The Annals of Mathematical Statistics
1971, Vol. 42, No. 5, 1769–1772

A NOTE ON THE WEAK CONVERGENCE OF STOCHASTIC PROCESSES[1]

BY M. J. WICHURA
University of Chicago

A simple method giving quick access to some important general weak convergence theorems is described and illustrated.

1. Introduction. We point out in this note a quick and easy method for establishing some general weak convergence theorems for probability distributions on a metric space. The only prerequisite needed is the so-called "portmanteau" theorem of Billingsley (1968), which gives various equivalent descriptions of weak convergence. More powerful methods are available and frequently yield more information. For instance, the method of Prohorov (1956), based on a characterization of the relative compactness of a family of probability distributions, enables one to deduce the existence of (as well as the weak convergence to) the limit distribution in the applications given below. The method of Skorohod (1956), which, roughly speaking, replaces a given sequence of processes having weakly convergent finite-dimensional distributions by a new sequence of processes having the same laws as the old processes but with sample paths converging almost surely at each time point, enables one to apply standard analytic arguments concerning pointwise convergent functions. But both these methods require a fair bit of effort for their development and application. A simple procedure that allows one to get into the thick of things quickly is therefore of some interest. Such a procedure is described and illustrated in what follows.

2. The method. Let (S, d) be a metric space, and let $\mathscr{S}$ be its Borel σ-algebra. Let X and X_k $(k \geqq 1)$ be S-valued random variables (measurable with respect to $\mathscr{S}$), each defined on some probability space. One says that the X_k's converge in distribution to X, and writes $X = d\lim_k X_k$, if $Ef(X_k) \to Ef(X)$ for each continuous bounded real-valued function f on S. The following result states that $X = d\lim_k X_k$ provided one has convergence in distribution for sufficiently small perturbations of the original variables.

PROPOSITION 1. *Let S, d, X, and X_k be as described above. For each $n \geqq 1$, let $A_n : S \to S$ be $\mathscr{S}$-measurable. Suppose that*

$$d\lim_k A_n X_k = A_n X \qquad \textit{for each } n \tag{1}$$

$$\operatorname{plim}_n \lim_k d(X_k, A_n X_k) = 0 \tag{2}$$

$$\operatorname{plim}_n d(X, A_n X) = 0. \tag{3}$$

Then $X = d\lim_k X_k$.

Received December 18, 1970.

[1] This paper was prepared with the partial support of NSF Grant GP 16071.

Conditions (2) and (3) make precise the sense in which $A_n x$ should approximate $x (x \in S)$; written out, the second condition says that $d(X_k, A_n X_k)$ converges to zero in probability as first $k \to \infty$ and then $n \to \infty$, i.e.

$$\lim_n \limsup_k P(\{d(X_k, A_n X_k) \geqq \varepsilon\}) = 0$$

for all $\varepsilon > 0$.

PROOF OF THE PROPOSITION. By Theorem 2.1 of Billingeley (1968), it suffices to show that

$$\lim_k Ef(X_k) = Ef(X) \tag{4}$$

for each uniformly-continuous bounded real-valued function f on S. For such f, one has, for each $\varepsilon > 0$,

$$\begin{aligned} &|Ef(X_k) - Ef(X)| \\ &\quad \leqq |E(f(X_k) - f(A_n X_k))| + |Ef(A_n X_k) - Ef(A_n X)| + |Ef(A_n X) - f(X))| \\ &\quad \leqq [2\|f\| P(\{d(X_k, A_n X_k) \geqq \varepsilon\}) + \delta_f(\varepsilon)] + |Ef(A_n X_k) - Ef(A_n X)| \\ &\qquad + [2\|f\| P(\{d(X, A_n X) \geqq \varepsilon\}) + \delta_f(\varepsilon)], \end{aligned}$$

where $\|f\| = \sup_{s \in S} |f(s)|$ and $\delta_f(\varepsilon) = \sup\{|f(t) - f(s)| : d(s, t) < \varepsilon\}$. Combining this with (1), (2), and (3) gives (4). □

From the proof, it is clear that one can replace sequences by nets throughout the above discussion.

3. Applications. We shall show how the method can be applied to establish three of the most useful general weak convergence theorems. Throughout X and the X_k's are S-valued random variables.

(a) $S = C[0, 1]$. Let S be the space of continuous real-valued functions on the unit interval, and let d be the usual sup-norm metric. Let $\omega_\delta : S \to [0, \infty)$ be the modulus of uniform continuity functional:

$$\omega_\delta(x) = \sup\{|x(t) - x(s)| : |t - s| < \delta\}$$

$(x \in S)$. Proposition 1 gives

COROLLARY 1. *If*

$$d \lim_k (X_k(t))_{t \in T} = (X(t))_{t \in T} \tag{5}$$

for each finite subset T of [0, 1]

$$\operatorname{plim}_{\delta \downarrow 0} \lim_k \omega_\delta(X_k) = 0, \tag{6}$$

then $X = d \lim_k X_k$.

PROOF. Set $A_n = i_n \pi_n$, where $\pi_n : S \to R^{n+1}$ sends x into the vector $(x(m/n))_{0 \leqq m \leqq n}$ and $i_n : R^{n+1} \to S$ sends the vector $(\xi_m)_{0 \leqq m \leqq n}$ into the function which has the value

ξ_m at m/n ($0 \leqq m \leqq n$) and which varies linearly between multiples of $1/n$. Since i_n is continuous, (5) implies (1), and (2) and (3) follow from the inequality

$$d(A_n x, x) \leqq \omega_{1/n}(x),$$

holding for all $x \in S$. □

(b) $S = l_2$. Let S be a separable Hilbert space, with the norm metric. Without loss of generality, we may take $S = l_2 = \{x = (x(n))_{n \geqq 1} : \sum_n x(n)^2 < \infty\}$. Define $J_n : S \to [0, \infty)$ by

$$J_n(x) = \textstyle\sum_{p>n} x(p)^2.$$

Here Proposition 1 gives

COROLLARY 2. *If*

$$d\lim_k (X_k(g))_{g \leqq h} = (X(g))_{g \leqq h} \qquad \textit{for each } h \geqq 1 \tag{7}$$

$$\operatorname{plim}_n \lim_k J_n(X_k) = 0, \tag{8}$$

then $X = d\lim_k X_k$.

PROOF. Let $A_n(x)$ be the sequence whose first n coordinates are those of x, the rest being zeroes. Then (7) implies (1), and (2) and (3) follow from the equality $d^2(x, A_n x) = J_n(x)$. □

(c) $S = D[0, 1]$. Let S be the space of real-valued functions on $[0, 1]$ which are right-continuous with left-limits everywhere, and let d be the Skorohod metric:

$$d(x, y) = \inf\{\max(\|x - y\lambda\|, \|\lambda - I\|) : \lambda \in \Lambda\},$$

where Λ consists of all continuous strictly-increasing maps of $[0, 1]$ onto itself, where $\|\cdot\|$ denotes the usual supremum norm, and where I is the identity function. Define $\omega_\delta' : S \to [0, \infty)$ by

$$\omega_\delta'(x) = \inf_{G \in \mathscr{G}_\delta} w_G(x),$$

where $\mathscr{G}_\delta$ consists of all finite partitions of $[0, 1)$ into left-closed right-open intervals of length exceeding δ, and

$$w_G(x) = \max_{E \in G} \sup_{s,t \in E} |x(t) - x(s)|$$

(confer (14.6) and (14.7) of Billingsley (1968)). Now Proposition 1 gives

COROLLARY 3. *If*

$$d\lim_k (X_k(t))_{t \in T} = (X(t))_{t \in T} \tag{9}$$

for all finite subsets T *of some dense subset* U *of* $[0, 1]$

$$\operatorname{plim}_{\delta \downarrow 0} \lim_k \omega_\delta'(X_k) = 0, \tag{10}$$

then $X = d\lim_k X_k$.

PROOF. The idea of the proof is clearest when $U = [0, 1]$, so we will make this simplifying assumption. Take $A_n = j_n \pi_n$, where π_n is defined as in the proof of (a) and $j_n : R^{n+1} \to S$ sends the vector $(\xi_m)_{0 \leqq m \leqq n}$ into the function whose value at m/n is ξ_m and which is constant on the left-closed right-open intervals separating the multiples of $1/n$. Then (9) and the continuity of j_n imply (1). And (2) and (3) follow (cf (14.8) of Billingsley (1968)) from (10) and the inequality

$$d(A_n x, x) \leqq 1/n + \omega'_{1/n}(x), \tag{11}$$

which we shall now establish.

Fix $x \in S$ and $n \geqq 1$. Let $G \in \mathscr{G}_{1/n}$ have division points

$$0 = t_0 < t_1 < \cdots < t_g = 1.$$

Let λ be that element of Λ which maps t_f onto the smallest multiple of $1/n$ at least as large as $t_f (0 \leqq f \leqq g)$ and which is linear on the intervals separating the t_f's. Let λ^{-1} be the inverse of λ; it is easy to check that $\lambda^{-1} \in \Lambda$ with

$$\|I - \lambda^{-1}\| \leqq 1/n$$

and

$$\|A_n x - x\lambda^{-1}\| \leqq w_G(x).$$

Consequently (11) holds. □

Sometimes in applying the method it is convenient to first embed S in a larger space. Suppose, for example, that we wish to generalize application (a) to the case in which $S = C(K, M)$ is the space of all continuous functions on some compact metric space K with values in an arbitrary metric space M. Embed S isometrically in F, the space of all bounded functions from K to M; in both spaces the metric is given by $d(x, y) = \sup_{k \in K} d_M(x(k), y(k))$. To construct A_n, partition K into finitely many subsets E_j of diameter not exceeding $1/n$, choose points $t_j \in E_j$ for each j, and set $A_n(x) = \sum_j x(t_j) I_{E_j}$. One has $d(A_n x, x) \leqq \omega_{1/n}(x)$, where now $\omega_\delta(x)$ is defined to be $\sup \{d_M(x(k), x(l)) : d_K(k, l) < \delta\}$. Under conditions corresponding to (5) and (6), we get the X_k's converging in distribution to X as F-valued variables, and therefore also as S-valued variables (cf Lemma 3, page 39 of Billingsley (1968)).

In all the above examples, the conditions stated are also necessary; this follows easily from the continuity of the functionals involved. The real difficulty lies in establishing that the necessary and sufficient conditions hold for specific processes, and in getting the rate of convergence.

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

PROHOROV, YU. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theor. Probability Appl.* **1** 157–214.

SKOROHOD, A. V. (1956). Limit theorems for stochastic processes. *Theor. Probability Appl.* **1** 261–290.