V.N. Sudakov's Work on Expected Suprema of Gaussian Processes

Richard M. Dudley

Dedicated to the memory of Evarist Giné.

Abstract It is noted that the late Volodya N. Sudakov (1934–2016) first published a statement in 1973 and proof in 1976 that the expected supremum of a centered Gaussian process is bounded above by a constant times a metric entropy integral. In particular, the present author (R.M. Dudley) defined such an integral but did not state nor prove such a bound.

Keywords Metric entropy

Mathematics Subject Classification (2010). Primary 60G15

1 Introductory Remarks

Vladimir N. Sudakov reached his 80th birthday in 2014. A rather well known fact, which I'll call a majorization inequality, says that the expected supremum of a centered Gaussian process is bounded above by a constant times a metric entropy integral. Who first (a) called attention to the expected supremum, (b) stated such an inequality, and (c) published a proof of it, when? My answer in all three cases is Sudakov (1973, for (a) and (b); 1976, for (c)) [19, 20]. I defined the metric entropy integral, as an equivalent sum in 1967, then explicitly in 1973, and showed that its finiteness implies sample continuity. Sudakov's work on Gaussian processes has perhaps been best known for his minoration; that he was first to state and give a proof for a majorization inequality seems to have passed almost unnoticed, and I hope to rectify that.

R.M. Dudley (✉)
Professor Emeritus, Mathematics, Massachusetts Institute of Technology, 92 Lewis St. Newton, MA 02458, USA
e-mail: rmd@math.mit.edu

C. Houdré et al. (eds.), *High Dimensional Probability VII*,
Progress in Probability 71, DOI 10.1007/978-3-319-40519-3_2

2 Sudakov and Strassen

At the International Congress of Mathematicians in Moscow in the summer of 1966, Sudakov gave a talk, in Russian, which applied metric entropy $\log(N(C, d, \varepsilon))$ (see Sect. 3), then called ε-entropy, of sets C in a Hilbert space H, to sample continuity and boundedness of the isonormal process L on H, the Gaussian process having mean 0 and covariance equal to the inner product, restricted to C. As far as I know this was the first presentation, oral or written, of such results by anyone, to an international audience. I attended the Moscow 1966 talk and took notes as best I could with my meager Russian. When I looked back at the notes later, I regretted not having absorbed the significance of Sudakov's talk at first. The notion of isonormal process on a Hilbert space originated, as far as I know, with Irving E. Segal, cf. Segal [12]. I did not give any talk at the 1966 Congress.

2.1 Strassen

Volker Strassen did give a talk at the 1966 Congress, on his then-new form of the law of the iterated logarithm. Whether he attended Sudakov's talk I don't recall, but he had been aware of ε-entropy by about 1964. Strassen was born in 1936. Like me, he got his doctorate in mathematics in 1962 and then spent several years in Berkeley, he in the Statistics Department (where probability resided) until 1968, and I in the Mathematics Department until the end of 1966; there was a seminar with probability topics organized by Jacob Feldman, a student of Segal. While we were both in Berkeley, Strassen and I talked about metric entropy. In the late 1960s Strassen began to work on speed of computation, on which he later won several prizes.

Strassen was invited to give a talk at a probability and information theory meeting in Canada which took place in 1968. He declined the invitation but kindly urged the organizers to invite me in his place, as they did; I went and presented the joint paper [16]. The paper gave two results: one, by Strassen, a central limit theorem in $C[0, 1]$ with a metric entropy hypothesis; and a counter-example, by me, showing that for i.i.d. variables X_j with values in $C[0, 1]$ having mean $EX_j = 0$ and being bounded: for some $M < \infty$, $\|X_1(\omega)\| \leq M$ for all ω, the central limit theorem can fail.

3 Early Papers on Metric Entropy and Gaussian Processes

Let (S, d) be a totally bounded metric space and for each $\varepsilon > 0$ let $N(S, d, \varepsilon)$ be the minimum number of points in an ε-net, within ε of each point of S. If d is a Hilbert space (e.g. L^2) metric it may be omitted from the notation. By "metric entropy integral" I mean

$$\int_0^u \sqrt{\log(N(S, d, \varepsilon))} d\varepsilon \tag{3.1}$$

for $u > 0$. The integrand is 0 for ε large enough, as $N(S, d, \varepsilon)$ is nonincreasing in ε and becomes equal to 1. Thus finiteness of (3.1) for some $u > 0$ implies it for all $u > 0$.

Fortunately for me, Irving Segal was one of the founding co-editors of *Journal of Functional Analysis* and solicited my 1967 paper for vol. 1 of the journal. The paper showed that finiteness of (3.1) for $u = 1$ (or an equivalent sum; formulated as an integral in 1973) for a subset S of a Hilbert space is sufficient for sample continuity and boundedness of the isonormal process restricted to S. Dudley [5] showed that if the metric entropy integral is finite for a Gaussian process, its indefinite integral gives a modulus of continuity for the process.

A weaker statement is that it suffices for sample continuity that for some r with $0 < r < 2$, as $\varepsilon \downarrow 0$,

$$\log N(S, d, \varepsilon) = O(\varepsilon^{-r}). \tag{3.2}$$

In my 1967 paper, p. 293, I wrote that "V. Strassen proved (unpublished) in 1963 or 1964" that condition (3.2) implies sample continuity of L on S. Sudakov stated the implication in his 1966 lecture, as I mentioned in Dudley [6, p. 87]. So before 1967, both Sudakov and Strassen had shown the sufficiency of (3.2) although neither had published a statement or proof of it. The abstract Sudakov [15] (in Russian) is quite short; it has two sentences, one about eigen element expansions as in its title, and the second, "For Gaussian distributions, new results are obtained." In Sudakov (1976, pp. 2–3 of the 1979 translation)[20] he reviews previous work, beginning with his 1966 talk.

In MathSciNet (*Mathematical Reviews* online) there is a gap in indexed reviews of Sudakov's publications. There are ten listed as published in the years 1958–1964, none for 1965–1970 (although at least one paper, Sudakov 1969, existed) and 20 for works published in 1971–1980, of which I was reviewer for 7, beginning with Sudakov [17]. Some of my reviews of Sudakov's works were not very perceptive. I had been a reviewer for Math. Reviews since April 1966. (In 1968 and 1971, I had the chance to review an announcement and then a paper by V.N. Vapnik and A.Ya. Chervonenkis.)

Sudakov [16] was as far as I can find his first publication on Gaussian processes. It made the connection with ε-entropy. Sudakov [17, 19] carried the work further. In particular in 1973 he gave an equivalent condition for sample-boundedness of a Gaussian process $\{X_t : t \in T\}$, namely that

$$E \sup_{t \in T} X_t := \sup\{E \sup_{t \in A} X_t : A \subset T, A \text{ countable}\} < +\infty. \tag{3.3}$$

Sudakov [20] gave a book-length presentation of his results on Gaussian processes (and also on doubly stochastic operators). An antecedent of the book is his doctoral dissertation Sudakov [18], which has the same title. In reviewing the book for *Math. Reviews* **MR0431359 (55 #4359)** I said I had learned about metric entropy "from V. Strassen, who wanted to give credit to 'someone' whose name we forgot." And so, I said, Sudakov was "too generous" in saying the application of such ideas to Gaussian processes came "independently" to several authors, although sufficiency of (3.2) for sample continuity seems to have been found independently by Strassen and Sudakov.

4 An Inequality: Majorization of E sup

For a Gaussian process $\{X_t,\ t \in T\}$ with mean 0, such an inequality says that

$$E \sup_{t \in T} X_t \leq K \int_0^{+\infty} \sqrt{\log N(\varepsilon, T, d_X)} d\varepsilon \tag{4.1}$$

for some $K < \infty$, where $d_X(s,t) := (E((X_s - X_t)^2))^{1/2}$. This has been attributed to me and called "Dudley's Theorem" by Ledoux and Talagrand, 1991, Theorem 11.17, p. 321. But in fact I am only responsible for the integral (3.1) over a finite interval and the fact that its finiteness implies sample continuity. In (4.1), $+\infty$ can clearly be replaced by

$$u = \operatorname{diam}(T) := \sup\{d_X(s,t) :\ s, t \in T\}.$$

(By the way, the left side of (4.1) may be finite while the right side is infinite.)

Sudakov [19] first defined the left-hand side (3.3) of (4.1). I was slow to appreciate it. My short review of Sudakov [19] in *Math. Reviews*, **MR0443059**, makes no explicit mention of the expected supremum; still less did I mention it in the earlier paper Dudley [4]. The bound (4.1) with $K = 24$ given by Ledoux and Talagrand had, as they say on p. 329, been proved by Pisier [11].

Ten years earlier, Sudakov (1973, Eq. (6))[19], had stated the inequality

$$E \sup_{x \in S} L(x) \leq CS_1 := C \sum_{k=-\infty}^{\infty} 2^{-k} \sqrt{\log_2(N(2^{-k}, S))} \tag{4.2}$$

for $C = 22/\sqrt{2\pi}$. Sudakov (1976, transl. 1979, Proposition 33)[20], gives a proof, pp. 54–56 of the translation. (If one is not convinced by Sudakov's proof, then the bound (4.1) might be attributed to Pisier, but in no case to me. Also Lifshits (2012, pp. 73–75)[10] gives further evidence that Sudakov's statement (or better) is correct.)

My review in Math. Revs. of Sudakov [20] also did not mention the quantity (3.3) and so neither (4.2) nor its proof.

We have straightforwardly for every integer k

$$2^{-k}\sqrt{\log_2 N(2^{-k},S)} = 2\int_{2^{-k-1}}^{2^{-k}} \sqrt{(\log 2)\log(N(2^{-k},S))}dx$$
$$\leq 2\sqrt{\log 2}\int_{2^{-k-1}}^{2^{-k}} \sqrt{\log(N(x,S))}dx.$$

It follows that $S_1 \leq 2\sqrt{\log 2}\int_0^{+\infty} \sqrt{\log(N(x,S))}dx$. This implies inequality (4.1) with the constant 24 improved to $K := 44\sqrt{\log 2}/\sqrt{2\pi} < 14.62$. As will be seen later, Lifshits [10] gave a still smaller constant. But I suggest that in view of Sudakov's priority, the inequality (4.1) for any (correct) finite K be called "Sudakov's majorization," by contrast with Sudakov's useful lower bound for $E\sup_{x\in S} L(x)$ based on metric entropy, well known as "Sudakov's minoration" (e.g., Ledoux and Talagrand, [8, pp. 79–84]). Chevet [2] gave, in a rather long paper, the first published proof of a crucial lemma in the Sudakov minoration.

According to Google Scholar, Sudakov [19] had only 15 citers as of May 19, 2015, but they did include Ledoux and Talagrand [8], also its chapter on Gaussian processes, and TalagrandŠs 1987 [26] paper on characterizing sample boundedness of Gaussian processes via majorizing measures. Sudakov (1976, transl. 1979)[20] had 228 citers (roughly half of them relating to optimal transportation and other non-Gaussian topics) as of June 12, 2015; it was from the list of citing works that I found Lifshits [10].

5 Books on Gaussian Processes

There are chapters on Gaussian processes in several books. For entire books, although there are some on applications such as machine learning, I will comment only on Lifshits [9, 10] and Bogachev [1].

5.1 Bogachev [1]

This book's Theorem 7.1.2, p. 334, states the Sudakov minoration and what I have called his majorization. For proof, Bogachev refers to Ledoux and Talagrand [8], Ledoux [7], and Lifshits [9]. Sudakov (1976, transl. 1979)[20] is not mentioned there; it is in the bibliography as ref. no. [733], p. 422, but I could not find a citation of it in the book.

5.2 Lifshits [9]

This book cites three works by Sudakov, [Sud1] = Sudakov [16], [Sud2] = Sudakov [17], and [Sud3] = Sudakov [20]. It gives an inequality (4.1), apparently as Theorem 14.1, although I have not seen the exact statement. Lifshits gives credit to Dmitrovskii [3] for the statement and proof.

5.3 Lifshits [10]

On p. 75 Lifshits gives the constant $K = 4\sqrt{2}$ in (4.1), which is the best I have seen. The proof seems to be self-contained. Lemma 10.1 on p. 73 says (correctly) that if $X_1, \ldots, X_N$ are centered jointly Gaussian variables and $E(X_j^2) \le \sigma^2$ for each j, then

$$E \max_{1 \le j \le N} X_j \le \sqrt{2 \log N} \sigma.$$

(Ledoux and Talagrand [8], (3.13), p. 71 have such an inequality with a factor of 3 instead of $\sqrt{2}$.) I was unable in a limited time to check Lifshits's proof of his version of (4.1) via the Lemma and induction.

The bibliography of Lifshits [10] lists 183 items, including Sudakov [16, 17, 20], but no works by Dmitrovskii. Sudakov (1976 and 1979) is his second most-cited work with 234 citations, Google Scholar, Nov. 14 (2015).

References[1]

1. *G V.I. Bogachev, *Gaussian Measures* (American Mathematical Society, Providence, RI, 1998). Translation of Gaussovskie Mery, Fizmatlit, "Nauka", Moscow (1997)
2. S. Chevet, Mesures de Radon sur $\mathbb{R}$ et mesures cylindriques. Ann. Fac. Sci. Univ. Clermont **43** (6), 91–158 (1970)
3. V.A. Dmitrovskii [or Dmitrovsky], On the distribution of the maximum and local properties of realizations of pregaussian fields, in *Teoriya Veroiatnost. i Matematichna Statistika*, vol. 25 (Kiev State University, Kiev, 1981), pp. 154–164. [English translation Theory of Probability and Mathematical Statistics]
4. R.M. Dudley, The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. J. Funct. Anal. **1**, 290–330 (1967)
5. R.M. Dudley, Sample functions of the Gaussian process. Ann. Probab. **1**, 66–103 (1973)
6. R.M. Dudley, *Uniform Central Limit Theorems* (Cambridge University Press, Cambridge, 1999), 2nd edn. 2014

[1] An asterisk * preceding an entry indicates a work I have not seen in the original. Sudakov's dissertations, 1962 and 1972, are cited in the long bibliography of Sudakov (1976, English transl. 1979). *G indicates a book I have seen only as a Google Book

7. M. Ledoux, Isoperimetry and Gaussian analysis, in *Lectures on Probability Theory and Statistics*. Lecture Notes in Mathematics, vol. 1648 (Springer, Berlin, 1996), pp. 165–204
8. M. Ledoux, M. Talagrand, *Probability in Banach Spaces* (Springer, Berlin, 1991). Michel Talagrand wrote on his web page that the book has been reprinted in the "affordable series" Classics in Mathematics
9. M.A. *G Lifshits, *Gaussian Random Functions*. Mathematics and its Applications, vol. 322 (Kluwer [later merged into Springer], Dordrecht, 1995)
10. M.A. *G Lifshits, *Lectures on Gaussian Processes* (Springer, Berlin, 2012)
11. G. Pisier, Some applications of the metric entropy bound in harmonic analysis, in *Banach Spaces, Harmonic Analysis and Probability Theory*. Proceedings University of Connecticut, 1980–1981, ed. by R.C. Blei, S.J. Sidney. Lecture Notes in Mathematics, vol. 995 (Springer, Berlin, 1983), 123–154.
12. I.E. Segal, Abstract probability spaces and a theorem of Kolmogoroff. Am. J. Math. **76**, 721–732 (1954)
13. V. Strassen, R.M. Dudley, The central limit theorem and ε-entropy, in *Probability and Information Theory*. Proceedings of International Symposia McMaster University, Hamilton, 1968 (Springer, Berlin, 1969), pp. 224–231
14. *V.N. Sudakov, Some problems connected with distributions in infinite-dimensional linear spaces Candidate's Dissertation, Leningrad (1962). State University, Leningrad, in Russian
15. V.N. Sudakov, (1968), The connection between a theorem on expansion in generalized eigen elements and a theorem on the extension of a weak distribution to a measure in Hilbert space. Abstracts of Brief Scientific Communications, Section 5, Functional Analysis, International Congress of Mathematicians, Moscow, 1966, p. 75 (in Russian)
16. V.N. Sudakov, Gaussian measures, Cauchy measures and ε-entropy. Soviet Math. Doklady **10**, 310–313 (1969). Russian original Doklady Akad. Nauk SSSR **185**(1), 51–53
17. V.N. Sudakov Gaussian random processes and measures of solid angles in Hilbert spaces. Soviet Math. Doklady **12**, 412–415 (1971). Russian original Doklady Akad. Nauk SSSR **197**(1), 43–45
18. *V.N. Sudakov, Geometrical problems in the theory of infinite-dimensional probability distributions. Doctoral Dissertation, Leningrad State University, Leningrad (1972) [in Russian]
19. V.N. Sudakov, A remark on the criterion of continuity of Gaussian sample functions, in *Proceedings of Second Japan–USSR Symposium on Probability Theory*, Kyoto, 1972, ed. by G. Maruyama, Y.V. Prohorov. Lecture Notes in Mathematics, vol. 330 (Springer, Berlin, 1973), pp. 444–454
20. V.N. Sudakov, Geometric problems in the theory of infinite-dimensional probability distributions (in Russian). Trudy Mat. Inst. Steklov **141**, 191 pp. (1976); English "cover to cover translation," Proc. Steklov Inst. Math. **1979**(2), American Mathematical Society
21. M. Talagrand, Regularity of Gaussian processes. Acta Math. (Sweden) **159**, 99–149 (1987)

http://www.springer.com/978-3-319-40517-9

High Dimensional Probability VII
The Cargèse Volume
Houdré, C.; Mason, D.M.; Reynaud-Bouret, P.; Rosiński, J. (Eds.)
2016, XXVIII, 461 p., Hardcover
ISBN: 978-3-319-40517-9
A product of Birkhäuser Basel