Z. Wahrscheinlichkeitstheorie verw. Gebiete
43, 173–181 (1978)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Weak Convergence of an Empirical Process Indexed by the Closed Convex Subsets of I^2

Erwin Bolthausen

Fachgruppe Statistik, Universität Konstanz, D-7750 Konstanz

Summary. Let I^2 be the unit cube of $\mathbb{R}^2$ and X_i be independent I^2-valued random variables that are distributed according to Lebesgue-measure. If S is the set of closed convex subsets of I^2 we consider the process $\{\mu_n(A)\}_{A\in S}$, where $\mu_n(A)=(1/n)\sum^{n} 1_A(X_i)$. It is proved that this process suitably normalized converges in a suitable weak sense to a Gaussian process.

§0. Introduction

Let I^2 be the unit cube in $\mathbb{R}^2$ and X_i, $i\in\mathbb{N}$, be a sequence of independent I^2-valued random variables that are distributed according to Lebesgue-measure λ. Let S be the set of closed convex subsets of I^2. We consider the process $\{\mu_n(A)\}_{A\in S}$ where $\mu_n(A)=\sum_{i=1}^{n} 1_A(X_i)$. It is the purpose of this paper to prove that this process, appropriately normalized, converges in a suitably weak sense to a Gaussian process.

S has two natural metrics: $\rho(A,B)=\lambda(A-B)+\lambda(B-A)$ and the Hausdorff metric $\delta(A,B)=\max(\sup_{x\in A}\inf_{y\in B} d(x,y),\ \sup_{y\in B}\inf_{x\in A} d(x,y))$ (d the Euclidean metric). It is known that (S,δ) is a compact metric space and that δ and ρ generate the same topology on the subset of S which consists of sets with nonvoid interior. There exists a constant $K>0$ such that

(0.1) $\rho(A,B)\leqq K\,\delta(A,B)$ for all $A,B\in S$ (see e.g. [5] (4.2)).

The first problem one encounters is that the paths of μ_n are not continuous on S. They are of course bounded and Borel-measurable functions on S, but μ_n as a mapping from the basic probability space to the nonseparable Banach space of bounded measurable functions on S (endowed with the sup-norm) cannot be Borel-measurable. The tool for solving these problems is provided by the theory

0044-3719/78/0043/$ 01.80

of weak convergence developed by Dudley [3]. In order to apply this theory one has to show that μ_n is measurable as a mapping into the space of bounded measurable functions on S equipped with the σ-algebra generated by the balls of the sup-norm. (This σ-algebra is smaller than the Borel σ-algebra). These measurability considerations are found in §1. After some preliminary work in §2, weak convergence in the sense of Dudley is proved in §3.

There is a paper by A. deHoyos in which such a theorem has been stated in incorrect form because it does not take into account the fact that the paths of μ_n are not continuous ([6], Cor. 3). Beside this the proof contains other defects, the most serious one being that deHoyos seems to make use of the following false statement:

If $\{Z_n^\varepsilon(t)\}_{t\in T}$, $\{Z^\varepsilon(t)\}_{t\in T}$, $n\in\mathbb{N}$, $\varepsilon>0$ are processes with some index set T and if $Z_n^\varepsilon \to Z^\varepsilon$ in some weak sense for each $\varepsilon>0$ and if $Z_n^\varepsilon(t)\to Z_n^0(t)$ and $Z^\varepsilon(t)\to Z^0(t)$ in L_2 as $\varepsilon\to 0$ for each t then $Z_n^0\to Z^0$ weakly. ([6], p. 161).

§1. Notations and Measurability Considerations

Let S, λ, δ, ρ and d be defined as in §0. Let further $\mathscr{B}(S)$ be the set of Borel-subsets of S (with respect to δ), $B(S)$ the set of real-valued bounded $\mathscr{B}(S)$-measurable functions on S and $C(S)$ the continuous functions on S. For $f\in B(S)$ we set $\|f\|=\sup\{|f(A)|\colon A\in S\}$. $(C(S), \|\ \|)$ is a separable and $(B(S), \|\ \|)$ a nonseparable Banach-space. Let $\mathscr{C}$ and $\mathscr{D}$ be the class of Borel-subsets of $C(S)$ and $B(S)$. Further let $\mathscr{S}$ be the σ-algebra on $B(S)$ which is generated by the balls $\{f\colon \|f-g\|<\varepsilon\}_{g\in B(S),\,\varepsilon>0}$.

Let $(\Omega,\mathscr{A},P)$ be a complete probability space and X_i, $i\in\mathbb{N}$, be a sequence of independent I^2-valued, λ-distributed random variables. Let

$$\mu_n(A,\omega)=(1/n)\sum_{i=1}^{n} 1_A(X_i(\omega)) \quad\text{and}\quad Y_n(A,\omega)=\sqrt{n}(\mu_n(A,\omega)-\lambda(A)).$$

We interpret the μ_n (and Y_n) as mappings from Ω to $B(S)$: $\mu_n(\omega)(A)=\mu_n(A,\omega)$. The aim of this section is to prove

Proposition 1. $\mu_n\colon (\Omega,\mathscr{A})\to(B(S).\mathscr{S})$ is measurable for each n.

Throughout this section ε, δ are always rational and >0. If $n\in\mathbb{N}$ let $\mathscr{B}_n^*$ be the class of universally measurable sets in $((I^2)^n, \mathscr{B}((I^2)^n))$, that is the subsets of $(I^2)^n$ which are in the completion of $\mathscr{B}((I^2)^n)$ relative to any finite Borel-measure on $(I^2)^n$.

For $A\in S$ let $A^1=A$ and $A^0=A^c$ (the complement of A)

Lemma 1. Let $\Gamma\in\mathscr{B}(S)$, $n\in\mathbb{N}$, $(i_1,\dots,i_n)\in\{0,1\}^n$. Then

$$\bigcup_{A\in\Gamma} A^{i_1}\times\cdots\times A^{i_n}\in\mathscr{B}_n^*.$$

Proof. Let Δ be a point which is separated from $(I^2)^n$. We define a map

$$\Phi\colon S\times(I^2)^n\to(I^2)^n\cup\{\Delta\}$$

$$\Phi(A,y_1,y_2,\dots,y_n)=\begin{cases}(y_1,\dots,y_n) & \text{if } y_k\in A^{i_k} \text{ for each } k\\ \Delta & \text{else.}\end{cases}$$

We claim that this map is Borel-measurable.

For $\delta(>0$, rational) let $x_1^\delta, \ldots, x_{m(\delta)}^\delta \in I^2$ be such that $\bigcup_{j=1}^{m(\delta)} U_\delta(x_j^\delta) = I^2$ where $U_\delta(x) = \{y \in I^2 : d(x,y) < \delta\}$. We will drop the upper index δ in x_i^δ when no confusion can arise.

Let $V_1, \ldots, V_n$ be open in I^2 with $\Delta \notin V_1 \times \cdots \times V_n$. If we show that

(1.1) $\quad \Phi^{-1}(V_1 \times \cdots \times V_n) \in \mathscr{B}(S) \otimes \mathscr{B}((I^2)^n)$

measurability will follow.

For δ and $1 \leqq j_1, \ldots, j_n \leqq m(\delta)$ let

$$A^{j_1, \ldots, j_n} = \{A \in S : A \cap U_\delta(x_{j_k}) \neq \emptyset \text{ for those } k \text{ with } i_k = 1,\ A \cap U_\delta(x_{j_k}) = \emptyset \text{ for those } k \text{ with } i_k = 0\}$$

$$C^{j_k} = \begin{cases} U_\delta(x_{j_k}) & \text{if } U_\delta(x_{j_k}) \subset V_k \\ \emptyset & \text{if not.} \end{cases}$$

It is easy to see that $A^{j_1, \ldots, j_n} \in \mathscr{B}(S)$. Now

(1.2) $\quad \Phi^{-1}(V_1 \times \cdots \times V_n)$

$$= \bigcup_{\varepsilon > 0} \bigcap_{\delta < \varepsilon} \bigcup_{j_1, \ldots, j_n = 1}^{m(\delta)} A^{j_1, \ldots, j_n} \times C^{j_1} \times \cdots \times C^{j_n}.$$

In fact: If $(A, y_1, \ldots, y_n) \in \Phi^{-1}(V_1 \times \cdots \times V_n)$ then $y_k \in A^{i_k} \cap V_k$ for all k. So there exists $\varepsilon > 0$ such that for $\delta < \varepsilon$ and $y_k \in U_\delta(x_{j_k})$, $U_\delta(x_{j_k}) \subset V_k$ for all k and $U_\delta(x_{j_k}) \subset A^{i_k}$ for those k with $i_k = 0$. Further $U_\delta(x_{j_k}) \cap A^{i_k} \neq \emptyset$ for k with $i^k = 1$. So $(A, y_1, \ldots, y_n) \in$right side of (1.2).

Conversely if $(A, y_1, \ldots, y_n) \in$right side of (1.2) then there is an ε such that for $\delta < \varepsilon$ there exists $j_1, \ldots, j_n$ with $y_k \in U_\delta(x_{j_k}) \subset V_k$ for all k, $A \cap U_\delta(x_{j_k}) \neq \emptyset$ if $i_k = 1$ and $A \cap U_\delta(x_{j_k}) = \emptyset$ if $i_k = 0$. By compactness of A we have $y_k \in A$ if $i_k = 1$. So

$$(A, y_1, \ldots, y_n) \in \Phi^{-1}(V_1 \times \cdots \times V_n)$$

So (1.2) and (1.1) are proved.

Now

$$\bigcup_{A \in \Gamma} A^{i_1} \times \cdots \times A^{i_n} = \Phi(\Gamma \times (I^2)^n) \cap (I^2)^n.$$

As $S \times (I^2)^n$ is a complete separable metric space and Φ is measurable, it follows from Theorem 3.4 Ch. I of [7] that $\bigcup_{A \in \Gamma} A^{i_1} \times \cdots \times A^{i_n}$ is analytic, so in $\mathscr{B}_n^*$.

The same proof gives the following corollary which is probably well known but has until now escaped my attention.

Corollary. *Let K be a compact metric space, 2^K the set of closed subsets of K equipped with the Hausdorff metric and further $\mathscr{B}(2^K)$ the class of Borel subsets of 2^K (with respect to the Hausdorff metric). If $\Gamma \in \mathscr{B}(2^K)$ then $\bigcup \Gamma$ is analytic in K.*

Proof of Proposition 1. Let

$\Psi\colon (I^2)^n \to B(S)$ be defined as follows

$$\Psi(y_1, \dots, y_n) = (1/n) \sum_{j=1}^{n} 1.(y_j)$$

Clearly $\mu_n(\omega, A) = \Psi(X_1(\omega), \dots, X_n(\omega))(A)$.

By the completeness of $(\Omega, \mathscr{A}, P)$ it suffices to prove that

$$\Psi\colon ((I^2)^n, \mathscr{B}_n^*) \to (B(S), \mathscr{S})$$

is measurable.

Let $g \in B(S)$ and $\sigma > 0$ be fixed. We will show that

$$(1.3) \quad D_{g,\sigma} = \{(y_1, \dots, y_n) \in (I^2)^n\colon \ \|\Psi(y_1, \dots, y_n) - g\| \leqq \sigma\} \in \mathscr{B}_n^*.$$

Let

$$\Gamma_k = \{A \in S\colon |g(A) - k/n| > \sigma\} \in \mathscr{B}(S).$$

As is easy to see

$$(1.4) \quad D_{g,\sigma} = \bigcap_{k=0}^{n} \bigcap_{A \in \Gamma_k} \Big[\bigcup_{\substack{i_1, \dots, i_n \in \{0,1\}^n \\ i_1 + \dots + i_n = k}} A^{i_1} \times \dots \times A^{i_n} \Big]^c.$$

By Lemma 1

$$\bigcup_{A \in \Gamma_k} \bigcup_{\substack{i_1, \dots, i_n \\ i_1 + \dots + i_n = k}} A^{i_1} \times \dots \times A^{i_n} \in \mathscr{B}_n^*.$$

So $D_{g,\sigma}$ which is the intersection of the complements of these $n+1$ sets is in $\mathscr{B}_n^*$. The proposition is proved.

§ 2. Approximating Convex Sets from Above

In order to prove weak convergence of our empirical process, we want to use a technique like that employed by Strassen and Dudley [8] for the case of Lipschitz-continuous independent summands in $C(K)$, where K is a compact metric space. Instead of continuity we use the fact that the summands in $\sum(1_A(X_i) - \lambda(A))$ up to the Lipschitz-continuous $\lambda(A)$ are monotone in A. To make this work, we have to approximate the elements in S from above by classes of not to many sets.

We have to blow up I^2 somewhat. Let $\hat{S}$ be the set of closed convex subsets of $[-1, 2]^2$. The following proposition is taken from Dudley [5] (Theorem 4.1)

Proposition 2. For $\varepsilon > 0$ there exist coverings $\tilde{S} = \bigcup_{j=1}^{N(\varepsilon)} \tilde{S}_j^\varepsilon$ with $\delta(\tilde{S}_j^\varepsilon) < \varepsilon$ and $\log(N(\varepsilon))$ $= o(\varepsilon^{-\beta})$ as $\varepsilon \downarrow 0$ for each $\beta > 1/2$, where

$$\delta(\tilde{S}_j^\varepsilon) = \sup_{A, B \in \tilde{S}_j^\varepsilon} \delta(A, B).$$

Let $B_j^{(m)}$ be elements in $\tilde{S}_j^{2^{-m}}$ for $1 \leqq j \leqq N(2^{-m})$. For $A \in S$ let $A_m = \{x \in \mathbb{R}^2: d(x, A) \leqq 2^{-m}\} \in \tilde{S}$. Then $A_m \in \tilde{S}_j^{2^{-m-1}}$ for some j. Let $\hat{A}_m = I^2 \cap B_j^{(m+1)}$. From the convexity of $A \in S$ and the above construction one has:

(2.1) for each $A \in S$, $m \in \mathbb{N}$ $\hat{A}_m \in S$

(2.2) $\hat{A}_m \supset A$

(2.3) $\delta(\hat{A}_m, A) \leqq 3 \cdot 2^{-m-1}$

Let $\lambda_m = \operatorname{card}\{\hat{A}_m: A \in S\}$. Then

(2.4) $\lambda_m \leqq N(2^{-m-1})$.

Let now $\check{A}_m = \{x \in I^2: d(x, (\hat{A}_m)^c) \geqq 2^{-m+1}\}$. Then one easily obtains

(2.5) $\check{A}_m \in S$

(2.6) $\check{A}_m \subset A$

(2.7) there is a constant $c > 0$ such that for each $m \in \mathbb{N}$, $A \in S$ $\rho(\hat{A}_m, \check{A}_m) \leqq c\, 2^{-m}$.

Note that one cannot have $\delta(\hat{A}_m, \check{A}_m) \leqq c\, 2^{-m}$ in (2.7).

§ 3. Weak Convergence of the Empirical Process

Let (M, d) be a metric space and $\mathscr{M}$ the class of Borel-sets. If $\mathscr{M}'$ is another σ-algebra of sets in M, ν a measure on $(M, \mathscr{M}')$ and F a bounded function on M we set

$$\int^{*} F\, d\nu = \inf\{\textstyle\int f\, d\nu: f \geqq F,\ f\ \mathscr{M}'\text{-measurable}\}$$

$$\int_{*} F\, d\nu = \sup\{\textstyle\int f\, d\nu: f \leqq F,\ f\ \mathscr{M}'\text{-measurable}\}.$$

We take the following definition from Dudley [3]:

If ν is a Borel-measure on M and ν_n is a sequence of measures (not necessarily Borel), we say

$$\nu_n \to \nu \quad (\text{weak}^*)$$

if for every bounded continuous function F on M

$$\lim \int_{*} F\, d\nu_n = \lim \int^{*} F\, d\nu_n = \int F\, d\nu.$$

A set $\mathscr{K}$ of measures on M is said to be weak*-sequentially relatively compact if for any sequence $\{\beta_n\} \subset \mathscr{K}$ there is a subsequence converging weak* to a Borel-measure on M.

Let μ be the Gaussian measure on $(C(S), \mathscr{C})$ such that for $A, B \in S$

$$\int_{C(S)} f(A)\, \mu(df) = 0, \qquad \int_{C(S)} f(A) f(B)\, \mu(df) = \lambda(A \cap B) - \lambda(A)\, \lambda(B).$$

Such a measure exists (see Dudley [4], Theorem 4.3). Clearly μ extends to a Borel-measure on $(B(S), \mathscr{D})$.

Let $\kappa_A: B(S) \to \mathbb{R}$, $\kappa_A(f) = f(A)$ and $\mathscr{S}'$ be the σ-algebra generated by $\mathscr{S}$ and the κ_A's. Further let μ^n be the measures on $(B(S), \mathscr{S}')$ induced by the maps

$$\Omega \to B(S)$$

$$\omega \to Y_n(\cdot, \omega) = \sqrt{n}(\mu_n(\cdot, \omega) - \lambda(\cdot)).$$

By Proposition 1 the μ^n are defined on $(B(S), \mathscr{S}')$.

We can now formulate our main result. An earlier unsuccessful attempt at such a theorem was made by deHoyos [6].

Theorem 1. $\mu^n \to \mu$ *(weak*).*

We quote the following version of the Bernstein inequality

Lemma 2. *Let ξ_i be i.i.d.r.v. with $E(\xi_i) = 0$, $E(\xi_i^2) = \sigma^2$ and $|\xi_i| \leqq 1$ a.s. Then* $P\left(\left|\sum_{i=1}^{n} \xi_i\right| > t\right) \leqq 2\exp(-t^2/(2n\sigma^2 + 2t/3))$.

For a proof see [1].

A straightforward corollary is

Lemma 3. *For any $a > 0$ and $A, B \in S$*

$$P(|Y_n(A) - Y_n(B)| > a) \leqq 2\exp(-n a^2/(2n\rho(A,B) + 2\sqrt{n}\, a/3)).$$

We state the following convention for the rest of the paper: m and n are always related by $m = [\log n/\alpha \log 2]$ where $\alpha < 2$ is fixed and sufficiently close to 2 ('sufficiency' being clear in what follows)

Lemma 4. *For each $\varepsilon > 0$ and $m \in \mathbb{N}$ $\{f \in B(S): \sup_{A \in S} |f(A) - f(\hat{A}_m)| > \varepsilon\} \in \mathscr{S}'$.*

Proof. Let $\hat{B}(S) = \{g \in B(S): g(A) = g(B)$ whenever $\hat{A}_m = \hat{B}_m\}$. Clearly $(\hat{B}(S), \| \ \|)$ is separable. Let Q be a countable dense subset in $\hat{B}(S)$. For $f \in B(S)$ let $\hat{f} \in \hat{B}(S)$ be defined by $\hat{f}(A) = f(\hat{A}_m)$. Then clearly $\sup_{A \in S} |f(A) - f(\hat{A}_m)| > \varepsilon$ if and only if there is a $\varepsilon' > \varepsilon$ such that for each $\delta > 0$ there exists a $g \in Q$ such that $\|f - g\| > \varepsilon' - \delta$ and $\|\hat{f} - g\| < \delta$. So the lemma follows.

Lemma 5. *For each $\varepsilon > 0$ $\lim_{n \to \infty} (\mu^n)^* \{f: \sup_{A \in S} |f(A) - f(\hat{A}_m)| > \varepsilon\} = 0$.*

Proof. From Lemma 4 it follows that

$$(\mu^n)^* \{f: \sup_{A \in S} |f(A) - f(\hat{A}_m)| > \varepsilon\} = P\{\omega: \sup_{A \in S} |Y_n(A, \omega) - Y_n(\hat{A}_m, \omega)| > \varepsilon\}.$$

We therefore have to show that

$$\lim_{n \to \infty} P\{\omega: \sup_{A \in S} |Y_n(A, \omega) - Y_n(\hat{A}_m, \omega)| > \varepsilon\} = 0. \tag{3.1}$$

Now $|\lambda(\hat{A}_m)-\lambda(\check{A}_m)|=O(2^{-m})$, $\sqrt{n}=o(2^m)$, so there is a $n_0(\varepsilon)$ such that for $n\geqq n_0(\varepsilon)$

$$(3.2)\qquad \begin{aligned}&\{\omega: \sqrt{n}\,|(\mu_n(A)-\lambda(A))-(\mu_n(\hat{A}_m)-\lambda(\hat{A}_m))|>\varepsilon\}\\ &\subset\{\omega: \sqrt{n}(\mu_n(\hat{A}_m)-\mu_n(A))>2\varepsilon/3\}\\ &\subset\{\omega: \sqrt{n}(\mu_n(\hat{A}_m)-\mu_n(\check{A}_m))>2\varepsilon/3\}\\ &\subset\{\omega: \sqrt{n}\,|\mu_n(\hat{A}_m-\check{A}_m)-\lambda(\hat{A}_m-\check{A}_m)|>\varepsilon/3\}.\end{aligned}$$

Let $\xi_i=1_{\hat{A}_m-\check{A}_m}(X_i)-\lambda(\hat{A}_m-\check{A}_m)$. Then $E(\xi_i)=0$, $E(\xi_i^2)=O(2^{-m})$ and $|\xi_i|\leqq 1$ a.s. By Lemma 2

$$(3.3)\qquad \begin{aligned}&P\{\omega: \sup_{A\in S}\sqrt{n}\,|\mu_n(\hat{A}_m-\check{A}_m)-\lambda(\hat{A}_m-\check{A}_m)|>\varepsilon/3\}\\ &<2\lambda_m\exp(-n\,\varepsilon^2/(c\,n\,2^{-m}+2\sqrt{n}\,\varepsilon))\qquad \text{for some } c>0\\ &=o(1).\end{aligned}$$

(3.2) and (3.3) prove (3.1).

Proof of Theorem 1. If $A_1,\ldots,A_k\in S$ are fixed, one has by the multivariate central limit theorem

$$(3.4)\qquad (\mu^n\kappa_{A_1}^{-1},\ldots,\mu^n\kappa_{A_k}^{-1})\to(\mu\kappa_{A_1}^{-1},\ldots,\mu\kappa_{A_k}^{-1})\qquad \text{weakly in } \mathbb{R}^k.$$

Let $\varepsilon>0$ be fixed. We shall prove that

$$(3.5)\qquad \lim_{\delta\to 0}\limsup_{n\to\infty}(\mu^n)^*\{f: \sup_{\delta(A,B)<\delta}|f(A)-f(B)|>\varepsilon\}=0.$$

By (0.1) it suffices to prove this relation with $\delta(\cdot,\cdot)$ replaced by $\rho(\cdot,\cdot)$. We have

$$(3.6)\qquad \begin{aligned}&\{f: \sup_{\rho(A,B)<\delta}|f(A)-f(B)|>\varepsilon\}\\ &\subset\{f: \sup_{A\in S}|f(A)-f(\hat{A}_m)|>\varepsilon/3\}\cup\{f: \sup_{\rho(A,B)<\delta}|f(\hat{A}_m)-f(\hat{B}_m)|>\varepsilon/3\}.\end{aligned}$$

Clearly the last event is in $\mathscr{S}'$. It suffices therefore to show that

$$(3.7)\qquad \lim_{\delta\to 0}\limsup_{n\to\infty}P(\omega: \sup_{\rho(A,B)<\delta}|Y_n(\hat{A}_m)-Y_n(\hat{B}_m)|>\varepsilon)=0.$$

Let $\zeta>0$, then (Lemma 3)

$$P(\sup_{A\in S}|Y_n(\hat{A}_j)-Y_n(\hat{A}_{j-1})|>\zeta)\leqq\lambda_j^2\,\rho_{j,n}(\zeta)$$

where

$$\rho_{j,n}(\zeta)=2\exp(-\sqrt{n}\,\zeta^2/(c\sqrt{n}\,2^{-j}+2\zeta/3))\quad (\text{note } \rho(\hat{A}_j,\hat{A}_{j-1})\leqq c'\,2^{-j} \text{ by (2.7)}).$$

Take $\zeta_j=j^{-2}$; then there exists m_0 such that for $m_0\leqq j\leqq m$

$$\rho_{j,n}(\zeta_j)\leqq 2\exp(-2^{(\alpha m/2)}2^{(2j/3)}/(2^{(\alpha m/2)}+2^{(5j/6)}))$$

(recall the relation between m and n: $2^{(\alpha m/2)} \leqq \sqrt{n} \leqq 2^{(\alpha(m+1)/2)}$), whence for $10/6 < \alpha < 2$ $\rho_{j,n}(\zeta_j) \leqq 2\exp(-2^{(2j/3-1)})$. Applying (2.4), for each $\sigma > 0$ there exists $m_1(\sigma) \geqq m_0$ such that for $m_1 \leqq m' \leqq m$ $\sum_{j=m'}^{m} \lambda_j^2 \rho_{j,n}(\zeta_j) < \sigma$, whence

$$\text{(3.8)} \quad \sum_{j=m'}^{m} P\{\sup_{A\in S} |Y_n(\hat{A}_j) - Y_n(\hat{A}_{j-1})| > j^{-2}\} < \sigma.$$

Let now $\varepsilon, \varepsilon' > 0$ be given and choose $m_2 \geqq m_1(\varepsilon'/3)$ such that

$$\text{(3.9)} \quad \sum_{j=m_2}^{\infty} j^{-2} < \varepsilon/3.$$

If $\rho(A,B) < \delta$ then from Lemma 3

$$\begin{aligned}&\limsup_{n\to\infty} \lambda_k^2 P\{|Y_n(\hat{A}_k) - Y_n(\hat{B}_k)| > \varepsilon/3\}\\ &\quad\leqq \limsup_{n\to\infty} \lambda_k^2 2\exp\left(-\frac{n\varepsilon^2}{3}\Big/(2n(\delta + c' 2^{-k}) + 2\sqrt{n}\,\varepsilon/3)\right)\\ &\quad< 3\exp(2^{\beta k} - c\,\varepsilon^2(1/(\delta + c' 2^{-k})))\end{aligned}$$

for some $c, c' > 0$, any $\beta > 1/2$ and $k \geqq k_1(\beta)$. We can therefore choose $\delta_1(\varepsilon, \varepsilon', \beta) > 0$, $m_3 > \max(m_2, k_1(\beta))$ and n_1 such that for all $0 < \delta < \delta_1$

$$\text{(3.10)} \quad \lambda_{m_3}^2 P\{|Y_n(\hat{A}_{m_3}) - Y_n(\hat{B}_{m_3})| > \varepsilon/3\} < \varepsilon'/3 \quad \text{if } n \geqq n_1, \quad \rho(A,B) < \delta.$$

(3.8) − (3.10) give

$$\begin{aligned}&P(\sup_{\rho(A,B)<\delta} |Y_n(\hat{A}_m) - Y_n(\hat{B}_m)| > \varepsilon)\\ &\quad\leqq P(\sup_{\rho(A,B)<\delta} |Y_n(\hat{A}_{m_3}) - Y_n(\hat{B}_{m_3})| > \varepsilon/3)\\ &\qquad+ 2\sum_{j=m_3+1}^{\infty} P(\sup_{A\in S} |Y_n(\hat{A}_j) - Y_n(\hat{A}_{j-1})| > j^{-2}) < \varepsilon'.\end{aligned}$$

for $n \geqq n_1$ and $m \geqq m_3$, that is for n sufficiently large. This yields (3.7) and with (3.6) and Lemma 5 (3.5).

It follows from Theorem 1 and Proposition 2 of [3] that μ^n is weak-sequentially relatively compact.

If μ^{n_k} is a subsequent converging to v then for $A_1, \ldots, A_n \in S$

$$(\mu^{n_k} \kappa_{A_1}^{-1}, \ldots, \mu^{n_k} \kappa_{A_n}^{-1}) \to (\mu \kappa_{A_1}^{-1}, \ldots, \mu \kappa_{A_n}^{-1}).$$

It follows that

$$(\mu \kappa_{A_1}^{-1} \ldots, \mu \kappa_{A_n}^{-1}) = (v \kappa_{A_1}^{-1}, \ldots, v \kappa_{A_n}^{-1}).$$

Now proposition 2 of [3] shows that v too sits on $C(S)$. But the finite dimensional separate the distributions on $(C(S), \mathscr{C})$. So it follows that $v = \mu$.

Acknowledgement. I should like to thank the refree for pointing out some minor errors and suggesting some improvements in the paper's presentation.

References

1. Bennet, G.: Probability inequalities for the sum of independent random variables. J. Amer. Statist. Assoc., **57**, 33–45 (1962)
2. Billingsley. P.: Convergence of probability measures. New York: Wiley 1968
3. Dudley, R.M.: Weak convergence of probability measures on nonseparabel metric spaces and empirical measures on Euclidean spaces. Illinois J. Math. **10**, 109–126 (1968)
4. Dudley, R.M.: Sample functions of Gaussian processes. Ann. Probability **1**, 66–103 (1973)
5. Dudley, R.M.: Metric entropy of some classes of sets with differentiable boundary. J. Approx. Th. **10**, 227–236
6. deHoyos, A.: Continuity and convergence of some process parametrized by the compact sets in R^s; Z. Wahrscheinlichkeitstheorie und verw. Gebiete **23**, 153–162 (1972)
7. Parthasaraty, K.R.: Probability measures on metric spaces. New York: Academic Press 1967
8. Strassen, V., Dudley, R.M.: The central limit theorem and ε-entropy; Lecture notes in Mathematics **89**. Berlin-Heidelberg-New York: Springer 1969

Received September 30, 1976; in revised form February 25, 1978