RELATIONS BETWEEN WEAK AND UNIFORM CONVERGENCE OF MEASURES WITH APPLICATIONS

By R. Ranga Rao

Indian Statistical Institute, Calcutta

Summary. In this paper the relation between weak convergence of a sequence of measures and uniform convergence over certain classes of continuity sets (or uniform convergence of the integrals over certain classes of continuous functions) is studied. These results are applied to obtain laws of large numbers for random functions and generalizations of the Glivenko-Cantelli lemma.

1. Introduction. A well known theorem of Pólya states that if $F_n(x)$ is a sequence of distribution functions on the real line converging weakly to a continuous distribution function $F(x)$ then

$$\lim_{n\to\infty} \sup_{-\infty<x<\infty} |F_n(x) - F(x)| = 0.$$

This is one of the simplest results where weak convergence is shown to imply uniform convergence over a certain class of sets. Similar assertions of the uniform convergence of the integrals $\int f(x)\, dF_n$ occur in probability theory in various disguised forms (see for instance Wald [23], Theorem 2.14, p. 49). However the study of the relationship between weak convergence and uniform convergence does not appear to have received much attention. More precisely, let X be a separable metric space and μ_n $(n = 1, 2, \cdots)$ a sequence of measures converging weakly to a measure μ. This is equivalent to each one of the following statements:

1. For each bounded continuous function f on X,

$$\lim_{n\to\infty} \int f(x)\, d\mu_n = \int f(x)\, d\mu. \tag{1.1}$$

2. For each continuity set A of μ

$$\lim_{n\to\infty} \mu_n(A) = \mu(A). \tag{1.2}$$

The questions studied in this paper are the uniformity of the convergences (1.1) and (1.2) and their applications.

The motivation for considering these problems is provided by their applications. This consists in interpreting several results in probability theory as merely assertions of uniformity of the convergences (1.1) or (1.2). For instance let $\xi_1, \xi_2, \cdots$ be a sequence of independent and identically distributed random variables taking values in a separable Banach space $\mathfrak{X}$. If we suppose that $E\|\xi_1\| < \infty$, then the strong law of large numbers of Mourier [14] asserts that,

Received April 25, 1960; revised January 20, 1962.

with probability one,

$$\lim_{n\to\infty} \|S_n - E\xi_1\| = 0, \tag{1.3}$$

where $S_n = (\xi_1 + \cdots + \xi_n)n^{-1}$. Clearly this is an assertion of the uniformity of the convergence (1.1) over the class of all linear functionals on $\mathfrak{X}$ with norm $\leqq 1$, where μ_n, μ represent the sample and true distributions. The connecting link here is a theorem of Varadarajan [21] stating that $P[\mu_n \Rightarrow \mu] = 1$. Exactly similar remarks apply to the convergence of sample distribution functions or the various generalizations of the Glivenko-Cantelli lemma given by Wolfowitz [24], Fortet and Mourier [9] and others.

The paper consists of seven sections. The general case of measures on an arbitrary separable metric space is considered in Section 3. It is shown here that the convergence (1.1) is uniform over a class $\mathcal{A}$ satisfying the two conditions: (1) $\mathcal{A}$ is uniformly bounded, and (2) $\mathcal{A}$ is equicontinuous (Theorem 3.1). This theorem in its essentials is known. In Section 4 we consider distributions on Euclidean spaces. The main theorem of this section is the following: If $\mathcal{C}$ is a continuity class for μ, then $\mu_n \Rightarrow \mu$, if and only if the convergence (1.2) is uniform over the class $\mathcal{C}$ (Theorem 4.2). Here $\mathcal{C}$ is the class of all convex subsets of the k-dimensional Euclidean space. A similar result is proved in Section 5 for a sequence of measures converging to the Brownian motion process (Theorem 5.1).

The applications of the main theorems to the convergence problems of a sequence of random measures are given in Section 6. If $\mu(A, \omega)$, $\mu_n(A, \omega)$ is a sequence of random measures possessing a certain "ergodic" property (cf. Definition 6.2), then under certain conditions on the class of functions $\mathcal{A}$ we have (a) $P[\eta_n \to 0] = 1$, and (b) $E\eta_n^{1+\alpha} \to 0$, as $n \to \infty$, where

$$\eta_n = \sup_{f\varepsilon\mathcal{A}} \left| \int f(x)\mu_n(dx, \omega) - \int f(x)\mu(dx, \omega) \right|$$

(Theorem 6.2). As an immediate consequence we obtain the laws of large numbers of random variables taking values in a separable Banach space. The last section is devoted to the generalizations of the Glivenko-Cantelli lemma to random vectors. The results of this section constitute a considerable improvement of the earlier results of Wolfowitz [24], Fortet and Mourier [9], and Tucker [19].

Some of the results of this paper have appeared earlier as an abstract [16].

2. Preliminaries. Let X be a separable metric space and $\mathcal{B}$ the class of all Borel subsets of X. By a measure we mean a finite, non-negative, countably additive set function on $\mathcal{B}$. The term distribution will be used to denote a measure μ for which $\mu(X) = 1$. $\mathfrak{M}$ denotes the collection of all measures on $\mathcal{B}$. If $A \varepsilon \mathcal{B}$ and $\mu \varepsilon \mathfrak{M}$, then μ_A denotes the restriction of μ to the set A, i.e., $\mu_A(E) = \mu(A \cap E)$ for each E. A measure μ is said to be continuous (or nonatomic) if the μ-measure of each single point set is zero. A set $A \varepsilon \mathcal{B}$ is said to be a continuity set for μ if A has μ-null boundary. If $\xi(x)$ is a measurable map into another metric space Y, then each measure μ on X induces a measure $\mu\xi^{-1}$ on the Borel subsets of Y, where $\mu\xi^{-1}(E) = \mu[\xi^{-1}(E)]$.

Let $C_k(X)$ denote the collection of all continuous maps of X into the k-dimensional Euclidean space R_k. The following convergence notion in $C_k(X)$, commonly known as uniform convergence on compacta (u.c.c. in short), will be used in the sequel. A sequence $\xi_n(x)$ of continuous maps of X into R_k converges uniformly on compacta to $\xi(x)$, if

$$\sup_{x \varepsilon K} |\xi_n(x) - \xi(x)| \to 0$$

as $n \to \infty$, for each compact subset K of X. (Here $|z|$, where

$$z = (z_1, \cdots, z_k) \; \varepsilon \; R_k,$$

denotes the quantity $(z_1^2 + \cdots + z_k^2)^{\frac{1}{2}}$). Characterization of compact subsets of $C_k(X)$ is well known and is given by the Ascoli theorem: A family $\mathfrak{A} \subset C_k(X)$ is conditionally compact (i.e., its closure is compact), if and only if the following two conditions are satisfied.

(i) for each x, $\sup [|f(x)|, f \; \varepsilon \; \mathfrak{A}] < \infty$.

(ii) $\mathfrak{A}$ is equicontinuous at each $x \; \varepsilon \; X$, i.e., for each $\epsilon > 0$ there exists a neighbourhood N of x such that $|f(y) - f(x)| < \epsilon$, for all $y \; \varepsilon \; N$ and all $f \; \varepsilon \; \mathfrak{A}$.

An immediate consequence of the Ascoli theorem is that compact subsets of $C_k(X)$ are also sequentially compact when X is a separable metric space. For a detailed discussion of the above topology we refer to Kelley ([10], Chapter 7).

A sequence of measures $\mu_n \; \varepsilon \; \mathfrak{M}$ $(n = 1, 2, \cdots)$ is said to converge weakly to $\mu \; \varepsilon \; \mathfrak{M}$ ($\mu_n \Rightarrow \mu$ in symbols) if for each bounded continuous function $f(x)$ on X, $\int f \, d\mu_n \to \int f \, d\mu$. A detailed study of this convergence has been carried out by Alexandrov in [1], Chapter 4. The following characterization of weak convergence is well known.

THEOREM 2.1. *The following statements are equivalent*: (1) $\mu_n \Rightarrow \mu$, (2) $\lim_{n\to\infty} \mu_n(A) = \mu(A)$ *for each continuity set* A *of* μ, *and* (3) *for each bounded and uniformly continuous function* $g(x)$ *on* X, $\lim_{n\to\infty} \int g \, d\mu_n = \int g \, d\mu$.

For the proof of this theorem we refer to [1], [2] and [15]. Other works on weak convergence bearing on the problems considered here are [12], [22], [27] and [28].

3. The main theorems. The results of this section are basic to the rest of the paper. In what follows X always denotes a separable metric space. Before proceeding to the main problems we begin with some lemmas.

LEMMA 3.1. *Let* μ *be a measure on* X *and* $\mathfrak{A}$ *be an equicontinuous family of functions on* X. *Then for each* $\epsilon > 0$ *there exists a sequence of sets* A_j $(j = 1, 2, \cdots)$ *such that* (1) $\bigcup_{j=1}^{\infty} A_j = X$, (2) $A_j \cap A_{j'} = \phi$, *for* $j \neq j'$; (3) *for each* j, A_j *is a continuity set for* μ, *and* (4) *for any* $x, y \; \varepsilon \; A_j$ *and any* $f \; \varepsilon \; \mathfrak{A}$, $|f(x) - f(y)| < \epsilon$.

PROOF. Let $d(x, y)$ be the metric on X. Since $\mathfrak{A}$ is equicontinuous, for each $x \; \varepsilon \; X$ there exists a $\delta = \delta(x)$ such that if $N_x = [y : d(y, x) < \delta]$, then

$$|f(y) - f(x)| < \tfrac{1}{2}\epsilon, \tag{3.1}$$

for any $f \varepsilon \mathcal{A}$ and any $y \varepsilon N_x$. It is clear that we may suppose without loss of generality that N_x is a continuity set for μ. The class of sets $\{N_x, x \varepsilon X\}$ constitutes an open covering of X. Since X is separable there exists a sequence of points x_j $(j = 1, 2, \cdots)$ such that $X = \bigcup_{j=1}^{\infty} N_{x_j}$. Let $A_1 = N_{x_1}$, $A_2 = N_{x_2} \cap N'_{x_1}$, and so on. If we observe that the class of continuity sets of a measure form a field of sets, then it is not difficult to verify that the sequence $\{A_j\}$ has all the required properties. This proves the lemma.

The following lemma is well known.

LEMMA 3.2. *Let $F(x_1, x_2, \cdots, x_k)$ be a distribution function* (d.f) *on R_k such that each marginal* d.f. *is continuous. Then a sequence of* d.f.'s *$F_n(x_1, \cdots, x_k)$ converges weakly to $F(x_1, \cdots, x_k)$ when and only when*

$$\lim_{n\to\infty} \sup_x |F_n(x_1, \cdots, x_k) - F(x_1, \cdots, x_k)| = 0, \tag{3.2}$$

where the supremum is taken over all the vectors $x = (x_1, \cdots, x_k) \varepsilon R_k$.

THEOREM 3.1.[1] *Let $\mathcal{A}$ be a class of continuous functions on X possessing the following properties*: (1) *$\mathcal{A}$ is uniformly bounded, i.e., there exists a constant M such that $|f(x)| \leqq M$ for all $f \varepsilon \mathcal{A}$, and $x \varepsilon X$.* (2) *$\mathcal{A}$ is equicontinuous.*

If $\mu_n, \mu \varepsilon \mathfrak{M}$ $(n = 1, 2, \cdots)$, then $\mu_n \Rightarrow \mu$ if and only if, for each family $\mathcal{A}$ satisfying (1) *and* (2) *above, we have*

$$\lim_{n\to\infty} \sup_{f\varepsilon\mathcal{A}} \left| \int f\, d\mu_n - \int f\, d\mu \right| = 0. \tag{3.3}$$

PROOF. Let A_j $(j = 1, 2, \cdots)$ be a sequence of sets with the properties stated in Lemma 3.1. Let $x_j \varepsilon A_j$ be any fixed sequence of points. For any measure λ on X, let λ^* be the discrete measure confined to the set $\{x_j : j = 1, 2, \cdots\}$ in such a way that the mass at x_j is equal to $\lambda(A_j)$. Then clearly

$$\left| \int f\, d\lambda - \int f\, d\lambda^* \right| \leqq \sum_{j=1}^{\infty} \left| \int_{A_j} f\, d\lambda - \int_{A_j} f\, d\lambda^* \right|$$

$$\leqq \sum_{j=1}^{\infty} \int_{A_j} |f(x)_j - f(x)|\, d\lambda \leqq \epsilon\lambda(X)$$

by the property (4) (of Lemma 3.1) of the sequence $\{A_j\}$. Thus for any measure λ on X

$$\sup_{f\varepsilon\mathcal{A}} \left| \int f\, d\lambda - \int f\, d\lambda^* \right| \leqq \epsilon\lambda(X). \tag{3.4}$$

Since A_j is a continuity set for μ, $\mu_n(A_j) \to \mu(A_j)$ for each j. Also

$$\mu_n(X) \to \mu(X).$$

Consequently

$$\lim_{n\to\infty} \sup_{f\varepsilon\mathcal{A}} \left| \int f\, d\mu_n^* - \int f\, d\mu^* \right| \leqq M \lim_{n\to\infty} \sum_{j=1}^{\infty} |\mu_n(A_j) - \mu(A_j)| = 0. \tag{3.5}$$

[1] Results closely related to this theorem have been obtained recently by Bartoszynski [28] (cf., in particular, Theorems 1 and 2).

From (3.4) and (3.5) it follows that

$$\lim_{n\to\infty} \sup_{f\varepsilon\mathfrak{A}} \left|\int f\, d\mu_n - \int f\, d\mu\right| \leqq 2\epsilon\mu(X).$$

This completes the proof of the theorem.

REMARK. It should be noted that in a sense Theorem 3.1 describes the most general class $\mathfrak{A}$ for which weak convergence implies the convergence (3.3). Suppose that $\mathfrak{A}$ is a family of continuous functions such that (3.3) holds whenever $\mu_n \Rightarrow \mu$. Then it is not difficult to show that (1) $\mathfrak{A}$ is equicontinuous at each point $x \varepsilon X$ and (2) the family $\mathfrak{A}$ is uniformly bounded. The first property is obvious if we take μ_n and μ to be degenerate measures in (3.3). If $\mathfrak{A}$ is not uniformly bounded then there exists a sequence $f_n \varepsilon \mathfrak{A}$ and points $x_n \varepsilon X$ such that $f_n(x_n) > n$. Clearly the sequence $\{x_n\}$ can have no limit point. It is therefore easy to construct μ_n, μ such that $\mu_n \Rightarrow \mu$ but $\int f_n\, d\mu_n \to \infty$ and $\int f_n(x)\, d\mu \to c < \infty$.

THEOREM 3.2. *Let $\mathfrak{A}$ be a family of continuous functions on X satisfying the following conditions: (i) there exists a continuous function $g(x)$ on X such that $|f(x)| \leqq g(x)$ for all $f \varepsilon \mathfrak{A}$ and $x \varepsilon X$.* (ii) *$\mathfrak{A}$ is equicontinuous.*

Now suppose that μ_n, μ is a sequence of measures on X such that (a) *$\mu_n \Rightarrow \mu$ and* (b) *$\int g\, d\mu_n \to \int g\, d\mu$ $(< \infty)$. Then we have*

$$\lim_{n\to\infty} \sup_{f\varepsilon\mathfrak{A}} \left|\int f\, d\mu_n - \int f\, d\mu\right| = 0. \tag{3.6}$$

PROOF. Let $E = [x: g(x) > 0]$. Let ν_n, ν be measures on E defined by $\nu_n(A) = \int_A g\, d\mu_n$ and $\nu(A) = \int_A g\, d\mu$. Since E is open in X the conditions (a) and (b) of the theorem imply that $\nu_n \Rightarrow \nu$. Further the family $\mathfrak{A}_1 = [f/g; f \varepsilon \mathfrak{A}]$ on E is uniformly bounded and equicontinuous. Thus (3.6) is an immediate consequence of (3.3). This completes the proof.

THEOREM 3.3. *Let μ_n, $\mu \varepsilon \mathfrak{M}$ and $\mu_n \Rightarrow \mu$. If $\xi_n(x)$ is a sequence of continuous maps of X into a metric space Y, converging uniformly on compacta to $\xi(x)$, then $\mu_n\xi_n^{-1} \Rightarrow \mu\xi^{-1}$.*

PROOF. Let $g(y)$ be any bounded uniformly continuous function on Y; then it is plain that $g(\xi_n) = g[\xi_n(x)]$ converges uniformly on compacta to $g(\xi)$. Therefore by Theorem 3.1.

$$\lim_{n\to\infty} \left|\int g(\xi_n)\, d\mu_n - \int g(\xi_n)\, d\mu\right| = 0. \tag{3.7}$$

Further

$$\lim_{n\to\infty} \int g(\xi_n)\, d\mu = \int g(\xi)\, d\mu. \tag{3.8}$$

Thus it follows from (3.7) and (3.8) that for each bounded uniformly continuous function $g(y)$ on Y

$$\lim_{n\to\infty} \int g(\xi_n)\, d\mu_n = \int g(\xi)\, d\mu.$$

By Theorem 2.1 this implies that $\mu_n\xi_n^{-1} \Rightarrow \mu\xi^{-1}$. This completes the proof.

REMARK. The above theorem has been obtained earlier by Prohorov (cf. [15], p. 186) when the metric space X is complete.

THEOREM 3.4. *Let $\mu \in \mathfrak{M}$ and $\mathfrak{a}$ a family of continuous mappings of X into R_k satisfying the following conditions*: (i) $\mathfrak{a}$ *is* (u.c.c) *compact and*, (ii) μf^{-1} *has continuous marginal distributions for each* $f \in \mathfrak{a}$. *If* $\mu_n \Rightarrow \mu$, *then*

$$\lim_{n\to\infty} \sup_A |\mu_n(A) - \mu(A)| = 0, \tag{3.9}$$

where the supremum in (3.9) *is taken over all sets A of the form* $A = [x: f_j(x) \leqq a_j$ *for* $j = 1, 2, \cdots, k]$ *where* $f(x) = (f_1(x), \cdots, f_k(x)) \in \mathfrak{a}$ *and* $(a_1, \cdots, a_k)$ *is an arbitrary vector of* R_k.

PROOF. For any two measures ν and λ on R_k let $\rho(\nu, \lambda) = \sup |\nu(E) - \lambda(E)|$, where the supremum is taken over all sets $E \subset R_k$ of the form $= [x: x_j \leqq a_j$ for $j = 1, 2, \cdots, k]$. Then obviously $\rho(\nu, \lambda)$ is a distance function.

Clearly the assertion (3.9) is equivalent to the following:

$$\lim_{n\to\infty} \sup_{f \in \mathfrak{a}} \rho(\mu_n f^{-1}, \mu f^{-1}) = 0.$$

Now suppose that Theorem 3.4 is not true. Then there exists a $\delta > 0$ and a sequence of functions $f_n \in \mathfrak{a}$ such that for all n

$$\rho(\mu_n f_n^{-1}, \mu f_n^{-1}) \geqq \delta > 0. \tag{3.10}$$

Since X is a separable metric space, a compact family $\mathfrak{a}$ is also sequentially compact. Thus we may suppose without loss of generality (restricting ourselves to a subsequence if necessary) that f_n converges uniformly on compacta to a $f \in \mathfrak{a}$. Theorem 3.3 then implies that (1) $\mu_n f_n^{-1} \Rightarrow \mu f^{-1}$, and (2) $\mu f_n^{-1} \Rightarrow \mu f^{-1}$. Hence by Lemma 3.2 we have

$$\lim_{n\to\infty} \rho(\mu_n f_n^{-1}, \mu f^{-1}) = 0, \qquad \lim_{n\to\infty} \rho(\mu f_n^{-1}, \mu f^{-1}) = 0.$$

Since $\rho(\cdot, \cdot)$ is a distance function it follows from the triangle inequality that that $\lim_{n\to\infty} \rho(\mu_n f_n^{-1}, \mu f_n^{-1}) = 0$, contradicting the assumption (3.10). This completes the proof.

REMARKS.

1. Let X be any completely regular Hausdorff space and consider measures defined on Baire subsets of X. If X is assumed to be a Lindelöf space, then the proof of Theorem 3.1 is valid without any essential change. Thus Theorems 3.1–3.4 hold for such spaces[2].

2. Let $I(X) = [f: f \in C(X)$ and $\|f\| \leqq 1]$. Then the content of Theorem 3.1 is that the following two statements are equivalent (when X is a separable metric space).[3]

(i) $\mu_n \Rightarrow \mu$,

[2] It was pointed out by the referee that Theorem 3.1 is valid for all completely regular Hausdorff spaces with a countable dense subset. In a private communication R. N. Dudley of Princeton University informs the author that he has obtained the same result.

[3] It was remarked by the referee (and also by R. N. Dudley) that the statement "(i) $\Rightarrow$ (ii) for all completely regular Hausdorff spaces" is compatible with all the axioms of set theory.

(ii) $\mu_n(f) \to \mu(f)$ uniformly for $f \in \mathfrak{a}$ where $\mathfrak{a}$ is an equicontinuous family $\subset I(X)$.

It follows from the work of LeCam [12] and Prohorov [15] that when X is a complete separable metric space (i) actually implies the stronger statement:

(iii) The family $[\mu_n(f)]$ is equicontinuous at the origin of $I(X)$ (Here $I(X)$ is endowed with the u.c.c. topology), or equivalently,

(iv) For each $\epsilon > 0$, there is a compact set K_ϵ such that $\mu_n^*(X - K_\epsilon) \leqq \epsilon$ for all n.

As is well known (i) does not in general imply (iii) or (iv). However under very general conditions (which subsumes the case of separable metric spaces) (i) actually implies

(v) For every $\epsilon > 0$ and for every separable metric topology weaker than the topology of X there is a totally bounded closed set S_ϵ such that $\mu_n(X - S_\epsilon) \leqq \epsilon$ for all n (cf. LeCam [12], p. 223).

In fact (i) and (v) together always imply (ii). Thus for instance for separable metric spaces, Theorem 3.1 could also be deduced from (v). However the proof given in this paper is more direct.

4. Applications to measures on a Euclidean space. Let R_k be k-dimensional Euclidean space. In this section we discuss uniform convergence of measures over certain classes of sets special to these spaces. The following notation will be employed in the sequel. A subset of the form $[x:l(x) < p]$ where $l(x)$ is a linear function on R_k and p a real number is called a half space. $\mathfrak{H}_m$ denotes the collection of all sets which are intersections of at most m half-spaces, i.e., all sets A of the form

$$A = [x:l_1(x) < p_1, \cdots, l_r(x) < p_r]$$

with $r \leqq m$, where $l_1(x), \cdots, l_r(x)$ are arbitrary linear functions and $p_1, \cdots, p_r$ are arbitrary real numbers.

THEOREM 4.1. *Let μ be a measure on R_k such that μl^{-1} is continuous for every linear function $l(x)$ on R_k. Then $\mu_n \Rightarrow \mu$ if and only if, for each integer m,*

$$\lim_{n\to\infty} \sup_{A \in \mathfrak{H}_m} |\mu_n(A) - \mu(A)| = 0. \tag{4.1}$$

PROOF. Consider the class $\mathfrak{a}_m$ of all linear maps of norm one of R_k into R_m. Then $\mathfrak{a}_m$ is clearly (u.c.c) compact. Since μl^{-1} is continuous for each linear function $l(x)$, it follows that for each $L \in \mathfrak{a}_m$, μL^{-1} has continuous marginal distributions. The assertion (4.1) is then an easy consequence of Theorem 3.4.

THEOREM 4.2. *Suppose μ is a measure on R_k such that every convex subset of R_k has μ-null boundary (i.e., each convex set is a continuity set for μ). Then $\mu_n \Rightarrow \mu$ if and only if*

$$\sup [|\mu_n(C) - \mu(C)|, C \in \mathfrak{C}] \to 0, \tag{4.2}$$

where $\mathfrak{C}$ denotes the class of all measurable convex[4] *sets. In particular* (4.2) *is valid if the measure μ is absolutely continuous with respect to Lebesgue measure.*

[4] A convex set need not be measurable. For instance, let C be the unit sphere and let A be a non-measurable subset of the boundary of C. Then, $C \cup A$ is convex but non-measurable. In the sequel by a convex set we always mean a measurable convex set.

Before we can prove this theorem it is necessary to consider the properties of certain functions associated with convex sets. The proof of the theorem will be given in Section 4.2. We first develop some lemmas on convex sets.

4.1. *Some lemmas on convex sets.* The content of most of the lemmas given here is well known and they are collected here for ease of reference. For proofs the reader may refer to [8], or [7] Chapter 5, p. 411.

Let C be a convex set with non-empty interior. Let a be an interior point of C. Then the gauge function (also known as the support function [7], or the distance function [8]), $F(x)$ of C with respect to a is defined as follows.

$$F(x) = \inf [\lambda : \lambda > 0, a + \lambda^{-1}(x - a) \varepsilon C].$$

It is clear that if $F(x)$ is the gauge function (g.f.) of C, then C^0 = interior of $C = [x:F(x) < 1]$ and $\bar{C}$ = the closure of $C = [x:F(x) \leqq 1]$ and the boundary of $C = [x:F(x) = 1]$. The following properties of a gauge function are well known and are easily deduced from its definition.

(α) If $F(x)$ is the g.f. of C with respect to the origin, then, (i) $F(x) \geqq 0$, and $F(x) = 0$ if and only if $x = 0$. (ii) $F(cx) = cF(x)$, for all $c > 0$, and (iii) $F(x + y) \leqq F(x) + F(y)$, for each pair x, y. Conversely any function $F(x)$ with the above properties is a gauge function of the convex set $[x:F(x) < 1]$.

(β) Suppose $F(x)$ is the g.f. of C with respect to the origin and that $[x:|x| < r] \subset C$. Then for each x, $F(x) \leqq |x|/r$. (Here $|x|$ denotes the norm in R_k).

For any convex set C, the *inradius* of C, denoted by $r(C)$, is defined to be

$$r(C) = \sup [r:S(a, r) \subset C, \text{ for some } a \varepsilon C],$$

where $S(a, r)$ denotes the sphere $[x:|x - a| \leqq r]$. Then it follows from the definition that $r(C) = 0$, if and only if C has an empty interior, and that $r(C) < \infty$, if C is bounded. Also it is obvious that $r(C) = r(\bar{C})$. The introduction of the inradius is necessary to choose a suitable gauge function to represent a convex set C. This choice is made clear by the following lemma.

LEMMA 4.1. *Let C be a bounded convex set with non-empty interior. Then there is a gauge function associated with C for which*

$$|F(x) - F(y)| \leqq |x - y|/r(C),$$

for all $x, y \varepsilon R_k$.

PROOF. From the definition of $r(C)$ it follows that we can choose a sequence $z_n \varepsilon C$ and a sequence r_n such that $S(z_n, r_n) \subset C$ and $r_n \uparrow r(C)$. Since C is bounded, let z be one of the limit points of $\{z_n\}$. Then it is easy to verify that z belongs to the interior of C and $r(C) = \sup [r:S(z, r) \subset C]$. Let $F(x)$ be the g.f. of C with respect to z. Transferring the origin to z if necessary, we can assume without loss of generality that $z = 0$. Then by the sub-additive property of a g.f. we have at once

$$F(x) - F(y) \leqq F(x - y), F(y) - F(x) \leqq F(y - x).$$

Consequently $|F(x) - F(y)| \leqq \max [F(x - y), F(y - x)]$. Now property ($\beta$) of a g.f., and the fact that $S(z, r) \subset C$ for all $r < r(C)$, at once implies the assertion of the lemma.

LEMMA 4.2. *Let C_n be a sequence of convex sets $\subset S = [x:|x| \leqq m]$. If $F_n(x)$ is a* g.f. *of C_n and $F_n(x)$ converges uniformly on compacta to F, then $F(x)$ itself is a* g.f. *of a convex set $C \subset S$.*

PROOF. Suppose F_n is the g.f. of C_n with respect to z_n. Since $z_n \varepsilon S$, they have a limit point $z \varepsilon S$. Restricting ourselves to a subsequence if necessary, we may suppose without loss of generality that $z_n \to z$. Then clearly $F_n(x + z_n) \to F(x + z)$. Since $F_n(x + z_n)$ is the g.f. of $C - z_n$ with respect to the origin it follows from (α) that $F(x + z)$ is the g.f. of the set $= [x:F(x + z) < 1]$. Let $C = [x:F(x) < 1]$. Then clearly F is a g.f. of C and $C \subset \liminf_n C_n \subset S$. This proves the lemma.

Another well known theorem on convex sets which we will be using in the sequel is a theorem of Blaschke ([8], p. 64). For any set A let $S(A, \rho)$ denote the union of all spheres of radius ρ with centers in A. For any two bounded closed convex sets A, B let

$$d(A, B) = \inf [\rho:S(A, \rho) \supset B, S(B, \rho) \supset A]. \tag{4.3}$$

It is easily verified that $d(A, B)$ has all the properties of a distance.

BLASCHKE'S THEOREM. The class of all closed convex sets contained in the sphere $= [x:|x| \leqq m]$ is a compact metric space under the distance $d(A, B)$.

4.2. *Proof of Theorem* 4.2. First of all it is clear that (4.2) trivially implies that $\mu_n \Rightarrow \mu$. To prove the converse implication let us first observe that $\mu_n[S(0, m)] \to \mu[S(0, m)]$ for each m. Consequently it is sufficient to show that for each m, as $n \to \infty$

$$\sup [|\mu_n(C) - \mu(C)|, C \text{ convex} \quad \text{and} \quad \subset S(0, m)] \to 0.$$

Note that we may confine our attention to closed convex sets. Let $\mathcal{C}_\alpha$ be the collection of all closed convex sets for which $r(C) \geqq \alpha$. Consider the family $\mathfrak{F}_\alpha$ of functions F satisfying the conditions:

(1) F is a g.f. of some convex set contained in $S(0, m)$

(2) for each pair x, y

$$|F(x) - F(y)| \leqq |x - y|/\alpha.$$

Then it follows from Lemma 4.2 that $\mathfrak{F}_\alpha$ is (u.c.c) compact. Also it is easy to see that μF^{-1} is continuous for each $f \varepsilon \mathfrak{F}_\alpha$. In view of Lemma 4.1. it follows that for each convex set $C \varepsilon \mathcal{C}_\alpha$, there is an $F \varepsilon \mathfrak{F}_\alpha$ such that $C = [x:F(x) \leqq 1]$. It therefore follows from Theorem 3.4 that

$$\lim_{n\to\infty} \sup_{C \varepsilon \mathcal{C}_\alpha} |\mu_n(C) - \mu(C)| = 0, \tag{4.4}$$

for each $\alpha > 0$.

Now suppose that the theorem is not true. Then there exists a sequence C_n of closed convex sets, $\subset S(0, m)$ for all n, and a positive ϵ such that

$$|\mu_n(C_n) - \mu(C_n)| > \epsilon > 0, \qquad \text{for all } n. \tag{4.5}$$

We proceed to show that this leads to a contradiction. From (4.4) it follows that $r(C_n) \to 0$. Further we may suppose on the strength of Blaschke's Theorem (restricting ourselves to a subsequence if necessary) that there is a closed convex set $C \subset S(0, m)$, such that $d(C_n, C) \to 0$, where the distance $d(\cdot, \cdot)$ is defined by (4.3). From the definition of this distance it follows that $C \subset S(C_n, 2\epsilon_n)$, where $\epsilon_n = d(C_n, C)$. Consequently $r(C) \leqq r(C_n) + 2\epsilon_n \to 0$ as $n \to \infty$; i.e., $r(C) = 0$ or C has empty interior. Hence there is a hyperplane $H = [x:l(x) = p]$ such that $C \subset H$. Since $\mu(H) = 0$ choose η so small, that $\mu(A) \leqq \epsilon/4$ where

$$A = [x:p - \eta \leqq l(x) \leqq p + \eta].$$

Again from the definition of the distance $d(\cdot, \cdot)$ we deduce that for large n

$$C_n \subset S(C, \eta) \subset A.$$

Hence $\limsup_{n\to\infty} \mu_n(C_n) \leqq \lim \mu_n(A) = \mu(A) \leqq \epsilon/4$. Further $\mu(C_n) \leqq \mu(A) \leqq \epsilon/4$. Thus we finally obtain

$$\limsup_{n\to\infty} |\mu_n(C_n) - \mu(C_n)| \leqq \epsilon/2,$$

contradicting (4.5). The proof of the theorem is thus complete.

REMARK. We give an example below to show that uniformity of convergence in (4.1) cannot be extended to the bigger class $\mathcal{K}_\infty = \bigcup_m \mathcal{K}_m$. Let us first observe that every closed convex set is the monotone limit of sets in $\mathcal{K}_\infty$ and consequently uniformity of convergence over $\mathcal{K}_\infty$ will imply that the same uniformity also extends to the class of all convex sets. Let μ_n $(n = 0, 1, \cdots)$ be the Lebesgue measure of arcs concentrated on the circle $S_n = [(x, y): x^2 + y^2 = r_n^2]$. Let $r_n \downarrow r_0$. Then it is easily verified that $\mu_n \Rightarrow \mu$ and the μ-measure of every hyperplane is zero. Let $C = [(x, y): x^2 + y^2 \leqq r_0^2]$. Then C is closed and convex, $\mu_n(C) = 0$ for all $n \geqq 1$ but $\mu_0(C) = 1$. This example also makes it clear the necessity (in Theorem 4.2) of the assumption that the class $\mathcal{C}$ is a continuity class.

5. Application to the convergence of stochastic processes. It is well known that a stochastic process can be thought of as a distribution in a function space. In many cases this function space can be made into a separable metric space. Thus for instance we may consider distribution in $C[0, 1]$—the space of continuous functions on the closed interval $[0, 1]$, with the distance

$$\|x - y\| = \sup_{0\leqq t\leqq 1} |x(t) - y(t)|.$$

Some of the other function spaces that have been studied are $L_2[0, 1]$, the space of square integrable functions on $[0, 1]$, and $D[0, 1]$, the space of all functions on $[0, 1]$ discontinuities only of the first kind[5]. It is clear that the results of Section 3 apply to all these spaces. The theorem below is a simple illustration in the case of the space $C[0, 1]$.

[5] Several interesting topologies have been introduced in the space $D[0, 1]$ by Skorohod (cf. [25]). These topologies convert $D[0, 1]$ into a separable metric space which is topologically complete (cf. [25], [26]).

For each $m > 0$ and δ, $0 < \delta < 1$, let $\mathfrak{A}_{\delta,m}$ denote the class of functions $\alpha(t) \varepsilon C[0, \delta]$ which are of the form

$$(5.1) \qquad \alpha(t) = \int_0^t \alpha'(t)\, dt \quad \text{where} \quad \int_0^\delta |\alpha'(t)|^2\, dt \leqq m.$$

Let $\mathfrak{B}_{\delta,m}$ denote the class of all subsets $E \subset C[0, 1]$ which are of the form

$$(5.2) \qquad \begin{aligned} E = [x(t) : b_1 + \beta_1(t) \leqq x(t) \leqq a_1 + \alpha_1(t) \quad &\text{for} \quad 0 \leqq t \leqq \delta; \\ \text{and} \quad b_2 + \beta_2(t) \leqq x(t) \leqq a_2 + \alpha_2(t) \quad &\text{for} \quad \delta \leqq t \leqq 1], \end{aligned}$$

where (i) α_1, $\beta_1 \varepsilon \mathfrak{A}_{\delta,m}$; (ii) α_2, $\beta_2 \varepsilon C[\delta, 1]$ with $\|\alpha_2\| \leqq m$, $\|\beta_2\| \leqq m$, and (iii) a_1, a_2, b_1 and b_2 are arbitrary real numbers.

Let W denote the distribution in $C[0, 1]$ corresponding to the Brownian motion process (with $\sigma^2 = 1$). Then we have the following:

THEOREM 5.1. *Let P_n, P be a ssquence of distributions in $C[0, 1]$. Suppose that $P \ll W$,* i.e., *P is absolutely continuous with respect to W. Then $P_n \Rightarrow P$, if and only if for each δ and m,*

$$(5.3) \qquad \lim_{n\to\infty} \sup\, [|P_n(A) - P(A)|, A \,\varepsilon\, \mathfrak{B}_{\delta,m}] = 0.$$

PROOF. Necessity. Let $\xi_1(x, \alpha)$ and $\eta_1(x, \alpha)$ be the functionals on $C[0, \delta]$ defined as follows:

$$(5.4) \quad \xi_1(x, \alpha) = \sup_{0 \leqq t \leqq \delta} [x(t) - \alpha(t)], \qquad \eta_1(x, \alpha) = \inf_{0 \leqq t \leqq \delta} [x(t) - \alpha(t)]$$

and let $\xi_2(x, \alpha)$ and $\eta_2(x, \alpha)$ be defined similarly on $C[\delta, 1]$. Let $\mathfrak{F}$ denote the collection of all maps of $C[0, 1]$ into the four dimensional Euclidean space of the form $T(x) = \{\xi_1(x, \alpha_1), \eta_1(x, \beta_1), \xi_2(x, \alpha_2), \eta_2(x, \beta_2)\}$ where α_1, $\beta_1 \varepsilon \mathfrak{A}_{\delta,m}$ and α_2, $\beta_2 \varepsilon C[\delta, 1]$ with norm $\leqq m$. Now observe that the set E defined by (5.2) can be written in the form

$$E = [x : \xi_1(x, \alpha_1) \leqq a_1, \eta_1(x, \beta_1) \geqq b_1, \quad \xi_2(x, \alpha_2) \leqq a_2 \quad \text{and} \quad \eta_2(x, \beta_2) \geqq b_2].$$

Thus the necessity of (5.3) will follow from Theorem 3.4 if we show that $\bar{\mathfrak{F}}$, the (u.c.c) closure of $\mathfrak{F}$, satisfies conditions (i) and (ii) of Theorem 3.4.

Now it is obvious that $\mathfrak{A}_{\delta,m}$ is a compact subset of $C[0, \delta]$. Also it is easily verified that the functions ξ_2 and η_2 satisfy the inequality $|f(x) - f(y)| \leqq \|x - y\|$ for each $x, y \,\varepsilon\, C[\delta, 1]$ where $f = \xi_2$ or η_2. Thus the family $\mathfrak{F}$ is conditionally compact in the u.c.c. topology. Its (u.c.c) closure $\bar{\mathfrak{F}}$ is therefore compact. Thus it is sufficient to prove that for each $T \,\varepsilon\, \bar{\mathfrak{F}}$, PT^{-1} has continuous marginal distributions. Since by assumption $P \ll W$, it is sufficient to prove the property for W.

Now it is clear that $W[\xi_1(x, \alpha) = a] = W_\alpha[\xi_1(x, 0) = a]$, where W_α is the translation of W by α, i.e., $W_\alpha(A) = W(A + \alpha)$. It is known that $W_\alpha \ll W$ if α satisfies (5.1) (cf. Cameron and Martin [4], and also Skorohod [18], p. 426). Also $\xi_1(x, 0)$ has an absolutely continuous distribution under W (cf. Doob [6], p. 392). Thus $W[\xi_1(x, \alpha) = a] = 0$ for α satisfying (5.1). A similar argument applies to the function $\eta_1(x, \beta)$. This takes care of the first two co-ordinates of all maps $T \varepsilon \bar{\mathfrak{F}}$. As for the function $\xi_2(x, \alpha)$ it is clear that

$$\xi_2(x, \alpha) = \sup_{\delta \leqq t \leqq 1} [(x(t) - x(\delta)) - \alpha(t)] + x(\delta)$$
$$= \xi_2' + x(\delta) \quad \text{say.}$$

Since ξ_2' and $x(\delta)$ are independent, it follows that $\xi_2(x, \alpha)$ and its limits have absolutely continuous distributions under W. Similar arguments apply to $\eta_2(x, \alpha)$ or its limits. Thus for each $T \varepsilon \bar{\mathfrak{F}}$, WT^{-1} has continuous marginal distributions. This proves the necessity.

Sufficiency. Let π_δ denote the map of $C[0, 1]$ into $C[\delta, 1]$ defined as follows: $\pi_\delta(x) = x(t)$, for $\delta \leqq t \leqq 1$. If $\mathcal{S}_\delta$ denotes the class of all sets $E \subset C[\delta, 1]$ of the form

$$E = [x(t) : \alpha(t) \leqq x(t) \leqq \beta(t) \quad \text{for} \quad \delta \leqq t \leqq 1]$$

where α and β are arbitrary elements of $C[\delta, 1]$, then (5.3) implies that

$$P_n\pi_\delta^{-1}(A) \to P\pi_\delta^{-1}(A) \tag{5.5}$$

for each $A \varepsilon \mathcal{S}_\delta$. Since $\mathcal{S}_\delta$ is closed under intersections and contains a basis of the topology of $C[\delta, 1]$ it follows from (5.5) that $P_n\pi_\delta^{-1} \Rightarrow P\pi_\delta^{-1}$, for each $\delta > 0$. Let $\delta_j \downarrow 0$, then there exists a compact set $K_j \subset C[\delta_j, 1]$ such that $P_n\pi_{\delta_j}^{-1}(K_j) \geqq 1 - \epsilon 2^{-j}$, for all n. Let $K = \bigcap_j \pi_{\delta_j}^{-1}(K_j)$. Then $P_n(K) \geqq 1 - \epsilon$, for all n. Since $P[x(0) = 0] = 1$ and

$$P_n[\sup_{0 \leqq t \leqq \delta} |x(t)| < a] \to P[\sup_{0 \leqq t \leqq \delta} |x(t)| < a],$$

for each δ and a, it is easy to deduce that there is a set $E \subset C[0, 1]$, of functions equicontinuous at $t = 0$ such that $P_n(E) \geqq 1 - \epsilon$ for all n. If $K_o = K \cap E$, then $P_n(K_o) \geqq 1 - 2\epsilon$ for all n and K_o is compact. Thus the sequence $\{P_n\}$ is weakly compact. Any limit of $\{P_n\}$ obviously coincides with P. This completes the proof.

6. Application to the convergence of sample distribution functions. Let (Ω, S, P) be a probability space. A mapping $\xi(w)$ of Ω into a complete separable metric space X is said to be an X-valued random variable (or a random element in X), if $\xi^{-1}(\mathcal{B}) \subset S$ where $\mathcal{B}$ is the σ-field of Borel subsets of X. If $\xi_1, \xi_2, \cdots$ is a sequence of random elements in X then the sample distribution (s.d. in short) based on a sample of size n is the probability measure with masses $(1/n)$ at the points $\xi_1, \cdots, \xi_n$. This s.d. will be denoted by μ_n or $\mu_n(A, \omega)$ where $A \varepsilon \mathcal{B}$ and $\omega \varepsilon \Omega$. In this section we utilize the results of Section 3 to discuss the convergence of μ_n's. We also deduce here the results of Mourier, and Fortet and Mourier on the law of large numbers (LLN) for random elements in a separable Banach space. The method presented here is essentially different from theirs and seems to be of some interest.

Since in problems of the type we are considering here, the nature of the dependence in the sequence $\{\xi_n\}$ plays its role only through the validity of the ergodic theorem for the sequence of real valued random variables $\{f(\xi_n)\}$, where $f(x)$ is a function on X, it is advantageous to consider the problem in a general setting. For this purpose we introduce the notion of a random measure.

Definition 6.1. A random measure on X is a function $\lambda(A, \omega)$ defined for each $A \varepsilon \mathcal{B}$ and $\omega \varepsilon \Omega$ possessing the following properties: (i) for each fixed ω, ω, $\lambda(A, \omega)$ is a countably additive measure on $\mathcal{B}$ and (ii) for each fixed $A \varepsilon \mathcal{B}$, $\lambda(A, \omega)$ is a measurable function in ω. It is clear from the definition that if $f(x)$ is a measurable function X such that $\int_X f(x)\lambda(dx, \omega) < \infty$ almost everywhere, then $\int_X f(x)\lambda(dx, \omega)$ is a real valued random variable. In the sequel we will write $\int f\, d\lambda$ in place of $\int_X f(x)\lambda(dx, \omega)$ when there is no danger of confusion. We denote the mean or expected value of real valued random variable ξ by $E\xi$.

A trivial example of a random measure is the sample d.f. $\mu_n(A, \omega)$ associated with a sequence $\{\xi_n\}$ of random elments in X. A non-trivial example is provided by the conditional probability distributions. This is made clear by the following well known lemma (cf. [11]).

Lemma 6.1. *Let $\xi(\omega)$ be a random element in X (a complete separable metric space) and $\mathfrak{F}$ a sub-σ-field of $\mathcal{S}$. Then there exists a random measure $\lambda(A, \omega)$ on X, such that for every function $f(x)$ on X with $E|f(\xi)| < \infty$, the following relation holds almost everywhere.*

$$E[f(\xi) \mid \mathfrak{F}] = \int_X f(x)\lambda(dx, \omega) \tag{6.1}$$

where $E[f(\xi) \mid \mathfrak{F}]$ denotes the conditional expectation of the random variable $f(\xi)$ with respect to $\mathfrak{F}$.

Let $\xi_1, \xi_2, \cdots$ be a strictly stationary sequence of random elements in X. It is well known that we may suppose without loss of generality that the sequence $\{\xi_n\}$ is generated from ξ_1 through a measure preserving transformation T, i.e., $\xi_n = \xi_1(T^n\omega)$. Let $\mathfrak{J}$ denote the invariant σ-field of measurable sets A such that $A = TA$. Then by Lemma 6.1 there exists a random measure $\mu(A, \omega)$ such that, if $E|f(\xi_1)| < \infty$, then

$$E[f(\xi_1) \mid \mathfrak{J}] = \int_X f(x)\mu(dx, \omega) \tag{6.2}$$

almost everywhere. Now the classical theorem of Birkhoff (cf. Loève [13], p. 421) states that if $E|f(\xi_1)| < \infty$, then

$$\begin{aligned} \lim_{n\to\infty} \frac{f(\xi_1) + \cdots + f(\xi_n)}{n} &= \lim_{n\to\infty} \int f(x)\mu_n(dx, \omega) \\ &= E[f(\xi_1) \mid \mathfrak{J}] = \int f(x)\mu\ (dx, \omega) \end{aligned} \tag{6.3}$$

almost surely. Since the validity of the ergodic theorem is an essential requirement in all our results we introduce for the sake of brevity the following definition.

DEFINITION 6.2. An arbitrary sequence of random measures $\{\lambda_n(A, \omega), \lambda(A, \omega); n = 1, 2, \cdots\}$ on X is said to possess the ergodic property if for each real valued function $g(x)$ on X, for which $E\int |g(x)|\lambda(dx, \omega) < \infty$

$$\lim_{n\to\infty} \int g(x)\, d\lambda_n = \int g(x)\, d\lambda$$

almost everywhere.

The following result is due essentially to Varadarajan ([21], p. 25).

THEOREM 6.1. *Let* $\{\lambda_n, \lambda; n = 1, 2, \cdots\}$ *be a sequence of random measures on* X *possessing the ergodic property. Then*

$$P[\lambda_n \Rightarrow \lambda] = 1. \tag{6.4}$$

The above theorem is an immediate consequence of the ergodic property and the following observation due to Varadarajan (cf. [21], p. 24): in a separable metric space there is a fixed sequence of bounded continuous functions $f_1, f_2, \cdots$ such that if $\int f_i\, d\nu_n \to \int f_i\, d\nu$ for each i, then $\nu_n \Rightarrow \nu$.

THEOREM 6.2. *Let* $\{\lambda_n, \lambda; n = 1, 2, \cdots\}$ *be a sequence of random measures on* X *possessing the ergodic property. Let* $\mathfrak{A}$ *be a family of continuous functions on* X *satisfying the following two conditions*: (i) *there exists a continuous function* $g(x)$ *on* X *such that* $|f(x)| \leqq g(x)$ *for each* $f \in \mathfrak{A}$ *and* $x \in X$; $E\int g(x)\lambda(dx, \omega) < \infty$; *and* (iii) $\mathfrak{A}$ *is equicontinuous. Then* $P[\eta_n \to 0] = 1$, *where*

$$\eta_n = \sup_{f\in\mathfrak{A}} \left| \int f\, d\lambda_n - \int f\, d\lambda \right|. \tag{6.5}$$

PROOF. Since the sequence $\{\lambda_n, \lambda\}$ possesses the ergodic property, it follows from Theorem 6.2 that with probability one

$$\lambda_n => \lambda \quad \text{and} \quad \int g\, d\lambda_n -> \int g d\lambda. \tag{6.6}$$

Theorem 6.2 is then an immediate consequence of (6.6) and Theorem 3.2.

The following lemma facilitates the passage from almost sure convergence of Theorem 6.2 to convergence in the mean of order α.

LEMMA 6.2. *Let* $\mathfrak{A}$ *be a class of measurable functions on* X, *and* $g(x)$ *a measurable function on* X *such that* (i) $|f(x)| \leqq g(x)$ *for each* $f \in \mathfrak{A}$ *and* $x \in X$; *and* (ii) $E\int [g(x)]^{1+\alpha}\lambda(dx, \omega) < \infty$ *for some* $\alpha \geqq 0$. *Suppose that* $\{\lambda_n, \lambda\}$ *possesses the ergodic property and*

$$\lim_{n\to\infty} P[\eta_n > \epsilon] = 0, \tag{6.7}$$

for each $\epsilon > 0$, *where* η_n *is defined by* (6.5). *Then* $\lim_{n\to\infty} E\eta_n^{1+\alpha} = 0$.

PROOF. Let h_n, h be real valued random variables defined as follows:

$$h_n = \int [g(x)]^{1+\alpha}\, d\lambda_n, \qquad h = \int [g(x)]^{1+\alpha}\, d\lambda,$$

for $n = 1, 2, \cdots$. It is clear that h_n and h possess the following properties:

(a) $h_n \geqq 0$, $h \geqq 0$; (b) $\lim_{n\to\infty} Eh_n = Eh$; and (c) $P[h_n \to h] = 1$. Hence it follows from a theorem of Scheffé ([17], p. 435) that

$$\lim_{n\to\infty} E|h_n - h| = 0. \tag{6.8}$$

From the property (i) of the family $\mathcal{A}$ it is easily deduced that

$$\eta_n^{1+\alpha} \leqq 2^{1+\alpha}\left[\left(\int g\, d\lambda_n\right)^{1+\alpha} + \left(\int g\, d\lambda\right)^{1+\alpha}\right] \leqq 2^{1+\alpha}(h_n + h) \tag{6.9}$$

where in the last step we have utilized Holder's inequality. Let $E_n = [\omega: \eta_n \geqq \epsilon]$. Then clearly

$$E\eta_n^{1+\alpha} \leqq \epsilon^{1+\alpha} + \int_{E_n} \eta_n^{1+\alpha}\, dP. \tag{6.10}$$

From (6.7)–(6.10) it is easy to deduce that $E\eta_n^{1+\alpha} \to 0$, as $n \to \infty$. This completes the proof.

An immediate consequence of Lemma 6.2 and Theorem 6.2 is the following

THEOREM 6.3. *With the same notation as in Theorem* 6.2, *we suppose in addition that* $E\int [g(x)]^{1+\alpha}\, d\lambda < \infty$, *for some* $\alpha \geqq 0$. *Then* $\lim_{n\to\infty} E\eta_n^{1+\alpha} = 0$ *where* η_n *is defined by* (6.5).

It is clear that the above results are valid in particular if we take $\lambda_n = \mu_n$ and $\lambda = \mu$ where μ_n is the sequence of s.d's and μ is defined by (6.2). The ergodic theorem, viz., (6.3), implies that the sequence $\{\mu_n, \mu\}$ possess the ergodic property. Thus we have

THEOREM 6.4. *Let* $\mathcal{A}$ *be an equicontinuous class of continuous functions and* $g(x)$ *a continuous function on* X *satisfying conditions* (i) *and* (iii) *of Theorem* 6.2. *Suppose* $E|g(\xi_1)|^{1+\alpha} < \infty$ *for some* $\alpha \geqq 0$, *then*

(a) $P[\lim_{n\to\infty} \eta_n = 0] = 1$,

(b) $\lim_{n\to\infty} E\eta_n^{1+\alpha} = 0$,

where

$$\eta_n = \sup_{f\varepsilon\mathcal{A}} \left|\frac{f(\xi_1) + \cdots + f(\xi_n)}{n} - \int f(x)\mu(dx, \omega)\right|.$$

The above theorem is an extension of the results of Fortet and Mourier ([9]). These authors consider only classes $\mathcal{F}$ of functions f such that

$$|f(x) - f(y)| \leqq Md(x, y)$$

where $d(x, y)$ is the metric in X, where it is supposed that $Ed(x_0, \xi_1) < \infty$ for some $x_0 \varepsilon X$. They deduce this result from the results of Mourier ([14]) on the law of large numbers for random elements in a separable Banach space. Since we have given an independent proof of Theorem 6.4, we can now proceed to deduce the strong law of large numbers for such general random elements.

Let $X = \mathfrak{X}$ be a separable Banach space. If ξ is a random element with values in $\mathfrak{X}$ then the expected value of ξ is defined to be its Pettis integral, i.e., an element $E\xi \varepsilon \mathfrak{X}$ such that $(x^*, E\xi) = E(x^*, \xi)$ for each $x^* \varepsilon \mathfrak{X}^*$—the adjoint of $\mathfrak{X}$. The Pettis integral of ξ is known to exist if $E\|\xi\| < \infty$. In an analogous manner conditional expectation can also be defined. It can also be shown (Chatterji [5]) that if $E\|\xi\| < \infty$, then there exists a random element to be denoted by $E(\xi|\mathfrak{Z})$ such that, for each $x^* \varepsilon \mathfrak{X}^*$,

$$(6.11) \qquad (x^*, E(\xi_1|\mathfrak{Z})) = \int (x^*, x)\mu(dx, \omega) = E[(x^*, \xi_1)|\mathfrak{Z}]$$

almost surely. Here the random measure $\mu(A, \omega)$ is defined by the relation (6.2). Since $\mathfrak{X}$ is assumed to be separable we may even suppose that (6.11) holds with probability one, simultaneously for all $x^* \varepsilon \mathfrak{X}^*$.

Let $S_n = (\xi_1 + \xi_2 + \cdots + \xi_n)n^{-1}$ where $\xi_1, \cdots, \xi_n, \cdots$ is a strictly stationary sequence for which $E\|\xi_1\| < \infty$. It is clear that

$$\|S_n - E(\xi_1 \mid \mathfrak{Z})\| = \sup_{\|x^*\| \leqq 1} |(x^*, S_n - E(\xi_1 \mid \mathfrak{Z}))|$$
$$= \sup_{\|x^*\| \leqq 1} \left| \int x^*, x)\, d(\mu_n - \mu) \right|.$$

As a consequence of Theorem 6.4 we have

THEOREM 6.5. *With the notations of the above paragraph we have*

$$\lim_{n\to\infty} \|S_n - E(\xi_1 \mid \mathfrak{Z})\| = 0$$

almost surely. If it is supposed in addition that $E\|\xi_1\|^{1+\alpha} < \infty$, *for some* $\alpha \geqq 0$, *then*

$$\lim_{n\to\infty} E\|S_n - E(\xi_1 \mid \mathfrak{Z})\|^{1+\alpha} = 0.$$

REMARK. This theorem has been essentially obtained by Mourier earlier ([14]). However the statement here is more precise and identifies the limit random element as $E(\xi_1 \mid \mathfrak{Z})$. It should be noted that the method employed by Mourier is entirely different from ours.

7. Generalization of the Glivenko-Cantelli lemma. The classical theorem of Glivenko-Cantelli ([13], p. 20) states that if $\xi_1, \xi_2, \cdots$ is a sequence of independent real valued random variables with common distribution μ, and μ_n's denote their s.d's, then

$$(7.1) \qquad P[\sup_{-\infty<x<\infty} |F_n(x) - F(x)| \to 0] = 1$$

where $F_n(x) = \mu_n(-\infty, x)$ and $F(x) = \mu(-\infty, x)$. This result was generalized to the case of k-dimensional random vectors by Wolfowitz [24] and also by Fortet and Mourier [9]. It was proved in [24] that with probability one

$$(7.2) \qquad \sup [|\mu_n(A) - \mu(A)|, A \varepsilon \mathfrak{H}] \to 0$$

where $\mathfrak{H}$ denotes the collection of all half-spaces. Recently Tucker [19] extended (7.1) to a strictly stationary sequence of random variables. In this section we apply the general theorems of Sections 3 and 4 to obtain results which constitute a considerable improvement over these results.

THEOREM 7.1. *Let* $\xi_1, \xi_2, \cdots$ *be a sequence of independent random vectors in* R_k *with a common distribution* μ. *Let* μ_n *be their* s.d. *Suppose that each convex set is a continuity set for the non-atomic component of* μ, *then*

$$(7.3) \qquad P[\sup |\mu_n(C) - \mu(C)| \to 0] = 1,$$

where the supremum in 7.3 *is taken over the class* $\mathcal{C}$ *of all convex sets.*

PROOF. Let A denote the set of atoms of μ, i.e., $A = [x\colon \mu\text{-measure of the single point } \{x\} > 0] = \{x_i : i = 1, 2, \cdots\}$ say. Then it is not difficult to deduce from the strong law of large numbers that with probability one

$$(7.4) \qquad \sum_{i=1}^{\infty} |\mu_n(x_i) - \mu(x_i)| \to 0$$

as $n \to \infty$. (Here $\mu(x_i)$ denotes the mass at x_i). If we introduce $\lambda_n = (\mu_n)_N$, $\lambda = (\mu)_N$ where $N = R_k - A$, then it is clear that (7.3) is true or false according as

$$(7.5) \qquad P[\sup_{C \varepsilon \mathcal{C}} |\lambda_n(C) - \lambda(C)| \to 0] = 1.$$

If we now suppose that each $C \varepsilon \mathcal{C}$ is a continuity set for λ then the assertion (7.5) is an immediate consequence of Theorem 4.2 and Theorem 6.1. This completes the proof.

THEOREM 7.2. *Let* $\xi_1, \xi_2, \cdots$ *be a strictly stationary sequence of random vectors in* R_k *and let* μ_n *denote their* s.d. *Then with probability one, for each integer* m,

$$(7.6) \qquad \lim_{n \to \infty} \sup [|\mu_n(A) - P(\xi_1 \varepsilon A | \mathfrak{J})|; A \varepsilon \mathcal{K}_m] = 0,$$

where $\mathfrak{J}$ *is the invariant* σ*-field associated with the sequence* $\{\xi_n\}$. *In particular when the* ξ_n*'s are independent,* (7.6) *is valid with* $P(\xi_1 \varepsilon A \mid \mathfrak{J}) = P(\xi_1 \varepsilon A)$.

The proof of this theorem will be given in Section 7.2. In the following section we develop the tools necessary for this purpose.

7.1. We first derive a decomposition theorem for measures on R_k generalizing the usual decomposition into atomic and non-atomic parts. For this purpose we introduce the following

DEFINITION. A measure λ on R_k is said to be r-dimensional ($0 \leqq r \leqq k$) if (i) there exists an r-dimensional subspace or a translate of such a subspace, say E, such that $\lambda(R_k - E) = 0$ and (ii) $\lambda(A) = 0$ for every set A which is a translate of a subspace of dimension less than r.

It is clear from the definition that in the case $k = 1$ a measure μ is 0-dimensional if it is degenerate and 1-dimensional if it is non-atomic. Similar remarks apply to the case $k > 1$.

LEMMA 7.1. *Let* $0 \leqq r \leqq k - 1$. *Let* $\mathcal{V}_r$ *be the collection of all sets A which are translates of r-dimensional subspaces of* R_k. *Then for any two sets* $A, B \varepsilon \mathcal{V}_r (1 \leqq r \leqq k - 1)$, *either* (i) $A = B$, *or* (ii) $A \cap B = \phi$, *or* (iii) $A \cap B \varepsilon \mathcal{V}_s$ *for some* $s < r$.

The proof is straight forward and will be omitted.

LEMMA 7.2. *Let* λ *be any measure and* $\{A_\alpha\}$ *be a family of sets such that* $\lambda(A_\alpha \cap$

$A_\beta) = 0$ *for* $\alpha \neq \beta$. *Then* $\lambda(A_\alpha) = 0$ *for all but a countable number of sets* A_α.

This result is well known and its proof is omitted.

LEMMA 7.3. *Let* λ *be an arbitrary measure on* R_k. *Then there exist measures* $\lambda_r^{(i)} (r = 0, 1, \cdots, k-1, i = 0, 1, \cdots,$ *and* $r = k, i = 0)$ *all mutually orthogonal such that*

(i) $\lambda = \sum_{r=0}^{k-1} \sum_{i=0}^{\infty} \lambda_r^{(i)} + \lambda_k^{(0)}$;

(ii) *for* $0 \leqq r \leqq k$, $\lambda_r^{(i)}$ *is an r-dimensional measure for all i.*

PROOF. Since the class $\mathcal{V}_0$ consists of all single point sets it follows from Lemma 7.2 that λ can be written in the form $\lambda = \sum_{i=0}^{\infty} \lambda_0^{(i)} + \nu_1$ where $\lambda_0^{(i)}$ are the atoms of λ and the ν_1-measure of each set $A \varepsilon \mathcal{V}_0$ is zero. Now consider the measure ν_1 and the class $\mathcal{V}_1$. For any two distinct sets $A, B \varepsilon \mathcal{V}_1$, by Lemma 7.1 $A \cap B = \phi$ or $A \cap B \varepsilon \mathcal{V}_0$ and consequently $\nu_1(A \cap B) = 0$ always. By Lemma 7.2 there exists a countable number of sets $V_{1i} \varepsilon \mathcal{V}_1 (i = 0, 1, \cdots)$ such that $\nu_1(A) = 0$ if $A \varepsilon \mathcal{V}_1$, and $A \neq A_i$ for some i. Define $\lambda_1^{(i)}$ to be the restriction of ν_1 to V_{1i}. Then it is clear that ν_1 can be written as

$$\nu_1 = \sum_{i=0}^{\infty} \lambda_1^{(i)} + \nu_2 .$$

It is obvious from the construction that all the measures μ_2 and $\{\lambda_1^{(i)}\}$ are all mutually orthogonal. Thus λ may be written as

$$\lambda = \sum_{i=0}^{\infty} \lambda_0^{(i)} + \sum_{i=0}^{\infty} \lambda_1^{(i)} + \nu_2$$

where $\nu_2(A) = 0$ for $A \varepsilon \mathcal{V}_1$ and $\lambda_0^{(i)}, \lambda_1^{(i)}$ are respectively zero and one-dimensional measures. It is then clear that the same argument can be repeated with ν_2 and the class $\mathcal{V}_2$ and so on.

Thus we obtain

$$\lambda = \sum_{r=0}^{k-1} \sum_{i=0}^{\infty} \lambda_r^{(i)} + \nu_k \tag{7.7}$$

where $\lambda_r^{(i)}$ are all mutually orthogonal and $\lambda_r^{(i)}$ is an r-dimensional measure. Also from (7.7), $\nu_k(A) = 0$ for each $A \varepsilon \mathcal{V}_{k-1}$. Thus ν_k is a k-dimensional measure and the lemma follows from (7.7) with $\nu_k = \lambda_k^{(0)}$. This completes the proof.

LEMMA 7.4. *Let E be a translate of an r-dimensional subspace, and* ν *an r-dimensional measure concentrated in E, i.e.,* $\nu(R_k - E) = 0$. *If* ν_n *is a sequence of measures for which* (a) $\nu_n \Rightarrow \nu$, *and* (b) $\nu_n(E) \to \nu(E)$ *as* $n \to \infty$, *then*

$$\lim_{n\to\infty} \sup_{A \varepsilon \mathcal{K}_m} |\nu_n(A) - \nu(A)| = 0. \tag{7.8}$$

PROOF. Let V be an r-dimensional subspace and z a point of R_k such that $E = V + z$. Consider the measures λ_n and λ defined by

$$\lambda_n(A) = \nu_n(A + z); \qquad \lambda(A) = \nu(A + z).$$

Clearly $\lambda_n \Rightarrow \lambda$. Since $\lambda_n(V) \to \lambda(V)$ it follows that the measures $(\lambda_n)_V$ and λ_V restricted to V and considered as measures on V are such that $(\lambda_n)_V \Rightarrow \lambda_V$. Since ν is r-dimensional it follows that $\nu(H + z) = \lambda(H) = 0$ for every hyperplane of the vector space V. Thus Theorem 4.1 applies to the sequence $(\lambda_n)_V$ and λ_V and yields

$$\lim_{n\to\infty} \sup_{A\in\mathfrak{F}_m} |\lambda_n(A \cap V) - \lambda(A \cap V)| = 0$$

for each m, where $\mathfrak{F}_m$ is the class of all sets $A \subset R_k$ such that $A \cap V$ is the intersection of at most m half-spaces of the vector space V. It is easily verified that $\mathfrak{K}_m \subset \mathfrak{F}_{m+n}$. Consequently

$$\lim_{n\to\infty} \sup_{A\in\mathfrak{K}_m} |\lambda_n(A \cap V) - \lambda(A \cap V)| = 0. \tag{7.9}$$

Since $\lambda(R_k - E) = \nu(R_k - E) = 0$, we have $\lambda_n(R_k - V) = 0$. In view of this and the fact that $\mathfrak{K}_m$ is invariant under translations, it now follows from (7.9) that (7.8) holds. This completes the proof.

Let ν be any arbitrary measure on R_k and let

$$\nu = \sum_{r=0}^{k-1} \sum_{i=1}^{\infty} \nu_r^{(i)} + \nu_k^{(0)} \tag{7.10}$$

be the decomposition asserted by Lemma (7.3). Since all the measures $\nu_r^{(i)}$ are all mutually orthogonal, there exist mutually disjoint sets $A_r^{(i)}$ ($r = 0, 1, \cdots, k-1$, $i = 1, \cdots$, and $r = k$, $i = 0$) such that

$$\nu_r^{(i)}(B) = \nu(B \cap A_r^{(i)}) \tag{7.11}$$

for each Borel set B.

LEMMA 7.5. *Let ν_n be a sequence of measures weakly converging to ν such that, in the notation of the above paragraph,*

(i) $\nu_n(A_r^{(i)}) \to \nu(A_r^{(i)})$, *for all r and i,*

(ii) $\nu_{nr}^{(i)} \Rightarrow \nu_r^{(i)}$, *for all r and i, where $\nu_{nr}^{(i)}$ is defined by the relation $\nu_{nr}^{(i)}(B) = \nu_n(B \cap A_r^{(i)})$ for each Borel set B. Then, for each m,* (7.8) holds.

PROOF. Since $\nu_r^{(i)}$ is r-dimensional there exists a set $E_r^{(i)}$ which is a translate of an r-dimensional subspace and such that $\nu_r^{(i)}(R_k - E_r^{(i)}) = 0$. We may also suppose that $A_r^{(i)} \subset E_r^{(i)}$. Thus the conditions (i) and (ii) of the theorem imply that $\nu_{nr}^{(i)} \Rightarrow \nu_r^{(i)}$ and $\nu_{nr}^{(i)}(E_r^{(i)}) \to \nu_r^{(i)}(E_r^{(i)})$. Therefore, by Lemma 7.4, we have

$$\lim_{n\to\infty} \sup_{A\in\mathfrak{K}_m} |\nu_{nr}^{(i)}(A) - \nu_r^{(i)}(A)| = 0, \tag{7.12}$$

for $0 \leqq r \leqq k - 1$ and all i. From Theorem 4.1 it follows that

$$\lim_{n\to\infty} \sup_{A\in\mathfrak{K}_m} |\nu_{nk}^{(0)}(A) - \nu_k^{(0)}(A)| = 0. \tag{7.13}$$

Now define $E_N = [\cup A_r^{(i)},\ 0 \leqq r \leqq k-1,\ 1 \leqq i \leqq N] \cup A_k^{(0)}$. It then follows from assumption (i) of the theorem that $\nu_n(E_N) \to \nu_n(E_N)$ as $n \to \infty$ for each N. Or equivalently,

$$\lim_{n\to\infty} \nu_n(R_k - E_N) = \nu(R_k - E_N). \tag{7.14}$$

Finally it is easily seen that for any set A

$$\begin{aligned} |\nu_n(A) - \nu(A)| \leqq & \sum_{r=0}^{k-1} \sum_{i=1}^{N} |\nu_{nr}^{(i)}(A) - \nu_r^{(i)}(A)| \\ & + |\nu_{nk}^{(0)}(A) - \nu_k^{(0)}(A)| + \nu_n(R_k - E_N) + \nu(R_k - E_N). \end{aligned} \tag{7.15}$$

Take the supremum in (7.15) over the class $\mathcal{K}_m$ and then let $n \to \infty$. It then follows from (7.12) and (7.13) that

$$\lim_{n\to\infty} \sup_{A \varepsilon \mathcal{K}_m} |\nu_n(A) - \nu(A)| \leqq 2\nu(R_k - E_N), \tag{7.16}$$

for each N. The assertion of the theorem then follows from (7.16) and (7.14). The proof is thus complete.

7.2. *Proof of Theorem* 7.2. Let λ denote the probability distribution of ξ_1, i.e. $\lambda(A) = P(\xi_1 \varepsilon A)$. By Lemma 7.3, there exist mutually orthogonal measures $\lambda_r^{(i)}$ such that

$$\lambda = \sum_{r=0}^{k-1} \sum_{i=1}^{\infty} \lambda_r^{(i)} + \lambda_k^{(0)}. \tag{7.17}$$

Let $A_r^{(i)}$ be mutually disjoint Borel sets such that the measures $\lambda_r^{(i)}$ are defined by

$$\lambda_r^{(i)}(E) = \lambda(E \cap A_r^{(i)}). \tag{7.18}$$

Let $\mu(A, \omega)$ and $\mu_r^{(i)}(A, \omega)$ be the random measures defined by the relations

$$\mu(A, \omega) = P[\xi_1 \varepsilon A | \mathfrak{J}] \qquad \mu_r^{(i)}(A, \omega) = P[\xi_1 \varepsilon A \cap A_r^{(i)} | \mathfrak{J}], \tag{7.19}$$

where $\mathfrak{J}$ is the invariant σ-field associated with a stationary sequence. It follows easily from (7.17), (7.18) and (7.19) that

$$\mu(A, \omega) = \sum_{r=0}^{k-1} \sum_{i=1}^{\infty} \mu_r^{(i)}(A, \omega) + \mu_k^{(0)}(A, \omega) \tag{7.20}$$

almost surely. It is not difficult to see from (7.18) and (7.19) that the measure $\mu_r^{(i)}$ is the restriction of μ to the set $A_r^{(i)}$, i.e., with probability one,

$$\mu_r^{(i)}(E, \omega) = \mu(E \cap A_r^{(i)}, \omega) \tag{7.21}$$

simultaneously for all Borel sets E and all r and i.

Now consider the sequence of sample distribution functions $\mu_n(A, \omega)$. Let the measures $\mu_{nr}^{(i)}$ be defined as follows. For each Borel set E,

$$\mu_{nr}^{(i)}(E) = \mu_n(E \cap A_r^{(i)}, \omega)$$

for all r and i. With this definition, Theorem 6.1 implies that

$$\text{(7.22)} \qquad \begin{aligned} &\text{(i)} && \mu_n \Rightarrow \mu, \\ &\text{(ii)} && \mu_{nr}^{(i)} \Rightarrow \mu_r^{(i)}, \qquad \text{for all } r \text{ and } i, \end{aligned}$$

with probability one. From Birkhoff's individual ergodic theorem it follows that with probability one

$$\lim_{n\to\infty} \mu_n(A_r^{(i)}, \omega) = \mu(A_r^{(i)}, \omega), \tag{7.23}$$

for all r and i. Thus (7.22) and (7.23) and Lemma 7.5 together imply (7.6) holds with probability one. This completes the proof.

REMARKS. We give an example to show that, in general, Theorem 7.2 cannot be improved to assert uniform convergence over the class $\bigcup_m \mathfrak{K}_m$. Let S denote the surface of the unit sphere in R_k and let S_o be the interior of the unit sphere. Let $\xi_1, \xi_2, \cdots$ be a sequence of independent random vectors with a non-atomic distribution μ concentrated in S. If μ_n denotes their s.d's, then with probability one

$$(7.24) \qquad \lim_{n\to\infty} \sup_{A \varepsilon \mathfrak{K}_m} |\mu_n(A) - \mu(A)| = 0$$

for each integer m, by Theorem 7.2. Now suppose A is any Borel subset of S. It is easy to see that $S_o \cup A$ is convex. Consequently

$$\sup_{A\varepsilon\mathfrak{C}} |\mu_n(A) - \mu(A)| = \sup_{A\varepsilon\mathfrak{F}} |\mu_n(A) - \mu(A)|$$

where $\mathfrak{F}$ is the class of all Borel subsets of S. Since μ_n and μ are concentrated in S and μ_n is atomic, while μ is not, it follows that

$$\sup_{A\varepsilon\mathfrak{F}} |\mu_n(A) - \mu(A)| = 1.$$

Thus the convergence (7.24) is not in general uniform with respect to m.

In the above example, suppose in addition that μ is concentrated in

$$S \subset [x: x_j \geqq 0\, j = 1, \cdots, k].$$

Then an argument similar to the one employed in the preceding paragraph shows that $\sup_{A\varepsilon\mathfrak{F}} |\mu_n(A) - \mu(A)| = 1$, where $\mathfrak{F}$ is the class of all sets A with the following property: $x \varepsilon A$ and y is such that $y_j \leqq x_j$ for all $j = 1, \cdots, k$, implies that y also εA. This contradicts a conjecture of Blum ([3]) stating that with probability one

$$\lim_{n\to\infty} \sup_{A\varepsilon\mathfrak{F}} |\mu_n(A) - \mu(A)| = 0.$$

Acknowledgment. The author is indebted to R. R. Bahadur for his encouragement and for valuable suggestions and discussions during the preparation of the paper.

REFERENCES

[1] ALEXANDROV, A. D. (1943). Additive set functions in abstract spaces. *Mat. Sbornik. New Series* **13** 169–234.

[2] BILLINGSLEY, P. (1956). The invariance principle for dependent random variables. *Trans. Amer. Math. Soc.* **83** 250–268.

[3] BLUM, J. R. (1955). On the convergence of empiric distribution functions. *Ann. Math. Statist.* **26** 527–529.

[4] CAMERON, R. H. AND MARTIN, W. T. (1944) Transformations of Wiener integrals under translations. *Ann. Math.* **45** 386–396.

[5] CHATTERJI, S. D. (1960). Conditional expectations of Banach valued random variables. (Abstract) *Ann. Math. Statist.* **31** 232.

[6] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[7] DUNFORD, N. AND SCHWARTZ, J. (1958). *Linear Operators*. **1**, Inter-Science Pub., New York.

[8] EGGLESTON (1958). Convexity. Cambridge tracts in mathematics and mathematical physics No. 47. Cambridge Univ. Press.

[9] FORTET, R. AND MOURIER, E. (1953). Convergence de la répartition empirique vers la répartition théorique. *Ann. Scientifiques de l'École Normale Supérieure* **60** 267–285.

[10] KELLEY, J. L. (1955). *General Topology*. Van Nostrand, New York.

[11] BLACKWELL, D. (1956). On a class of probability spaces. *Proc. Third Berkeley Symp. Math. Stat.* Prob. **2** 1–6. Univ. of California Press, Berkeley.

[12] LECAM, L. (1957). Convergence in distribution of stochastic processes. *Univ. of California Publications in Statistics*. Vol. 2, No. 11 (1957) p. 207–236.

[13] LOÈVE, M. (1960). *Probability Theory*. 2nd ed. Van Nostrand, New York.

[14] MOURIER, E. (1953). Éléments aléatoires dans un espace de Banach. *Ann. Inst. Henri Poincaré* **13** 161–244.

[15] PROHOROV, YU. V. (1956). Convergence of random processes and limit theorems of the theory of probability (Russian). *Teor. Veroyatnost. i Primenen.* **1** 177–238.

[16] RANGA RAO, R. (1960). Generalization of a theorem of Pólya and applications (Abstract). *Ann. Math. Statist.* **31** 241.

[17] SCHEFFÉ, H. (1947). A useful convergence theorem for probability distributions. *Ann. Math. Statist.* **18** 434–438.

[18] SKOROHOD, A. V. (1957). On the differentiability of measures which correspond to stochastic processes (Russian). *Teor. Veroyatnost. i Primenen.* **2** 417–443.

[19] TUCKER, H. G. (1959). A Generalization of the Glivenko-Cantelli theorem. *Ann. Math. Statist.* **30** 828–830.

[20] VARADARAJAN, V. S. (1958). Weak convergence of measures on separable metric spaces. *Sankhya* **19** 15–22.

[21] VARADARAJAN, V. S. (1958). On the convergence of probability distributions, *Sankhya*, **19** 23–26.

[22] VARADARAJAN, V. S. (1958). Convergence in distribution of stochastic processes. Thesis submitted to Calcutta University.

[23] WALD, A. (1950). *Statistical Decision Functions*. Wiley, New York.

[24] WOLFOWITZ, J. (1954). Generalization of the theorem of Glivenko-Cantelli. *Ann. Math. Statist.* **25** 131–138.

[25] SKOROHOD, A. V. (1956). Limit theorems for stochastic processes (Russian). *Teor. Veroyatnost. i Primenen.* **1** 89–319.

[26] KOLMOGOROV, A. N. (1956). On Skorohod convergence (Russian). *Teor. Veroyatnost. i Primenen.* **1** 239–247.

[27] VARADARAJAN, V. S. (1961). Convergence of stochastic processes. *Bull. Amer. Math. Soc.* **67** 276–280.

[28] BARTOSZYŃSKI, R. (1961). A characterization of weak convergence of measures. *Ann. Math. Statist.* **32** 561–567.