*The Annals of Mathematical Statistics*
1970, Vol. 41, No. 1, 284–291

# ON THE CONSTRUCTION OF ALMOST UNIFORMLY CONVERGENT RANDOM VARIABLES WITH GIVEN WEAKLY CONVERGENT IMAGE LAWS[1]

BY MICHAEL J. WICHURA

*University of Chicago*

**1. Introduction.** Let $S$ be an arbitrary metric space, with distance function $d$, and let $\mathscr{S}$ be its Borel $\sigma$-algebra. Denote by $\mathscr{P}(S)$ the class of all probability distributions on $(S, \mathscr{S})$. A net $(P_\gamma)_{\gamma\in\Gamma}$ of probabilities $P_\gamma \in \mathscr{P}(S)$ is said to converge weakly to a probability $P \in \mathscr{P}(S)$ if $P(f) = \lim_\gamma P_\gamma(f)$ for each real-valued bounded continuous function $f$ on $S$; here $P(f) = \int f\,dP$, $P_\gamma(f) = \int f\,dP_\gamma$. Let $\mathscr{P}_s(S)$ denote the subclass of $\mathscr{P}(S)$ consisting of those probabilities $P$ for which there exists a separable subset of $S$ in $\mathscr{S}$ of $P$-probability one. $\mathscr{P}_s(S)$ includes the so-called tight probabilities i.e. probabilities $P$ such that $\sup\{P(K): K \text{ compact}\} = 1$ ([5] page 29). The chief result of this paper is stated in the following.

THEOREM 1. *Let $(S, d)$ be a metric space and let $(P_\gamma)_{\gamma\in\Gamma}$ be a net of probabilities $P_\gamma \in \mathscr{P}(S)$ converging weakly to a probability $P \in \mathscr{P}_s(S)$. Then there exists a probability space $(\Omega, \mathscr{B}, \mu)$ and $\mathscr{B}$–$\mathscr{S}$ measurable, $S$-valued functions $X$ and $X_\gamma(\gamma \in \Gamma)$ defined on $\Omega$ such that the distributions $\mu X^{-1}$ of $X$ and $\mu X_\gamma^{-1}$ of $X_\gamma$ are respectively $P$ and $P_\gamma(\gamma \in \Gamma)$ and such that $X_\gamma$ converges to $X$ almost uniformly.*

One sometimes ([1], [8]) has occasion to consider the weak convergence of probability distributions $P_\gamma$ which are defined only on certain sub-$\sigma$-algebras of $\mathscr{S}$, and it is therefore of interest to know that the requirement in Theorem 1 that the $P_\gamma$ belong to $\mathscr{P}(S)$ can be weakened. To make this precise, let us say that a net $(P_\gamma)_{\gamma\in\Gamma}$ of probabilities $P_\gamma$ defined on sub-$\sigma$-algebras $\mathscr{A}_\gamma$ of $\mathscr{S}$ converges weakly to a probability $P \in \mathscr{P}(S)$ if $\lim_\gamma \bar{P}_\gamma(f) = P(f) = \lim_\gamma \underline{P}_\gamma(f)$ for each real-valued bounded continuous function $f$ on $S$; here $\bar{P}_\gamma$ and $\underline{P}_\gamma$ denote respectively the upper and lower probabilities associated with $P_\gamma$:

$$\bar{P}_\gamma(f) = \inf\{P_\gamma(g): f \leqq g,\ P_\gamma(g) \text{ defined}\}$$

$$\underline{P}_\gamma(f) = \sup\{P_\gamma(g): f \geqq g,\ P_\gamma(g) \text{ defined}\}$$

(for equivalent formulations of this definition see Theorem 1 of [8]). It is clear that this definition of weak convergence reduces to the usual one if all the $\mathscr{A}_\gamma$ equal $\mathscr{S}$. Let $\mathscr{S}_0$ denote the sub-$\sigma$-algebra of $\mathscr{S}$ generated by the open balls of $S$. We then have the following extension of Theorem 1:

THEOREM 2. *Let $S, \mathscr{S}$, and $\mathscr{S}_0$ be defined as above and let $(P_\gamma)_{\gamma\in\Gamma}$ be a net of probabilities $P_\gamma$, defined on $\sigma$-algebras $\mathscr{A}_\gamma$ containing $\mathscr{S}_0$ and contained in $\mathscr{S}$, which*

Received October 23, 1968; revised August 7, 1969.

[1] This research was supported in part by Research Grant No. NSF-GP-8026 from the Division of Mathematical, Physical and Engineering Sciences of the National Science Foundation, and, in part from the Statistics Branch, Office of Naval Research.

*converges weakly to a probability $P \in \mathscr{P}_s(S)$. Then there exists a probability space $(\Omega, \mathscr{B}, \mu)$ and S-valued variables $X$ and $X_\gamma(\gamma \in \Gamma)$ defined on $\Omega$ such that*:

(1) $X$ *is* $\mathscr{B}-\mathscr{S}$ *measurable*, $X_\gamma$ *is* $\mathscr{B}-\mathscr{A}_\gamma$ *measurable* $(\gamma \in \Gamma)$

(2) $$\mu X^{-1} = P, \qquad \mu X_\gamma^{-1} = P_\gamma(\gamma \in \Gamma)$$

(3) $X_\gamma \to X$ *almost uniformly.*

We remark that it is consistent with all the usual axioms of set theory to assume that $\mathscr{P}_s(S) = \mathscr{P}(S)$ (see [2] page 252). In this sense, the requirement in Theorems 1 and 2 that $P \in \mathscr{P}_s(S)$ can be replaced by the trivial one that $P \in \mathscr{P}(S)$.

In the construction used to validate Theorems 1 and 2, $\Omega$ is the product space $S \times \prod_{\gamma\in\Gamma} S_\gamma$, where each $S_\gamma$ is a copy of $S$, $\mathscr{B}$ is a $\sigma$-algebra which contains the product $\sigma$-algebra $\mathscr{A} = \mathscr{S} \times \prod_{\gamma\in\Gamma}\mathscr{A}_\gamma$, $\mu$ is the prolongation to $(\Omega, \mathscr{B})$ of a mixture of product probabilities on $(\Omega, \mathscr{A})$, and $X$ and $X_\gamma$ are the canonical projections of $\Omega$ onto $S$ and $S_\gamma(\gamma \in \Gamma)$. When $\Gamma$ is countable and $S$ separable, one has $\mathscr{B} = \mathscr{A}$. Other constructions have been used to validate special cases of Theorem 1. Working with sequences (for which almost uniform convergence is equivalent to almost sure convergence by Egoroff's theorem) instead of nets, Skorokhod ([7], Theorem 3.1.1) has proved Theorem 1 for $S$ separable and complete; in his construction $\Omega = [0, 1]$, the unit interval, $\mathscr{B}$ is the $\sigma$-algebra of its Borel sets, and $\mu$ is Lebesgue measure. Again working with sequences, Dudley ([3], Theorem 3) has proved Theorem 1 for $S$ separable; in his construction, $\Omega$ is a countable product of copies of $S \times [0, 1]$, $\mathscr{B}$ is the product $\sigma$-algebra on $\Omega$, and $\mu$ is a mixture of product probabilities on $(\Omega, \mathscr{B})$. For applications and other constructions of almost surely convergent processes which are of interest in the theory of weak convergence, see the survey paper by Pyke [6].

**2. Proof for $\Gamma$ countable and $S$ finite.** The simplicity of our construction is obscured in the general case by several technical considerations; in order to illustrate the general idea we will in this section prove Theorem 1 under the assumptions that $\Gamma$ is countable and $S$ is finite. To this end, let $(S_\gamma, \mathscr{S}_\gamma)$ be a copy of $(S, \mathscr{S})$ for each $\gamma$, and let $(\Omega, \mathscr{B}) = (S \times \prod_{\gamma\in\Gamma} S_\gamma, \mathscr{S} \times \prod_{\gamma\in\Gamma}\mathscr{S}_\gamma)$ be the product of the measurable spaces $(S, \mathscr{S})$ and $(S_\gamma, \mathscr{S}_\gamma)(\gamma\in\Gamma)$. Let the canonical coordinate mappings $X$ and $X_\gamma(\gamma \in \Gamma)$ be defined on $\Omega$ by

(4) $$X((s, (s_\theta)_{\theta\in\Gamma})) = s, \qquad X_\gamma((s, (s_\theta)_{\theta\in\Gamma})) = s_\gamma(\gamma \in \Gamma).$$

The required measurability properties clearly hold.

Let $k : \gamma \to k(\gamma)$ be any function from $\Gamma$ to $\{0, 1, 2, \cdots, \infty\}$ such that

(5) $$\lim_{\gamma\in\Gamma} k(\gamma) = \infty$$

($k(\gamma)$ should be thought of as a measure of the largeness of $\gamma$ and later will be further specified). For $1 \leqq k < \infty$, set

(6) $$U_k = \bigcap_{\gamma : k(\gamma) \geqq k} \{X_\gamma = X\}.$$

Observe that each $U_k \in \mathscr{B}$ since $\Gamma$ is countable and that $X_\gamma \to X$ uniformly over each $U_k$ in view of (5).

Let $Q_\gamma (\gamma \in \Gamma)$ be any family of probabilities on $(S, \mathscr{S})$. It later will be further specified. Letting $\delta_s$ denote the probability giving mass one to the point $s \in S$, let

$$\mu_{j,s} = \delta_s \times \prod_{\gamma \in \Gamma} \mu_{j,s,\gamma} \tag{7}$$

$(1 \leqq j < \infty, s \in S)$ denote the product probability ([4] page 166) on $(\Omega, \mathscr{B})$ whose components are respectively: $\delta_s$, defined on $(S, \mathscr{S})$, and

$$\begin{aligned} \mu_{j,s,\gamma} &= Q_\gamma \quad &\text{if} \quad 0 \leqq k(\gamma) < j, \\ &= \delta_s \quad &\text{if} \quad j \leqq k(\gamma) \leqq \infty, \end{aligned}$$

defined on $(S_\gamma, \mathscr{S}_\gamma)$. Clearly $\mu_{j,s} X^{-1} = \delta_s$, $\mu_{j,s} X_\gamma^{-1} = \mu_{j,s,\gamma}$; moreover, since $\Gamma$ is countable, $\mu_{j,s}(U_k) = 1$ for $j \leqq k$. Next, define probabilities $\mu_j (1 \leqq j < \infty)$ on $(\Omega, \mathscr{B})$ by

$$\mu_j = \sum_{s \in S} P\{s\} \mu_{j,s}. \tag{8}$$

Clearly $\mu_j X^{-1} = P$,

$$\begin{aligned} \mu_j X_\gamma^{-1} &= Q_\gamma \quad &\text{if} \quad 0 \leqq k(\gamma) < j \\ &= P \quad &\text{if} \quad j \leqq k(\gamma) \leqq \infty, \end{aligned}$$

and $\mu_j(U_k) = 1$ if $j \leqq k$.

Finally, let $(w_k)_{1 \leqq k < \infty}$ be any sequence of numbers $w_k$ satisfying

$$w_k \geqq 0, \qquad \sum_k w_k = 1, \qquad \sum_{j \leqq k} w_k < 1 (1 \leqq k < \infty) \tag{9}$$

and put

$$\omega_k = \sum_{1 \leqq j \leqq k} w_j (0 \leqq k \leqq \infty). \tag{10}$$

Note $\omega_0 = 0$, $\omega_\infty = 1$. Define the probability $\mu$ on $(\Omega, \mathscr{B})$ by

$$\mu = \sum_j w_j \mu_j. \tag{11}$$

Clearly $\mu X^{-1} = P$

$$\mu X_\gamma^{-1} = \omega_{k(\gamma)} P + (1 - \omega_{k(\gamma)}) Q_\gamma \tag{12}$$

and

$$\mu(U_k) \geqq \omega_k (1 \leqq k < \infty). \tag{13}$$

Since $\lim_{k \to \infty} \omega_k = 1$, (13) implies that $X_\gamma \to X$ almost uniformly with respect to $\mu$. To complete the proof in this special setting it suffices, in view of (12), to show that the weak convergence of $P_\gamma$ to $P$ implies the existence of $k(\gamma)$'s satisfying (5) and probabilities $Q_\gamma$ satisfying

$$P_\gamma = \omega_{k(\gamma)} P + (1 - \omega_{k(\gamma)}) Q_\gamma \tag{14}$$

for all $\gamma \in \Gamma$. Now if $k(\gamma) = \infty$, there exists a $Q_\gamma$ satisfying (14) if and only if

$$P_\gamma = P, \tag{15}$$

and then any $Q_\gamma$ will do. On the other hand, if $0 \leqq k(\gamma) < \infty$, we see, after setting

$$q_{k,s,\gamma} = P_\gamma\{s\} + (\omega_k/(1-\omega_k))(P_\gamma\{s\} - P\{s\}) \tag{16}$$

$$m_{k,\gamma} = \min_{s \in S} q_{k,s,\gamma}, \tag{17}$$

that there exists a probability $Q_\gamma$ satisfying (14) if any only if $m_{k(\gamma),\gamma} \geqq 0$ and $\sum_{s \in S} q_{k(\gamma),s,\gamma} = 1$, and then one must take

$$Q_\gamma = \sum_{s \in S} q_{k(\gamma),s,\gamma}\, \delta_s. \tag{18}$$

We note that $\sum_{s \in S} q_{k,s,\gamma} = 1$ for all $k (0 \leqq k < \infty)$ and that $m_{0,\gamma} \geqq 0$. Thus it suffices to show that (5) is satisfied and (15) holds for $k(\gamma) = \infty$ if we put

$$k(\gamma) = \sup\{j \geqq 0 \colon m_{j,\gamma} \geqq 0\}. \tag{19}$$

Now since $P_\gamma \to P$, we have

$$P_\gamma\{s\} \to P\{s\} \tag{20}$$

for each $s \in S (I_{\{S\}}$ being a continuous bounded function on the discrete space $S$). Hence $\lim_{\gamma \in \Gamma} q_{k,s,\gamma} = P\{s\}$ for each $s \in S$, $1 \leqq k < \infty$; this, together with the fact that $q_{k,s,\gamma} \geqq 0$ if $P\{s\} = 0$, implies that $q_{k,s,\gamma}$ is ultimately nonnegative for each $k$ $(1 \leqq k < \infty)$. Thus since $S$ is finite, there exists for each $k$ an index $\gamma_k \in \Gamma$ such that $m_{k,\gamma} \geqq 0$ for all $\gamma \geqq \gamma(k)$. Since $m_{k,\gamma} \geqq 0$ implies $k(\gamma) \geqq k$, (5) is satisfied. Next, if $k(\gamma) = \infty$, we have (recall $\sum_{s \in S} q_{k,s,\gamma} = 1$)

$$0 \leqq P_\gamma\{s\} + (\omega_k/(1-\omega_k))(P_\gamma\{s\} - P\{s\}) \leqq 1 \tag{21}$$

for each $s \in S$ and arbitrarily large $k$; since $\omega_k/(1-\omega_k) \to \infty$, it follows that $P_\gamma\{s\} = P\{s\}$ for each $s \in S$, i.e., that (15) holds. This completes the proof of Theorem 1 for $\Gamma$ countable and $S$ finite.

**3. Proof of Theorem 2 in the general case.** Let $P$, $P_\gamma(\gamma \in \Gamma)$, $\mathscr{S}$, $\mathscr{S}_0$, and $\mathscr{A}_\gamma(\gamma \in \Gamma)$ be as in Theorem 2. Let $\mathscr{C}(P) = \{C \in \mathscr{S} \colon P(\text{boundary of } C) = 0\}$ be the class of $P$-continuity sets. We recall ([5] page 50) that $\mathscr{C}(P)$ is an algebra and that for each $s \in S$, the open ball

$$\{t \colon d(t,s) < r\} \in \mathscr{C}(P) \tag{22}$$

for all but at most countably many values of $r$. The following lemma shows that the analogue of (20) holds for sets $C \in \mathscr{C}(P)$ (confer $T1.1$ of [8]):

LEMMA 1. *In the present context, $C \in \mathscr{C}(P)$ implies*

$$\lim_{\gamma \in \Gamma} \underline{P}_\gamma(C) = P(C) = \lim_{\gamma \in \Gamma} \bar{P}_\gamma(C).$$

PROOF. Let $F$ be a closed subset of $S$. Since the continuous bounded functions $f_n \colon s \to \max((1-nd(s,F)), 0)$ decrease to the indicator function of $F$, the weak convergence of $P_\gamma$ to $P$ implies that $\limsup_\gamma \bar{P}_\gamma(F) \leqq \limsup_\gamma \bar{P}_\gamma(f_n) = P(f_n) \downarrow P(F)$. The dual relation for open sets is seen to hold by taking complements; thus for any $C \in \mathscr{S}$ we have

$$P(\dot{C}) \leqq \liminf_\gamma \underline{P}_\gamma(C) \leqq \limsup_\gamma \bar{P}_\gamma(C) \leqq P(\bar{C}), \tag{23}$$

where $\dot{C}$ (resp. $\bar{C}$) denotes the interior (resp. closure) of $C$. When $C \in \mathscr{C}(P)$, the extreme members of (23) are equal. $\square$

We shall need a sequence of "finite approximations" to $S$. For this, choose and fix any two numerical sequences $(\Delta_k)_{1 \leqq k < \infty}$ and $(\varepsilon_k)_{1 \leqq k < \infty}$ such that

$$\Delta_k > 0, \qquad \lim_{k \to \infty} \Delta_k = 0 \tag{24}$$

$$\varepsilon_k > 0, \qquad \textstyle\sum_k \varepsilon_k < \infty. \tag{25}$$

Letting $d(C) = \sup\{d(y, z): y, z \in C\}$ denote the diameter of a subset $C$ of $S$, we then have

LEMMA 2. *In the present context, there exist positive integers* $n_k (1 \leqq k < \infty)$ *and disjoint subsets* $C_{m_1, \cdots, m_k} (0 \leqq m_j \leqq n_j, 1 \leqq j \leqq k)$ *of* $S$ *such that*

$$C_{m_1, \cdots, m_{k-1}} = \textstyle\sum_{0 \leqq m_k \leqq n_k} C_{m_1, \cdots, m_{k-1}, m_k} \tag{26}$$

$$\max_{0 \leqq m_j \leqq n_j, 1 \leqq j < k} \max_{1 \leqq m_k \leqq n_k} d(C_{m_1, \cdots, m_k}) \leqq \Delta_k \tag{27}$$

$$\textstyle\sum_{0 \leqq m_j \leqq n_j, 1 \leqq j < k} P(C_{m_1} \cdots, {}_{m_{k-1}, 0}) \leqq \varepsilon_k \tag{28}$$

$$C_{m_1, \cdots, m_k} \in \mathscr{C}(P) \cap \mathscr{S}_0 (0 \leqq m_j \leqq n_j, 1 \leqq j \leqq k). \tag{29}$$

PROOF. Let $E$ be a separable subset of $S$ such that $P(E) = 1$ and let $\{s_n, n \geqq 1\}$ be a countable dense subset of $E$. In view of (22), there exists for each $n \geqq 1$ an open ball in $S$, call it $E_n$, centered at $s_n$ with radius greater than $\frac{1}{2}\Delta_1$ but less than $\Delta_1$, such that $E_n \in \mathscr{C}(P)$. Since the union of these balls covers $E$ and hence has $P$-probability one, there exists a positive integer $n_1$ such that $P(\cup_{n \leqq n_1} E_n) \geqq 1 - \varepsilon_1$. Setting $C_{m_1} = E_{m_1} - \sum_{1 \leqq m < m_1} C_m (1 \leqq m_1 \leqq n_1)$, $C_0 = S - \cup_{n \leqq n_1} E_n = S - \sum_{1 \leqq m_1 \leqq n_1} C_{m_1}$, we get $S = \sum_{0 \leqq m_1 \leqq n_1} C_{m_1}$, $\max_{1 \leqq m_1 \leqq n_1} d(C_{m_1}) \leqq \Delta_1$, $P(C_0) \leqq \varepsilon_1$, $C_0, C_1, \cdots, C_{n_1} \in \mathscr{C}(P) \cap \mathscr{S}_0$. The proof is completed by induction on $k$. $\square$

Let $\prod_k (1 \leqq k < \infty)$ be the finite partition of $S$ whose members are the $C_{m_1, \cdots, m_k}$, and put $\prod_0 = \{S\}$. Choose and fix numbers $w_k$ satisfying (9) and define $\omega_k$ by (10). For $0 \leqq k < \infty$, $C \in \prod_k$, and $\gamma \in \Gamma$, set (confer (16), (17), and (19))

$$\begin{aligned} q_{k,C,\gamma} &= P_\gamma(C) + (P_\gamma(C) - P(C))(\omega_k/(1-\omega_k)) \\ m_{k,\gamma} &= \min\nolimits_{E \in \Pi_k} q_{k,E,\gamma} \\ k(\gamma) &= \sup\{j \geqq 0 : m_{j,\gamma} \geqq 0\}. \end{aligned} \tag{30}$$

In view of (29) and Lemma 1, the convergence of $P_\gamma$ to $P$ implies (see the argument following (20)) that

$$\lim\nolimits_{\gamma \in \Gamma} k(\gamma) = \infty. \tag{31}$$

For $\gamma$ such that $0 \leqq k(\gamma) < \infty$, put (confer (18))

$$Q_\gamma = \textstyle\sum_{C \in \Pi_{k(\gamma)}} q_{k(\gamma), C, \gamma} P_\gamma(\cdot \mid C), \tag{32}$$

where $P_\gamma(\cdot \mid C)$ denotes the probability on $(S, \mathscr{A}_\gamma)$ obtained from $P_\gamma$ by conditioning on the occurrence of the event $C$. It is easy to see that $Q_\gamma$ is itself a probability on $(S_\gamma, \mathscr{A}_\gamma)$ and that (confer (14))

$$\omega_{k(\gamma)} \left(\textstyle\sum_{C \in \Pi_{k(\gamma)}} P_\gamma(\cdot \mid C) P(C)\right) + (1 - \omega_{k(\gamma)}) Q_\gamma = P_\gamma. \tag{33}$$

Now for each $\gamma \in \Gamma$, let $S_\gamma$ be a copy of $S$, and let $(\Omega, \mathscr{A}) = (S \times \prod_{\gamma \in \Gamma} S_\gamma, \mathscr{S} \times \prod_{\gamma \in \Gamma} \mathscr{A}_\gamma)$ be the product of the measurable spaces $(S, \mathscr{S})$ and $(S_\gamma, \mathscr{A}_\gamma)(\gamma \in \Gamma)$. Let $C_{k,s}$ denote the element of $\prod_k$ containing $s \in S$, and let (confer (7))

$$\nu_{j,s} = \delta_s \times \prod_\gamma \nu_{j,s,\gamma} \tag{34}$$

be the product probability on $(\Omega, \mathscr{A})$ whose components are respectively: $\delta_s$, defined on $(S, \mathscr{S})$, and

$$\begin{aligned} \nu_{j,s,\gamma} &= Q_\gamma && \text{if} \quad 0 \leqq k(\gamma) < j, \\ &= P_\gamma(\cdot \mid C_{k(\gamma),s}) && \text{if} \quad j \leqq k(\gamma) < \infty, \\ &= \delta_s && \text{if} \quad k(\gamma) = \infty, \end{aligned}$$

defined on $(S_\gamma, \mathscr{A}_\gamma)$. For each $j$, the mapping $s \to \nu_{j,s}(A)$ is a random variable on $(S, \mathscr{S})$ whenever $A \in \mathscr{A}$ is a cylinder set with a finite-dimensional base (since for each $\gamma \in \Gamma$, each of the finitely many $C$ in $\prod_{k(\gamma)}$ belongs to $\mathscr{S}$), and hence ([4] page 74) this mapping is a random variable for each $A \in \mathscr{A}$. Thus ([4] page 76) we may define a probability $\nu_j$ on $(\Omega, \mathscr{A})$ by the formula (confer (8)) $\nu_j = \int_S \nu_{j,s} P(ds)$. Finally (confer (11)), define the probability $\nu$ on $(\Omega, \mathscr{A})$ by

$$\nu = \sum_{1 \leqq j < \infty} w_j \nu_j. \tag{35}$$

Once again, let the coordinate mappings $X$ and $X_\gamma (\gamma \in \Gamma)$ be defined on $\Omega$ by (4). We have

LEMMA 3. *In the present context,*

$$X \text{ is } \mathscr{A} - \mathscr{S} \text{ measurable}, X_\gamma \text{ is } \mathscr{A} - \mathscr{A}_\gamma \text{ measurable} (\gamma \in \Gamma) \tag{36}$$

$$\nu X^{-1} = P \tag{37}$$

$$\nu X_\gamma^{-1} = P_\gamma (\gamma \in \Gamma). \tag{38}$$

PROOF. Relations (36) and (37) follow directly from the definitions. For (38), observe that

$$\begin{aligned} \nu X_\gamma^{-1} &= \omega_{k(\gamma)} \left( \textstyle\sum_{C \in \Pi_{k(\gamma)}} P_\gamma(\cdot \mid C) P(C) \right) + (1 - \omega_{k(\gamma)}) Q_\gamma && \text{if} \quad 0 \leqq k(\gamma) < \infty, \\ &= P \text{ restricted to } \mathscr{A}_\gamma && \text{if} \quad k(\gamma) = \infty. \end{aligned}$$

In view of (33), (38) holds when $0 \leqq k(\gamma) < \infty$. It remains to show that (38) holds when $k(\gamma) = \infty$; the argument here is similar to, but more complicated than, that at (21). Put

$$D_k = \textstyle\sum_{0 \leqq m_j \leqq n_j, 1 \leqq j < k} \sum_{1 \leqq m_k \leqq n_k} C_{m_1, \cdots, m_k} (k \geqq 1); \qquad D = \liminf_k D_k \tag{39}$$

and observe that (28) and (25) imply

$$P(D) = 1. \tag{40}$$

Let $\mathscr{C}_k$ be the sub-algebra of $\mathscr{S}_0$ made up of sums of members of $\prod_k$ and put $\mathscr{C} = \bigcup_{k \geqq 1} \mathscr{C}_k$; since (in view of (26))

$$\mathscr{C}_1 \subset \mathscr{C}_2 \subset \cdots \subset \mathscr{C}_k \subset \mathscr{C}_{k+1} \subset \cdots \tag{41}$$

$\mathscr{C}$ itself is a sub-algebra of $\mathscr{S}_0$. Let $\sigma\langle\mathscr{C}\rangle\cap D$ (resp. $\mathscr{A}_\gamma\cap D, \mathscr{S}\cap D$) be the trace on $D$ ([4] page 19) of $\sigma\langle\mathscr{C}\rangle$ (resp. $\mathscr{A}_\gamma$, $\mathscr{S}$), and let $\mathscr{S}_D$ denote the Borel $\sigma$-algebra of $D$. In view of (24), (27), and (39), each open subset of $D$ is a union, necessarily countable, of sets of the form $C\cap D$ with $C\in\mathscr{C}$; it follows that $\mathscr{S}_D\subset\sigma\langle\mathscr{C}\cap D\rangle = \sigma\langle C\rangle\cap D$. Since $D_k$ belongs to $\mathscr{C}_k$, we have $D\in\sigma\langle\mathscr{C}\rangle$, and since ([5] page 5) $\mathscr{S}_D = \mathscr{S}\cap D$, we have

$$D\in\mathscr{A}_\gamma\cap D\subset\mathscr{S}\cap D = \mathscr{S}_D\subset\sigma\langle\mathscr{C}\rangle. \tag{42}$$

In view of (41) and the additivity of $P$ and $P_\gamma$, the condition $k(\gamma) = \infty$ implies (see (30) and (21)) that for each $C\in\mathscr{C}$ the inequalities $0\leqq P_\gamma(C)+(P_\gamma(C)-P(C))(\omega_k/(1-\omega_k))\leqq 1$ hold for arbitrarily large values of $k$; since $\lim_{k\to\infty}\omega_k/(1-\omega_k) = \infty$, it follows that $P_\gamma$ and $P$ coincide over $\mathscr{C}$, hence over $\sigma\langle\mathscr{C}\rangle$, and hence, in view of (42), over $\mathscr{A}_\gamma\cap D$. But then, in view of (40), we have $P_\gamma(D) = P(D) = 1$, so that $P$ and $P_\gamma$ coincide over $\mathscr{A}_\gamma$. This completes the proof of the lemma. $\square$

Now put $\Delta_\infty = 0$ and set (confer (6) and (24))

$$U_k = \textstyle\bigcap_{\gamma:k(\gamma)\geqq k}\{d(X_\gamma, X)\leqq\Delta_{k(\gamma)}\}. \tag{43}$$

The $U_k$ need not belong to $\mathscr{A}$ in general, although they will if $\Gamma$ is countable and $S$ is separable (so that $d(X_\gamma, X)$ is $\mathscr{A}$-measurable (confer [5] page 6)). For any subset $\Omega_\circ$ of $\Omega$, let $\nu^*(\Omega_\circ) = \inf\{\nu(A):\Omega_\circ\subset A\in\mathscr{A}\}$ denote the outer probability of $\Omega_\circ$ under $\nu$.

LEMMA 4. *In the present context,*

$$X_\gamma\to X \quad \textit{uniformly over each} \quad U_k \tag{44}$$

$$\lim_{k\to\infty}\nu^*(U_k) = 1. \tag{45}$$

PROOF. We get (44) from (24), (31), and (43). For (45) put $E_k = \inf_{m\geqq k}D_m(1\leqq k<\infty)$, where $D_m$ is defined by (39). Suppose that $U_k\subset A\in\mathscr{A}$. Then there exists ([4] page 81) a countable subset $\Gamma_A$ of $\Gamma$ such that $A$ depends only on $X$ and the $X_\gamma$ with $\gamma\in\Gamma_A$; it follows that

$$\textstyle\bigcap_{\gamma\in\Gamma_A;k(\gamma)\geqq k}\{d(X_\gamma, X)\leqq\Delta_{k(\gamma)}\}\subset A$$

(the set on the left need not belong to $\mathscr{A}$). Thus for $j\leqq k$ we have (confer (34) and (27))

$$\begin{aligned}\nu_j(A) &= \textstyle\int_S\nu_{j,s}(A\cap\{X = s\})P(ds)\\ &\geqq \textstyle\int_{E_k}\nu_{j,s}(\bigcap_{\gamma\in\Gamma_A;k(\gamma)\geqq k}\{d(X_\gamma, s)\leqq\Delta_{k(\gamma)}\})P(ds)\\ &= P(E_k).\end{aligned}$$

By (35), $\nu(A)\geqq\sum_{j\leqq k}w_j\nu_j(A)\geqq\omega_kP(E_k)$; it follows that $\nu^*(U_k)\geqq\omega_kP(E_k)$. But (28) implies $P(E_k)\geqq 1-\sum_{m\geqq k}\varepsilon_m$; (45) now follows from (9) and (25). $\square$

We note that the $U_k$ increase with $k$. In view of Lemmas 3 and 4, to complete the proof of Theorem 2 it suffices to establish

LEMMA 5. *Let* $(\Omega, \mathscr{A}, \nu)$ *be any probability space and let* $(U_k)_{k \geqq 1}$ *be an increasing sequence of subsets of* $\Omega$ *of outer probabilities* $\nu^*(U_k)$. *Let* $\mathscr{B}$ *be the* $\sigma$*-algebra generated by* $\mathscr{A}$ *and the* $U_k (1 \leqq k < \infty)$. *Then* $\nu$ *may be prolonged to a probability* $\mu$ *on* $(\Omega, \mathscr{B})$ *such that*

$$\mu(U_k) = \nu^*(U_k) \tag{46}$$

*for each* $k$.

PROOF. Put $B_k = U_k - U_{k-1} (1 \leqq k < \infty)$, put $B_\infty = (\sup_k U_k)^c$, and choose $B_k^* \in \mathscr{A}$ such that $\nu(B_k^*) = \nu^*(B_k)(1 \leqq k \leqq \infty)$. According to [4] page 43, $\mathscr{B}$ coincides with the class of sets of the form $\sum_k A_k B_k$, where $A_k \in \mathscr{A}$; moreover the formula

$$\mu(\textstyle\sum_k A_k B_k) = \sum_k \int_{A_k} f_k \, d\nu \tag{47}$$

defines a probability $\mu$ on $(\Omega, \mathscr{B})$, whose restriction to $(\Omega, \mathscr{A})$ is $\nu$, provided that each $f_k$ is a nonnegative, $\mathscr{A}$-measurable random variable vanishing off of $B_k^*(1 \leqq k \leqq \infty)$ and that $\sum_k f_k = 1$.

Let $f_k$ be the indicator function of $B_k^* - \bigcup_{j<k} B_j^* (1 \leqq k \leqq \infty)$ and define $\mu$ by (47). Then

$$\mu(U_k) = \mu(\textstyle\sum_{j \leqq k} B_j) = \sum_{j \leqq k} \int f_j \, d\nu = \nu(\bigcup_{j \leqq k} B_j^*). \tag{48}$$

Since $\bigcup_{j \leqq k} B_j^*$ is an $\mathscr{A}$-measurable set containing $U_k$, we have

$$\nu(\textstyle\bigcup_{j \leqq k} B_j^*) \geqq \nu^*(U_k). \tag{49}$$

On the other hand, suppose $U_k \subset A \in \mathscr{A}$, so that $B_j \subset A$ for $j \leqq k$. Then each $B_j^*$, and hence also $\bigcup_{j \leqq k} B_j^*$, is contained in $A$ up to a $\nu$-equivalence. It follows that $\nu(A) \geqq \nu(\bigcup_{j \leqq k} B_j^*)$ and that

$$\nu^*(U_k) \geqq \nu(\textstyle\bigcup_{j \leqq k} B_j^*). \tag{50}$$

Together (48), (49), and (50) imply (46). □

## REFERENCES

[1] DUDLEY, R. M. (1966a). Weak convergence of probabilities on non-separable metric spaces and empirical measures on Euclidean spaces. *Ill. J. Math.* **10** 109–126.

[2] DUDLEY, R. M. (1966b). Convergence of Baire measures. *Studia Math.* **27** 251–268.

[3] DUDLEY, R. M. (1967). Distances of probability measures and random variables. *Ann. Math. Statist.* **39** 1963–1572.

[4] NEVEU, J. (1965). *Mathematical Foundations of the Calculus of Probability.* Holden-Day, Inc., San Francisco.

[5] PARATHASARATHY, D. R. (1967). *Probability Measures on Metric Spaces.* Academic Press, New York.

[6] PYKE, R. (1968). Applications of almost surely convergent constructions of weakly convergent processes. *Proc. Internat. Symp. Prob. Inform. Theor.* Springer-Verlag, Berlin.

[7] SKOROKHOD, A. V. (1956). Limit theorems for stochastic processes (*translated by SIAM*). *Th. Probability Appl.* **1** 261–290.

[8] WICHURA, MICHAEL J. (1968). On the weak convergence of non-Borel probabilities on a metric space. Unpublished dissertation, Columbia Univ.