The Annals of Probability
1987, Vol. 15, No. 4, 1327–1338

DONSKER CLASSES AND RANDOM GEOMETRY

BY MICHEL TALAGRAND

University of Paris VI and The Ohio State University

Let $\mathscr{F}$ be a class of square integrable functions. We give necessary and sufficient random geometric conditions for the empirical process indexed by $\mathscr{F}$ to satisfy the CLT. These conditions roughly mean that the trace of $\mathscr{F}$ on a random sample is a small (for the l^1 norm) perturbation of a set which is nice for the l^2 norm.

1. Introduction. The remarkable recent paper of Giné and Zinn (1984) has given a new impetus to the study of the central limit theorem (CLT) for empirical processes indexed by families of functions. The present paper will present a refinement of the results of Giné and Zinn (1984), Section 5.a, concerning the CLT under random entropy conditions. We shall give an example showing that the sharpest necessary random entropy conditions obtained by Giné and Zinn are not sufficient. We shall give necessary and sufficient random conditions. In the general case, we will make use of the methods of Giné and Zinn, as well as of the description of pregaussian sets recently obtained by the author (1985). This description allows easy manipulation of these sets. It could actually provide significant simplications to the proofs of some of the results of Giné and Zinn. When the class $\mathscr{F}$ of functions satisfies the metric entropy condition, our main result becomes stronger and can be proved entirely by the methods of Giné and Zinn. We will, however, in that case use a device that we introduced in our work (1985). It allows us to get rid of the usual chaining argument. The results presented here seem to be sharp for classes of functions. For classes of sets, a further and more precise description is possible. It, however, uses very different methods and will appear elsewhere [Talagrand (1987d)].

2. Notation and results. For convenience, Giné and Zinn (1984) will be referred to as GZ. We will use notation close to that of GZ. Let $(S, \mathscr{S}, P)$ be a probability space. Denoting Lebesgue's measure by Q, let

$$(\Omega, \Sigma, \Pr) = \left([0,1] \times S^{\mathbf{N}}, \mathscr{B} \otimes \mathscr{S}^{\mathbf{N}}, Q \otimes P^{\mathbf{N}}\right).$$

We denote P by P_X. Integration with respect to Pr (resp. Q; P_X) is denoted by E (resp. E_Q, E_X). On $\Omega = [0,1] \times S^{\mathbf{N}}$, we consider the coordinate functions X_i on S. We use the factor $[0,1]$ to define on Ω an i.i.d. Bernoulli sequence (ε_i) and an i.i.d. $N(0,1)$ sequence (g_i), both independent of all (X_i), so integration with respect to Q applies to those variables. The empirical measures P_n on S are the random measures $P_n = (1/n)\sum_{i \le n} \delta_{X_i}$. For f in $\mathscr{L}^2(P)$, let $\|f\| = (Ef^2)^{1/2}$ the L^2 norm of f.

Received March 1985; revised August 1986.
AMS 1980 *subject classifications*. Primary 60F17, 60B12; secondary 60F05, 62E20.
Key words and phrases. Central limit theorems, empirical processes, functional Donsker classes, metric entropy.

Let $\mathscr{F}$ be a class of square integrable functions on S. We denote

$$\mathscr{F}_\delta = \{f - g;\ f, g \in \mathscr{F}, \|f - g\| \le \delta\},$$

$$\mathscr{F}_{\varepsilon, n} = \mathscr{F}_\alpha, \quad \text{for } \alpha = \varepsilon n^{-1/4}.$$

For each $\lambda > 0$, let $N(\lambda)$ be the λ-covering number of $\mathscr{F}$, that is, the smallest number of balls of radius λ that covers $\mathscr{F}$, i.e.,

$$N(\lambda) = \min\Big\{n;\ \exists\, f_1, \dots, f_n \in \mathscr{F}, \forall\, g \in \mathscr{F},\ \inf_i \|g - f_i\| \le \lambda\Big\}.$$

Let $H(\lambda) = \log N(\lambda)$. We say that $\mathscr{F}$ satisfies the (metric) entropy condition if $\int_0^1 H(\lambda)^{1/2}\, d\lambda < \infty$ or, equivalently, if $\sum_{k \ge 0} 2^{-k} H(2^{-k})^{1/2} < \infty$. Consider the random norms

$$M_{n,1}(f) = \sum_{i \le n} \frac{1}{n} |f(X_i)|,$$

$$M_{n,2}(f) = \left(\sum_{i \le n} \frac{1}{n} |f(X_i)|^2\right)^{1/2}.$$

We denote by $B_{n,1}$ (resp. $B_{n,2}$) the unit ball for $M_{n,1}$ (resp. $M_{n,2}$) in the space of all functions on Ω. For a set of functions G, we denote by $N_{n,2}(\lambda, G)$ and $N_{n,1}(\lambda, G)$ the λ-covering numbers for G corresponding to $M_{n,2}$ and $M_{n,1}$, respectively.

We refer to GZ for the meaning of "$\mathscr{F}$ pregaussian" and "$\mathscr{F}$ is a Donsker class." Given a class of functions $\mathscr{G}$, we write $\|\sum_{i \le n} f(X_i)\|_{\mathscr{G}}$ for the quantity $\sup_{f \in \mathscr{G}} |\sum_{i \le n} f(X_i)|$ and expressions of the like. The study of empirical processes runs into bothersome measurability problems. It is sometimes possible, at the expense of some complication, to work without any measurability assumption [Talagrand (1987a) and (1987b)], but our theorems here do require some measurability. One approach to measurability problems has been proposed by Dudley (1984). The methods of Giné and Zinn require use of Fubini's theorem at crucial steps; they require measurability of expressions of the type $\|\sum_{i \le n} \varepsilon_i f(X_i)\|_{\mathscr{F}_\delta}$. These conditions are carefully spelled out in GZ. We do not feel it is appropriate here to raise any measurability question, so we shall assume $\mathscr{F}$ to be countable. [A less restrictive hypothesis, with identical proofs, would be the separability of the processes $(P_n(f))_{f \in \mathscr{F}}$ for each n.] However, the use of outer probability at the obvious places and assumption of joint measurability when a Fubini type argument is used are enough for the proofs to carry over to a much more general setting.

For easy reference, we summarize the results of Giné and Zinn concerning random entropy for classes of functions.

THEOREM A (GZ, Remark 8.11). *If $\mathscr{F}$ is a Donsker class, then*

$$\lim_{\delta \to 0} \limsup_n E\Big(1 \wedge \sup_\lambda \lambda^2 \log N_{n,2}(\lambda, \mathscr{F}_\delta)\Big) = 0. \tag{1}$$

THEOREM B (GZ, Theorem 8.11). *Let $\mathscr{F} \subset \mathscr{L}^2$ be such that $\sup\{|f|;\ f \in \mathscr{F}\}$ is square integrable. Assume*

$$(2) \qquad \lim_{\delta \to 0} \limsup_{n \to \infty} E\left(1 \wedge \int_0^{\delta} (\log N_{n,2}(\lambda, \mathscr{F}))^{1/2}\, d\lambda\right) = 0.$$

Then $\mathscr{F}$ is a Donsker class.

THEOREM C (GZ, Theorems 5.1 and 5.4). *If $\mathscr{F}$ is pregaussian and uniformly bounded, then it is a Donsker class under either of the following conditions*:

(3) *For all $\varepsilon > 0$,*

$$\lim_n E\left[1 \wedge \int_0^{n^{-1/4}} (\log N_{n,2}(\lambda, \mathscr{F}_{\varepsilon,n}))^{1/2}\, d\lambda\right] = 0.$$

(4) *There exist $\delta, \sigma, \varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$,*

$$\lim_n \Pr\left\{\log N_{n,1}\left(\delta\varepsilon n^{-1/2}, \mathscr{F}_{\varepsilon,n}\right) > \sigma\varepsilon n^{1/2}\right\} = 0.$$

For two sets of functions B, G, let

$$B + G = \{f + g;\ f \in B,\ g \in G\}.$$

Let $\operatorname{conv} B$ be the set of *infinite* sums $\Sigma\alpha_n f_n$ for f_n in B and $\alpha_n > 0$, $\Sigma\alpha_n = 1$.

For simplicity, we set $l_1 = 1$, $l_n = (\log n)^{-1/2}$ for $n \geq 2$. The philosophy of our main result is very simple. It essentially states that if the quantities $\|n^{-1/2}\Sigma_{i=1}^n \varepsilon_i f(X_i)\|_{\mathscr{F}_\delta}$ are small, then for each n, one can find two classes of functions $\mathscr{U}$ and $\mathscr{W}$, such that each f in $\mathscr{F}_\delta$ is a sum $u + w$, u in $\mathscr{U}$, w in $\mathscr{W}$, and that the following hold, except on a set of small probability:

(i) $\|n^{-1/2}\Sigma_{i=1}^n |f|(X_i)\|_{\mathscr{U}}$ is small.

(ii) The metric structure of $(\mathscr{W}, M_{n,2})$ is close enough to the metric structure of $(\mathscr{F}_\delta, L^2)$ that one can conclude that $E_Q\|n^{1/2}\Sigma_{i=1}^n \varepsilon_i f(X_i)\|_{\mathscr{W}}$ is small because $\mathscr{F}$ is pregaussian.

We note the absolute values in (i), so cancellations, which are at the heart of the CLT, play no role here. The precise formulation of (ii) depends on which characterization of pregaussian sets one uses. We have given several rather different characterizations of these sets, and it is too early to tell which one will be the most useful. Hence, we just choose the simplest (as in Lemma 10), but other rather different formulations of our main result are possible.

THEOREM 1. *The following are equivalent*:

(I) *$\mathscr{F}$ is a Donsker class.*

(II) *$\mathscr{F}$ is totally bounded in $L^2(P)$, and the following condition holds, where K is a constant:*

$\forall\, \alpha > 0, \exists\, \delta > 0,$

$$\liminf_n \Pr\Big(\big\{\exists\, (h_q) \in L^2, \forall\, q \geq 1,\ M_{n,2}(h_q) \leq \inf(K\delta, \alpha l_q),$$
$$\mathscr{F}_\delta \subset \alpha n^{-1/2} B_{n,1} + \operatorname{conv}\{h_q;\ q \geq 1\}\big\}\Big) \geq 1 - \alpha.$$

(III) $\mathscr{F}$ is totally bounded in $L^2(P)$, and the following condition holds, where K is a constant:

$$\forall\, \alpha > 0, \forall\, \delta > 0, \exists\, n_0, \forall\, n \geq n_0, \exists\, (h'_q) \in L^2(P),$$
$$\forall\, f \in \mathscr{F}_\delta, \exists\, \alpha_q(f) \geq 0,\ \sum_{q \geq 1} \alpha_q(f) = 1,$$
$$\Pr\Bigg(\bigg\{\forall\, q \geq 1,\ M_{n,2}(h'_q) \leq \inf(K\delta, l_q), \forall\, f \in \mathscr{F}_\delta,$$
$$f - \sum_{q \geq 1} \alpha_q(f) h'_q \in \alpha n^{-1/2} B_{n,1}\bigg\}\Bigg) \geq 1 - \alpha.$$

When $\mathscr{F}$ satisfies the entropy condition [that is, when $\int_0^1 H(\lambda)^{1/2}\, d\lambda < \infty$], the result becomes more precise.

Theorem 2. *The following are equivalent*:

(I) *$\mathscr{F}$ is a Donsker class and satisfies the entropy condition.*

(II) *$\mathscr{F}$ is totally bounded in $L^2(P)$ and there exists a summable sequence (β_k) such that $\forall\, \alpha > 0, \forall\, \delta > 0$,*

$$\liminf_n \Pr\Big\{\exists\, G \text{ finite};\ G \subset 2^5 \delta B_{n,2}, \forall\, q \geq 1,$$
$$2^{-q}(\log N_{n,2}(2^{-q}, G))^{1/2} \leq \beta_q;\ \mathscr{F}_\delta \subset \alpha n^{-1/2} B_{n,1} + G\Big\} \geq 1 - \alpha.$$

We shall prove these results in Section 3. When $\mathscr{F}$ is also uniformly bounded, we can use the methods of GZ to give weaker sufficient conditions. Several formulations are possible; we give one that generalises Theorem C.

Theorem 3. *Suppose that $\mathscr{F}$ is pregaussian and uniformly bounded. Then it is a Donsker class whenever the following condition holds*:

$$\lim_{\alpha \to 0} \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \Pr\bigg\{\exists\, G \text{ finite},\ \mathscr{F}_{\varepsilon, n} \subset \alpha^2 n^{-1/2} B_{n,1} + G;$$

$$\int_0^{20\alpha n^{-1/4}} (\log N_{n,2}(\lambda, G))^{1/2}\, d\lambda \leq \gamma;\ \log N_{n,2}(2\alpha n^{-1/4}, G) \leq \alpha^2 n^{1/2}\bigg\} = 1, \tag{5}$$

for each $\gamma > 0$.

Using the fact that $N_{n,2}(f) \leq N_{n,1}(f)^{1/2} \|f\|_\infty^{1/2}$, it is possible to see that condition (5) is weaker than either conditions (3) or (4). The proof of Theorem 3 is essentially the same as the proof of Theorem 5.4 of GZ, so we shall not give it. [One needs to use the fact that the proof of Theorem 3.2 of GZ shows that hypothesis (3.6) can be weakened to

$$\forall\, \gamma > 0, \qquad \lim_{\varepsilon \to 0} \limsup_n \Pr\bigg\{\bigg\| \sum_{i \leq n} \varepsilon_i f(X_i) \bigg\|_{\mathscr{F}_{\varepsilon, n}} > \gamma n^{1/2}\bigg\} = 0.]$$

In Section 4, we shall show that even for bounded classes that satisfy the entropy condition, condition (1) does not imply that $\mathscr{F}$ is a Donsker class; we shall give an example of a Donsker class that fails condition (5).

3. Proof of Theorems 1 and 2. Before we start the proof, we list the main tools that we shall use. The proof of the following lemma is similar to that of Lemma 2.7 of GZ.

LEMMA 4. *Assume* $\sup_{f \in \mathscr{F}} E|f - Ef| = \gamma < \infty$. *Then for* $t > 2\gamma n^{1/2}$, *we have*

$$\Pr\left\{\left\| \sum_{i\le n} f(X_i) - Ef \right\|_{\mathscr{F}} > tn^{1/2}\right\} \le 4\Pr\left\{\left\| \sum_{i\le n} \varepsilon_i f(X_i)\right\|_{\mathscr{F}} \ge (tn^{1/2} - 2n\gamma)/2\right\}.$$

For a class of functions G, let $|G| = \{|f|,\ f \in G\}$.

LEMMA 5. *For a class of functions* G, *we have*

$$E_Q\left(\left\|\sum g_i f(x_i)\right\|_{|G|}\right) \le 2E_Q\left(\left\|\sum g_i f(X_i)\right\|_G\right).$$

PROOF. We can assume that G contains zero. We note that $M_{n,2}(|f| - |f'|) \le M_{n,2}(f - f')$. Fernique's theorem as in GZ, (2.28), shows

$$E_Q\left(\sup_{f\in G} \sum g_i |f|(X_i)\right) \le E_Q\left(\sup_{f\in G} \sum g_i f(X_i)\right) \le E_Q\left(\left\|\sum g_i f(X_i)\right\|_G\right)$$

and similarly,

$$E_Q\left(\sup_{f\in G} \left(-\sum g_i |f|(X_i)\right)\right) \le E_Q\left(\left\|\sum g_i f(X_i)\right\|_G\right).$$

The result follows since

$$\sup_{f\in G} \sum g_i |f|(X_i) \ge 0, \qquad \sup_{f\in G} -\sum g_i |f|(X_i) \ge 0. \qquad \square$$

LEMMA 6 [GZ, (2.9)].

$$E_Q\left(\left\| \sum_{i\le n} \varepsilon_i f(X_i)\right\|_G\right) \le \left(\frac{\pi}{2}\right)^{1/2} E_Q\left(\left\| \sum_{i\le n} g_i f(X_i)\right\|_G\right).$$

The following is a minor variation on GZ, (2.14).

LEMMA 7. (a) *If* $\mathscr{F}$ *is a Donsker class, then*

$$\lim_{\delta\to 0} \limsup_n E\left\| n^{-1/2} \sum_{i\le n} g_i f(X_i)\right\|_{\mathscr{F}_\delta} = 0. \tag{6}$$

(b) *If* $\mathscr{F}$ *is totally bounded in* L^2, *and if for each* $\varepsilon > 0$,

$$\lim_{\delta\to 0} \limsup_n \Pr\left\{\left\| \sum_{i\le n} \varepsilon_i f(X_i)\right\|_{\mathscr{F}_\delta} \ge \varepsilon n^{1/2}\right\} = 0, \tag{7}$$

then $\mathscr{F}$ *is a Donsker class.*

The following is a specialised version of Dudley's theorem. See GZ, (2.24).

LEMMA 8. *There is a universal constant K_1 such that for each n, each $\gamma > 0$, and each class of functions $G \subset \gamma B_{n,2}$, we have*

$$E_Q\Big(n^{-1/2}\big\|\textstyle\sum g_i f(X_i)\big\|_G\Big) \le K_1\Big(\gamma + \sum_q 2^{-q}(\log N_{n,2}(2^{-q}, G))^{1/2}\Big).$$

The following is a consequence of a standard computation.

LEMMA 9. *There is a universal constant K_2 such that whenever $G = \{h_q,\ q \ge 1\}$, where $M_{n,2}(h_q) \le l_q$, we have*

$$E_Q\Big(n^{-1/2}\big\|\textstyle\sum g_i f(X_i)\big\|_G\Big) \le K_2.$$

The study of Donsker classes, when the entropy condition is not assumed, will rely on the following, which is an adaptation of the results of the author (1985). A complete proof will appear in our (1987d).

LEMMA 10. *Suppose $\mathscr{F}$ is pregaussian. Fix $\varepsilon > 0$. Then for each $\delta > 0$ small enough there is a sequence (h_q) in $L^2(P)$, with $\|h_q\| < inf(K\delta, \varepsilon l_q)$ and $\mathscr{F}_\delta \subset \operatorname{conv}\{h_q;\ q \ge 1\}$.*

Of essential use will also be Bernstein's inequality (GZ, page 983). For a function g, with $Eg = 0$,

$$\Pr\Big\{\sum_{i\le n} g(X_i) \ge t\Big\} \le \exp\big\{-t^2/\big(2nEg^2 + 2t\|g\|_\infty/3\big)\big\}.$$

LEMMA 11. *Let h be in $L^2(P)$. Let $n, \varepsilon > 0$. Denote by h' the truncation of h at levels $w = \pm n^{1/2}\ Eh^2/\varepsilon$ [that is, $h' = \min(w, \max(h, -w))$] and let $h'' = h - h'$. Then we have $E|h''| \le \varepsilon n^{-1/2}$ and*

$$\Pr\big(\{M_{n,2}(h') \ge 2\|h\|\}\big) \le \exp\big(-3\varepsilon^2/2Eh^2\big).$$

PROOF. We first note that, since $|h| \ge w$ whenever $h'' \ne 0$, we have $w|h''| \le h^2$, so $E|h''| \le Eh^2/w$. Let $g = h'^2 - Eh'^2$. We note that $\|g\|_\infty < 2w^2$ and that $Eg^2 \le Eh'^4 \le w^2 Eh^2$. The result then follows from the fact that

$$\Pr\big(\{M_{n,2}(h') \ge 2\|h\|\}\big) \le \Pr\Big(\Big\{\sum_{i\le n} g(X_i) \ge 3nE(h^2)\Big\}\Big)$$

and from Bernstein's inequality. □

PROOF OF THEOREM 1. We first prove that (II) ⇒ (I). Suppose that $\mathscr{F}_\delta \subset \alpha n^{-1/2}B_{n,1} + \operatorname{conv}\{h_q;\ q \ge 1\}$, where $M_{n,2}(h_q) \le \alpha l_q$. [We note that the extra information $M_{n,2}(h_q) \le K\delta$ is not actually needed.] Let $G = \operatorname{conv}\{h_q;\ q \ge 1\}$.

Since $\mathcal{F}_\delta \subset \alpha n^{-1/2}B_{n,1} + G$, we have

$$E_Q\Big(n^{-1/2}\Big\|\sum \varepsilon_i f(X_i)\Big\|_{\mathcal{F}_\delta}\Big) \le \alpha + E_Q\Big(n^{-1/2}\Big\|\sum \varepsilon_i f(X_i)\Big\|_G\Big).$$

Using Lemmas 9 and 6, we get that

$$E_Q\Big(n^{-1/2}\Big\|\sum \varepsilon_i f(X_i)\Big\|_G\Big) \le \alpha K_2$$

and $\mathcal{F}$ is a Donsker class from lemma 7.

We now prove that (I) ⇒ (III). Fix $\alpha > 0$. Let $\varepsilon = 10^{-2}\alpha$. Let $\eta < \varepsilon^2/2K_2$ be small enough that

$$\sum_{q \ge 1} \exp\big(-3\varepsilon^2/2\eta^2 l_q^2\big) < \varepsilon. \tag{8}$$

From Lemmas 7 and 10, there is $\delta > 0$ such that

$$\exists\, n_0, \forall\, n \ge n_0, \qquad P_X\left(\left\{E_Q\Big\|\sum_{i \le n} g_i f(X_i)\Big\|_{\mathcal{F}_\delta} \ge \varepsilon^2 n^{1/2}\right\}\right) \le \varepsilon \tag{9}$$

and such that there is a sequence (h_q) in $L^2(P)$, with $\|h_q\| \le \inf(K\delta, \eta l_q)$ and $\mathcal{F}_\delta \subset \operatorname{conv}\{h_q;\ q \ge 1\}$.

The construction will depend on n. To simplify notation, we consider $n \ge n_0$ fixed. Denote by h'_q the truncation of h_q at levels $\pm n^{1/2}Eh_q^2/\varepsilon$, and let $h''_q = h_q - h'_q$. Consider the event $A = \{\forall\ q,\ M_{n,2}(h'_q) \le 2\|h_q\|\}$. It then follows from Lemma 11 and (8) that $\Pr(A) \ge 1 - \varepsilon$. From Lemma 9 it follows that if $G = \operatorname{conv}\{h'_q;\ q \ge 1\}$,

$$A \subset \left\{E_Q\Big\|\sum_{i \le n} g_i f(X_i)\Big\|_G < \varepsilon^2 n^{1/2}\right\}.$$

For each f in $\mathcal{F}_\delta$, we fix coefficients $\alpha_q(f)$ such that $f = \sum_{q \ge 1}\alpha_q(f)h_q$ and $\alpha_q(f) \ge 0$, $\sum\alpha_q(f) = 1$. Let D be the set of functions $\sum_{q \ge 1}\alpha_q(f)h''_q$. We have $f_\delta - \sum_{q \ge 1}\alpha_q(f)h'_q \subset D$ so it remains to show that with probability $\ge 1 - \alpha + \varepsilon$, we have $D \subset \alpha n^{-1/2}B_{n,1}$. Since $E|h''_q| \le \varepsilon n^{-1/2}$ for each q, we have $E|f| \le \varepsilon n^{-1/2}$ for f in D. Since $D \subset F_\delta - G$, we have

$$P_X\left(\left\{E_Q\Big\|\sum_{i \le n} g_i f(X_i)\Big\|_D \ge 2\varepsilon^2 n^{1/2}\right\}\right) \le 2\varepsilon.$$

It follows from Lemma 5 that

$$P_X\left(\left\{E_Q\Big\|\sum_{i \le n} g_i f(X_i)\Big\|_{|D|} \ge 4\varepsilon^2 n^{1/2}\right\}\right) \le 2\varepsilon.$$

It follows from Lemma 6 that

$$P_X\left(\left\{E_Q\Big\|\sum_{i \le n} \varepsilon_i f(X_i)\Big\|_{|D|} \ge 4\varepsilon^2 n^{1/2}\right\}\right) \le 2\varepsilon,$$

so

$$\Pr\left(\left\{\Big\|\sum_{i \le n} \varepsilon_i f(X_i)\Big\|_{|D|} \ge \varepsilon n^{1/2}\right\}\right) \le 6\varepsilon.$$

We now use Lemma 4 with $\gamma = 2\varepsilon n^{-1/2}$, $t = 6\varepsilon$, to get

$$\Pr\left(\left\{\left\|\sum_{i\le n} f(X_i) - Ef\right\|_{|D|} \ge 6\varepsilon n^{1/2}\right\}\right) \le 24\varepsilon.$$

Since $Ef \le \varepsilon n^{-1/2}$, we have

$$\Pr\left(\left\{\sup_{f\in D} M_{n,1}(f) \ge 7\varepsilon n^{-1/2}\right\}\right) \le 24\varepsilon.$$

The proof is complete since (III) $\Rightarrow$ (II) is obvious. □

PROOF OF THEOREM 2. We first prove that (II) $\Rightarrow$ (I). The proof that (II) implies that $\mathscr{F}$ is a Donsker class goes as in Theorem 1, using Lemma 8 instead of Lemma 9. We now show that $\mathscr{F}$ satisfies the entropy condition. Let δ corresponding to the choice of $\alpha = 1/2$ in condition (II). Since $\mathscr{F}$ is totally bounded in L^2, we can cover it by a finite family $B_1, \dots, B_k$ of balls of radius δ. For $\lambda > \delta$, we have $N(\lambda) \le k$. We show that for $2^{-q} < \delta$ we have $N(2^{-q}) \le k\exp(2^{2q+8}\beta_{q+4}^2)$; this will obviously suffice. let F be a maximal subset of $\mathscr{F}$ such that any two distinct elements of $\mathscr{F}$ are at distance $\ge 2^{-q}$. Fix $m \le k$. Let g be the center of B_m. Enumerate as $h_1, \dots, h_l$ the elements of the type $f - g$ for f in $F \cap B_m$. Note that they belong to $\mathscr{F}_\delta$. Fix a large enough that if for $i < j \le l$, we denote by $h_{i,j}$ the truncation of $h_i - h_j$ at levels $\pm a$, we have $\|h_{i,j}\| \ge 2^{-q-1}$ for each $i < j \le l$. The law of large numbers shows that

$$\lim_n \Pr\left\{\forall\ i < j \le l,\ M_{n,2}(h_{i,j}) \ge 2^{-q-2}\right\} = 1.$$

So there is n large enough such that with positive probability $M_{n,2}(h_{i,j}) \ge 2^{-q-2}$ for $i, j \le l$ and $\mathscr{F}_\delta \subset \frac{1}{2}n^{-1/2}B_{n,1} + G$, where for each q, $2^{-q-4}(\log N_{n,2}(2^{-q-4}, G))^{1/2} \le \beta_{q+4}$. And we can assume n large enough that $n^{-1/2} \le 2^{-3q-12}/a^2$. For $i \le l$, write $h_i = u_i + v_i$, where $M_{n,1}(u_i) \le \frac{1}{2}n^{-1/2}$, $v_i \in G$. For $i < j$, we have

$$v_i - v_j = h_i - h_j - (u_i - u_j),$$

where $M_{n,1}(u_i - u_j) \le n^{-1/2}$. We now fix $i < j$ and set $B = \{|u_i - u_j| \ge 2^{-q-4}\}$, so we have $P_n(B) \le n^{-1/2}/2^{-q-4} \le 2^{-2q-8}/a^2$.

We have, since $\|h_{i,j}\|_\infty \le a$,

$$\begin{aligned} M_{n,2}(v_i - v_j) &\ge M_{n,2}\big((h_i - h_j - (u_i - u_j))1_{\Omega\setminus B}\big) \\ &\ge M_{n,2}\big((h_i - h_j)1_{\Omega\setminus B}\big) - M_{n,2}\big((u_i - u_j)1_{\Omega\setminus B}\big) \\ &\ge M_{n,2}\big((h_i - h_j)1_{\Omega\setminus B}\big) - 2^{-q-4} \\ &\ge M_{n,2}\big(h_{i,j}1_{\Omega\setminus B}\big) - 2^{-q-4} \\ &\ge M_{n,2}\big(h_{i,j}\big) - M_{n,2}\big(h_{i,j}1_B\big) - 2^{-q-4} \\ &\ge M_{n,2}\big(h_{i,j}\big) - 2^{-q-3} \ge 2^{-q-3}. \end{aligned}$$

It follows that $N_{n,2}(2^{-q-4}, G) \ge l$. This concludes the proof.

We now prove (I) $\Rightarrow$ (II). Surely $\mathscr{F}$ is totally bounded if it satisfies the entropy condition.

Since $\mathscr{F}$ satisfies the entropy condition, so does $\mathscr{F}' = \mathscr{F} - \mathscr{F}$. So, for each $q \geq 0$, one can find a subset F_q of $\mathscr{F}'$, such that each element of $\mathscr{F}'$ is at L^2 distance $\leq 2^{-q}$ of an element of F_q, and that $\sum_{q \geq 1} \gamma_q < \infty$, where $\gamma_q = 2^{-q}(\log \operatorname{card} F_q)^{1/2}$. Let $\xi_q = \gamma_q + \gamma_{q-1}$, so $\sum_{q \geq 1} \xi_q < \infty$. For each $q \geq 1$, let

$$C_q = \left\{x - y;\, x \in F_q,\, y \in F_{q-1},\, \|x - y\| \leq 2^{-q+1}\right\},$$

so $\|z\| \leq 2^{-q+1}$ for z in C_q, and 2^{-q} $(\log(\operatorname{card} C_q))^{1/2} \leq \xi_q$. Also, the usual chaining argument shows that if $x \in \mathscr{F}'$, $\|x\| \leq 2^{-l-2}$, we can write $x = \sum_{q \geq l} x_q$ where $x_q \in C_q$. Now let $\alpha > 0$. Let $\varepsilon = 10^{-2}\alpha$ and take $\eta < \varepsilon^2/2K_2$ small enough that (8) holds. Take l large enough that (9) holds for $\delta = 2^{-l-2}$ and $\sum_{q \geq l} \xi_q \leq \eta/4$, $\sum_{q \geq l} 2^{-q}\sqrt{q+1} \leq \eta/4$. For $q \geq l$, set

$$a_q = 2^{-q+1}\left(\log\left(2^{q+1}\operatorname{card} C_q\right)\right)^{1/2}. \tag{10}$$

We have

$$\sum_{q \geq l} a_q \leq \sum_{q \geq l} 2^{-q+1}\left(\log\left(\operatorname{card} C_q\right)\right)^{1/2} + \sum_{q \geq l} 2^{-q+1}\sqrt{q+1} \leq \eta.$$

Set $a = \sum_{q \geq l} a_q$ so $a \leq \eta$. Set $H_q = (a/a_q)C_q = \{(a/a_q)y;\ y \in C_q\}$. Set $H = \bigcup_{q \geq l} H_q$. Let $t > 0$. If for some q and some y in H_q we have $\|y\| \geq t$, then $2^{-q+1}a \geq a_q t$. This implies first that $t \leq a$. Using (10) this also shows that $\operatorname{card} H_q = \operatorname{card} C_q \leq 2^{-q-1}\exp(a^2/t^2)$. It follows that we have

$$\operatorname{card}\{y \in H;\ \|y\| \geq t\} \leq \exp\left(a^2/t^2\right). \tag{11}$$

Let us enumerate H as a sequence (h_n) with $\|h_{n+1}\| \leq \|h_n\|$. It follows from (11) that $\|h_n\| \leq a l_n \leq \eta l_n$. Since each x in $\mathscr{F}_\delta$ can be written $x = \sum_{q \geq l} x_q$ for x_q in C_q, it can be written $x = \sum_q b_q y_q(x)$, where $b_q = a_q/a$ and $y_q(x) = (a/a_q)x_q \in H_q$. So $\sum b_q = 1$. We now fix $n \geq n_0$. For a function h in H, write h' its truncation at levels $\pm n^{1/2}Eh^2/\varepsilon$ and $h'' = h - h'$. Consider the event

$$A = \left\{\forall\, h \in H,\ M_{n,2}(h') \leq 2\|h\|,\ \forall\, x \in \mathscr{F}_\delta,\ M_{n,1}\left(\sum_q b_q y_q''(x)\right) \leq \alpha n^{-1/2}\right\}.$$

Inspection of the proof of Theorem 1 shows that we have $\Pr(A) \geq 1 - \alpha$. Let $G = \{\sum_{q \geq l} b_q y_q'(x);\ x \in \mathscr{F}_\delta\}$. We know that on A, $M_{n,2}(y_q'(x)) \leq 2\|y_q(x)\| \leq 2^{-q+2}a/a_q$, so $M_{n,2}(b_q y_q'(x)) \leq 2^{-q+2}$. It is now routine to check that $G \subset 2^{-q+3}B_{n,2}$ and that $2^{-q}(\log N_{n,2}(2^{-q}, G))^{1/2} \leq \beta_q$, where $\beta_q = \sum_{i \leq q+3} \xi_i 2^{i-q}$ (so $\sum \beta_q < \infty$). The proof is complete. □

4. Examples.

THEOREM 12. *There exists a countable uniformly bounded class of functions that satisfies the entropy condition, and condition* (1), *but is not a Donsker class.*

The proof contains a simple idea, but checking all the routine details would be tiresome. We shall hence just give a sketch. The basic measure space is [0, 1]; Lebesgue's measure is denoted by P. The basic construction is as follows. Let k be an integer, $k \geq 3$. Let $n = 2^{4k+1}$. Fix disjoint subintervals $(A_j)_{j \leq k}$ with $P(A_j) = 2^{j-2k-1}$. Note that $P(\cup A_j) \leq 2^{-k}$. Consider the family F_k of functions of the type

$$f = \sum_{j \leq k} k^{-1} 2^{-j} 1_{B_j}$$

where $B_j \subset A_j$ and $\operatorname{card} B_j = 2^{2k+j-1}$. Let

$$a_k = P^n\{(x_1, \dots, x_n); \forall j \leq k, \operatorname{card}\{x_1, \dots, x_n\} \cap A_j \geq nP(A_j)/2\}.$$

Standard estimates show that $\lim_{k \to \infty} a_k = 1$. When

$$\forall j \leq k, \qquad \operatorname{card}\{x_1, \dots, x_n\} \cap A_j \geq nP(A_j)/2 = 2^{j+2k-1},$$

there exists an f in F_k such that

$$\sum_{i \leq n} f(x_i) \geq \sum_{j \leq k} k^{-1} 2^{2k-1} = 2^{2k-1} = 2^{-3/2} n^{1/2}.$$

So we have

$$\Pr\left\{\left\| \sum_{i \leq n} f(x_i) \right\|_{F_k} \geq n^{1/2}/4 \right\} \geq a_k. \tag{12}$$

We show now that there are sets $U_{k,m} \subset [0,1]^m$, with

$$P^m(U_{k,m}) \geq 1 - b_{k,m}, \quad \text{where } b_{k,m} \leq 3k^{-1/2}, \ \lim_{m \to \infty} b_{k,m} = 0, \tag{13}$$

such that

$$U_{k,m} \subset \left\{ \sup_{\lambda} \lambda^2 \log N_{m,2}(\lambda, F_k) \leq 1/k \right\}.$$

Let $A'_j = \bigcup_{i \leq j} A_i$, so $P(A'_j) \leq 2^{j-2k}$. Let

$$U_{k,m} = \{(x_1, \dots, x_m); \forall j \leq k, \operatorname{card}\{x_1, \dots, x_m\} \cap A'_j \leq mkP(A'_j)\}.$$

For any set A, Bernstein's inequality gives

$$P^m\left(\left\{ \frac{1}{m} \sum_{i \leq m} 1_A(x_i) \geq kP(A) \right\}\right) \leq \exp(-m(k-1)P(A)).$$

Also, the left hand side is less than $mP(A)$. So we have

$$P^m\left(\left\{ \frac{1}{m} \sum_{i \leq m} 1_A(x_i) \geq kP(A) \right\}\right) \leq \inf(mP(A), \exp(-m(k-1)P(A))),$$

so

$$P^m(U_{k,m}) \geq 1 - \sum_{j \leq k} \inf\left(mP(A'_j), \exp\left(-m(k-1)P(A'_j)\right)\right).$$

Let $b_{k,m} = \sum_{j \leq k} \inf(mP(A'_j), \exp(-m(k-1)P(A'_j)))$. Then $\lim_m b_{k,m} = 0$.

Moreover,

$$b_{k,m} \le \sum_{j \le p} mP(A'_j) + \sum_{j>p} \exp\big(-m(k-1)P(A'_j)\big),$$

where p is the largest value of j such that $mP(A'_j) \le k^{-1/2}$. Since for each j we have $P(A'_j) \ge 2P(A'_{j-1})$, we get that for all m, we have

$$b_{k,m} \le 2k^{-1/2} + k\exp\big(-(k-1)k^{1/2}\big) \le 3k^{-1/2}.$$

We now fix $x_1, \dots, x_m$ in $U_{k,m}$, and $\lambda > 0$. For $1 \le l \le k$ let

$$a_l = \left(\frac{1}{mk^2} \sum_{l \le j \le k} 2^{2k-j-1}\right)^{1/2}.$$

Suppose first $\lambda \le a_k$. Let G be the trace of F_k on $\{x_1, \dots, x_m\}$. We have $N_{m,2}(\lambda, F_k) \le \operatorname{card} G$. An element of G is zero outside A'_k. Suppose first that $mk2^{-k} \ge 2^{3k}$. We have $\operatorname{card}\{x_1, \dots, x_m\} \cap A'_k \le mkP(A'_k) \le mk2^{-k}$; and a function in g is determined by a subset of $\{x_1, \dots, x_n\} \cap A'_k$ of cardinality $\le 2^{3k}$; so we have

$$N_{m,2}(\lambda, F_k) \le \sum_{i \le 2^{3k}} \binom{mk2^{-k}}{i} \le \big(emk2^{-4k}\big)^{2^{3k}}$$

and

$$\lambda^2 \log N_{m,2}(\lambda, F_k) \le \frac{1}{2mk^2} 2^{4k} \log\big(emk2^{-4k}\big),$$

so easily $\lambda^2 \log N_{m,2}(\lambda, F_k) \le 1/2k$. Suppose now $mk2^{-k} \le 2^{3k}$. Then $\operatorname{card} G \le 2^{mk2^{-k}}$ and

$$\lambda^2 \log N_{m,2}(\lambda, F_k) \le \frac{mk}{2mk^2} \log 2 \le 1/k.$$

We now investigate the case $a_k \le \lambda$. Let l be the smallest with $a_l \le \lambda$. If $l = 1$, since each function of F_k has $M_{m,2}$norm $\le a_1$, we get $N_{m,2}(\lambda, F_k) = 1$. If $l > 1$, for $f = \sum_{j \le k} 2^{-k} 1_{B_j}$, write $f' = \sum_{l \le j \le k} 2^{-j} 1_{B_j}$, $f'' = f - f'$. We have $M_{m,2}(f') \le a_l \le \lambda$, so we can bound $N_{m,2}(\lambda, F_k)$ by the number of different traces on $\{x_1, \dots, x_m\}$ of functions of the type f''. Note that $\lambda \le a_{l-1}$; the method is similar to the case $\lambda \le a_k$ so the details are left to the reader. We thus have

$$\sup_\lambda N_{m,2}(\lambda, F_k) \le 10/k \tag{14}$$

on $U_{k,m}$. This completes the basic construction. Now let (k_p) be an increasing sequence; let $\mathscr{F}' = \bigcup_p F_{k_p}$. This class satisfies the entropy condition, since each function in $\mathscr{F}'$ is zero a.e. Condition (12) shows that $\mathscr{F}'$ is not a Donsker class. It is easy to check that if the sequence (k_p) increases fast enough, (13) and (14) imply that for each $\gamma > 0$

$$\lim_m \Pr\Big\{\sup_\lambda \lambda^2 M_{m,2}(\lambda, \mathscr{F}') > \gamma\Big\} = 0.$$

The only problem is that $\mathscr{F}'$ is not countable. Let (q_p) be any sequence. Each function of F_{k_p} is a finite sum $\Sigma \alpha_i 1_{\{x_i\}}$. Replace each function $1_{\{x_i\}}$ by 1_J, where J is the dyadic interval of length 2^{-q_p} containing x_i. Let $\mathscr{F}$ be the class obtained this way. It is routine to see that $\mathscr{F}$ will satisfy the conditions of Theorem 12 provided (q_p) increases fast enough.

THEOREM 13. *There exists a countable Donsker class that does not satisfy condition* (5).

PROOF. Consider a sequence (a_n) and the class $\mathscr{F} = \{a_n r_n;\ n \geq 0\}$, where (r_n) is an independent Bernoulli sequence. Then it is easy to see that it is a Donsker class if and only if the following holds:

(15) For each $t > 0$, the series $\sum_n \exp(-t^2/a_n^2)$ converges.

Consider the sequence (a_n) constructed in the following way: For each k, let $n(k) = 2^{2k^2}$. For each k and each j with $1 \leq j \leq k$, one repeats $2^{2^{2j} n(k)^{1/2}}$ times the number $1/(k 2^j n(k)^{1/4})$. This sequence satisfies (15). For one given k, the total number of terms involved is of order $\exp c_k n(k) m$, where $c_k \to 0$, so with probability going to one, any two of the functions r_n involved for a given k are at distance $\geq 1/3$ for $M_{n(k),2}$. It follows that with probability close to one,

$$\forall\, j \leq k, \qquad \log N_{n(k),2}\Big(\big(3k2^j n(k)^{1/4}\big)^{-1}, \mathscr{F}_{1/k,\, n(k)}\Big) \geq 2^{2j} n(k)^{1/2} \log 2.$$

It follows that if $\delta^{-1} = 3kn(k)^{1/4}$,

$$\int_0^\delta \big(\log N_{n(k),2}(\lambda, \mathscr{F}_{1/k,\, n(k)})\big)^{1/2} d\lambda \geq (\log 2)^{1/2}/6.$$

Since the functions $a_n r_n$ have a constant absolute value, one can check that a small perturbation in l^1 norm is essentially irrelevant and that (5) fails. □

Acknowledgments. The author thanks E. Giné and J. Zinn for introducing him to their work and for many conversations.

REFERENCES

DUDLEY, R. M. (1984). A course on empirical processes. *École d'Été de Probabilités de Saint Flour XII–1982. Lecture Notes in Math.* **1097** 1–142. Springer, New York.

GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** 929–989.

TALAGRAND, M. (1985). La description des processus gaussiens bornés. *C. R. Acad. Sci. Paris Sér. I Math.* **301** 751–753.

TALAGRAND, M. (1987a). Measurability problems for empirical processes. *Ann. Probab.* **15** 204–212.

TALAGRAND, M. (1987b). The Glivenko–Cantelli problem. *Ann. Probab.* **15** 837–870.

TALAGRAND, M. (1987c). Donsker classes of sets. *Probab. Theory Related Fields.* To appear.

TALAGRAND, M. (1987d). Regularity of Gaussian processes. *Acta. Math.* To appear.

EQUIPE D'ANALYSE-TOUR 46
U.A. AU C.N.R.S. No. 754
4 PLACE JUSSIEU
75230 PARIS CEDEX 05
FRANCE

DEPARTMENT OF MATHEMATICS
THE OHIO STATE UNIVERSITY
231 WEST 18TH AVENUE
COLUMBUS, OHIO 43210