The Annals of Mathematical Statistics
1971, Vol. 42, No. 6, 2050–2055

GENERALIZATIONS OF THE GLIVENKO-CANTELLI THEOREM

BY J. DEHARDT
California State College at Los Angeles

0. Introduction. Let μ be a probability measure on the Borel sets, $\mathscr{B}$, in k dimensional Euclidean space E_k and $X_1, X_2, \cdots$ a sequence of independent random vectors with values in E_k such that $P[X_i \in A] = \mu(A)$ for every A in $\mathscr{B}$, $i = 1, 2, \cdots$. A necessary and sufficient condition is given for

$$(*) \qquad \sup_{f \in \mathscr{M}} \left| n^{-1} \sum_{i=1}^{n} f(X_i) - \int f \, d\mu \right| \rightarrow_{a.s.} 0,$$

where $\mathscr{M}$ is the class of all monotone functions on E_k with a uniform bound. (*) is shown to hold with no restriction on μ for several classes of functions, one of which is the class of characteristic functions of half-spaces in E_k. This result strengthens the theorem of Wolfowitz (1954). I am obliged to H. D. Brunk for many invaluable conversations.

1. Sufficient conditions on $\mathscr{M}$ and μ for (*). Let $\mathscr{M}$ denote a class of real-valued, measurable, uniformly bounded functions defined on E_k. For f in $\mathscr{M}$ let

$$S_n(f) = n^{-1} \sum_{i=1}^{n} f(X_i).$$

If $\mathscr{M}$ and μ are such that

$$P[\lim_{n \to \infty} \sup_{f \in \mathscr{M}} |S_n((f) - \int f \, d\mu| = 0] = 1,$$

it will be said that (*) holds. It will be assumed that $\mathscr{M}$ is such that $\sup_{f \in} \quad S_n(f)$ is measurable. The particular classes $\mathscr{M}$ discussed in Section 2 have this property.

Lemmas 1–6 give sufficient conditions on $\mathscr{M}$ and μ for (*), while the remainder of the results are concerned with (*) holding for particular classes $\mathscr{M}$.

LEMMA 1. *If corresponding to each positive number ε there is a finite class of functions $\mathscr{M}_\varepsilon$ such that for each f in $\mathscr{M}$ there are f_1 and f_2 in $\mathscr{M}_\varepsilon$ with $f_1 \leqq f \leqq f_2$ and $\int f_2 d\mu - \int_1 f d\mu < \varepsilon$, then (*) holds.*

PROOF. Corresponding to each positive integer k, let $\{f_1{}^k, f_2{}^k, \cdots, f^k{}_m\}$ be the finite class $\mathscr{M}_{1/k}$ which corresponds to the positive number $1/k$ by the hypothesis. If

$$A_i^k = [S_n(f_i^k) \to \int f_i^k \, d\mu] \qquad i = 1, 2, \cdots, m; k = 1, 2, \cdots,$$

then

$$P(A_i^k) = 1,$$

by the Law of Large Numbers.
If $A = \bigcap_{k=1}^{\infty} \bigcap_{i=1}^{\infty} A_i^k$, then $PA = 1$.

Received September 3, 1970.

If $f_j^k \leqq f_i^k$ and $0 \leqq \int f_i^k d\mu - \int f_j^k d\mu \leqq 1/k$, then $\int f_j^k d\mu \leqq \int f d\mu \leqq \int f_i^k d\mu$ and $S_n(f_j^k \leqq S_n(f) \leqq S_n(f_i^k)$ for every f in $\mathcal{M}$ such that $f_j^k \leqq f \leqq f_i^k$.
Then for these f,

$$S_n(f) - \int f d\mu \leqq S_n(f_j^k) - \int f_j^k d\mu \leqq S_n(f_i^k) - \int f_i^k d\mu + k^{-1}$$

and

$$S_n(f) - \int f d\mu \geqq S_n(f_i^k) - \int f_i^k d\mu \geqq S_n(f_j^k) - \int f_j^k d\mu - k^{-1}.$$

It follows that for every f in $\mathcal{M}$ and all k,

$$|S_n(f) - \int f d\mu| \leqq \sup_{1 \leqq i \leqq m} |S_n(f_i^k) - \int f_i^k d\mu| + k^{-1},$$

and so

$$\Delta = \limsup_{n\to\infty} \sup_{f \in \mathcal{M}} |S_n(f) - \int f d\mu|$$
$$\leqq \limsup_{n\to\infty} \sup_{1 \leqq i \leqq m} |S_n(f_i^k) - \int f_i^k d\mu| + k^{-1}.$$

If ω is in A then

$$\lim_{n\to\infty} \sup_{1 \leqq i \leqq m} |S_n(f_i^k)(\omega) - \int f_i^k d\mu| = 0 \quad \text{for all} \quad k,$$

and so

$$\Delta = \lim_{n\to\infty} \sup_{f \in \mathcal{M}} |S_n(f) - \int f d\mu| = 0 \quad \text{for that} \quad \omega.$$

Therefore, $[\Delta = 0] \supset A$ which implies that $P[\Delta = 0] = 1$, which concludes the proof.

LEMMA 2. *If there is a compact metric on $\mathcal{M}$ such that the mapping $T(f) = \int f d\mu$ is continuous and every neighborhood $N(f)$ of f in $\mathcal{M}$ contains g_1 and g_2 in its closure with $g_1 \leqq g \leqq g_2$ for every g in $N(f)$, then (*) holds.*

PROOF. Given $\varepsilon > 0$, for each f in $\mathcal{M}$ select a neighborhood $N(f)$ such that g in $N(f)$ implies that

$$|\int f d\mu - \int g\, d\mu| < \varepsilon/4.$$

This class of neighborhoods covers $\mathcal{M}$ so there exists a finite subclass covering $\mathcal{M}$, say $N(f_1), N(f_2), \cdots, N(f_n)$. Corresponding to $N(f_i)$ there are g_{i1} and g_{i2} in the closure of $N(f_i)$ with $g_{i1} \leqq g \leqq g_{i2}$ for all g in $N(f_i)$, $i = 1, 2, \cdots, n$. It follows that

$$|\int g_{i1} d\mu - \int g_{i2} d\mu| \leqq |\int g_{i1} d\mu - \int f_i d\mu| + |\int f_i d\mu - \int g_{i2} d\mu| < \varepsilon.$$

Hence, corresponding to $\varepsilon > 0$ is the finite class of functions $\{g_{i1}, g_{i2}\}$ $i = 1, 2, \cdots, n$, which satisfies the hypothesis of Lemma 1, and so (*) holds.

Denote $n^{-1}\sum_{i=1}^n f(X_i) I_{[X_i \in D]}$ by $S_n(f, D)$, where D is in β and $I_{[X_i \in D]}$ is the indicator random variable of $[X_i \in D]$. Let $A_1, A_2, \cdots$ be a partition of E_k. Define $\mu_j(A)$ as $\mu(A\, A_j)$ for all A in β, $j = 1, 2, \cdots$.

LEMMA 3. *If*

$$P[\lim_{n\to\infty} \sup_{f \in \mathcal{M}} |S_n(f, A_j) - \int f d\mu_j| = 0] = 1$$

for $j = 1, 2, \cdots$, then (*) holds.

PROOF. If M is a uniform bound for $\mathscr{M}$ then corresponding to each positive integer k there exists a positive integer $N(k)$ such that $\sum_{j=N(k)}^{\infty}\mu(A_j) < 1/(4kM)$. $A^k = \bigcup_{j=N(k)}^{\infty} A_j$ and $B_k = [\lim_{n\to\infty} |\mu^n(A^k)-\mu(A^k)| = 0]$, where $\mu^n(A^k) = n^{-1}\sum_{i=1}^n I_{[X_i \in A^k]}$.

If $B = \bigcap_{k=1}^{\infty} B_k$ then $P(B) = 1$.

Let $C_j = [\lim_{n\to\infty} \sup_{f\in\mathscr{M}} S_n|(f, A_j)-\int f d\mu_j| = 0]$. If $C = \bigcap_{j=1}^{\infty} C_j$ then $P(C) = 1$ and so $P(BC) = 1$. From

$$\begin{aligned}|S_n(f)-\textstyle\int f\,d\mu| &\leqq \textstyle\sum_{j=1}^{N(k)-1} |S_n(f, A_j)-\int f\,d\mu_j| + |S_n(f, A^k)-\int_{A^k} f\,d\mu| \\ &\leqq \textstyle\sum_{j=1}^{N(k)-1} |S_n(f, A_j)-\int f\,d\mu_j|| + M\mu(A)+\mu(A^k)|\end{aligned}$$

it follows that

$$\begin{aligned}&\limsup_{n\to\infty} \sup_{f\in\mathscr{M}} |S_n(f)-\textstyle\int f\,d\mu| \\ &\quad\leqq \textstyle\sum_{j=1}^{N(k)-1} \limsup_{n\to\infty} \sup_{f\in\mathscr{M}} |S_n(f, A_j)-\int f\,d\mu_j| + M\lim |\mu^n(A^k)+\mu(A^k)|.\end{aligned}$$

On CB the left-hand member is then less than or equal to $1/k$ for every positive integer k and hence (*) holds.

Let $\mathscr{M}(R)$ denote the class of functions obtained from $\mathscr{M}$ by restricting the domain of each function in $\mathscr{M}$ to a rectangle R in E_k. If

$$P[\lim_{n\to\infty} \sup_{f\in\mathscr{M}} |S_n(f, R)-\textstyle\int_R f\,d\mu| = 0] = 1,$$

it will be said that (*) holds for $\mathscr{M}(R)$.

LEMMA 4. *If* (*) *holds for* $\mathscr{M}(R)$, *for every* R, *then* (*) *holds for* $\mathscr{M}$.

PROOF. Corollary of Lemma 3.

LEMMA 5. *If "a" is a point in* E_k *with* $\mu(a) > 0$ *then*

$$P[\lim_{n\to\infty} \sup_{f\in\mathscr{M}} |S_n(f, a)-\textstyle\int_a f\,d\mu| = 0] = 1.$$

The proof is immediate.

LEMMA 6. *If* (*) *holds for* $\mathscr{M}$ *then* (*) *holds for the class of functions of the form* $f-g$, f *and* g *in* $\mathscr{M}$.

The proof is immediate.

2. Necessary and sufficient conditions for (*) on classes, $\mathscr{M}$, of monotone and related functions. A real-valued function defined on E_k is to be called monotone if it is monotone in the separate variables. The following definition and Lemmas 7, 8 and 9 will be given for $\mathscr{M}$ the class of all real-valued functions on E_k which are non-increasing in each variable and have a fixed uniform bound. A similar definition and lemmas could be given for a class of functions of mixed monotonicity or which are non-decreasing in each variable.

A metric on $\mathscr{M}$, due to P. Lévy in the one-dimensional case, is defined as follows: At each discontinuity point $a = (a_1, a_2, \cdots, a_k)$ of each f in $\mathscr{M}$ adjoint to the graph of f the line segment joining the points $(a_1, a_2, \cdots, a_k, f(a-0)$ and $(a_1, a_2, \cdots, a_k, f(a+0))$. Call the resulting graph G_f. For f and g in $\mathscr{M}$, at each point $p = (p_1, \cdots p_{k+1})$ on G_f let $d_p\ (f, g)$ be the distance from p to G_g along the line through p with parametric equations $x_i = t/(k+1)^{\frac{1}{2}}+p_i$. The distance $d(f, g) = \sup_{p \in G_f} d_p\ (f, g)$ is a metric on $\mathscr{M}$ if functions with 0 distance are identified.

LEMMA 7. *$\mathscr{M}(R)$ is compact in the metric $d(f, g)$, for every R in E_k.*

PROOF. Let $\{f_n\}$ be a sequence of functions in $\mathscr{M}(R)$. There is a subsequence $\{f_n'\}$ of $\{f_n\}$ and a function f in $\mathscr{M}(R)$ such that $f_n'(x) \to f(x)$ for every x with rational coordinates in R. It follows from a direct generalization of a theorem in Gnedenko-Kolmogorov ((1954), page 33) that $d(f_n', f) \to 0$, which completes the proof.

LEMMA 8. *$d(f_n, f) \to 0$ implies that $f_n(x) \to f(x)$ at each continuity point of f.*

PROOF. A direct generalization of the theorem in Gnedenko-Kolmogorov (1954) referred to in the proof of Lemma 7.

Corresponding to a point b in E_k let $A_b^{1} = [x \in E_k : x_i \geqq b_i,\ i = 1, \cdots, n]$. By reversing the inequalities in the definition one at a time, sets A_b^{i} are defined $i = 1, 2, 3, \cdots, 2^k$. A Borel set B is a strictly monotone graph if there exists i such that for all b in B, $A_b^{i} \cap B = \{b\}$. If $A_b^{i} \cap B$ is contained in the boundary of A_b^{i} then B is a monotone graph.

LEMMA 9. *If μ is 0 on every monotone graph in E_k then the mapping $T(f) = \int f d\mu$ on $\mathscr{M}$ is continuous in the metric $d(f, g)$.*

PROOF. If $d(f_n, f) \to 0$ then $f_n \to f$ on the continuity points of f by Lemma 8. The discontinuities of f lie on a countable number of monotone graphs in E [Brunk, Ewing, Utz (1956)], so $\mu[f_n \to f] = 1$ and $\int f_n - f d\mu \to 0$.

Let L be the set of points x in E_k with $\mu(x) > 0$. Define $\mu^L(A)$ as $\mu(A\bar{L})$ for all A in β.

THEOREM 1. *If $\mathscr{M}$ is a class of uniformly bounded monotone functions on E_1 then (*) holds.*

PROOF. The proof will be given for a class of non-increasing functions. With the appropriate change in the definition of the metric and in Lemmas 7 and 8 the proof for a non-decreasing class would be the same.

By Lemma 4 it suffices to show that (*) holds for $\mathscr{M}(R)$ where R is an arbitrary finite interval. By Lemmas 3 and 5 (*) will hold for μ if it holds for μ^L. Lemmas 7 and 8 provide a metric on $\mathscr{M}(R)$ which, by Lemma 9, satisfies the hypotheses of Lemma 2 for μ^L. Hence (*) holds for $\mathscr{M}(R)$ and μ^L and the proof is completed.

A set in E_k of the form $[(x_1, \cdots x_k) : x_{j_1} = c_1, \cdots, x_{j_{k-i}} = c_{k-i}]$ will be called an i-dimensional flat in E_k. Let B_0 denote the union of the 0-dimensional flats of E_k, each of which has positive μ-measure. If $\mu^{B_0}(A) = \mu(A\bar{B}_0)$ for all A in β, then there are at most a countable number of 1-dimensional flats of E_k, each of which has positive μ^{B_0}-measure.

Let B_1 be the union of these intersected with $\bar{B}_0$. If $\mu^{B_1}(A) = \mu(A\bar{B}_1)$ for all A in β, then there are at most a countable number of 2-dimensional flats of E_k, each of which has positive μ^{B_1}-measure. Let B_2 be the union of these intersected with $\overline{B_0 \cup B_1}$. Continuing in this way, disjoint sets $B_0, B_1, \cdots, B_{k-1}$ are defined. Let $V = \bigcup_{i=0}^{k-1} B_i$ and $\mu^V(A) = \mu(A\bar{V})$ for all A in β. μ^V is 0 on every i-dimensional flat of E_k, $i = 1, 2, \cdots, k-1$.

THEOREM 2. *If $\mathscr{M}$ is the class of all monotone functions on E_k with any fixed uniform bound then* (*) *holds if and only if μ^L is* 0 *on every strictly monotone graph in E_k.*

PROOF. The necessity is given in DeHardt (1970).

Sufficiency. The proof will be given for the case in which the functions in $\mathscr{M}$ are monotone non-increasing in each variable. With the appropriate change in the definition of the metric and in Lemmas 7 and 8, the proof for the other 2^k-1 classes of monotone functions would be the same, and if the sufficiency is true for each of these classes, clearly it is true for their union.

The proof is by induction. The condition on μ is sufficient for (*) in E_1 by Theorem 1. Suppose the sufficiency holds in $E_j, j = 1, 2, \cdots, k-1$. By Lemma 4, it suffices to give the proof for $\mathscr{M}(R)$, where R is an arbitrary rectangle in E_k. Let $B_0, B_1, \cdots, B_{k-1}$ be the sets whose union defines V. Each B_i is a countable union of i-dimensional flats of E_k, and hence can be put equal to $\bigcup_{j=1}^{\infty} A_j^{\,i}$ where $A_j^{\,i}A_k^{\,i}$ is empty for $j \neq k$ and $A_j^{\,i}$ is contained in an i-dimensional flat of the type described in the definition of B_i. The assumed condition on μ is inherited by these flats so by the inductive hypothesis,

$$P[\lim_{n\to\infty} \sup_{f\in\mathscr{M}} |S_n(f, A_{ji}) - \textstyle\int_{A_{ji}} f\,d\mu| = 0] = 1,$$
$$j = 1, 2, \cdots, \quad i = 1, 2, \cdots, k-1.$$

Therefore, after applying Lemma 3 twice, it remains to show that (*) holds for μ^V. The discontinuities of each f in $\mathscr{M}(R)$ lie on a countable number of monotone graphs in E_k [2]. Since μ^V is 0 on every monotone graph in E_k, it follows from Lemmas 7 and 9 that the hypotheses of Lemma 2 are satisfied for $\mathscr{M}(R)$ and μ^V. Hence (*) holds for $\mathscr{M}(R)$ and μ^V on E_k and the sufficiency is completed.

A set $\{(x_1, \cdots, x_k) : \sum_{i=1}^{k} a_i x_i \leqq c\}$ or its complement will be called a hyper-space in E_k.

THEOREM 3. *If $\mathscr{M}$ is the class of characteristic functions of hyper-spaces in E_k then* (*) *holds.*

PROOF. In the metric $d(f, g)$, $\mathscr{M}(R)$ is a closed subclass of the class of all monotone functions with a uniform bound greater than 1. It follows from Lemma 7 that $\mathscr{M}(R)$ is compact in this metric. Let W be the union of the hyper-planes E_k which have positive μ-measure. Let $\mu^W(A) = \mu(A\,\overline{W})$ for all A in β. Since the discontinuities of each $f \in \mathscr{M}$ are on a hyper-plane of E_k, it follows from Lemmas 9 and 2 that (*) holds for μ^W. Since W can be considered as a countable union of disjoint subsets (not necessarily proper) of hyper-planes, it follows from Lemma 3 that it remains to prove that (*) holds when μ has all its mass on a hyper-plane of E_k. This case, when $k = 2$, reduces to a special case of Theorem 1, and holds for arbitrary k by induction.

With the assumption that the components of the X_i be independent, Wolfowitz (1954) obtained the conclusion of Theorem 3. He pointed out later that his proof will go through with no assumption on μ. If the class of hyper-spaces is expanded so that the boundaries may include proper subsets of the hyper-plane, then the conditions of Theorem 2 are necessary and sufficient for (*).

THEOREM 4. (*) *holds for the class of functions on* E_k *which have first and second differences of constant sign.*

PROOF. According to Hobson (1950) each function in this class has its discontinuities on a countable number of graphs of the form $x_i = C$. Hence μ^V is 0 on the discontinuity set of each function in this class and (*) holds by an argument similar to that in the proof of the sufficiency of Theorem 2.

COROLLARY. *If* $F(x_1, \cdots, x_k)$ *is the common distribution function of the* X_i, *and* $F_n(x_1, \cdots, x_k)$ *is the corresponding empiric distribution function, then*

$$P[\limsup_{n\to\infty} {}_{-\infty < x_i < \infty i=1,\cdots k} |F_n(x_1, \cdots, x_k) - F(x_1, \cdots, x_k)| = 0] = 1.$$

The definition of a function of bounded variation used for the following theorem is given in Hobson (1950).

THEOREM 5. *If* $\mathscr{M}$ *is the class of all functions on* E_k *of uniformly bounded variation corresponding to any fixed bound, then* (*) *holds for* $\mathscr{M}$ *if and only if* μ^L *is* 0 *on every strictly monotone graph in* E_k.

PROOF. The sufficiency holds by Lemma 6 from Theorem 2. The necessity is proved in DeHardt (1970).

REFERENCES

BRUNK, H. D., EWING, G. M. and UTZ, W. R. (1956). Some Helly theorems for monotone functions. *Proc. Amer. Math. Soc.* **7** 776–783.

DEHARDT, J. H. (1970). A necessary condition for Glivenko-Cantelli convergence in E_k. *Ann. Math. Statist.* **41** 2177–2178.

GNEDENKO, B. V. and KOLMOGOROV, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables.* Addison-Wesley, Reading, Mass.

HOBSON, E. W. (1950). *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, **1** (3rd ed.). Cambridge Univ. Press.

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