$ is the class of all sets $\pi_{\boldsymbol{t}}^{-1}(E)$, where $\boldsymbol{t}$ ranges over all finite subsets of T and E ranges over all Borel subsets of the proper Euclidean spaces. Then $\mathscr{A}$ is a field separating points T_2.*

Proof. Clearly, $\mathscr{A}$ is a field. We shall prove that $\mathscr{A}$ separates points T_2. Let x and y be distinct functions in D. This means that, for some s in $[0,1]$, $x(s)$ is distinct from $y(s)$. We shall assume that $x(s) < y(s)$ holds.

If $s=1$, then the set $A = \pi_1^{-1}((-\infty, m))$, where m is the midpoint of $[x(1), y(1)]$ lies in $\mathscr{A}$ and by Lemma 2 we also find that $x \in \mathring{A}$ and $y \notin \bar{A}$.

If $s<1$, we argue as follows. First choose three real numbers m_1, m and m_2 such that $x(s) < m_1 < m < m_2 < y(s)$. Then, by the right continuity, we can find a positive δ with $s+\delta < 1$ such that $x(t) \leqq m_1$ and $y(t) \geqq m_2$ hold for any t in $(s, s+\delta)$. Since T is dense in $[0,1]$, we can find a t from T in $(s, s+\delta)$. Now put $A = \pi_t^{-1}((-\infty, m))$. A lies in $\mathscr{A}$ and by Lemma 1 we also find that $x \in \mathring{A}$ and $y \notin \bar{A}$. Thus $\mathscr{A}$ separates points T_2.

Proof of the weak convergence in Theorem 2. We begin by remarking that the family of measures $(P_{\boldsymbol{t}})$ where $\boldsymbol{t}$ ranges over all finite subsets of T is consistent. Now, let Q_1 be any limit measure for (P_n), say $P_{n'} \to Q_1$. Consider a finite subset $\boldsymbol{t} = \{t_1, \ldots, t_k\}$ of T and a k-dimensional Borel set E. Then

$$Q_1(\widehat{\pi_{\boldsymbol{t}}^{-1}E}^{\circ}) \leqq \liminf P_{n'}(\widehat{\pi_{\boldsymbol{t}}^{-1}E}^{\circ}) \leqq \liminf P_{n'}(\pi_{\boldsymbol{t}}^{-1}\mathring{E}) \leqq \limsup P_{n'}\pi_{\boldsymbol{t}}^{-1}(\bar{E}) \leqq P_{\boldsymbol{t}}(\bar{E}),$$

that is, we have

$$Q_1(\widehat{\pi_{\boldsymbol{t}}^{-1}E}^{\circ}) \leqq P_{\boldsymbol{t}}(\bar{E}). \tag{2}$$

If Q_2 is another limit measure for (P_n) then we find in an analogous manner

$$(3) \qquad P_{\boldsymbol{t}}(\overset{\circ}{E}) \leqq Q_2(\overline{\pi_{\boldsymbol{t}}^{-1}E}) .$$

Consider the class $\mathscr{A}$ of those subsets A of $D[0,1]$ for which there exist a finite subset $\boldsymbol{t}=\{t_1,\ldots,t_k\}$ of T and a k-dimensional Borel set E such that $A=\pi_{\boldsymbol{t}}^{-1}E$ and $P_{\boldsymbol{t}}(\partial E)=0$. Here ∂E is the boundary in R^k of E. In checking which sets A belong to $\mathscr{A}$ it does not matter which representation we use for A. Clearly, $\mathscr{A}$ is a field. The proof of Lemma 3 shows that $\mathscr{A}$ separates points T_2. By (2) and (3), the inequality $Q_1(\overset{\circ}{A})\leqq Q_2(\bar{A})$ holds for any set A in $\mathscr{A}$. By Theorem 1, Q_1 and Q_2 are identical.

We shall now prove (1). In case the limit measure P has only finitely many fixed points of discontinuity, this formula obviously holds. In the general case we use a rather elaborate argumentation and we begin with a lemma.

LEMMA 4. *Let x be a function in $D[0,1]$ and t a point in $[0,1]$ such that $|x(t)-x(t-)|<\varepsilon$. Then there exists a positive δ and a positive h such that any element in the open sphere $S(x,\delta)$ with center x and radius δ oscillates by less than ε in the interval $[t-h,t+h]$.*

This result is obvious.

PROOF OF (1) OF THEOREM 2. To ease the notation, we shall assume that the finite subset $\boldsymbol{t}$ that we consider in fact only contains one point t_0. We may also assume that $t_0<1$. What we have given is a sequence (s_k) of points in T with $s_k>t_0$ for all k and $s_k \downarrow t_0$ as $k\to\infty$. We want to prove that $P_{s_k}\to P\pi_{t_0}^{-1}$. Fix, for some time, two positive numbers ε and η. Since there are at most finitely many points t for which

$$P\{x:\ |x(t)-x(t-)|\geqq\varepsilon\}\ \geqq\ \eta$$

holds, we can find an integer k_0 such that

$$PA_k\ >\ 1-\eta \qquad \text{for all } k\geqq k_0\,,$$

where we have put

$$A_k\ =\ \{x:\ |x(s_k)-x(s_k-)|<\varepsilon\}\,.$$

Choose a compact set K_k with $K_k\subset A_k$ and $PK_k>1-\eta$. To any function x in K_k we choose two positive numbers δ_x and h_x such that the oscillation of any function from $S(x,\delta_x)$ over the interval $[s_k-h_x,\,s_k+h_x]$ is less than ε. Finitely many of the spheres, say $S(x_1,\delta_{x_1}),\ldots,S(x_r,\delta_{x_r})$, cover K_k. Put

$$G_k = S(x_1, \delta_{x_1}) \cup \ldots \cup S(x_r, \delta_{x_r}) \quad \text{and} \quad h_k = \min\{h_{x_1}, \ldots, h_{x_r}\}.$$

Then G_k is open, G_k contains K_k and any function in G_k oscillates by less than ε in the interval $[s_k - h_k, s_k + h_k]$. Now choose a point t_k in the interval $[s_k - h_k, s_k + h_k]$ such that $t_k > t_0$, $|t_k - s_k| < 1/k$ and such that $t_k \in T_P$. Let E be any Borel subset of R. Then, for $k \geqq k_0$ the following inclusion holds:

$$\{x : x(s_k) \in E\} \subset \{x : x(t_k) \in E^\varepsilon\} \cup G_k{}^c .$$

Here E^ε denotes the ε-neighbourhood of E, that is, the set of points within distance less than ε from E, and ^{c} indicates complementation. Considering a $k \geqq k_0$ and using the weak convergence $P_n \pi_{t_k}{}^{-1} \to P\pi_{t_k}{}^{-1}$ (which follows since $t_k \in T_P$), we now find that

$$P_{s_k}(\overset{\circ}{E}) \leqq \liminf_{n\to\infty} P_n \pi_{s_k}{}^{-1}(E) \leqq \limsup_{n\to\infty} P_n \pi_{t_k}{}^{-1}(\overline{E^\varepsilon}) + \limsup_{n\to\infty} P_n(\overline{G_k{}^c})$$
$$\leqq P\pi_{t_k}{}^{-1}(\overline{E^\varepsilon}) + P(\overline{G_k{}^c}) \leqq P\pi_{t_k}{}^{-1}(\overline{E^\varepsilon}) + \eta .$$

Since $t_k \downarrow t_0$, the projections π_{t_k} converge everywhere to π_{t_0} as $k \to \infty$, hence

$$P\pi_{t_k}{}^{-1} \to P\pi_{t_0}{}^{-1} \quad \text{as } k \to \infty .$$

From the above inequality we thus find

$$\limsup_{k\to\infty} P_{s_k}(\overset{\circ}{E}) \leqq \limsup_{k\to\infty} P\pi_{t_k}{}^{-1}(\overline{E^\varepsilon}) + \eta \leqq P\pi_{t_0}{}^{-1}(\overline{E^\varepsilon}) + \eta$$
$$\leqq P\pi_{t_0}{}^{-1}(E^{2\varepsilon}) + \eta .$$

Since this holds for all positive ε and η, we finally find that

$$\limsup P_{s_k}(\overset{\circ}{E}) \leqq P\pi_{t_0}{}^{-1}(\bar{E}) .$$

This being so for any E, we conclude by Theorem 1 that $P_{s_k} \to P\pi_{t_0}{}^{-1}$.

Acknowledgement.

I have had helpful discussions with K. R. Parthasarathy who suggested to me the two problems stated in the introduction; he was led to these problems by observing an error in the book by Gihman and Skorohod (p. 578 in [2]), where these authors seem to rely on a result like Theorem 2 without observing the difficulties involved.

REFERENCES

1. P. Billingsley, *Weak convergence of probability measures*, New York, 1968.
2. I. I. Gihman and A. V. Skorohod, *Introduction to the theory of random processes*, Izdat. "Nauka", Moscow, 1965. (Russian.)

3. G. W. Mackey, *Borel structure in groups and their duals,* Trans. Amer. Math. Soc. 85 (1957), 134–165.
4. K. R. Parthasarathy, *Probability measures on metric spaces*, New York, 1967.
5. Yu. V. Prohorov, *Convergence of random processes and limit theorems in probability,* Theor. Probability Appl. 1 (1956), 157–214.
6. A. V. Skorohod, *Limit theorems for stochastic processes,* Theor. Probability Appl. 1 (1956), 261–290.

UNIVERSITY OF COPENHAGEN, DENMARK