Z. Wahrscheinlichkeitstheorie verw. Gebiete
61, 355–368 (1982)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Empirical and Poisson Processes on Classes of Sets or Functions Too Large for Central Limit Theorems*

R. M. Dudley

Massachusetts Institute of Technology, Dept. of Mathematics, Room 2-245, Cambridge, Mass. 02139, USA

Summary. Let P be the uniform probability law on the unit cube I^d in d dimensions, and P_n the corresponding empirical measure. For various classes $\mathscr{C}$ of sets $A \subset I^d$, upper and lower bounds are found for the probable size of $\sup\{|P_n - P)(A)|: A \in \mathscr{C}\}$. If $\mathscr{C}$ is the collection of lower layers in I^2, or of convex sets in I^3, an asymptotic lower bound is

$$((\log n)/n)^{1/2} (\log\log n)^{-\delta-1/2} \quad \text{for any } \delta > 0.$$

Thus the law of the iterated logarithm fails for these classes.

If $\alpha > 0$, β is the greatest integer $< \alpha$, and $0 < K < \infty$, let $\mathscr{C}$ be the class of all sets $\{x_d \leqq f(x_1, \ldots, x_{d-1})\}$ where f has all its partial derivatives of orders $\leqq \beta$ bounded by K and those of order β satisfy a uniform Hölder condition $|D^p(f(x) - f(y))| \leqq K|x-y|^{\alpha-\beta}$. For $0 < \alpha < d-1$ one gets a universal lower bound $\delta n^{-\alpha/(d-1+\alpha)}$, for a constant $\delta = \delta(d, \alpha) > 0$. When $\alpha = d-1$ the same lower bound is obtained as for the lower layers in I^2 or convex sets in I^3. For $0 < \alpha \leqq d-1$ there is also an upper bound equal to a power of $\log n$ times the lower bound, so the powers of n are sharp.

1. Introduction

First let us define a collection of sets in d-dimensional Euclidean space $\mathbb{R}^d$ whose boundaries are given by functions with derivatives up through some order β bounded and satisfying a Hölder condition of order $0 < \alpha - \beta \leqq 1$. Specifically, let $\alpha > 0$ and $K > 0$. Let β be the greatest integer $< \alpha$. Let

$$D^p = \partial^{[p]} / \partial x_1^{p_1} \ldots \partial x_d^{p_d}, \qquad [p] = p_1 + \ldots + p_d,$$

for p_i integers $\geqq 0$, $p = (p_1, \ldots, p_d)$. For a function f on $\mathbb{R}^d$ such that $D^p f$ is continuous whenever $[p] \leqq \beta$, let

* This research was partially supported by National Science Foundation Grant MCS-79-04474

0044-3719/82/0061/0355/$02.80

$$\|f\|_\alpha = \max_{[p]\leqq\beta} \sup\{|D^p f(x)|: x\in\mathbb{R}^d\} + \max_{[p]=\beta} \sup_{x\neq y}\{|D^p f(x)-D^p f(y)|/|x-y|^{\alpha-\beta}\}$$

where $|u|=(u_1^2+\ldots+u_k^2)^{1/2}$, $u\in\mathbb{R}^k$.

Let I^d be the unit cube $\{x\in\mathbb{R}^d: 0\leqq x_j\leqq 1, j=1,\ldots,d\}$. Let $x_{(d)}=(x_1,\ldots,x_{d-1})$, $x\in\mathbb{R}^d$. Then let $\mathscr{C}(\alpha,d,K)$ be the collection of all sets of the form

$$\{x\in I^d: 0\leqq x_d\leqq f(x_{(d)})\}$$

for all f on $\mathbb{R}^{d-1}$ with $\|f\|_\alpha\leqq K$.

As the most difficult result in this paper (Theorem 3) gives an asymptotic *lower* bound for empirical and Poisson processes over classes $\mathscr{C}(d-1,d,K)$, the relatively small class $\mathscr{C}(\alpha,d,K)$ just defined will suffice here. Larger classes with the same degree of boundary smoothness, but allowing unions and intersections of a fixed (for each class) number of such sets have been defined and studied previously ([5, 6, 14], Révész (1976a–b)). Corresponding asymptotic results will also hold for them.

Throughout, let P be the uniform Lebesgue measure on I^d. Let $X(1)$, $X(2),\ldots$, be independent and identically distributed with law P, $P_n=n^{-1}\sum_{j=1}^{n}\delta_{X(j)}$, and $\nu_n=n^{1/2}(P_n-P)$. Let $\mathscr{A}$ be the σ-algebra of measurable sets completed for P.

The main results of the paper will be stated in this section and proved in later sections.

For any collection $\mathscr{C}$ of measurable sets and finite signed measure μ let $\|\mu\|_{\mathscr{C}}=\sup_{A\in\mathscr{C}}|\mu(A)|$. If $\mathscr{C}=\mathscr{C}(\alpha,d,K)$ let

$$\|\mu\|_{\alpha,d,K}=\|\mu\|_{\mathscr{C}(\alpha,d,K)}.$$

Theorem 1. *If $0<\alpha<d-1$ then for any $K>0$ there is a $\delta>0$ such that for all possible values of P_n, we have*

$$\|P_n-P\|_{\alpha,d,K}\geqq\delta n^{-\alpha/(d-1+\alpha)}.$$

Remarks. Theorem 1 is related to a result of N.S. Bakhvalov (1959) and will be proved, by a technique similar to his, in Sect. 2.

Specifically, Bakhvalov (1959, Teorema 1, p. 6) implies that for each $d=1,2,\ldots$, and $\alpha>0$, there is a $\gamma=\gamma(d,\alpha)>0$ such that for all possible values of P_n, we have

$$\sup\{\textstyle\int f\,d(P_n-P): \|f\|_\alpha\leqq 1\}\geqq\gamma n^{-\alpha/d}. \tag{1.1}$$

This gives information for empirical measures mainly when $\alpha\leqq d/2$, since for $\alpha>d/2$ the supremum even over $\pm f$ for a single $f\neq 0$ tends to be of order $n^{-1/2}$.

In (1.1) one can, replacing γ by $\gamma/2$ if necessary, restrict the supremum to those f with $\int f\,dP=0$, implying the results of Kaufman and Philipp (1978, Sect.

4) with some improvements: $\delta=1$, $\varepsilon=0$, and no independence, lacunarity or other probabilistic assumption is needed.

The next result applies to a general probability space $(X, \mathscr{A}, Q)$. Here $\nu_n = n^{1/2}(Q_n - Q)$. For a collection $\mathscr{C} \subset \mathscr{A}$ and $\varepsilon > 0$ let [6]

$$N_I(\varepsilon, \mathscr{C}, Q) = \inf\{m\colon \exists A_1, \ldots, A_m \in \mathscr{A}\colon \forall A \in \mathscr{C}\ \exists i, j\colon A_i \subset A \subset A_j \text{ and } Q(A_j \setminus A_i) < \varepsilon\}.$$

Let $\Pr^*(A) = \inf\{\Pr(U)\colon U \supset A\}$.

Theorem 2. *If $\mathscr{C} \subset \mathscr{A}$ and for some constants ζ, $1 \leqq \zeta < \infty$, and $K < \infty$, $N_I(\varepsilon, \mathscr{C}, Q) \leqq \exp(K\varepsilon^{-\zeta})$ for $0 < \varepsilon \leqq 1$, then*

$$\Pr^*\{\|\nu_n\|_{\mathscr{C}} > n^{\theta}(\log n)^{\eta}\} \to 0$$

as $n \to \infty$, where $\theta = (\zeta - 1)/(2\zeta + 2)$ and for any $\eta > 2/(\zeta + 1)$.

Remarks. The classes $\mathscr{C}(\alpha, d, K)$ will satisfy the hypothesis of Theorem 2 for $\zeta = (d-1)/\alpha$ (Kolmogorov and Tikhomirov, 1959, Sect. 5, Theorems XIII–XV; [5, (3.2), as corrected, 1979, with its proof]). Then in Theorem 2, we get $\theta = \frac{1}{2} - \frac{\alpha}{d-1+\alpha}$. This θ cannot be reduced, by Theorem 1. Conversely, the exponent $\alpha/(d-1+\alpha)$ in Theorem 1 cannot be improved. It remains to find the best exponent η for $\log n$ in Theorem 2, a problem left open here.

The condition $\zeta \geqq 1$ in Theorem 2 is necessary, as we clearly cannot have $\theta < 0$ even for $\mathscr{C}$ consisting of a single set A with $0 < Q(A) < 1$.

Theorem 2 will also be proved in Sect. 2.

On I^d, for each $\lambda > 0$ let X_λ be the Poisson point process with intensity measure λP. That is, for each measurable set $A \subset I^d$, $X_\lambda(A)$ has a Poisson distribution with parameter $\lambda P(A)$, and for disjoint measurable sets A_i, the $X_\lambda(A_i)$ are independent. Let Y_λ be the centered process

$$Y_\lambda(A, \omega) = X_\lambda(A, \omega) - \lambda P(A),$$

which has mean 0 and still has independent values on disjoint sets. In Theorems 3 and 4 P is again uniform on I^d; $\ln = \log$.

Theorem 3. *If $0 < \alpha = d - 1$ then for any $K > 0$ and $\delta > 0$ there is a $c > 0$ such that as $\lambda \to +\infty$*

$$\Pr\{\|Y_\lambda\|_{\alpha,d,K} > c(\lambda \ln \lambda)^{1/2}(\ln\ln\lambda)^{-0.5-\delta}\} \to 1$$

and as $n \to \infty$,

$$\Pr\{n\|P_n - P\|_{\alpha,d,K} > c(n \ln n)^{1/2}(\ln\ln n)^{-0.5-\delta}\} \to 1.$$

Theorem 3 will be proved in Sects. 3–4. Kaufman (1980) proved that $\|\nu_n\|_{d-1,d,K}$ is unbounded in probability.

If $\alpha > d-1$, then $\|\nu_n\|_{\alpha,d,K}$ is bounded in probability (and further, central limit theorems and laws of the iterated logarithm hold: Révész, 1976b; Sun and Pyke, 1982; [6, 14]).

If $\alpha < d-1$, then it was known (without Theorem 1 above) that the central limit theorem must fail since the limiting Gaussian process is almost surely unbounded over $\mathscr{C}(\alpha, d, K)$ [4, Theorem 4.2]. For $\alpha = d-1$, the power η of $\log n$, between 1/2 (Theorem 3) and 1 (Theorem 2) remains to be settled.

A *lower layer* in $\mathbb{R}^2$ is a set B such that if $(x, y) \in B$, $u \leqq x$ and $v \leqq y$, then $(u, v) \in B$.

Steele (1978, Sect. 7), Wright (1981), and others they cite, prove laws of large numbers (Glivenko-Cantelli theorems) uniformly over the lower layers, for suitable P, as had R. Ranga Rao (1962) for convex sets, in all dimensions.

For the central limit theorem or law of the iterated logarithm the critical dimension is 2 for the lower layers and 3 for the convex sets. For these classes the central limit theorem fails since $\|v_n\|_{\mathscr{C}}$ is unbounded in probability [7]. The next result, apparently for the first time, allows us to conclude that $\|v_n\|_{\mathscr{C}}/(\log\log n)^{1/2}$ is also almost surely unbounded for $\mathscr{C}$ = lower layers in $\mathbb{R}^2$ or convex sets in $\mathbb{R}^3$ (for previous results see e.g. Stute, 1977).

Theorem 4. *The conclusion of Theorem 3 also holds if $d=2$ and $\mathscr{C}(1, 2, K)$ is replaced by the collection of all lower layers, or by the set of all lower layers in $\mathscr{C}(1, 2, K)$. For $d=3$, the conclusion of Theorem 3 holds if $\mathscr{C}(2, 3, K)$ is replaced by the collection of all convex sets in $\mathbb{R}^3$.*

Theorem 4 will be proved in Sect. 5.

Acknowledgment. Thanks to Walter Philipp for several conversations, specifically for the case $\zeta = 1$ in Theorem 2, where he has proved a more refined result, and for pointing out W. Schmidt's work. Thanks also to Rae Shortt for a helpful comment.

2. Proof of Theorems 1 and 2

First, to prove Theorem 1, let f be a C^∞ function on $\mathbb{R}^{d-1}$, 0 outside I^{d-1}, with $f > 0$ in the interior of I^{d-1}. For any ε with $0 < \varepsilon \leqq 1$, set $f_\varepsilon(x) = \varepsilon^\alpha f(x/\varepsilon)$. Then

$$\|f_\varepsilon\|_\alpha \leqq \|f\|_\alpha. \tag{2.1}$$

Given n, let $j = j_n = [(2nc)^{1/(\alpha+d-1)}]$ where $[x]$ is the greatest integer $\leqq x$ and $c = K\int f\,dx/(2\|f\|_\alpha)$. Then $1/(2n) \leqq c j^{1-d-\alpha}$. We have

$$\sup_{n \geqq 1} 2nc j_n^{1-d-\alpha} = M < \infty.$$

Let $\theta_n = j_n^{\alpha+d-1}/(2nc)$. Then $1/M \leqq \theta_n \leqq 1$.

Let $h = K\theta_n f/(2\|f\|_\alpha)$. We decompose I^{d-1} into j^{d-1} disjoint cubes C_i, $i = 1, \ldots, j^{d-1}$, of side $1/j$. Let $c_i \in I^{d-1}$ be the vertex of C_i closest to 0. Let

$$A_i = \{x \in I^d\colon 0 < x_d \leqq j^{-\alpha} h(j(x_{(d)} - c_i))\}.$$

Then the sets A_i are disjoint. Since by (2.1)

$$K\theta_n \|f_{1/j}\|_\alpha/(2\|f\|_\alpha) \leqq K/2,$$

the union of any set of the A_i, together with $\{x \in I^d\colon x_d = 0\}$ forms a set in $\mathscr{C}(\alpha, d, K)$. We have for each i

$$P(A_i) = j^{1-d-\alpha} \int h\, dx = j^{1-d-\alpha}\, \theta_n c = 1/(2n),$$

and $P_n(A_i) = 0$ or $P_n(A_i) \geqq 1/n$. Either at least half the A_i have $P_n(A_i) = 0$, or at least half have $P_n(A_i) \geqq 1/n$. In either case,

$$\|P_n - P\|_{\alpha, d, K} \geqq j^{d-1}/(4n) \geqq c j^{-\alpha}/(2M) \geqq \delta n^{-\alpha/(\alpha+d-1)}$$

for some $\delta = \delta(\alpha, d, K) > 0$, proving Theorem 1.

Proof of Theorem 2. The proof is similar to that of [6, Theorem 5.1]. Given n, let

$$k(n) = [(\tfrac{1}{2} - \theta) \cdot \log_2 n - \eta \log_2 \log n] \sim (\tfrac{1}{2} - \theta) \cdot \log_2 n$$

as $n \to \infty$ ($\log_2 = \log$ to base 2).

For $k = 1, 2, \ldots$, let $N(k) = N_I(2^{-k}, \mathscr{C}, Q)$. By its definition choose sets $A_{ki} \in \mathscr{A}$, $i = 1, \ldots, N(k)$, such that for all $A \in \mathscr{C}$ there are i and j with $A_{ki} \subset A \subset A_{kj}$ and $Q(A_{kj} \setminus A_{ki}) \leqq 2^{-k}$. Let $A_{01} = \emptyset$ (empty set) and $A_{02} = X$. For each $A \in \mathscr{C}$ and $k = 0, 1, \ldots$, choose such an $i = i(k, A)$ and $j = j(k, A)$. Then for $k \geqq 1$,

$$Q(A_{k, i(k,A)}\, \Delta A_{k-1, i(k-1, A)}) \leqq 2^{2-k},$$

where Δ denotes symmetric difference.

Let $\mathscr{B}(k)$ be the collection of all sets $B = A_{ki} \setminus A_{k-1,j}$ or $A_{k-1,j} \setminus A_{ki}$ or $A_{kj} \setminus A_{ki}$ such that $Q(B) \leqq 2^{2-k}$ $(k \geqq 1)$. Then $\operatorname{card}(\mathscr{B}(k)) \leqq 2N(k-1)N(k) + N(k)^2 \leqq 3\exp(2K\, 2^{k\zeta})$.

For each $B \in \mathscr{B}(k)$ we have by Bernstein's inequality (Bennett, 1962) for any $t > 0$

$$\text{(2.2)} \qquad \Pr\{|v_n(B)| > t\} \leqq \exp(-t^2/(2^{3-k} + t n^{-1/2})).$$

Set $t = t_{n,k} = n^{\theta} (\log n)^{\eta} c k^{-1-\delta}$ for a δ such that $0 < \delta < 1$ and $\eta > (2 + 2\delta)/(1 + \zeta)$, as is possible since $\eta > 2/(1+\zeta)$, and where $c = \delta/(1+\delta)$. Then for $1 \leqq k \leqq k(n)$,

$$2^{3-k} \geqq 8 n^{\theta - 1/2} (\log n)^{\eta} \geqq t_{n,k} n^{-1/2}.$$

Then (2.2) gives

$$\Pr\{|v_n(B)| > t_{n,k}\} \leqq \exp(-t_{n,k}^2/2^{4-k}).$$

Hence

$$\begin{aligned} p_{nk} &= \Pr\{\sup_{B \in \mathscr{B}(k)} |v_n(B)| > t_{n,k}\} \\ &\leqq 3\exp(2K\, 2^{k\zeta} - 2^{k-4} t_{n,k}^2) \\ &= 3\exp(2K\, 2^{k\zeta} - 2^{k-4} c^2 n^{2\theta} (\log n)^{2\eta} k^{-2-2\delta}). \end{aligned}$$

We have $2\theta/(\frac{1}{2} - \theta) = \zeta - 1 \geqq 0$. For $k \leqq k(n)$ we have $n^{0.5-\theta} \geqq 2^k (\log n)^{\eta}$, so $n^{2\theta} \geqq 2^{k(\zeta-1)} (\log n)^{\eta(\zeta-1)}$, and

$$p_{nk} \leqq 3\exp(2K\, 2^{k\zeta} - 2^{k\zeta-4} c^2 (\log n)^{\eta(\zeta+1)} k^{-2-2\delta}).$$

Now $k \leqq \log n$ since $\frac{1}{2} - \theta < \log 2$, so

$$p_{nk} \leqq 3 \exp(2^{k\zeta}(2K - c^2 2^{-4}(\log n)^\gamma)$$

where $\gamma = \eta(\zeta+1) - 2 - 2\delta > 0$ by choice of δ. Thus for n large enough so that $(\log n)^\gamma > 32\, K/c^2$,

$$\sum_{k=1}^{k(n)} p_{nk} \leqq 3(\log n) \exp(4K - c^2 2^{-3}(\log n)^\gamma) \to 0$$

as $n \to \infty$.

Let $\mathscr{E}_n$ be the event that $|v_n(B)| \leqq t_{n,k}$ for all $B \in \mathscr{B}(k)$, $k = 1, \ldots, k(n)$. Then $\Pr(\mathscr{E}_n) \to 1$ as $n \to \infty$. On $\mathscr{E}_n$, for each $A \in \mathscr{C}$ and $i = i(k(n), A)$, $j = j(k(n), A)$,

$$\begin{aligned} |v_n(A_{k(n),i})| &\leqq 2 \sum_{k=1}^{k(n)} t_{n,k} = 2cn^\theta (\log n)^\eta \sum_{k \geqq 1} k^{-1-\delta} \\ &\leqq 2n^\theta (\log n)^\eta, \end{aligned}$$

and $|v_n(A_{k(n),j} \setminus A_{k(n),i})| \leqq n^\theta (\log n)^\eta$. Now

$$n^{1/2} Q(A_{k(n),j} \setminus A_{k(n),i}) \leqq n^{1/2}/2^{k(n)} < 2n^\theta (\log n)^\eta.$$

Thus $n^{1/2} Q_n(A_{k(n),j} \setminus A_{k(n),i}) \leqq 3n^\theta (\log n)^\eta$, so $|v_n(A \setminus A_{k(n),i})| \leqq 3n^\theta (\log n)^\eta$. Thus on $\mathscr{E}_n$, $|v_n(A)| \leqq 5n^\theta (\log n)^\eta$ for our arbitrary $A \in \mathscr{C}$. Using a smaller η we can drop the 5, proving Theorem 2.

If $\alpha = d - 1$, the method of proof of Theorem 1 shows that for larger n there are smaller and smaller sets on which v_n is not small. But Theorem 3 gives better information.

The proof of (1.1) above by Bakhvalov (1959) is like the proof of Theorem 1 here, replacing $d-1$ by d, and letting $j = [(2n)^{1/d}] + 1$. On the $j^d \gtrsim 2n$ little cubes C_i let $g = 0$ on those with $P_n(C_i) > 0$. On all other C_i let $g(x) = -f_{1/j}(x - c_i)$. Then $\|g\|_\alpha \leqq \|f\|_\alpha$ and using g gives (1.1). Bakhvalov notes, in turn, that Kolmogorov had used a similar method to prove a lower bound for the metric entropy of classes $\{f\colon \|f\|_\alpha \leqq K\}$ in the supremum norm [13]. Later, W. Schmidt (1975) applied such a method to classes of convex sets.

3. Lemmas: Poissonization and Random Sets

First, let us relate empirical and Poisson processes ("Poissonization"). Consider the following property of a function f defined for large enough $x > 0$:

(*) for any $\varepsilon > 0$ there is a $\delta > 0$ such that whenever

$$x > 1/\delta \quad \text{and} \quad \left|1 - \frac{y}{x}\right| < \delta \quad \text{then} \quad \left|1 - \frac{f(y)}{f(x)}\right| < \varepsilon.$$

If f is continuous, and slowly varying (Karamata), i.e. for all $k > 0$, $f(kx)/f(x) \to 1$ as $x \to +\infty$; or if f is regularly varying, i.e. for some real r, $f(x) \equiv x^r L(x)$ where L is slowly varying, then (*) holds (see e.g. Feller, Vol. II, VIII.8, Lemma 2). The following Lemma will treat more general situations than are needed in this paper. For any probability space $(S, \mathscr{A}, P)$ we can define the Poisson processes X_λ and Y_λ, as in Sect. 1 for the uniform P, and the empirical processes P_n and ν_n.

(3.1) **Lemma.** *Let $(S, \mathscr{A}, P)$ be a probability space and $\mathscr{C} \subset \mathscr{A}$, where we assume that for each n and constant t, $\sup_{A \in \mathscr{C}} (P_n - tP)(A)$ is measurable. Let f be a function satisfyng (*) such that $f(\lambda) \to +\infty$ and $f(\lambda)/\lambda \to 0$ as $\lambda \to +\infty$. Suppose that for some constant $c > 0$ and some κ with $0 \leqq \kappa \leqq 1$, we have*

$$\liminf_{\lambda \to \infty} \Pr\{\sup_{A \in \mathscr{C}} Y_\lambda(A) \geqq c f(\lambda)\,\lambda^{1/2}\} = \kappa.$$

Then for $0 < K < c$ we have

$$\liminf_{n \to \infty} \Pr\{\sup_{A \in \mathscr{C}} \nu_n(A)/f(n) \geqq K\} \geqq \kappa.$$

Proof. Let $n(\lambda)$ be a Poisson variable with parameter λ, independent of the $X(i)$. Then $n(\lambda) P_{n(\lambda)}$ has the properties of X_λ, as is well known (Kac, 1949; Csörgő and Révész, 1981, pp. 250–251). Also, $n(\lambda)$ can be defined from X_λ or Y_λ by $n(\lambda) = X_\lambda(S) = (Y_\lambda + \lambda P)(S)$. We can then write

$$\begin{aligned} Y_\lambda &= n(\lambda) P_{n(\lambda)} - \lambda P = n(\lambda)(P_{n(\lambda)} - P) + (n(\lambda) - \lambda) P, \\ Y_\lambda/\lambda^{1/2} &= (n(\lambda)/\lambda)^{1/2} \nu_{n(\lambda)} + (n(\lambda) - \lambda)\lambda^{-1/2} P, \end{aligned} \tag{3.2}$$

where if $n(\lambda) = 0$, we replace $P_{n(\lambda)}$ and $\nu_{n(\lambda)}$ by 0. As $\lambda \to \infty$, $\Pr(n(\lambda) > 0) \to 1$, $n(\lambda)/\lambda^{1/2} \to 1$ in probability, and $(n(\lambda) - \lambda)\lambda^{-1/2}$ is bounded in probability.

From this construction we see that $\sup_{A \in \mathscr{C}} Y_\lambda(A)$ is also measurable.

If the Lemma fails, then there is a $\theta < \kappa$ and an infinite sequence of values $m = m_k \to +\infty$ such that

$$\Pr(\sup_{A \in \mathscr{C}} \nu_m(A) \geqq K f(m)) \leqq \theta.$$

Take $0 < \varepsilon < 1/3$ such that $K(1 + 7\varepsilon) < c$. Then take a $\delta > 0$ such that (*) holds for f. We may assume $\delta < 1/2$ and $(1+\delta)(1+5\varepsilon) < 1 + 6\varepsilon$. We may also assume that for all k, $m = m_k \geqq 2/\delta$ and $1 + 2\varepsilon < K\varepsilon f(m)^{1/2}$.

Let $\delta_m = (f(m)/m)^{1/2}$. Then since $f(m)/m \to 0$ we may assume that $\delta_m < \delta/2$ for all $m = m_k$. For any $m = m_k$, if $(1 - \delta_m) m \leqq n \leqq m$ and $A \in \mathscr{A}$, then $m P_m(A) \geqq n P_n(A)$, so $m(P_m - P)(A) \geqq n(P_n - P)(A) - m\delta_m$, and

$$\nu_m(A) \geqq (n/m)^{1/2} \nu_n(A) - f(m)^{1/2} \geqq (1+\delta)^{-1} \nu_n(A) - f(m)^{1/2},$$

or $\nu_n(A) \leqq (1+\delta)(\nu_m(A) + f(m)^{1/2})$. Now

$$\left|1 - \frac{m}{n}\right| = \frac{m}{n} - 1 < 2\delta_m < \delta \quad \text{implies} \quad \left|\frac{f(m)}{f(n)} - 1\right| < \varepsilon$$

and $1/f(n) < (1+\varepsilon)/f(m)$, so if $\nu_n(A) \geqq 0$ then

$$\nu_n(A)/f(n) \leqq (1 + 2\varepsilon)(\nu_m(A) f(m)^{-1} + f(m)^{-1/2}).$$

So since $(1+2\varepsilon)f(m)^{-1/2}<K\varepsilon$,

$$\Pr\{\sup_{A\in\mathscr{C}} \nu_n(A)\geqq Kf(n)(1+3\varepsilon)\}\leqq\theta.$$

For each $m=m_k$, set $\lambda=(1-\delta_m/2)m$. Then as $k\to\infty$, since $f(m_k)\to\infty$ we will have

$$\Pr((1-\delta_m)m\leqq n(\lambda)\leqq m)\to 1.$$

Then for any γ with $\theta<\gamma<\kappa$ and k large enough, since the $X(i)$ are independent of $n(\lambda)$,

$$\Pr\{\sup_{A\in\mathscr{C}} \nu_{n(\lambda)}(A)\geqq Kf(n(\lambda))(1+3\varepsilon)\}<\gamma.$$

By (*), for $(1-\delta_m)m\leqq n\leqq m$ we have $\left|1-\frac{f(n)}{f(\lambda)}\right|<\varepsilon$, so that for k large enough we may assume

$$\Pr\{\sup_{A\in\mathscr{C}} \nu_{n(\lambda)}(A)\geqq Kf(\lambda)(1+5\varepsilon)\}<\gamma.$$

By (3.2), we then have for k large enough

$$\Pr\{\sup_{A\in\mathscr{C}} Y_\lambda(A)\geqq K\lambda^{1/2}f(\lambda)(1+7\varepsilon)\}<(\gamma+\kappa)/2<\kappa,$$

a contradiction, proving Lemma 3.1.

Note. Pyke (1968) has related estimates with and without Poissonization.

For any real x let $x^+=\max(x,0)$.

(3.3) **Lemma.** *There is a constant $c>0$ such that whenever z has a Poisson law with parameter $m\geqq 1$ then*

$$E(z-m)^+\geqq cm^{1/2}.$$

Proof. Let j be the greatest integer $\leqq m$. Then by a telescoping sum and Stirling's formula with an error bound (e.g. Feller, Vol. I, Sect. II.9, p. 54),

$$\begin{aligned}E(z-m)^+ &= \sum_{k>m} e^{-m}m^k(k-m)/k!\\ &\geqq me^{-m}m^j/((j/e)^j(2\pi j)^{1/2}e^{1/(12j)})\\ &\geqq (m^{j+1}/j^{j+1/2})e^{-13/12}(2\pi)^{-1/2}\geqq cm^{1/2}\ \text{(with } c=0.135).\quad \text{Q.E.D.}\end{aligned}$$

Now the Poisson process X_λ has the property that for any two disjoint measurable sets A, $B\in\mathscr{A}$, given the σ-algebra generated by X_λ on B and its measurable subsets, or any sub-σ-algebra $\mathscr{G}$ of that σ-algebra, the conditional distribution of $X_\lambda(A)$ given $\mathscr{G}$ is Poisson with parameter $\lambda P(A)$. This property will be extended to suitable random sets $A=C_\omega$ and $B=L_\omega$ where $P(A)$ is $\mathscr{G}$-measurable.

More generally, let $(X,\mathscr{A})$ be a measurable space. Let $\mathscr{A}_f\subset\mathscr{A}$ be such that for any $A,B\in\mathscr{A}_f$ and $C\in\mathscr{A}$, $A\cup B\in\mathscr{A}_f$ and $A\cap C\in\mathscr{A}_f$. (For example if $(X,\mathscr{A},\mu)$ is a σ-finite measure space, we can take $\mathscr{A}_f=\{A\in\mathscr{A}\colon \mu(A)<\infty\}$.) Let

$Y: \langle A,\omega\rangle \to Y(A)(\omega)$ be a real-valued stochastic process indexed by $A \in \mathscr{A}_f$, with $\omega \in \Omega$ for a probability space $(\Omega, \mathscr{B}, \Pr)$, such that for any disjoint $A_1, \dots, A_n \in \mathscr{A}_f$, $Y(A_j)$ are independent, $1 \leqq j \leqq m$, and $Y(A_1 \cup A_2) = Y(A_1) + Y(A_2)$. Then we say Y has *independent pieces*. Clearly each Y_λ is such a process.

For each $C \in \mathscr{A}$ let $\mathscr{B}_C$ be the smallest σ-algebra for which all $Y(A)(\cdot)$ are measurable for $A \subset C$, $A \in \mathscr{A}_f$.

Let G be a function from Ω into $\mathscr{A}$. Then (by analogy with stopping times) we call $G(\cdot)$ a *stopping set* iff for all $C \in \mathscr{A}$, $\{\omega: G(\omega) \subset C\} \in \mathscr{B}_C$.

Given a stopping set $G(\cdot)$, let B_G denote the σ-algebra of all sets $B \in \mathscr{B}$ such that for all $C \in \mathscr{A}$, $B \cap \{G \subset C\} \in \mathscr{B}_C$.

(3.4) **Lemma.** *Suppose Y has independent pieces and we have random sets $G_j(\omega) = G(j)(\omega)$, $j = 0, \dots, m$, and $A(\omega)$ in $\mathscr{A}_f$ such that:*

i) *$G_0(\omega)$ is a fixed set $G_0 \in \mathscr{A}_f$;*

ii) *for all ω, $G_0 \subset G_1(\omega) \subset \dots \subset G_m(\omega)$, and $G_m(\omega) \cap A(\omega) = \emptyset$;*

iii) *each $G_j(\omega)$, and $A(\omega)$, has only countably many possible values $G(j,i) = G_{ji}$ and $C_i = C(i)$ respectively;*

(iv) *for all i and for $1 \leqq j \leqq m$, $\{G_j(\cdot) = G_{ji}\} \in \mathscr{B}_{G(j-1)}$ and $\{A(\cdot) = C_i\} \in \mathscr{B}_{G(m)}$.*

Then the $G_j(\cdot)$ are all stopping sets, and the conditional probability

$$\Pr\{Y(A)(\cdot)) \leqq t \mid \mathscr{B}_{G(m)}\} = \sum_i 1_{\{A(\omega) = C(i)\}} \Pr\{Y(C_i) \leqq t\}$$

almost surely, for each $t \in \mathbb{R}$.

Proof. First it will be shown by induction on j that $G_j(\cdot)$ are stopping sets. Clearly G_0 is. For the induction step, given $C \in \mathscr{A}$ and $j \geqq 1$,

$$\{G_j \subset C\} = \bigcup_i \{\{G_j = G_{ji}\}: G_{ji} \subset C\}.$$

For each i, $\{G_j = G_{ji}\} \in \mathscr{B}_{G(j-1)}$. If $G_{ji} \subset C$, then by ii),

$$\{G_j = G_{ji}\} = \{G_j = G_{ji}\} \cap \{G_{j-1} \subset C\} \in \mathscr{B}_C$$

by definition of $\mathscr{B}_{G(j-1)}$ and the induction hypothesis. Thus $\{G_j \subset C\}$, as a countable union of sets in $\mathscr{B}_C$, is in $\mathscr{B}_C$, so G_j is a stopping set.

If $A(\omega) = C_i$ and $G_m(\omega) = G_{mj}$ for some ω, then $C_i \cap G_{mj} = \emptyset$ by ii), so $Y(C_i)$ is independent of $\mathscr{B}_{G_{mj}}$. Let $B_i = \{A(\cdot) = C_i\} \in \mathscr{B}_{G(m)}$ by iv). For each j, $\{G_m = G_{mj}\} \in \mathscr{B}_{G(m-1)}$, by iv), so by ii),

$$\{G_m = G_{mj}\} = \{G_m = G_{mj}\} \cap \{G_{m-1} \subset G_{mj}\} \in \mathscr{B}_{G_{mj}}.$$

For any $B \in \mathscr{B}_{G(m)}$, $B \cap \{G_m = G_{mj}\} = B \cap \{G_m \subset G_{mj}\} \cap \{G_m = G_{mj}\} \in \mathscr{B}_{G_{mj}}$. So a.s.

$$\begin{aligned}\Pr(Y(A) \leqq t \mid \mathscr{B}_{G_m}) &= \sum_{i,j} \Pr(Y(A) \leqq t \mid \mathscr{B}_{G_m}) 1_{B_i} 1_{\{G_m = G_{mj}\}} \\ &= \sum_{i,j} \Pr(Y(C_i) \leqq t \mid \mathscr{B}_{G_{mj}}) 1_{B_i} 1_{\{G_m = G_{mj}\}} \\ &= \sum_i \Pr(Y(C_i) \leqq t) 1_{B_i}. \quad \text{Q.E.D.}\end{aligned}$$

Note. Evstigneev (1977, Theorem 1) proves a strong Markov property for suitable random fields, indexed by closed subsets of a Euclidean space, which might also be used, with some work, in place of Lemma 3.4 above.

4. Proof of Theorem 3

$\|P_n - P\|_{\alpha,d,K}$ is a measurable random variable [6, proof of (5.12)]. Writing $X_\lambda = n(\lambda) P_{n(\lambda)}$ as in the proof of Lemma (3.1) we see that $\|Y_\lambda\|_{\alpha,d,K}$ is also measurable.

Theorem 3 for Y_λ implies it for $P_n - P$, taking a smaller value of c and using (3.1). Let us prove it for Y_λ.

By assumption $d \geqq 2$. Let $J = [0, 1[$, so

$$J^{d-1} = \{x \in \mathbb{R}^{d-1}\colon 0 \leqq x_j < 1, j = 1, \dots, d-1\}.$$

Let $x_{(d)} = \langle x_1, \dots, x_{d-1} \rangle$, $x \in \mathbb{R}^d$.

Let f be a C^∞ function on $\mathbb{R}^{d-1}$ with $f(x) = 0$ outside the unit cube J^{d-1} and $f(x) > 0$ for x in the interior of J^{d-1}. We choose f so that

$$\sup_x \sup_{[p] \leqq d-1} |D^p f(x)| \leqq 1$$

and such that at all points x of the sub-cube

$$\{x\colon 1/3 \leqq x_j \leqq 2/3, j = 1, \dots, d-1\}, \quad f(x) = \sup\{f(t)\colon t \in I^{d-1}\} = \gamma,$$

say, $\gamma < 1$.

A sequence of sets and functions, some of them random, will be defined recursively as follows. As the jth stage, $j = 1, 2, \dots$, J^{d-1} is decomposed into $3^{j(d-1)}$ disjoint sub-cubes A_{ji} of side 3^{-j} for $i = 1, \dots, 3^{j(d-1)}$, where each A_{ji} is also a Cartesian product of left closed, right open intervals.

Let B_{ji} be a cube of side 3^{-j-1}, concentric with and parallel to A_{ji}. Let x_{ji} be the point of A_{ji} closest to 0 and y_{ji} the point of B_{ji} closest to 0. For $\delta > 0$ and $j = 1, 2, \dots$, let

$$c_j = C j^{-1} (\log(j+1))^{-1-2\delta}$$

where the contant $C > 0$ is chosen so that $\sum_{j \geqq 1} c_j \leqq 1$.

For $x \in \mathbb{R}^{d-1}$ let

$$f_{ji}(x) = c_j 3^{-j(d-1)} f(3^j(x - x_{ji})),$$
$$g_{ji}(x) = c_j 3^{-(j+1)(d-1)} f(3^{j+1}(x - y_{ji})).$$

Then the supports of f_{ji} and g_{ji} are the closures of A_{ji} and B_{ji} respectively. Note that on B_{ji}, $f_{ji}/g_{ji} \geqq 3^{d-1} > 1$.

Let $S_0 = 1/2$. We will define recursively a sequence of random variables $s_{ji}(\omega) = \pm 1$ and let

$$S_k = S_0 + \sum_{j=1}^{k} \sum_{i=1}^{3^{j(d-1)}} s_{ji}(\omega) f_{ji}. \tag{4.0}$$

Then since $\sum_{j \geqq 1} c_j 3^{-j(d-1)} \leqq 1/3$, $d \geqq 2$, we have $0 < S_k < 1$ for all k.

Given S_{j-1}, $j\geqq 1$, let $C_{ji}=C_{ji}(\omega)=\{x\in J^d\colon |x_d-S_{j-1}(x_{(d)})(\omega)|<g_{ji}(x_{(d)})\}$. Let $s_{ji}(\omega)=+1$ if $Y_\lambda(C_{ji}(\omega))>0$, otherwise $s_{ji}(\omega)=-1$. This completes the recursive definition of the s_{ji} and so of the S_k.

Recursively, one sees that each C_{ji} has only finitely many possible values, each on a measurable event, so the s_{ji} and S_k are all measurable.

Since the interiors of the A_{ji} are disjoint, we have for any $k\geqq 1$

$$\sup_{[p]\leqq d-1}\sup_x |D^p S_k(x)|\leqq \sum_{j=1}^{k}\sup_{[p]\leqq d-1}|D^p f_{j1}(x)|\leqq \sum_{j=1}^{k} c_j<1.$$

The volume of $C_{ji}(\omega)$ is always

$$P(C_{ji}(\omega))=2\int g_{ji}\,dx=2\mu c_j/9^{(j+1)(d-1)} \tag{4.1}$$

where $0<\mu=\int f\,dx<1$.

Now let us show that $C_{ji}(\omega)$ for different i,j are always disjoint. For $1\leqq j\leqq k$ and any ω, i and $x\in B_{ji}$ we have

$$\begin{aligned}|S_k(x)-S_{j-1}(x)|(\omega)&\geqq \gamma c_j 3^{-j(d-1)}-\sum_{r>j}\gamma c_r 3^{-r(d-1)}\\ &\geqq \gamma c_j 3^{-j(d-1)}(1-\tfrac{1}{2})\\ &\geqq \gamma c_j 3^{-j(d-1)}/2\geqq \sup_y (g_{ji}+\sup_r g_{k+1,r})(y).\end{aligned}$$

Thus if $s_{ji}(\omega)=+1$, then for any r and any $x\in B_{ji}$,

$$S_k(x)(\omega)-g_{k+1,r}(x)\geqq S_{j-1}(x)(\omega)+g_{ji}(x),$$

so $C_{ji}(\omega)$ is disjoint from $C_{k+1,r}(\omega)$. They are likewise disjoint if $s_{ji}(\omega)=-1$, interchanging $+$ and $-$, $\geqq$ and $\leqq$.

For the same j and different i, the $C_{ji}(\omega)$ are disjoint since they project into disjoint B_{ji}.

Given $\lambda>0$ let $r=r(\lambda)$ be the largest j, if one exists, such that $2\lambda\mu c_j\geqq 9^{(j+1)(d-1)}$. Then as $\lambda\to+\infty$, $r(\lambda)\sim(\log\lambda)/((d-1)\log 9)$.

Let $G_0(\omega)=\emptyset$. For $m=1,2,\dots$, let

$$G(m)(\omega)=G_m(\omega)=\bigcup\{C_{ji}(\omega)\colon j\leqq m\}.$$

Let $H_m(\omega)=\{x\colon 0\leqq x_d\leqq S_m(x_{(d)})(\omega)\}$, so that for all ω, $H_m\in\mathscr{C}(d-1,d,2d)$. Let $A_m(\omega)=H_m(\omega)\setminus G_m(\omega)$. Then from the disjointness proof,

$$H_m(\omega)\cap G_m(\omega)=\bigcup\{C_{ji}(\omega)\colon j\leqq m,\ s_{ji}=+1\}.$$

For each m, each of these sets has finitely many possible values, each on a measurable event. For each m, we apply Lemma 3.4 to these G_j, with $A=A_m$. Then each G_j is $\mathscr{B}_{G(j-1)}$ measurable, $A_m(\cdot)$ is $\mathscr{B}_{G(m)}$ measurable, and the other hypotheses of Lemma (3.4) clearly hold. Thus, conditional on $\mathscr{B}_{G(m)}$, $X_\lambda(A)$ is Poisson with parameter $\lambda P(A_m(\omega))$. Also,

$$Y_\lambda((H_m\cap G_m)(\omega))=\sum_{j=1}^{m}\sum_{i=1}^{3^{j(d-1)}} Y_\lambda(C_{ji})^+. \tag{4.2}$$

Now $P(C_{ji})$ does not depend on ω nor i, and $C_{ji}(\cdot)$ is $\mathscr{B}_{G(j-1)}$ measurable. Thus by Lemma (3.4), applied to $A=C_{mi}$, replacing G_m by $G_m \setminus C_{mi}$, the $Y_\lambda(C_{ji})$ for different i or j are jointly independent, and each has the law of $Y_\lambda(C)$ for a fixed set C with $P(C)=2\int g_{j1}\,dx$.

Taking $m=r(\lambda)$, (4.2) is a sum of independent nonnegative parts of centered Poisson variables with parameters $\lambda P(C_{ji}) \geqq 1$, by (4.1). Thus by Lemma (3.3), for a constant $c>0$,

$$\begin{aligned}
\lambda^{-1/2} E Y_\lambda((H_m \cap G_m)(\omega)) &\geqq c \sum_{j=1}^{r(\lambda)} \sum_{i=1}^{3^{j(d-1)}} P(C_{ji})^{1/2} \\
&= c \sum_{j=1}^{r(\lambda)} 3^{j(d-1)} (2\mu c_j / 9^{(j+1)(d-1)})^{1/2} \qquad \text{by (4.1)} \\
&= c\, 3^{1-d} (2\mu)^{1/2} \sum_{j=1}^{r(\lambda)} (C j^{-1} (\log(j+1))^{-1-2\delta})^{1/2} \qquad \text{(def. of } c_j) \\
&= c (2\mu C)^{1/2} 3^{1-d} \sum_{j=1}^{r(\lambda)} j^{-1/2} (\log(j+1))^{-0.5-\delta} \\
&\geqq a_d (\log(r(\lambda)+1))^{-0.5-\delta} \sum_{j=1}^{r(\lambda)} j^{-1/2} \\
&\geqq 2 a_d (r(\lambda)^{1/2} - 1)(\log(r(\lambda)+1))^{-0.5-\delta}
\end{aligned}$$

for some constant $a_d>0$. For λ large the above is

$$\geqq 3 b_d (\log \lambda)^{1/2} (\log\log \lambda)^{-0.5-\delta}$$

for some $b_d>0$.

By independence of the $Y_\lambda(C_{ji})$, the variance of $Y_\lambda((H_{r(\lambda)} \cap G_{r(\lambda)})(\omega)$ is less than

$$\sum_{j=1}^{r(\lambda)} \sum_{i=1}^{3^{j(d-1)}} \lambda P(C_{ji}) = \lambda \sum_{j=1}^{r(\lambda)} 3^{j(d-1)} 2\mu C j^{-1} (\log(j+1))^{-1-2\delta} q^{-(j+1)(d-1)} < \lambda.$$

Thus by Chebyshev's inequality,

$$\Pr\{\lambda^{-1/2} Y_\lambda(H_{r(\lambda)} \cap G_{r(\lambda)})(\omega) \geqq 2 b_d (\log \lambda)^{1/2} (\log\log \lambda)^{-0.5-\delta}\} \to 1$$

as $\lambda \to +\infty$.

By Lemma (3.4), the conditional distribution of $Y_\lambda(A_{r(\lambda)}(\omega))(\omega)$ given $\mathscr{B}_{G(r(\lambda))}$ is that of $Y_\lambda(D)$ for $P(D)=P(A_{r(\lambda)}(\omega))$, where $E Y_\lambda(D)^2 \leqq \lambda$ for all D. Thus $E Y_\lambda (D_{r(\lambda)})^2 \leqq \lambda$, $Y_\lambda(A_{r(\lambda)})/\lambda^{1/2}$ is bounded in probability, and

$$\Pr\{Y_\lambda(H_{r(\lambda)}) \geqq b_d (\lambda \log \lambda)^{1/2} (\log\log \lambda)^{-0.5-\delta}\} \to 1$$

as $\lambda \to +\infty$. This proves Theorem 3 (for Y_λ) if $K \geqq 2d$. For smaller values of $K>0$ we can just multiply the constant C (in c_j, f_{ji} and g_{ji}) by $K/(2d)$, completing the proof of Theorem 3.

5. Convex Sets and Lower Layers

We have measurability of the relevant norms as in [6, (4.3), (4.4), (5.13)]. Theorem 4 will be proved first for the convex sets in case $d=3$. In the proof of Theorem 3 in the last section, let us take $d=3$ and define S_k, $k \geqq 1$, by (4.0) with a *new* definition of S_0:

$$S_0 = \tfrac{3}{4} - (x_1 - \tfrac{1}{2})^2 - (x_2 - \tfrac{1}{2})^2.$$

Then since $\sum_{j \geqq 1} c_j/9^j < 1/4$, $0 < S_k < 1$ for all k. The rest of the proof remains the same, to prove Theorem 4 for the convex sets, except for the consideration of second derivatives, as follows. We now get

$$\sup_{[p] \leqq 2} \sup_x |D^p(S_k - S_0)(x)| \leqq 1.$$

Let $H(f)$ denote the Hessian matrix $H_{ij} = \partial^2 f/\partial x_i \partial x_j$ for the function f. Then $H(S_0) \equiv \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$ and $H(S_k - S_0)$ is symmetric with all its entries in $[-1, 1]$. Hence $-H(S_k)$ is everywhere nonnegative definite. Since S_k is C^∞, it is concave (Roberts and Varberg, 1973, pp. 100, 103) so that the set

$$H_{r(\lambda)}(\omega) = \{x: 0 \leqq x_d \leqq S_{r(\lambda)}(x_{(d)})(\omega)\}$$

is now convex. Thus Theorem 4 is proved for the convex sets in $\mathbb{R}^3$.

To prove Theorem 4 for lower layers, we take $d=2$ in the proof in Sect. 4 and make the following changes. Choose C now so that $\sum_j c_j < 1/2$. Let $S_0 = \dfrac{3}{4} - \dfrac{x}{2}$. Then $0 < S_k < 1$ for all k, since $\sum_{j \geqq 1} c_j 3^{-j} < 1/6$. Also

$$S_k'(x) \leqq -\tfrac{1}{2} + \sum_{j \geqq 1} c_j < 0$$

for all k and x. Thus each H_k is a lower layer. The rest of the proof works as before, proving Theorem 4.

References

1. Bakhvalov [Bahvalov], N.S.: On approximate calculation of multiple integrals (in Russian). Vestnik Mosk. Univ. Ser. Mat. Mekh. Astron. Fiz. Khim. **1959**, no. 4, 3–18 (1959)
2. Bennett, G.: Probability inequalities for sums of independent random variables. J. Amer. Statist. Assoc. **57**, 33–45 (1962)
3. Csörgő, M., Révész, P.: Strong Approximations in Probability and Statistics. Akadémiai Kiadó, Budapest (1981)
4. Dudley, R.M.: Sample functions of the Gaussian process. Ann. Probab. **1**, 66–103 (1973)
5. Dudley, R.M.: Metric entropy of some classes of sets with differentiable boundaries. J. Approximation Theory **10**, 227–236 (1974); Correction, ibid. **26**, 192–193 (1979)
6. Dudley, R.M.: Central limit theorems for empirical measures. Ann. Probab. **6**, 899–929 (1978); Correction, ibid. **7**, 909–911 (1979)

7. Dudley, R.M.: Lower layers in $\mathbb{R}^2$ and convex sets in $\mathbb{R}^3$ are not *GB* classes. Lecture Notes in Math. no. **709**, 97–102. Berlin-Heidelberg-New York: Springer 1979
8. Evstigneev, I.V.: "Markov times" for random fields. Theor. Probability Appls. **22**, 563–569 = Teor. Verojatnost. i Primenen. **22**, 575–581 (1977)
9. Feller, W.: An Introduction to Probability Theory and its Applications, vol. I, 3d ed.; vol. II, 2d ed. New York: Wiley 1968 and 1971
10. Kac, M.: On deviations between theoretical and empirical distributions. Proc. Nat. Acad. Sci. USA **35**, 252–257 (1949)
11. Kaufman, R., Walter Philipp: A uniform law of the iterated logarithm for classes of functions. Ann. Probab. **6**, 930–952 (1978)
12. Kaufman, Robert: Smooth functions and Gaussian processes. Approximation Theory III, ed. E.W. Cheney, p. 561–564. N.Y.: Academic Press 1980
13. Kolmogorov, A.N., Tikhomirov, V.M.: ε-entropy and ε-capacity of sets in functional spaces. Amer. Math. Soc. Transl. (Ser. 2) **17**, 277–364 (1961) = Uspekhi Mat. Nauk. **14**, vyp. 2(86), 3–86 (1959)
14. Kuelbs, J., Dudley, R.M.: Log log laws for empirical measures. Ann. Probab. **8**, 405–418 (1980)
15. Pyke, R.: The weak convergence of the empirical process with random sample size. Proc. Cambridge Philos. Soc. **64**, 155–160 (1968)
16. Rao, R. Ranga: Relations between weak and uniform convergence of measures with applications. Ann. Math. Statist. **33**, 659–680 (1962)
17. Révész, P.: On strong approximation of the multidimensional empirical process. Ann. Probab. **4**, 729–743 (1976)
18. Révész, P.: Three theorems of multivariate empirical process. Empirical Distributions and Processes, ed. P. Gaenssler and P. Révész. Lecture Notes in Math. **566**, 106–126. Berlin-Heidelberg-New York: Springer 1976
19. Roberts, A.W., Varberg, D.E.: Convex Functions. New York: Academic Press 1973
20. Schmidt, W.: Irregularities of distribution IX. Acta Arith. **27**, 385–396 (1975)
21. Steele, J. Michael: Empirical discrepancies and subadditive processes. Ann. Probab. **6**, 118–127 (1978)
22. Stute, W.: Convergence rates for the isotrope discrepancy. Ann. Probab. **5**, 707–723 (1977)
23. Sudakov, V.N.: Gaussian and Cauchy measures and ε-entropy. Soviet Math. Doklady **10**, 310–313 (1969)
24. Sun, Tze-Gong, Pyke, R.: Weak convergence of empirical processes. Technical Report, Dept. Statist. Univ. Washington, Seattle (1982)
25. Wright, F.T.: The empirical discrepancy over lower layers and a related law of large numbers. Ann. Probab. **9**, 323–329 (1981)

Received April 29, 1982; in final form June 15, 1982