Probab. Th. Rel. Fields 78, 169–191 (1988)

Probability Theory and Related Fields

Donsker Classes of Sets

Michel Talagrand*

Université de Paris VI, Equipe d'Analyse, Place Jussieu, F-75252, Paris Cedex 05, France
and Department of Mathematics, Ohio State University, Columbus, OH 43210, USA

Summary. We study the central limit theorem (CLT) and the law of large numbers (LLN) for empirical processes indexed by a (countable) class of sets $\mathscr{C}$. The main result, of purely measure-theoretical nature, relates different ways to measure the "size" of $\mathscr{C}$. It relies on a new rearrangement inequality that has been inspired by techniques used in the local theory of Banach spaces. As an application, we give sharp necessary conditions for the CLT, that are in some sense the best possible. We also obtain a way to compute the rate of convergence in the LLN.

1. Introduction

Consider a complete probability space (Ω, Σ, P). If the space is a model for a real world phenomenon, one often has no knowledge of P. The best that one can do is to get independent samples $(x_1, \ldots, x_n)$, and consider the empirical measure $Q_n = \frac{1}{n} \sum_{i \leq n} \delta_{x_i}$. Of special interest is thus the problem of recovering the law of P from the random laws Q_n. A natural way to compare Q_n and P is to study the uniform convergence of Q_n to P over a class of measurable sets $\mathscr{C}$. If $Q_n - P$ converges uniformly to zero over $\mathscr{C}$, we say that $\mathscr{C}$ is a Glivenko-Cantelli class. The classical Glivenko-Cantelli theorem states that for any probability P on $\mathbb{R}$, the class of intervals $]-\infty, t]$ is a Glivenko-Cantelli class. If $n^{1/2}(Q_n - P)$ converges in law over $\mathscr{C}$ to a Gaussian process, we say that $\mathscr{C}$ is a Donsker class. The classical Kolmogorov-Smirnov theorem states that for any probability P on $\mathbb{R}$, the class of intervals $]-\infty, t]$ is a Donsker class. Donsker classes have been extensively studied by R.M. Dudely. In their recent paper [6], E. Giné and J. Zinn have made clever use of Gaussian processes techniques and have achieved remarkable progress. The present paper will introduce new techniques. Some originate in the measure theoretical work of [8] and its application to the new description of Glivenko-Cantelli classes given in [10]. The main tool (Theorem 11) is a rearrangement inequality that has been inspired by some techniques used in the local theory of Banach spaces

[1, 5]. This quantitative result will allow us to give a complete description of Donsker classes and to compute the rate of convergence of $Q_n - P$ to zero over any Glivenko-Cantelli class.

Finite classes are Donsker classes; and a subclass of a Donsker class is Donsker. The property of being a Donsker class is thus a smallness property. So the first issue is what sensible description of smallness one should use. Let us say that $\mathscr{C}$ shatters a set $\{x_1, \ldots, x_n\}$ if each subset of $\{x_1, \ldots, x_n\}$ is the trace of an element of $\mathscr{C}$. Say that $\mathscr{C}$ shatters $(x_1, \ldots, x_n)$ if the points $x_1, \ldots, x_n$ are distinct, and if $\mathscr{C}$ shatters $\{x_1, \ldots, x_n\}$. A widely used approach to measure the size of $\mathscr{C}$ is to study the trace of $\mathscr{C}$ on a random sample $\{x_1, \ldots, x_n\}$. One can consider the cardinal T_n of the cardinal Δ_n of the trace of $\mathscr{C}$ on $\{x_1, \ldots, x_n\}$ as in [6]. (Also entropy conditions have been considered, but we shall not use them.) One then evaluates the most common values of T_n and Δ_n. The definition of these quantities T_n, Δ_n is not simple. It is, however, easier to find which sequences $(x_1, \ldots, x_n)$ are shattered by $\mathscr{C}$. Let V_n be the set of these sequences. Let $r_n = (P^n(V_n))^{1/n}$. We will show the surprising fact that the speed of convergence of the sequence (r_n) to zero almost completely determines the properties of $\mathscr{C}$ for the empirical measure. Our main result (Theorem 2) relates the sequence (r_n) and the behavior of Δ_n and T_n; it is of purely measure-theoretical nature. The description of Donsker classes will follow by combining this result with the results of Giné and Zinn.

2. Notations and Results

We denote by (Ω', Σ', P') the space $(\Omega^{\mathbb{N}}, \Sigma^{\mathbb{N}}, P^{\mathbb{N}})$. The generic point of Ω' will be denoted by $x = (x_1, x_2, \ldots)$. Denote by $X = l^{\infty}_{\mathscr{C}}$ the Banach space of uniformly bounded sequences indexed by $\mathscr{C}$. We say that $\mathscr{C}$ is pregaussian if there is a centered Radon gaussian measure μ on X such that for each A, B in $\mathscr{C}$, we have

$$\int e_A^*(t)\, e_B^*(t)\, d\mu(t) = P(A \cap B) - P(A)\, P(B),$$

where e_A^* is the coordinate function of index A on X. Consider the map $\theta\colon \Omega \to X$ given by $\theta(t) = (1_A(t) - P(A))_{A \in \mathscr{C}}$. We say that $\mathscr{C}$ is a Donsker class if $\frac{1}{\sqrt{n}} \sum_{i \leq n} \theta(x_i)$ converges in law to μ. This means that for each bounded norm-continuous function g on X, we have

$$\lim_{n \to \infty} \int^* g\left(\frac{1}{\sqrt{n}} \sum_{i \leq n} \theta(x_i)\right) dP'(x) = \int_X g\, d\mu.$$

(This elegant definition, due to J. Hoffmann-Jørgensen, avoids the difficulty of defining the law of θ). Only essentially standard arguments are needed to see that this is equivalent to the usual definition of Donsker Classes. They are carried out, e.g., in [2].

A Donsker class is in particular pregaussian. It was considered until recently a difficult problem to recognise if a class of sets is pregaussian. However, the recent results of the author on Gaussian processes [13] show that this is equivalent to the following simple condition, on the metric space $(\mathscr{C}, d)$, where $d(A, B) = P(A \triangle B)$: there exists a probability measure m on $(\mathscr{C}, d)$ such that

$$\lim_{\varepsilon \to 0} \sup_{A \in \mathscr{C}} \int_{0+}^{\varepsilon} (-\ln(m\{B \in \mathscr{C}; P(A \triangle B) \leqq t\}))^{1/2} t^{-1/2} \, dt = 0.$$

It is also equivalent to the existence of a sequence $(f_n)_{n \geqq 2}$ of $L^2(P)$ such that $\lim_n (\ln n)^{1/2} \|f_n\|_2 = 0$, and that for each A in $\mathscr{C}$ we can write (in $L^2(P)$) $1_A = \sum_n \alpha_n(A) f_n$, where $\sum_n |\alpha_n(A)| \leqq 1$. We shall not use these results here.

The study of empirical measures runs into tedious measurability problems. It is often possible to dispense of any measurability condition [10, 11]. This, however, requires extra work. We wish here to avoid measurability problems. We shall often assume $\mathscr{C}$ countable; (weaker assumptions, as in [6], could also be used). When no extra work is required, we shall only assume the following very general condition.

Definition 1. We say that $\mathscr{C}$ satisfies condition (M) if for each integers k, l, the set

$$\bigcup_{A \in \mathscr{C}} A^k \times (\Omega \setminus A)^l = \{(x_1, \ldots, x_{k+l}) \colon \exists A \in \mathscr{C}, \forall i \leqq k, x_i \in A, \forall k < i \leqq k+l, x_i \notin A\},$$

is measurable.

Condition (M) implies that V_n is measurable for each n. For x in Ω', let $\Delta_n(x)$ be the cardinal of the trace of $\mathscr{C}$ on $\{x_1, \ldots, x_n\}$. Let $T_n(x)$ be the largest integer k such that $\mathscr{C}$ shatters a subset of $\{x_1, \ldots, x_n\}$ of cardinal k. Condition (M) implies that T_n and Δ_n are measurable.

The sequence (r_n) might decrease in an irregular way. To avoid technical difficulties, we shall compare r_n with an auxilliary function ϕ. This function ϕ is non-increasing continuous in $[1, \infty[$ with $0 < \phi \leqq 1$. We shall always assume it satisfies the following condition

$$\exists c > 0 \ \forall t \geqq 1 \qquad \phi(t) \leqq c\phi(2t). \tag{$*$}$$

This condition is satisfied for example if $\phi(t) = t^{-\alpha}$, $\alpha \geqq 0$. The most important case is $\phi(t) = 1/t$. We denote by a_n the root of the equation $t/\phi(t) = n$ (so if $\phi(t) = 1/t$, $a_n = n^{1/2}$).

Our main result is the following.

Theorem 2. *Let $\mathscr{C}$ be a class of sets that satisfies condition (M). Then the following conditions are equivalent:*

(I) $\lim_n r_n/\phi(n) = 0$.

(II) $\forall \gamma>0, \lim_n P'(\{T_n(x) \geqq \gamma a_n\})=0.$

(III) $\forall \gamma>0, \lim_n P'(\{\ln \Delta_n(x) \geqq \gamma a_n\})=0.$

The equivalence of (II) and (III) is surprising. Indeed the class F of subsets of $\{x_1, \ldots, x_n\}$ of cardinal $\leqq T_n(x)$ does not shatter a subset of $\{x_1, \ldots, x_n\}$ of cardinal $>T_n(x)$. However, $\operatorname{card} F = \sum_{i \leqq T_n(x)} \binom{n}{i}$. For $T_n(x)$ significantly smaller than n, ln (card F) is of the order $T_n(x) \ln(n/T_n(x))$, that is much bigger than $T_n(x)$. So it seems that in condition (III), one should not be able to do better than

$$\lim_n P'(\{\ln \Delta_n(x) \geqq \gamma a_n \ln(n/a_n)\})=0.$$

The elimination of the extra logarithm requires a careful global analysis.

Theorem 2 will be established in Sect. 4. It relies on a symmetrisation (= rearrangement) procedure (in the sense of Steiner) that may be of independent interest, and that is described in Sect. 3. The application of Theorem 2 to Donsker classes yields the following:

Theorem 3. *A countable class of measurable sets is a Donsker class if and only if it is pregaussian and satisfies the conditions of Theorem 2 for $\phi(t)=1/t$.*

We shall give in Sect. 7 an example showing that the conditions of Theorem 2 for $\phi(t)=1/t$ do not imply in general that $\mathscr{C}$ is pregaussian. However, $\mathscr{C}$ is pregaussian whenever the series $\Sigma 2^{-i} m_i^{1/2}$ converges, where m_i is the smallest integer m for which $r_m \leqq 2^{-2i}$. A typical case is $r_m = O(1/m(\ln m)^\alpha)$ for $\alpha>2$.

Theorem 3 will be proved in Sect. 5. We shall also give a "local" version of this result.

In the case that (a_n) increases significantly faster that $n^{1/2}$, Theorem 2 can be used to compute the rate of convergence of $Q_n - P$ over $\mathscr{C}$. Consider the condition

$$\exists b>0, \ \forall k, \quad \sum_{l \leqq k} a_{2^l} 2^{-l/2} \leqq b a_{2^k} 2^{-k/2}. \tag{$**$}$$

In Sect. 6, we shall prove the following:

Theorem 4. *Let $\mathscr{C}$ be a class of sets that satisfies condition (M), and assume that (a_n) satisfies condition $(**)$. Then the conditions of Theorem 2 are equivalent to*

$$\forall \gamma>0, \quad \lim_n P'\left(\left\{\sup_{A \in \mathscr{C}} \left|\frac{1}{n} \sum_{i \leqq n} 1_A(x_i) - P(A)\right| \geqq \gamma a_n/n\right\}\right)=0.$$

3. Symmetrisation Results

We start by proving a known extension of Sauer's lemma. This result was shown to me by V. Milman. It is due to N. Alon. The proof of our crucial symmetrisation result will use an adaptation of this elegant idea. (See also [1] for other applications).

We consider a finite set I, and $F \subset \{0, 1\}^I$.

Proposition 5. *The set of subsets of I shattered by F has cardinal at least* card F.

The idea is to find a simple operation (called symmetrisation) that will make the set F more regular, while at the same time not decreasing the number of sets shattered by F. One then applies this operation until the set F is so regular that the result is obvious. In the case where F is hereditary (that is $B \in F$ whenever $B \subset A$, for some $A \in F$), F shatters each set it contains, so the result is obvious. The idea of the symmetrisation is to transform F in a set whose elements contain as few points as possible.

Given i in I, we define $T_i(F) = \{T_i(H); H \in F\}$, where for H in F, $T_i(H)$ is defined in the following way:

If $i \notin H$, then $T_i(H) = H$.
If $i \in H$, and $H \setminus \{i\} \in F$, $T_i(H) = H$.
If $i \in H$, and $H \setminus \{i\} \notin F$, $T_i(H) = H \setminus \{i\}$.

Lemma 6. (i) Card $T_i(F) =$ Card F.
(ii) *If $T_i(F)$ shatters A, F shatters A.*

Proof. (i) We show that T_i is one to one on F. Suppose $T_i(H_1) = T_i(H_2)$ for H_1, H_2 in F. Then obviously $H_1 \setminus \{i\} = H_2 \setminus \{i\}$. If $i \in T_i(H_1)$, $i \in H_1, H_2$, so $H_1 = H_2$. If $i \notin T_i(H_1)$, and $H_1 \neq H_2$, we can assume $i \in H_1$, $i \notin H_2$. But since $H_2 = H_1 \setminus \{i\} \in F$, then $T_i(H_1) = H_1$, so $i \in T_i(H_1)$, a contradiction.

(ii) Suppose $T_i(F)$ shatters A. If $i \notin A$, $T_i(F)$ and F have the same trace on A. If $i \in A$, for $B \subset A \setminus \{i\}$, there is $H \in T_i(F)$ with $H \cap A = B \cup \{i\}$. Since $H \in T_i(F)$, we have $H = T_i(G)$ for some $G \in F$. Since $i \in H$, both G and $G \setminus \{i\}$ belong to F, so F shatters A. Q.E.D.

We now prove Proposition 5. Let $w(F) = \sum_{H \in F} \operatorname{card} H$. Let F' such that F' is obtained from F by application of some transformations T_i, and such that $w(F')$ is minimal. Then for each H in F', i in H, we have $H \setminus \{i\} \in F'$, for otherwise $w(T_i(F')) < w(F')$. This means that F' is hereditary. Lemma 6 shows that card F' = card F, and that F shatters more sets than F'. The proposition is proved..

We suppose now $I = \{1, \ldots, N\}$. Denote by $[I]^q$ the set of the subsets of I of cardinal q. Let $l \leqq n$, $m \leqq N$. Let $G \subset [I]^n$. Let $G(l, m) = \{A \in [I]^m;\ \exists H \in G, \operatorname{card}(A \cap H) \geqq l\}$. We are interested in finding lower bounds for card $G(l, m)$ when card G is given. To this end we will compare G with a set with simpler structure.

Definition 7. A set $G \subset [I]^n$ is called left hereditary if whenever $i_1 < \ldots < i_n$ and $\{i_1, \ldots, i_n\} \in G$, then $\{j_1, \ldots, j_n\} \in G$ whenever $j_q \leqq i_q$ for each $1 \leqq q \leqq n$.

Proposition 8. *Let $G \subset [I]^n$. Then there is a left hereditary set G' with* card G' = card G *and* card $G'(l, m) \leqq$ card $G(l, m)$ whenever $l \leqq n$, $m \leqq N$.

The natural idea is to try to move the sets of G toward the left of I. For $i < j$, $H \subset I$, let $S_{ij}(H) = (H \setminus \{j\}) \cup \{i\}$. We consider the symmetrisation $T_{ij}(G) = \{T_{ij}(H); H \in G\}$, where $T_{ij}(H)$ is defined in the following way:

If $i \in H$ or $j \notin H$, then $T_{ij}(H) = H$.
If $i \notin H$, $j \in H$, $S_{ij}(H) \in G$, then $T_{ij}(H) = H$.
If $i \notin H$, $j \in H$, $S_{ij}(H) \notin G$, then $T_{ij}(H) = S_{ij}(H)$.

We note that $T_{ij}(G) \subset [I]^n$. The following lemma contains the basic facts.

Lemma 9. (i) Card $T_{ij}(G) = \text{Card}\, G$.
(ii) Card $T_{ij}(G)(l, m) \leqq \text{card}\, G(l, m)$.

Proof. (i) Suppose $T_{ij}(H_1) = T_{ij}(H_2)$ for H_1, H_2 in G. If either $i \notin T_{ij}(H_1)$ or $j \in T_{ij}(H_1)$, then $H_1 = H_2 = T_{ij}(H_1)$. Suppose now $i \in T_{ij}(H_1)$ and $j \notin T_{ij}(H_1)$, but $H_1 \neq H_2$. Let $U = T_{ij}(H_1)$. Since $H_1, H_2 \in \{U, S_{ji}(U)\}$, we can suppose $H_1 = U$, $H_2 = S_{ji}(U)$, so $U = S_{ij}(H_2) \in G$. This shows that $T_{ij}(H_2) = H_2 \neq U = T_{ij}(H_1)$, a contradiction.

(ii) Let $K = T_{ij}(G)(l, m)$. The idea is to associate to each set A in $K \setminus G(l, m)$ a set B in $G(l, m) \setminus K$, the correspondence being one to one.

Step 1. Let A in $K \setminus G(l, m)$. We show that $i \in A$, $j \notin A$. Since $A \in K$, there is H in $T_{ij}(G)$ with card $H = m$ and $\text{card}(A \cap H) \geqq l$. Let $H = T_{ij}(H')$, where $H' \in G$. Since $A \notin G(l, m)$, we have $H \notin G$. This shows that $i \in H'$, $j \notin H'$ and $H = S_{ij}(H')$, $H' = S_{ji}(H)$. Since $A \notin G(l, m)$, we have $\text{card}(A \cap H') < l$. We have $A \cap (H' \setminus \{i\}) \supset A \cap (H \setminus \{i\})$ so this forces $i \in A$. Also, since $j \in H'$, $j \notin H$, $H \cap \{i, j\} = H' \cap \{i, j\}$, we have $j \notin A$ (otherwise, $\text{card}(A \cap H) = \text{card}(A \cap H')$).

Step 2. Let $B = (A \setminus \{i\}) \cup \{j\} = S_{ji}(A)$. We note that the correspondence $A \to B$ is one-to-one. We show that $B \in G(l, m) \setminus K$. First, $\text{card}(B \cap H') = \text{card}(A \cap H)$, so $B \in G(l, m)$. Suppose, if possible, that $B \in K$. Then there is L in $T_{ij}(G)$ such that $\text{card}(L \cap B) \geqq l$. Let $L' \in G$ with $L = T_{ij}(L')$. Suppose first that $L \neq L'$. This can happen only if $i \notin L'$, $j \in L'$, $L = S_{ij}(L')$. In that case

$$\text{card}(A \cap L') = \text{card}((A \cap L') \setminus \{i, j\}) = \text{card}((B \cap L) \setminus \{i, j\}) = \text{card}(B \cap L) \geqq l$$

and this is impossible since $A \notin G(l, m)$. So we have $L = L'$. Since $\text{card}(A \cap L) < \text{card}(B \cap L)$, we must have $i \notin L$, $j \in L$. Since $L = T_{ij}(L)$, this can happen only if $S_{ij}(L) \in G$. But we have

$$\text{card}(A \cap S_{ij}(L)) = \text{card}(B \cap L) \geqq l$$

and this contradicts $A \in G(l, m)$. We have shown that $B \notin K$. This concludes the proof of the lemma.

We now prove Proposition 8. For $G \subset [I]^n$, let $w(G) = \sum_{H \in G} \sum_{i \in H} i$. Note that for $i < j$, $w(T_{ij}(G)) \leqq w(G)$. Let $G' \in [I]^n$ be a set obtained from G by successive applications of some operations T_{ij} for $i < j$, and for which $w(G')$ is minimal. Then Lemma 9 shows that card $G' = \text{card}\, G$ and card $G'(l, m) \leqq \text{card}\, G(l, m)$ for $l \leqq m$, $n \leqq N$. Let $H = \{i_1, \ldots, i_n\} \in G'$ with $i_1 < \ldots < i_n$. Let $1 \leqq q \leqq n$ and $i_{q-1} < i < i_q$. Let $j = i_q$. Then $S_{ij}(H) \in G'$, for otherwise $w(T_{ij}(G')) < w(G')$. This means that $\{i_1, \ldots, i_{q-1}, i, i_{q+1}, \ldots, i_n\} \in G'$. From this follows that G' is left hereditary. Proposition 8 is proved.

Let us say that a subset G of Ω^n is symmetric if it is invariant by permutation of the coordinates. For a symmetric subset G of Ω^n, and $l \leqq m, n$ let

$$G(l, m) = \{(z_1, \ldots, z_m) \in \Omega^m;\ \exists (x_1, \ldots, x_n) \in G,\ \operatorname{card}\{x_1, \ldots, x_n\} \cap \{z_1, \ldots, z_m\} \geqq l\}.$$

Definition 10. A symmetric subset G of $]0, 1[^n$ is left hereditary if whenever $(x_1, \ldots, x_n) \in G$ and for $1 \leqq i \leqq n$, $0 < y_i < x_i$, then $(y_1, \ldots, y_n) \in G$.

It is easy to see that a left hereditary subset of $[0, 1]^n$ has the same measure as its interior. If G is left hereditary, then $G(l, m)$ is also left hereditary; so $G(l, m)$ is measurable; but in general the measurability of G does not imply the measurability of $G(l, m)$.

Our basic measure theoretic result is a continuous version of Proposition 8. We denote by P_*^m the inner measure associated to P^m.

We denote by λ Lebesgue's measure on $]0, 1[$.

Theorem 11. *Let $G \subset \Omega^n$ be a symmetric measurable set. Then there is a left hereditary set $\hat{G} \subset]0, 1[^n$ such that $\lambda^n(\hat{G}) = P^n(G)$ and for $l \leqq m, n$, $\lambda^m(\hat{G}(l, m)) \leqq P_*^m(G(l, m))$.*

Proof. It is tedious routine to deduce this result from Proposition 8. The first step of the proof is a discretisation procedure (that will allow the use of Proposition 8), the second step is a compactness argument.

Step 1. For $k \geqq 0$, we construct a left hereditary set $G_k \subset [0, 1]^n$, with $\lambda^n(G_k) \leqq P^n(G) + 2^{-k+2}$ and

$$\lambda^m(G_k(l, m)) \leqq P_*^m(G(l, m)) + 2^{-k+2}$$

for $m \leqq 2^k$.

Let $\mathscr{A}$ be a finite subalgebra of Σ such that there exists a $\mathscr{A}^n$ measurable set V with $P^n(G \triangle V) < 2^{-2k}$. Let K be the union of those atoms U of $\mathscr{A}^n$ for which

$$P^n(G \cap U) \geqq (1 - 2^{-k})\, P^n(U).$$

We note that K is symmetrical. If U is an atom of $\mathscr{A}^n$ that is contained in $V \setminus K$, then $P^n(U \cap G) < (1 - 2^{-k})\, P^n(U)$, so $P^n(U \setminus G) \geqq 2^{-k} P^n(U)$. Summation over these atoms U gives $2^{-k} P^n(V \setminus K) \leqq P^n(V \setminus G) \leqq 2^{-2k}$, so $P^n(V \setminus K) \leqq 2^{-k}$, so $P^n(G \setminus K) < 2^{-k+1}$, so $P^n(G) < P^n(K) + 2^{-k+1}$.

For a subset $I = \{i_1, \ldots, i_l\}$ of $\{1, \ldots, n\}$ (resp. $\{1, \ldots, m\}$) let Q_I denote the corresponding projection of Ω^n (resp. Ω^m) on Ω^l. The set $K(l, m)$ is $\mathscr{A}^m$ measurable. Fix an atom W of $\mathscr{A}^m$ such that $W \subset K(l, m)$. So there is $I \subset \{1, \ldots, n\}$ and $J \subset \{1, \ldots, m\}$ with $\operatorname{card} I = \operatorname{card} J = l$, and an atom U of $\mathscr{A}^n$ contained in K with $Q_J(W) = Q_I(U)$. It follows that

$$W \setminus G(l, m) \subset \{(y_1, \ldots, y_m) \in W;\ Q_J(y) \notin Q_I(G \cap U)\}.$$

Since

$$P_*^l(Q_I(G \cap U)) \geqq (1 - 2^{-k})\, P^l(Q_I(U))$$

we get

$$(P^m)^*(W \setminus G(l, m)) \leqq 2^{-k} P^m(W)$$

so

$$P^m(W) \leqq P^m_*(W \cap G(l, m)) + 2^{-k} P^m(W).$$

Summation over $W \subset K(l, m)$ gives $P^m(K(l, m)) \leqq P^m_*(G(l, m)) + 2^{-k}$. It is clear now that there exists N such that if $\mathscr{B}$ denotes the algebra of $]0, 1[$ generated by the partition of $]0, 1[$ in N equal intervals, there exists a symmetric $\mathscr{B}^n$-measurable set L of $]0, 1[^n$ such that $P^n(G) \leqq \lambda^n(L) + 2^{-k+1}$ and $\lambda^m(L(l, m)) \leqq P^m_*(G(l, m)) + 2^{-k+1}$ whenever $l \leqq n$, m, and $m \leqq 2^k$.

We now let $J = \{1, \ldots, N\}$, and denote by $I_1, \ldots, I_N$ the subintervals of $[0, 1]$ that are atoms of $\mathscr{B}$. We denote

$$X = \{\{i_1, \ldots, i_n\} \in [J]^n;\ \prod_{k \leqq n} I_{i_k} \subset L\}.$$

We have

$$L \subset \{\prod_{k \leqq n} I_{i_k};\ \{i_1, \ldots, i_n\} \in X\} \cup \{\prod_{k \leqq n} I_{i_k};\ \exists\, k, l \leqq n,\ i_k = i_l\}$$

so $\lambda^n(L) \leqq \dfrac{n!}{N^n} \operatorname{card} X + \dfrac{n(n-1)}{2} \dfrac{N^{n-1}}{N^n}$.

We can assume N large enough that $\dfrac{n(n-1)}{2N} \leqq 2^{-k}$, so we have $\dfrac{n!}{N^n} \operatorname{card} X \geqq P^n(G) - 2^{-k+2}$. We now use Proposition 8 to find a left hereditary set $Y \subset [J]^n$ with $\operatorname{card} Y = \operatorname{card} X$ and $\operatorname{card}(Y(l, m)) \leqq \operatorname{card}(X(l, m))$ whenever $l \leqq n, m$. We define G_k as the union of the atoms $\prod_{p \leqq n} I_{i_p}$ of $\mathscr{B}^n$ for which there exists $j_1 < \ldots < j_n \leqq N$ such that $\{j_1, \ldots, j_n\} \in Y$ and $i_p \leqq j_p$ for each $p \leqq n$. It is clear that G_k is left hereditary. We have

$$\lambda^n(G_k) \geqq \frac{n!}{N^n} \operatorname{card} Y \geqq P^n(G) - 2^{-k+2}.$$

If an atom $\prod_{p \leqq m} I_{i_p}$ of $\mathscr{B}^m$ belongs to $G_k(l, m)$, then if $i_1, \ldots, i_m$ are distinct, we have $\{i_1, \ldots, i_m\} \in Y(l, m)$. It follows that

$$\begin{aligned}
\lambda^m(G_k(l, m)) &\leqq \frac{m!}{N^m} \operatorname{card} Y(l, m) + \frac{m(m-1)}{2N} \\
&\leqq \frac{m!}{N^m} \operatorname{card} X(l, m) + \frac{m(m-1)}{2N} \\
&\leqq \lambda^m(L(l, m)) + \frac{m(m-1)}{2N} \\
&\leqq P^m_*(G(l, m)) + \frac{m(m-1)}{2N} + 2^{-k+1}.
\end{aligned}$$

Since we can assume N large enough that $m(m-1)/2N \leqq 2^{-k}$, the first step is complete.

Step 2. By extracting a subsequence, we can assume that 1_{G_k} converges weakly in L^2 to some function f.

For $x=(x_i)\in]0,1[^n$, denote S_x the set of $y=(y_i)$ in $]0,1[^n$ such that $x_i<y_i$ for $i\leqq n$. Consider the set $\hat{G}$ of those x for which $\int_{S_x} f\,d\lambda^n>0$. Then $\hat{G}$ is open and left hereditary. If $x\in\hat{G}$, then for k large enough we have $S_x\cap G_k\neq\phi$, and, since G_k is left hereditary, we have $x\in G_k$. This shows that $\hat{G}(l,m)\subset\liminf_k G_k(l,m)$, so we have

$$\lambda^m(\hat{G}(l,m))\leqq\liminf_k \lambda^m(G_k(l,m))\leqq P_*^m(G(l,m))$$

and

$$\lambda^n(\hat{G})\leqq\liminf_k \lambda^n(G_k)\leqq P^n(G).$$

On the other hand, it is clear that $\int_{]0,1[^n\setminus\hat{G}} f\,d\lambda^n=0$. Since $0\leqq f\leqq 1$, $\int f\,d\lambda^n =\lim_k \lambda^n(G_k)=P^n(G)$, we have $\lambda^n(\hat{G})\geqq P^n(G)$ and $\lambda^n(\hat{G})=P^n(G)$. The proof is complete.

4. Compatible Sequences of Sets

For each n consider a measurable subset V_n of Ω^n.

Definition 12. We say that the sequence (V_n) is compatible if each set V_n is measurable, invariant by permutation of the coordinates, and if for $m<n$ the projection of V_n on Ω^m is contained in V_m.

Compatible sequences of sets where considered by D. Fremlin in an unpublished paper, for purposes somewhat similar to those of this work. If $\mathscr{C}$ is a class of sets that satisfies condition (M), and V_n is the set of n-tuples shattered by $\mathscr{C}$, then the sequence (V_n) is compatible.

If (V_n) is a compatible sequence of sets, let

$$\mathscr{F}_n=\{1_{\{x_1,\ldots,x_n\}};(x_1,\ldots,x_n)\in V_n\}$$

and $\mathscr{F}=\cup\mathscr{F}_n$. Then a n-tuple $(y_1,\ldots,y_n)$ is shattered by $\mathscr{F}$ if and only if it belongs to V_n. This remark shows that the sequences (V_n) that arise as (V_n) being the set of n-tuples shattered by a class $\mathscr{C}$ are exactly the compatible sequences.

Proposition 13. *If (V_n) is a compatible sequence of sets, then the sequence $r_n=(P^n(V_n))^{1/n}$ is non-increasing.*

Proof. This proof was shown to me by J. Bourgain. It is enough to show that $P^n(V_n)\geqq P^{n+1}(V_{n+1})^{n/n+1}$. For simplicity we shall assume $n=2$. The general case uses the same method. Let $f=1_{V_2}$, $g=1_{V_3}$. The compatibility condition implies

$g(x_1, x_2, x_3) \leqq f(x_1, x_2) f(x_2, x_3) f(x_1, x_3)$. It is hence enough to show that for f_1, f_2, f_3 in $L^2(\Omega^2)$, all positive, we have

$$E = \int f_1(x_2, x_3) f_2(x_1, x_3) f_3(x_1, x_2) \, dP(x_1) \, dP(x_2) \, dP(x_3) \leqq \prod_{i \leqq 3} \|f_i\|_2.$$

Cauchy-Schwartz's inequality for the variables x_1, x_3, x_2 in that order gives

$$\begin{aligned} E &\leqq \int f_1(x_2, x_3) (\int f_2^2(x_1, x_3) \, dP(x_1))^{1/2} (\int f_3^2(x_1, x_2) \, dP(x_1))^{1/2} \, dP(x_2) \, dP(x_3) \\ &\leqq \|f_2\|_2 \int (\int f_1^2(x_2, x_1) \, dP(x_3))^{1/2} (\int f_3^2(x_1, x_2) \, dP(x_1))^{1/2} \, dP(x_2) \\ &\leqq \|f_1\|_2 \, \|f_2\|_2 \, \|f_3\|_2. \end{aligned}$$

The proof is complete.

For x in Ω', denote by $T_n(x)$ the largest integer k such that there is $i_1 < i_2 < \ldots < i_k \leqq n$ for which $(x_{i_1}, \ldots, x_{i_k}) \in V_k$. Our first objective is the following.

Theorem 14. *For a compatible sequence of sets, the following are equivalent*

(I) $\lim_{n \to \infty} r_n/\phi(n) = 0$.

(II) $\forall \gamma > 0, \lim_{n \to \infty} P'(\{T_n(x) \geqq \gamma a_n\}) = 0$.

We recall the following inequality proved in [6], 4–7: $\binom{n}{m} \leqq \left(\frac{en}{m}\right)^m$ whenever $m \leqq n$.

Proof that (I)$\Rightarrow$(II). We show that if for some m, k we have

$$r_m \leqq 2^{-2^k} (2c)^{-k} \, \phi(m)/e$$

then whenever n satisfies $2^{-k} a_n \leqq m \leqq 2^{-k+1} a_n$ we have

$$P'(\{T_n(x) \geqq 2^{-k+1} a_n\}) \leqq 2^{-a_n}.$$

We have

$$P'(\{T_n(x) \geqq m\}) \leqq \binom{n}{m} P^m(V_m) \leqq \left(\frac{e n r_m}{m}\right)^m \leqq \left(\frac{n}{m} \, 2^{-2^k} (2c)^{-k} \, \phi(m)\right)^m.$$

We note that from condition (*),

$$\phi(m)/m \leqq \phi(2^{-k} a_n)/2^{-k} a_n \leqq (2c)^k \, \phi(a_n)/a_n.$$

Moreover, $\phi(a_n)/a_n = 1/n$, so we get

$$P'(\{T_n(x) \geqq m\}) \leqq (2^{-2^k})^m = 2^{-m 2^k} \leqq 2^{-a_n}.$$

The proof is complete.

Since condition (*) implies that $\phi(t) \geqq \alpha t^{-\beta}$ for some $\alpha, \beta > 0$, we have that $\lim_n a_n = \infty$. It is then obvious that (II) follows from (I).

Proof that (II)$\Rightarrow$(I). The difficulty here is that even if $r_m/\phi(m)$ is not small $P'(\{T_n(x)\geqq\gamma a_n\})$ can be small when γa_n is of the order of m. Let $\varepsilon>0$. Let $d=c^a$, where $a=\sum_{i>0} i2^{-i}$. We fix $k_0>0$ such that for $m\geqq 2^{k_0}$ we have

$$P'\left(\left\{T_m(x)\geqq\frac{\varepsilon 2^{-6}}{cd}\,a_m\right\}\right)\leqq 1/2. \tag{1}$$

We now let p large enough that for $k\geqq p/2-1$, we have

$$\left(\frac{\varepsilon\phi(1)}{dc^{k+1}}\right)^{2^{k_0}}>2^{-2^k+2^{k_0}}. \tag{2}$$

We prove that for $n\geqq p$ we have $r_n/\phi(n)<\varepsilon$.

Suppose, if possible, that $r_n\geqq\varepsilon\phi(n)$ for some $n\geqq p$. Let k be the largest integer with $2^k\leqq n$, so $n<2^{k+1}$. We get

$$r_{2^k}\geqq r_n\geqq\varepsilon\phi(n)\geqq\varepsilon\phi(2^{k+1})\geqq\frac{\varepsilon}{c}\,\phi(2^k). \tag{3}$$

We have, using condition (*):

$$\phi(2^k)^{2^k}=\phi(2^k)^{2^{k_0}}\prod_{k_0\leqq i<k}\phi(2^k)^{2^i}\geqq\left(\frac{\phi(1)}{c^k}\right)^{2^{k_0}}\prod_{k_0\leqq i<k}c^{-(k-i)2^i}\,\phi(2^i)^{2^i}.$$

We have

$$\sum_{k_0\leqq i<k}-(k-i)\,2^i\geqq 2^k\sum_{i\geqq 0}-i2^{-i}=-2^k a$$

so

$$\prod_{k_0\leqq i<k}c^{-(k-i)2^i}\geqq d^{-2^k}.$$

It follows that

$$\phi(2^k)^{2^k}\geqq\left(\frac{\phi(1)}{dc^k}\right)^{2^{k_0}}\prod_{k_0\leqq i<k}d^{-2^i}\,\phi(2^i)^{2^i}.$$

From (3) we get

$$P^{2^k}(V_{2^k})\geqq\left(\frac{\varepsilon\phi(1)}{dc^{k+1}}\right)^{2^{k_0}}\prod_{k_0\leqq i<k}\left(\frac{\varepsilon}{cd}\,\phi(2^i)\right)^{2^i}.$$

Using (2) we get

$$P^{2^k}(V_{2^k})>\prod_{k_0\leqq i<k}\left(\frac{\varepsilon}{2cd}\,\phi(2^i)\right)^{2^i}. \tag{4}$$

We now use Theorem 11 to find a left hereditary set $Y\subset\,]0,1[^{2^k}$ such that $\lambda^{2^k}(Y)=P^{2^k}(V_{2^k})$ and $\lambda^m(Y(l,m))\leqq\lambda^m(V_{2^k}(l,m))$ whenever $l\leqq m,n$. We first prove that there exists $k_0\leqq i<k$ and $y=(y_1,\ldots,y_{2^k})$ in Y such that $y_{2^k}<\ldots<y_1$ and

$y_{2^i} > \frac{e}{8cd}\,\phi(2^i)$. Otherwise for each y in Y, one can find for $k_0 \leqq i < k$ disjoint sets A_i of $\{1, \ldots, 2^k\}$, with card $A_i = 2^i$, such that for l in A_i we have $y_l < \frac{\varepsilon}{8cd}\,\phi(2^i)$. It follows that

$$\lambda^{2^k}(Y) \leqq \prod_{k_0 \leqq i < k} \binom{2^{i+1}}{2^i} \left(\frac{\varepsilon}{8cd}\,\phi(2^i)\right)^{2^i}.$$

Since $\binom{2^{i+1}}{2^i} \leqq 2^{2^{i+1}} = 4^{2^i}$, we get

$$\lambda^{2^k}(Y) \leqq \prod_{k_0 \leqq i < k} \left(\frac{\varepsilon}{2cd}\,\phi(2^i)\right)^{2^i}.$$

This contradicts (4), and proves the claim. The crucial point now is that since Y is left hereditary, for each $z = (z_1, \ldots, z_m)$ such that

$$\operatorname{card}\left(\{z_1, \ldots, z_m\} \cap \left[0, \frac{\varepsilon}{2cd}\,\phi(2^i)\right]\right) \geqq 2^i$$

we have $z \in Y(2^i, m)$. Letting $b = \frac{\varepsilon}{2cd}\,\phi(2^i)$, we have (using the Remark 4–7 of [5]) that for each m,

$$1 - \lambda^m(Y(2^i, m)) \leqq \sum_{j < 2^i} \binom{m}{j} b^j (1-b)^{m-j} \leqq 2^i \left(\frac{bem}{(1-b)\,2^i}\right)^{2^i} (1-b)^m.$$

Take m the integer part of $2^{i+5}/b$. We can assume k_0 large enough that $mb \geqq 2^{i+4}$. Note that $(1-b)^m \leqq e^{-mb}$ so $(1-b)^m \leqq (e^{-16})^{2^i}$. Since $b \leqq 1/2$, we get

$$\lambda^m(Y(2^i, m)) \geqq 1 - 2^i (2^6 e^{-15})^{2^i} \geqq 1 - 2^i e^{-9.2^i} > 1/2.$$

So we also have

$$P_*^m(V_{2^k}(2^i, m)) \geqq 1/2.$$

For $x \in V_{2^k}(2^i, m)$, we have $T_m(x) \geqq 2^i$. Since

$$m \leqq 2^{i+5}/b \leqq 2^{i+6}\, cd/(\varepsilon\phi(2^i)) \leqq 2^{i+6}\, cd/(\varepsilon\phi(2^{i+6}\, cd/\varepsilon))$$

and since we have $a_m/\phi(a_m) \leqq m$, we have $a_m \leqq 2^{i+6}\, cd/\varepsilon$, so $2^i \geqq \varepsilon 2^{-6}\, a_m/cd$. This violates condition (1) and finishes the proof.

Theorem 14 implies that in Theorem 2, (I) and (II) are equivalent. If $\mathscr{C}$ shatters a set of cardinal m, its trace on this set has cardinal 2^m; so it is obvious that (III) implies (II). We shall now prove the converse. For further use, we state a more precise result.

Proposition 15. *Let $\gamma>0$. Then there is $\varepsilon>0$ such that if*

$$r_n=(P^n(V_n))^{1/n}\leqq\varepsilon\phi(n) \quad \text{for } n\geqq m$$

then there is a sequence (u_n) going to zero, that depends on γ and m but not on the probability space, such that

$$P^n(\{\ln\Delta_n(x)\geqq\gamma a_n\})\leqq u_n.$$

Proof. For x in Ω', denote by $R_n^q(x)$ the set of subsets of $\{x_1,\ldots,x_n\}$ of cardinal q that are shattered by $\mathscr{C}$. It follows from Proposition 5 that $\Delta_n(x)\leqq\sum_{i\leqq n}\operatorname{card}R_n^i(x)$. Let

$$\mu_{n,q}(x)=\binom{n}{q}^{-1}\sum_{i_1<\ldots<i_q}\delta_{x_{i_1}}\ldots\delta_{x_{i_q}} \tag{5}$$

where the summation is taken over all choices $1\leqq i_1<\ldots<i_q\leqq n$. We note that for a measurable set $A\subset\Omega^q$, we have

$$P^q(A)=\int\mu_{n,q}(x)(A)\,dP'(x). \tag{6}$$

Let l be large enough that $\gamma>2^{-l}\ln(e^2c^{1/\ln 2})$. Let $\varepsilon=(2c)^{-l}/2e$. Let $m'\geqq m$. Since $\phi(t)\geqq\alpha t^{-\beta}$ for some $\alpha,\beta>0$, there is $n_0\geqq m'$ such that $n^{m'}\leqq e^{2^{-l}a_n}$ for $n\geqq n_0$. We fix $n\geqq n_0$. For $m'<q\leqq n$, let

$$B_q=\{\mu_{n,q}(x)(V_q)\leqq(2\varepsilon\phi(q))^q\}.$$

It follows from (6) that $P'(B_q)\geqq 1-2^{-q}$. Let $B=\bigcap_{m'<q\leqq n}B_q$, so $P'(B)\geqq 1-2^{-m'}$. We show that for x in B, we have $\Delta_n(x)\leqq e^{\gamma a_n}$. For $i\leqq m'$, we have

$$\operatorname{card}R_n^i(x)\leqq\binom{n}{i}\leqq n^i\leqq n^{m'}\leqq e^{2^{-l}a_n}. \tag{7}$$

For $i>m'$, since $x\in B_q$, we have

$$\operatorname{card}R_n^i(x)\leqq\binom{n}{i}\mu_{n,i}(x)(V_i)\leqq(2\varepsilon en\phi(i)/i)^i. \tag{8}$$

Let t be the root of the equation $2\varepsilon en\phi(t)/t=1$. We have (since $2\varepsilon e=(2c)^{-l}$)

$$n\phi(a_n)/a_n=1=(2c)^{-l}n\phi(t)/t\leqq n\phi(2^l t)/2^l t$$

so $2^l t\leqq a_n$, and $t\leqq 2^{-l}a_n$. It follows from condition $(*)$ that for $i\leqq t$ we have $\phi(i)\leqq c^{1+\log_2(t/i)}\phi(t)$. Moreover,

$$1+\log_2(t/i)\leqq(1+\ln(t/i))/\ln 2\leqq t/i\ln 2,$$

so

$$(\phi(i))^i\leqq c^{t/\ln 2}\phi(t)^i. \tag{9}$$

We note that for $i>t$, we have card $R_m^i(x)<1$, so card $R_m^i(x)=0$. It follows from (7), (8), (9), that

$$\sum_{i\leqq n} \operatorname{card} R_m^i(x) \leqq m' e^{2^{-l} a_n} + \sum_{m'<i\leqq t} c^{t/\ln 2}(2\varepsilon e n \phi(t)/i)^i.$$

Now $(2\varepsilon e n \phi(t)/i)^i \leqq e^{2\varepsilon n \phi(t)} = e^{t/e}$, so

$$\begin{aligned}\sum_{i\leqq n} \operatorname{card} R_m^i(x) &\leqq m' e^{2^{-l} a_n} + (t-m')\, c^{t/\ln 2}\, e^{t/e}\\ &\leqq m' e^{2^{-l} a_n} + (t-m')(c^{1/\ln 2}\, e)^{2^{-l} a_n}\\ &\leqq t(c^{1/\ln 2}\, e)^{2^{-l} a_n} \leqq (c^{1/\ln 2}\, e^2)^{2^{-l} a_n} \leqq e^{\gamma a_n}.\end{aligned}$$

The proof is complete.

5. Donsker Classes

The main aim of this section is the proof of Theorem 3.

Proposition 16. *If $\mathscr{C}$ satisfies condition (M) and is a Donsker class, then the conditions of Theorem 2 hold.*

Proof. The proof is a adaptation of the argument of [6], Remark 6.5. Suppose that $\mathscr{C}$ is a Donsker class and satisfies condition (M). Let $\varepsilon>0$. Then Dudley's equicontinuity criterion [4] shows that there is $\delta>0$ such that

$$\limsup_{n\to\infty} P'^*(\{x;\ \operatorname{Sup}_{f\in Z} |\sum_{i\leqq n} f(x_i)| \geqq \varepsilon n^{1/2}\}) < \varepsilon \tag{10}$$

where $Z=\{1_A - 1_B;\ A, B\in\mathscr{C},\ P(A\triangle B)\leqq\delta\}$.

We know also that $\mathscr{C}$ is precompact in measure, so there exists N, and for $j\leqq N$ there exists (A_j) in $\mathscr{C}$ such that

$$\forall B\in\mathscr{C},\ \exists j\leqq N,\ P(A_j\triangle B)\leqq\delta.$$

Let $(Y, \equiv, Q)$ be another probability space, and ε_i be an independent Bernoulli sequence on X. An obvious adaptation of the argument of [6], Lemma 2–7, a), to the non-measurable case shows that for $t>0$,

$$\begin{aligned}&(P'\otimes Q)^*(\{(x,y);\ \operatorname{Sup}_{f\in Z} |\sum_{i\leqq n} \varepsilon_i(y) f(x_i)| \geqq 2t\})\\ &\quad\leqq 2 \operatorname{Max}_{k\leqq n} P'^*(\{x;\ \operatorname{Sup}_{f\in Z} |\sum_{i\leqq n} f(x_i)| \geqq t\}).\end{aligned}$$

It follows that for n large enough,

$$(P'\otimes Q)^*(\{(x,y);\ \operatorname{Sup}_{f\in Z} |\sum_{i\leqq n} \varepsilon_i(y) f(x_i)| \geqq 2\varepsilon n^{1/2}\}) < 2\varepsilon. \tag{11}$$

Suppose now that $T_n(x) \geqq 12\varepsilon n^{1/2}$, that is, there is a subset I of $\{x_1, \ldots, x_n\}$ of cardinal $\geqq 12\varepsilon n^{1/2}$ that is shattered by $\mathscr{C}$. Then there is $j \leqq N$ such that the trace over I of $\mathscr{C}' = \{B \in \mathscr{C},\ P(A_j \triangle B) \leqq \delta\}$ has a cardinal $\geqq 2^{\operatorname{card} I}/N$. For n large enough, we have

$$2^{\operatorname{card} I}/N \geqq \sum_{i \leqq 4\varepsilon n^{1/2}} \binom{\operatorname{card} I}{i}.$$

It follows from Sauer's lemma (Proposition 5) that there is a subset J of I with $\operatorname{card} J \geqq 4\varepsilon n^{1/2}$ that is shattered by $\mathscr{C}'$. One then sees easily that for each y,

$$\sup_{B \in \mathscr{C}'} \Big|\sum_{i \leqq n} \varepsilon_i(y)(1_A - 1_B)(x_i)\Big| \geqq 2\varepsilon n^{1/2}.$$

It follows that (11) implies that

$$P'(\{T_n(x) \geqq 12\varepsilon n^{1/2}\}) \leqq 2\varepsilon$$

for n large enough. This concludes the proof.

The rest of Theorem 3 is a consequence of [6], Theorem 5.1. Given $A \in \mathscr{C}$ and $\delta > 0$, denote $T_n(A, \delta, x)$ the largest integer k such that there is a subset of $\{x_1, \ldots, x_n\}$ of cardinal k that is shattered by the class $\mathscr{C}(A, \delta) = \{B \in \mathscr{C};\ P(A \triangle B) \leqq \delta\}$. Denote by $V_n(A, \delta)$ the set of n-tuples $(x_1, \ldots, x_n)$ that are shattered by the class $\mathscr{C}(A, \delta)$. The following result has been inspired by [6], Theorem 4–8. It provides weaker sufficient conditions for Theorem 3.

Proposition 17. *Let $\mathscr{C}$ be a countable class of measurable sets. Assume that $\mathscr{C}$ is pregaussian. Then $\mathscr{C}$ is a donsker class under either of the following conditions:*

(a) $$\lim_{\tau \to 0} \limsup_{n \to \infty} n\Big(P'\Big(\bigcup_{A \in \mathscr{C}} V_n(A, \tau/n)\Big)\Big)^n = 0$$

(b) $$\forall \gamma > 0,\ \lim_{\tau \to 0} \limsup_{n \to \infty} P'(\{\sup_{A \in \mathscr{C}} T_n(x, A, \tau n^{-1/2}) \geqq \gamma n^{1/2}\}) = 0.$$

Proof. The crucial point is the following. Denote by $N(\delta)$ the cardinal of a smallest subfamily F_δ of $\mathscr{C}$ such that for each B in $\mathscr{C}$, there is A in F_δ with $P(A \triangle B) \leqq \delta$. Then if $\mathscr{C}$ is pregaussian $\lim_{\delta \to 0} \delta \ln N(\delta) = 0$ (Sudakov's minoration, [6], 2–26). We first prove that (b) is sufficient.

Fix $\gamma > 0$. Let $\varepsilon > 0$, and let $\tau > 0$ be such that

$$\limsup_{n \to \infty} P'(\{\sup_{A \in \mathscr{C}} T_n(x, A, \tau n^{-1/2}) \geqq \gamma n^{1/2}\}) \leqq \varepsilon.$$

It is enough to show that for n large enough one has

$$\sup_{A \in \mathscr{C}} T_n(x, A, \tau n^{-1/2}) \leqq \gamma n^{1/2} \Rightarrow T_n(x) \leqq 3\gamma n^{1/2}.$$

We fix n large enough that $\tau n^{-1/2}\ln N(\tau n^{-1/2})\leqq\gamma\tau\ln 2$, so $\ln N(\tau n^{-1/2})\leqq\gamma n^{1/2}\ln 2$. So, there is a set $F\subset\mathscr{C}$, of cardinal $\leqq 2^{\gamma n^{1/2}}$, such that for each B in $\mathscr{C}$, there is A in F with $P(A\triangle B)\leqq\tau n^{-1/2}$. Suppose now that $T_n(x)\geqq 3\gamma n^{1/2}$. Then there is a subset T of $\{x_1,\dots,x_n\}$ such that $\operatorname{card} I\geqq 3\gamma n^{1/2}$ and the trace of $\mathscr{C}$ on I has cardinal $2^{\operatorname{card} I}$. It follows that there is A in F such that the trace on I of $\mathscr{C}(A,\tau n^{-1/2})$ has cardinal $\geqq 2^{\operatorname{card} I-\gamma n^{1/2}}\geqq 2^{2\operatorname{card} I/3}$. Now for n large enough Sauer's lemma implies that $\mathscr{C}(A,\tau n^{-1/2})$ shatters a set of cardinal $\geqq\operatorname{card} I/3$. This completes the proof that (b) is sufficient. To prove that (a) is sufficient, one shows, as in the proof of theorem 14, I$\Rightarrow$II, that (a)$\Rightarrow$(b).

It is of interest to note the following necessary conditions, that are formally much stronger than the conditions of Proposition 17.

Proposition 18. *Let $\mathscr{C}$ be a countable Donsker class. Then the following holds:*

(a) $\exists\tau>0,\ \forall\varepsilon>0,\ \lim_{n\to 0} P'(\{\exists A, B\in\mathscr{C}, P(A\triangle B)\leqq\tau\varepsilon n^{-1/2}, \sum_{i\leqq n} 1_{A\triangle B}(x_i)\geqq\varepsilon n^{1/2}\})=0,$

(b) $\lim_{\gamma\to 0}\limsup_{n\to\infty} n P^n(\cup\{(A\triangle B)^n;\ A, B\in\mathscr{C}, P(A\triangle B)\leqq\gamma/n\})^{1/n}=0.$

Proof. (a) is proved in [6], Remark 4–5.

(b) Suppose, if possible, that

$$\lim_{\gamma\to 0}\limsup_{m\to\infty} n P^n(\cup\{(A\triangle B)^n;\ A, B\in\mathscr{C}, P(A\triangle B)\leqq\gamma/n\})^{1/n}\geqq\alpha>0.$$

Then for each $\gamma>0$,

$$\limsup_{n\to\infty} n P^n(V_n(\gamma))^{1/n}\geqq\alpha$$

where

$$V_n(\gamma)=\cup\{(A\triangle B)^n;\ A, B\in\mathscr{C},\ P(A\triangle B)\leqq\gamma/n\}.$$

The sequence $(V_n(\gamma))_n$ is compatible. Inspection of the proof of Theorem 14 shows that there is ε, depending of α but not of γ, such that for each $\gamma>0$,

$$\limsup_n P'(\{T_n(x,\gamma)\geqq\varepsilon n^{1/2}\})>0.$$

where $T_n(x,\gamma)$ is the largest integer k such that there exists $i_1<\dots<i_k\leqq n$ with $(x_{i_1},\dots,x_{i_k})\in V_k(\gamma)$. Let m be the smallest integer $\geqq\varepsilon n^{1/2}$. If $T_n(x,\gamma)\geqq\varepsilon n^{1/2}$, then $T_n(x,\gamma)\geqq m$, so there exists A, B in $\mathscr{C}$ with $P(A\triangle B)\leqq\gamma/m\ll\gamma/\varepsilon n^{1/2}$ and $\sum_{i\leqq n} 1_{A\triangle B}(x_i)\geqq\varepsilon n^{1/2}$. For $\gamma=\tau\varepsilon^2$, this violates (a) and concludes the proof.

6. Glivenko-Cantelli Classes

The author has recently given a characterisation of Glivenko-Cantelli classes of functions, that seems more precise than any other known characterisation

[10]. This approach of [10] is purely qualitative, and the arguments are purely measure theoretic. The use of the quantitative arguments of Sect. 3 will allow very precise estimates of the rate of convergence for classes of sets. Let ϕ, a_n be as in Sect. 4.

Let $\mathscr{C}$ be a class of sets. We shall always assume that $\mathscr{C}$ satisfies condition (M). The arguments of Proposition 16 show that

Proposition 19. *Suppose that for each* $\gamma>0$,

$$\lim_n P'^*\left(\left\{\operatorname{Sup}_{A\in\mathscr{C}}\left|\frac{1}{n}\sum_{i\leq n} 1_A(x_i)-P(A)\right|\geqq\gamma\, a_n/n\right\}\right)=0.$$

Then for each $\gamma>0$

$$\lim_{n\to\infty} P'(\{T_n(x)\geqq\gamma\, a_n\})=0.$$

In particular, the conditions of Theorem 2 hold.

Lemma 20. *If condition* (**) *holds,* $\lim_{k\to\infty} a_{2^k}2^{-k/2}=\infty$.

Proof. Otherwise, condition (**) shows that the series $\sum_l a_{2^l}2^{-l/2}$ is summable so $\lim_{k\to\infty} a_{2^k}2^{-k/2}=0$, and this contradicts (**).

To complete the proof of Theorem 4 we show that under condition (**), the converse of Proposition 19 holds.

Step 1. Let $\gamma>0$. Let ε be as in Proposition 15. Let m be large enough that $r_n\leqq\varepsilon\,\phi(n)/2$ for $n\geqq m$, and that $2^{-m+1}\leqq\gamma$.

Let $n\geqq m$, and for $m\leqq q\leqq n$, let

$$B_q=\{x\in\Omega';\ \mu_{n,q}(x)(V_q)\leqq(\varepsilon\,\phi(q))^q\}.$$

As seen in the proof of Proposition 15, we have $P'(B_q)\geqq 1-2^{-q}$. Let $A_n=\bigcup_{m<q<n} B_q$, so $P'(A_n)\geqq 1-\gamma$. For x in A_n, and $m\leqq q\leqq n$, we have $\mu_{n,q}(x)(V_q)\leqq(\varepsilon\,\phi(q))^q$. Let $\nu_n(x)=\frac{1}{n}\sum_{i\leq n}\delta_{x_i}$. Since V_q consists of q-tuples of points that are all distincts, we have

$$\nu_n^q(x)(V_q)\leqq\frac{\binom{n}{q}}{n^q}\,\mu_{n,q}(x)(V_q)\leqq(\varepsilon\,\phi(q))^q$$

whenever x is in A_n and $m\leqq q\leqq n$. For $q\geqq n$, we have $\nu_n^q(x)(V_q)=0$. We now apply Proposition 15 to the probability $\nu_n(x)$. There is k_1 that depends on γ and m, but not on n or x, such that whenever $l\geqq 2^{k_1}$, we get

$$\nu_n'(x)(\{y\in\Omega';\ \ln\Delta_l(y)>\gamma\, a_l\})\leqq 1/2, \tag{12}$$

where $v'_n(x)$ is the power of $v_n(x)$ on Ω'. We can also assume that k_1 is large enough that $\sum_{l \geqq k_1} 2\exp(-\gamma a_{2^l}) \leqq \gamma$.

For each $\delta > 0$, denote $N_{n,x}(\delta)$ the largest p such that there are elements $C_1, \ldots, C_p$ of $\mathscr{C}$ such that $v_n(x)(C_i \triangle C_j) \geqq \delta$ for $i < j \leqq p$. Fix δ, let $p = N_{n,x}(\delta)$, and let $C_1, \ldots, C_p$ elements of $\mathscr{C}$ with $v_n(x)(C_i \triangle C_j) \geqq \delta$ for $i < j \leqq p$.

Let s be an integer with $p(p-1)(1-\delta)^s < 1$. Then

$$v'_n(x)(\{y \in \Omega', \forall i < j \leqq p, \exists k \leqq s, y_k \in C_i \triangle C_j\})$$
$$\geqq 1 - \frac{p(p-1)}{2}(1-\delta)^s > 1/2.$$

If $s \geqq 2^{k_1}$, it follows from (12) that $\ln p \leqq \gamma a_s$, since if $\{y_1, \ldots, y_s\}$ meets each set $C_i \triangle C_j$, the trace of $\mathscr{C}$ on $\{y_1, \ldots, y_s\}$ has cardinal $\geqq p$. In other words, we have shown that for $s \geqq 2^{k_1}$, we have

$$N_{n,x}(\delta)^2 (1-\delta)^s < 1 \Rightarrow N_{n,x}(\delta) \leqq \exp \gamma a_s,$$

or

$$N_{n,x}(\delta) > \exp \gamma a_s \Rightarrow N_{n,x}(\delta)^2 (1-\delta)^s \geqq 1.$$

Since $1 - \delta \leqq \exp - \delta$, we get

$$N_{n,x}(\delta) > \exp \gamma a_s \Rightarrow N_{n,x}(\delta) \geqq \exp \delta s/2.$$

If s is the largest such that $N_{n,x}(\delta) > \exp \gamma a_s$, we get $\exp \delta s/2 \leqq N_{n,x}(\delta) \leqq \exp \gamma a_{s+1}$, so $\delta \leqq 2\gamma a_{s+1}/s$.

So we have shown that for $s > 2^{k_1}$, we have

$$\delta \geqq 2\gamma a_{s+1}/s \Rightarrow N_{n,x}(\delta) \leqq \exp \gamma a_s.$$

Since as easily seen, $2a_s \geqq a_{2s} \geqq a_{s+1}$, we have for $s \geqq 2^{k_1}$ that

$$\delta \geqq 4\gamma a_s/s \Rightarrow N_{n,x}(\delta) \leqq \exp \gamma a_s.$$

Step 3. Let b as in condition (**), so $\sum_{l \leqq k} a_{2^l} 2^{-l/2} \leqq b a_{2^k} 2^{-k/2}$. Recall that $(\varepsilon_i)_{i \leqq n}$ is an independent sequence of Bernoulli random variables defined on the space $(Y, \equiv, Q)$. Fixing n and $x \in P$, we estimate by a chaining argument the Q probability that

$$\sup_{C \in \mathscr{C}} \left|\sum_{i \leqq n} \varepsilon_i 1_C(x_i)\right| \geqq 8\gamma a_n (2+b).$$

Let k be the largest integer with $2^k \leqq n$. For $k_1 \leqq l \leqq k$, let $\delta_l = 2^{2-l} \gamma a_{2^l}$, and let F_l be a subset of $\mathscr{C}$ of card$\leqq \exp \gamma a_{2^l}$ such that for each C in $\mathscr{C}$, there is B in F_l with $v_n(x)(B \triangle C) \leqq \delta_l$.

The heart of the matter is that the process $C \to \sum_{i \leq n} \varepsilon_i 1_C(x_i)$ is subgaussian for the distance $d(A, B) = (n\, v_n(x)(A \triangle B))^{1/2}$, and that

$$\sum_{k_1 \leq l \leq k} \delta_l^{1/2} (\log \operatorname{card} F_l)^{1/2} \leq 2\gamma \sum_{l \leq k} 2^{-l/2} a_{2^l} \leq 2\gamma\, b\, a_{2^k} 2^{-k/2}.$$

The bound we need could be deduced from a suitable version of Dudley's metric entropy bound for the supremum of a Gaussian process, but unfortunately the version we need does not seem to appear in the literature. It can also be deduced from the usual version of Dudley's metric entropy bound and the known concentration properties of the supremum of a Gaussian process, as follows from Borell's inequality or weaker principles as in [13]. It is, however, just as short in that case to give the complete argument.

For C in $\mathscr{C}$ there is B in F_k with

$$v_n(x)(B \triangle C) \leq 4\gamma\, a_{2^k}/2^k \leq 8\gamma\, a_n/n$$

that is card $\{i \leq n,\, x_i \in B \triangle C\} \leq 8\gamma\, a_n$. So we have

$$\operatorname*{Sup}_{C \in \mathscr{C}} |\sum \varepsilon_i 1_C(x_i)| \leq 8\gamma\, a_n + \operatorname*{Sup}_{C \in F_k} \div \sum_{i \leq n} \varepsilon_i 1_C(x_i)|.$$

For $k_1 < l \leq k$ denote G_l the family of functions $1_C - 1_B$, for C in F_l, where B depends of C and is an element of F_{l-1} such that $v_n(x)(C \triangle B) \leq \delta_{l-1} \leq 2\delta_l$. We note that

$$\operatorname{card} G_l \leq \operatorname{card} F_l \leq \exp \gamma\, a_{2^l},$$

while for f in G_l, we have that $|f|$ takes only the values 0 and 1, and $v_n(x)(f) \leq \delta_l$.

We have

$$\operatorname*{Sup}_{C \in F_k} |\sum_{i \leq n} \varepsilon_i 1_C(x_i)| \leq \sum_{k_1 < l \leq k} \operatorname*{Sup}_{f \in G_l} |\sum_{i \leq n} \varepsilon_i f(x_i)| + \operatorname*{Sup}_{C \in F_{k_1}} |\sum_{i \leq n} \varepsilon_i 1_C(x_i)|.$$

It follows that

$$Q(\{\operatorname*{Sup}_{C \in F_k} |\sum_{i \leq n} \varepsilon_i 1_C(x_i)| \geq 8\gamma\, a_n(2+b)\}) \leq \sum_{k_1 \leq l \leq k} \alpha_l$$

where

$$\alpha_{k_1} = Q(\{\operatorname*{Sup}_{C \in F_{k_1}} |\sum_{i \leq n} \varepsilon_i 1_C(x_i)| \geq 8\gamma\, a_n\})$$

and, for $k_1 < l \leq k$

$$\alpha_l = Q(\{\operatorname*{Sup}_{f \in G_l} |\sum_{i \leq n} \varepsilon_i f(x_i)| \geq 8\gamma\, a_n\, 2^{(k-l)/2} a_{2^l}/a_{2^k}\}).$$

Since $a_{2^k} \leq a_n$ and $n \leq 2^{k+1}$, we have

$$\alpha_l \leq \operatorname{card} G_l \operatorname*{Sup}_{f \in G_l} Q(\{|\sum_{i \leq n} \varepsilon_i f(x_i)| \geq 4\gamma\, a_{2^l} n^{1/2} 2^{-l/2+1/2}\}).$$

We note now that $|\sum_{i\leqq n}\varepsilon_i f(x_i)|$ has the same distribution as $|\sum_{i\leqq n}\varepsilon_i |f|(x_i)|$. We note that there are at most $n2^{3-l}\gamma a_{2^l}$ indexes i for which $|f|(x_i)=1$. Using the standard subgaussian inequality

$$Q(\{|\sum_{i\leqq r}\varepsilon_i|\geqq t\})\leqq 2\exp -t^2/2r,$$

we get that

$$\alpha_l\leqq 2\exp\left(\gamma a_{2^l}-\frac{32\gamma^2 a_{2^l}^2 n 2^{-l}}{2n\gamma a_{2^l}2^{3-l}}\right)\leqq 2\exp(-\gamma a_{2^l}).$$

Moreover

$$\alpha_{k_1}\leqq \operatorname{card} F_{k_1}\sup_{C\in F_{k_1}} Q(\{|\sum_{i\leqq n}\varepsilon_i 1_C(x_i)|\geqq 8\gamma a_n\})$$

$$\leqq \exp\left(\gamma a_{2^{k_1}}-\frac{32\gamma^2 a_n^2}{n}\right).$$

Since k_1 is independent of n and x, we can find $n_1>m$ such that for $n\geqq n_1$ we have $32\gamma^2 a_n^2/n\geqq 2\gamma a_{2^{k_1}}$.

So, for $n\geqq n_1$, we have $\alpha_{k_1}\leqq 2\exp(-\gamma a_{2^{k_1}})$. It follows that

$$\sum_{k_1\leqq l\leqq k}\alpha_l\leqq \sum_{l\geqq k_1} 2\exp(-\gamma a_{2^l})\leqq\gamma.$$

Step 4. We have shown that for $n\geqq n_1$ and $x\in A_n$ where $P'(A_n)\geqq 1-\gamma$, we have

$$Q(\{\sup_{C\in\mathscr{C}}|\sum_{i\leqq n}\varepsilon_i 1_C(x_i)|\geqq 8\gamma a_n(2+b)\})\leqq\gamma.$$

We note that condition (M) implies that for each t the function

$$x\to Q(\{\sup_{C\in\mathscr{C}}|\sum_{i\leqq n}\varepsilon_i 1_C(x_i)|\geqq t\})$$

is P'-measurable. It follows that for $n\geqq n_1$

$$(P'\otimes Q)(\{\sup_{C\in\mathscr{C}}|\sum_{i\leqq n}\varepsilon_i 1_C(x_i)|\geqq 8\gamma a_n(2+b)\})\leqq 2\gamma.$$

One finishes the proof with [6], Lemma 2–7, b.

Using Theorem 4 for $\phi(t)=t$ (so $a_n=n$) we see that a class of sets $\mathscr{C}$ that satisfies condition (M) is a Glivenko-Cantelli class of sets if and only if $\lim_{n\to\infty} r_n=0$.

The following beautiful lemma and its proof are due to D. Fremlin.

Lemma 21. *Let (V_n) be a compatible sequence of sets, and $a=\lim_n r_n$. Then there is a measurable set A with $P(A)=a$ and $P^n(V_n\setminus A^n)=0$ for each n.*

Proof. We note first that if for each k, $(V_n^k)_n$ is a compatible sequence of sets with $V_n^{k+1} \subset V_n^k$ for each k, $P^n(V_n^k) \geqq a^n$, then $W_n = \bigcap_k V_n^k$ is a compatible sequence of sets with $P^n(W_n) \geqq a^n$ for each n. One can hence assume that V_n has the property that for each compatible sequence W_n with $W_n \subset V_n$, $P^n(W_n) \geqq a^n$, one has $P^n(V_n \setminus W_n) = 0$ for each n. For x in Ω, let

$$V_{n,x} = \{y \in \Omega^n;\ (x, y) \in V_{n+1}\}.$$

Since the sequence (V_n) is compatible, we have $V_{n,x} \subset V_n$, and the sequence $(V_{n,x})_n$ is compatible for each x. Let

$$a(x) = \inf_n P^n(V_{n,x})^{1/n} = \lim_n P^n(V_{n,x})^{1/n}$$

so $a(x) \leqq a$. Let $A = \{x;\ a(x) = a\}$. We prove that $P(A) \geqq a$. Otherwise, there is $a' < a$ such that if $A' = \{x;\ a(x) > a'\}$, we have $P(A') < a$. Let $P(A') < a_1 < a$, and $a' < a_2 < a$.

For n large enough, if

$$B = \{x \in \Omega;\ P^n(V_{n,x}) \leqq a_2^n\}$$

we have $P(\Omega \setminus B) < a_1$. It follows that

$$P^{n+1}(V_{n+1}) = \int P^n(V_{n,x})\, dP(x) \leqq \int_B + \int_{\Omega \setminus B} \leqq a_2^n + a_1 P^n(V_n).$$

Let a_3 such that $a_2, a_1 < a_3 < a$. For n large enough, $a_2^n + a_1 r_n^n \leqq a_3 P^n(V_n)$, so $P^{n+1}(V_{n+1}) \leqq a_3 P^n(V_n)$. This shows that $\limsup (P^n(V_n))^{1/n} \leqq a_3 < a$, a contradiction.

We note that for x in A, we have $P^n(V_{n,x}) \geqq a^n$ for each n, so for each n, $P^n(V_n \setminus V_{n,x}) = 0$ by minimality of V_n. This shows (by induction) that for almost all each $(x_1, \ldots, x_k)$ in A^k,

$$P^n(V_n \setminus \{y \in \Omega^n;\ (x_1, \ldots, x_k, y) \in V_{n+k}\}) = 0$$

and the compatibility property of (V_n) implies $P^n(A^k \setminus V_k) = 0$ for each k. The proof is complete.

As a consequence, we get a new (and very different) proof of one of the results of [10].

Theorem 22. *If a class $\mathscr{C}$ of sets satisfies condition* (M) *and is not a Glivenko-Cantelli class, there is a set A with* $P(A) > 0$ *such that for all such k, and almost each* $(x_1, \ldots, x_k)$ *in* A^k, *the class $\mathscr{C}$ shatters* $\{x_1, \ldots, x_k\}$.

It should be noted that the proof of the above theorem used only the easy implication (I)$\Rightarrow$(II) of Theorem 2. The approach used here can be adapted to characterise, under measureability, Glivenko-Cantelli classes of functions, but considerable work would still be required to reach the generality of [10].

7. An Example

We let $\Omega=\mathbb{N}$, $\Sigma=2^{\mathbb{N}}$. We let $\mathscr{C}$ be the class of finite sets. We let $p_n=P(\{n\})$. It is known [3, 4] that $\mathscr{C}$ is a Donsker class if and only if $\sum p_n^{1/2}<\infty$. We now evaluate r_n. There is no loss of generality to assume that (p_n) is non increasing. We will show that $n r_n \to 0$ if and only if $n^2 p_n \to 0$. This in particular give examples that the conditions of Theorem 2 for $\phi(t)=1/t$ do not imply that $\mathscr{C}$ is pregaussian. However, let m_i be the smallest m for which $r_m \leqq 2^{-2i}$. If $\sum 2^{-i} m_i^{1/2}<\infty$, then $\mathscr{C}$ is pregaussian. This remark follows from the proof of Theorem 2, I$\Rightarrow$II and the method of [5], Theorem 2.23. (In fact, $\mathscr{C}$ then satisfies the so-called entropy condition.) A typical case where this relation is satisfied is when $r_m \leqq 1/m(\ln m)^\alpha$ for $\alpha>2$.

Let us go back to our example. We first note that

$$P^n(V_n) \geqq n! \prod_{i \leqq n} p_i \geqq n!\, p_n^n$$

so

$$r_n \geqq (n!)^{1/n} p_n, \qquad n r_n \geqq n (n!)^{1/n} p_n.$$

Since $(n!)^{1/n} \sim n/e$, we see that $n r_n \to 0$ implies $n^2 p_n \to 0$. To show the converse, we use a dazzling trick due to J. Zinn, who kindly let us reproduce it here. We note that

$$V_n \subset \{(x_1, \ldots, x_k);\ \prod_{i \leqq n} (x_i \wedge n) \geqq n!\}$$

so

$$P(V_n) \leqq (n!)^{-2} E\Big(\prod_{i \leqq n} (x_i \wedge n)^2\Big) = (n!)^{-2} E((x_1 \wedge n)^2)^n.$$

We have

$$\begin{aligned} r_n &\leqq (n!)^{-2/n} E((x_1 \wedge n)^2)^n \\ &\leqq (n!)^{-2/n} \Big(\sum_{i \leqq n} i^2 p_i + n^2 \sum_{i \geqq n} p_i\Big). \end{aligned}$$

Since $(n!)^{-2/n} \leqq a n^{-2}$ for some $a>0$, we get

$$n r_n \leqq a\left(\frac{1}{n} \sum_{i \leqq n} i^2 p_i + n \sum_{i \geqq n} p_i\right).$$

Since $p_n = O(1/n^2)$, $\sum_{i>n} p_i = o(1/n)$, and Kronecker's lemma shows that $\lim_{n\to\infty} n r_n = 0$.

Using similar ideas, if $\phi(n)$ satisfies condition $(*)$ one can how that $\lim_n r_n/\phi(n)=0$ if and only if $\lim_n n p_n/\phi(n)=0$ and $\lim_n (\sum_{i \geqq n} p_i)/\phi(n)=0$.

Acknowledgements. The author thanks E. Giné and J. Zinn for many conversations about their work, and for asking the questions that led to this work.

References

1. Alon, N., Milman, V.: Embedding of l^k_∞ in finite dimensional Banach spaces. Isr. J. Math. **45**, 265–280 (1983)
2. Andersen, N.T., Dobric, V.: The central limit theorem for Stochastic processes I and II. Math. Inst. Aarhus. U. research report Nos. 124 and 133, (1984)
3. Borisov, I.S.: Some limit theorems for empirical distributions. Abstracts of reports, Third Vilnus Conf., Prob. Th. Math. Statist. 1, pp. 71–72. (1981)
4. Dudley, R.M.: A course on empirical processes. In: Hennequin, P.L. (ed.) Ecole d'Eté de Probabilités de Saint-Floor XII-1982. (Lect. Notes Math. vol. 1097, pp. 1–142) Berlin Heidelberg New York: Springer 1984
5. Figiel, T., Lindenstrauss, J., Milman, V.: The dimension of almost spherical sections of convex bodies. Acta Math. **139**, 53–94 (1977)
6. Giné, E., Zinn, J.: Some limit theorems for empirical processes. Ann. Probab. **12**, 929–983 (1984)
7. Sauer, N.: On the density of families of sets. J. Combin. Theory Ser. A **13**, 145–147 (1972)
8. Steele, J.M.: Empirical discrepancies and subadditive processes. Ann. Probab. **6**, 118–127 (1978)
9. Talagrand, M.: Pettis integral and measure theory. Mem. Am. Math. Soc. **307** (1984)
10. Talagrand, M.: The Glivenko-Cantelli problem. Ann. Probab. **15**, 837–870 (1987)
11. Talagrand, M.: Measurability problems for empirical processes. Ann. Probab. **15**, 204–212 (1987)
12. Talagrand, M.: Donsker classes and random geometry. Ann. Probab. **15**, 1327–1338 (1987)
13. Talagrand, M.: Regularity of Gaussian processes. Acta Math. **159**, 99–149 (1987)

Received February 20, 1986