N. H. Bingham, The work of A. N. Kolmogorov on strong limit theorems,
Теория вероятн. и ее примен., 1989, том 34, выпуск 1, 152–164

https://www.mathnet.ru/tvp1240

ТЕОРИЯ ВЕРОЯТНОСТЕЙ
Том XXXIV И ЕЕ ПРИМЕНЕНИЯ Выпуск 1
1989

THE WORK OF A. N. KOLMOGOROV ON STRONG LIMIT THEOREMS

BINGHAM N. H.

1. Historical setting. For probabilists, the modern era begins with Kolmogorov's classic book of 1933, and his work on strong limit theorems culminates in the famous strong law of large numbers and law of the iterated logarithm which bear his name. We begin by assessing the historical context from which these crowning achievements emerged. For an appreciation of Kolmogorov's work in this area on the occasion of his eightieth birthday, see Prokhorov [78].

The prehistory of the law of large numbers, like that of probability itself, is obscure and debatable. Hints of the idea of probability as linked with limiting frequency can be found in the written record as early as the work of Cardano (1501—1576), in his book *De Ludo Aleae* (written in 1526, and published posthumously in 1663, after the Pascal — Fermat correspondence of 1654). Thus «there can be no doubt that he had a fairly good idea of the rule which is now called the law of large numbers» (Ore [72], p. 120; cf. Maistrov [59], p. 19, Sheynin [83]). However, as both games of chance and the basics of combinatorics go back to antiquity, it is surprising that such ideas did not emerge much earlier; for a range of views on such questions see the books of David [12], Hacking [27] and the articles of Sheynin [84] and Garger & Zabell [25].

The first theorem recognisable as a precise form of a limiting-frequency statement (or «law of large numbers» in the terminology introduced later by Poisson) is the famous «weak law of large numbers for Bernoulli trials» of James Bernoulli (1654—1705): if S_n is the number of successes observed in n independent trials with success probability p,

$$S_n/n \to p \text{ in probability } (n \to \infty).$$

This is the most important single result in Bernoulli's classic book *Ars Conjectandi*, published posthumously in 1713 (and honoured in the bicentennial address by A. A. Markov to the St. Petersburg Academy of Sciences in 1913: Appendix 3 in Ondar [72]). One can hardly overstate the importance of Bernoulli's theorem: Kolmogorov, writing in 1986 in the preface to a book on Bernoulli, describes work prior to *Ars Conjectandi* as only the prehistory of probability proper, and its true history as beginning with Bernoulli's theorem.

The next major advance was made by S. D. Poisson (1781—1840), who in his book of 1837 extended Bernoulli's theorem to the case of varying success probabilities p_n. In 1867, P. L. Chebyshev (1821—1894) proved his famous inequality, which yields, in the hands of A. A. Markov (1856—1922), the weak law

$$\frac{1}{n}\sum_{1}^{n}(X_k - \mathbf{E}X_k) \to 0 \text{ in probability}$$

for independent X_n for which

$$\frac{1}{n^2}\sum_{1}^{n}\operatorname{var} X_k \to 0.$$

Hilbert included, as part of Problem 6 of his famous problem list of 1900, an axiomatic treatment of probability. The development of measure theory by Lebesgue around 1901—1904 created for the first time the language in which results on convergence with probability one could be formulated and the machinery for proving them. Emile Borel (1871—1956) is usually credited with the first almost-sure convergence theorem, his well-known normal number theorem of 1909. Note however that Borel's proofs are incomplete (for critical commentary see Fréchet [24], Barone & Novikoff [2]), and that Borel was anticipated by the American mathematician Van Vleck [89] (cf. Novikoff & Barone [71]). Cantelli [8] extended the theorem to general distributions with bounded fourth moments: if X_n are independent with $\mathbf{E}(|X_n - \mathbf{E}x_n|^4) \leqslant C < \infty$,

$$\frac{1}{n}\sum_{1}^{n}(X_k - \mathbf{E}X_k) \to 0 \text{ a. s.}$$

2. The weak law of large numbers. It was by now necessary to distinguish in-probability and almost-sure laws of large numbers; we owe the modern terminology of weak and strong laws to Khinchin [37]. Here Khinchin proves an early strong law for the dependent case.

The first textbook based on measure-theoretic ideas appears to be that of P. Lévy [56]. Here, Lévy systematically uses methods of characteristic functions, and in particular proves the continuity theorem for them that bears his name. Using Lévy's continuity theorem, it is nowadays a simple matter to prove the two «distributional» limit theorems of probability theory — the weak law of large numbers and the central limit theorem.

Khinchin [38] gives two proofs of the weak law in the independent and identically distributed (iid) case: if $X, X_1, X_2, \ldots$ are independent copies with mean μ,

$$\frac{1}{n}\sum_{1}^{n} X_k \to \mu \quad (n \to \infty) \text{ in probability.}$$

His first proof uses truncation of X_n at level n, his second that, writing φ for the characteristic function of X, $\frac{1}{n}\sum_{1}^{n} X_k$ has characteristic function

$$\mathbf{E}\exp\left\{it\sum_{1}^{n} X_k/n\right\} = \varphi(t/n)^n = \left(1 + \frac{i\mu t}{n} + o(1/n)\right)^n \to e^{i\mu t}$$

and the continuity theorem.

The (independent) non-identically distributed case of the weak law was considered by Kolmogorov [44] in a pioneering paper to which we return below, and in its sequel Kolmogorov [45]. In particular (Satz XII) Kolmogorov refined Khinchin's result above on the iid case by showing that there exist constants c_n with

$$\frac{1}{n}\sum_{1}^{n} X_k - c_n \to 0 \text{ in probability}$$

((X_n) are 'stable') if and only if

$$n\mathbf{P}(|X| > n) \to 0.$$

The constants c_n can then be taken as

$$c_n = \int_{-n}^{n} x\, dF(x),$$

writing F for the law of X (Feller [17]). For necessary and sufficient conditions in the general case, see e. g. the book of Lévy [57], Marcinkiewicz [63], or the survey Feller [19, § 9]. We note in passing that this condition is strictly weaker than existence of the mean. For,

$$\mathbf{E}(|X|) < \infty \Leftrightarrow \sum \mathbf{P}(|X| > n) < \infty,$$

while by Olivier's theorem (Knopp [42]), for $a_n \downarrow 0$

$$\sum a_n < \infty \Rightarrow n\, a_n \to 0,$$

but not conversely.

Although existence of the mean is a sufficient condition, the necessary and sufficient condition for the weak law is differentiability of the characteristic function $\varphi(\cdot)$ at the origin. The following are equivalent:

(i) There exists $\varphi'(0) = i\mu$,

(ii) $x\mathbf{P}(|X| > x) \to 0$ and $\int_{-x}^{x} y\, dF(y) \to \mu \quad (x \to \infty)$,

(iii) $\frac{1}{n}\sum_{1}^{n} X_k \to \mu \quad (n \to \infty)$ in probability

(Feller [23], XVII. 2a); further, $x\mathbf{P}(|X| > x) \to 0$ implies convergence of the truncated mean $\int_{-x}^{x} y\, dF(y)$ (Feller [23, VII.7]).

Pursuing the iid case further, one may ask whether norming functions a_n more general than the n above may yield convergence in probability to a constant limit. Simplest is the case of X_n non-negative; to avoid the degeneracy of an excessively large norming function one takes the limit positive, and by scaling one asks for limit one. In modern terminology, one says that (X_n) is relatively stable if constants a_n exist with

$$X_n/a_n \to 1 \text{ in probability.}$$

The classical instance of this is the St. Petersburg game, going back to the work of Daniel Bernoulli in the eighteenth century. For modern work, see Khinchin [34], Feller [20], [22, X. 3], [23, VII.7], Rogozin [81], Maller [61], [62]. The necessary and sufficient condition for relative stability, due to Rogozin and Maller, is

$$x\mathbf{P}(|X| > x) \Big/ \int_{-x}^{x} y\, dF(y) \to 0 \quad (x \to \infty),$$

generalising that for the weak law above. This may be compared with the (stronger) necessary and sufficient condition

$$x^2\mathbf{P}(|X|>x)\Big/\int_{-x}^{x} y^2\,dF(y)\to 0$$

for the iid case of the central limit theorem (Feller [23, IX.8]).

We note that independence is not needed for the weak law: pairwise independence suffices (Chung, [11, Th. 5.2.2]).

3. Random series. If $X_1, X_2, \ldots$ are independent random variables, when does the series $\sum_1^\infty X_n$ converge (almost surely, or in some other mode)? In Khinchin & Kolmogorov [41] — Kolmogorov's first work in probability theory — this question is answered in full for a. s. convergence, the criterion being given by the famous three-series theorem.

Here, and in Kolmogorov [44], Satz I, one finds the powerful generalisation of Chebyshev's inequality:

$$\mathbf{P}\left(\max_{j\leqslant n}\sum_1^j |X_k-\mathbf{E}X_k|>\varepsilon\right)\leqslant\frac{1}{\varepsilon^2}\sum_1^n \operatorname{var} X_k$$

(«Kolmogorov's inequality», or «Kolmogorov's first inequality»). This may be applied immediately to convergence of series (Sats VII): if $\sum_1^\infty \operatorname{var} X_n<\infty$,

$$\sum_1^\infty (X_n-\mathbf{E}X_n)\text{ converges a. s.}$$

In particular,

$$\sum_1^\infty \operatorname{var} X_n<\infty \text{ and } \sum_1^\infty \mathbf{E}X_n \text{ converges}\Rightarrow\sum_1^\infty X_n \text{ converges a. s.}$$

This is the basis for most subsequent proofs of the three-series theorem; see e. g. Doob [14, III.2], Loève [58, § 16].

Closely related is the «equivalence theorem» (Lévy [57]; Loève [58, § 17]): for independent random variables, convergence of $\sum_1^\infty X_n$ almost surely, in probability and in distribution are equivalent.

More generally, one may consider independent random vectors taking values in a real separable Banach space (the «Banach-valued case»): see Itô & Nisio [32], Maurey & Pisier [66].

4. The strong law of large numbers. In Kolmogorov [47], we find the following strong law: if the X_n are independent with mean zero and variance σ_n^2, and

$$\sum_1^\infty \sigma_n^2/n^2<\infty,$$

then

$$\frac{1}{n}\sum_1^n X_k\to 0\ (n\to\infty)\text{ a. s. (and in mean square).}$$

The original proof is obtained from Kolmogorov's (first) inequality and «filling in» from the subsequence (2^m) to the full sequence.

The converse is also given: if $\sum \sigma_n^2/n^2$ diverges, then there exist independent random variables X_n, with mean 0 and variance σ_n^2, for which the above convergence fails.

Kolmogorov returns to the strong law in the last section of his famous book, Kolmogorov [48]. Here he states (without proof) the following result — «Kolmogorov's strong law of large numbers»: if $X, X_1, \ldots$ are iid,

$$\text{(i)} \quad \mathbf{E}|X| < \infty \ \& \ \mathbf{E}X = \mu \Rightarrow \frac{1}{n}\sum_1^n X_k \to \mu \text{ a.s.},$$

$$\text{(ii)} \quad \mathbf{E}|X| = \infty \Rightarrow \frac{1}{n}\left|\sum_1^n X_k\right| \to \infty \text{ a.s.}$$

Equivalently, one may combine both statements:

$$\mathbf{E}|X| < \infty \ \& \ \mathbf{E}X = \mu \Leftrightarrow \frac{1}{n}\sum_1^n X_k \to \mu \text{ a.s.}$$

Full proofs (in greater generality) are to be found in the classic paper of Marcinkiewicz & Zygmund [64, Théorèmes 9, 10]. For the direct half, one uses truncation and Kronecker's lemma to reduce to the strong law in Kolmogorov [47]. For the converse half, writing S_n for $\sum_1^n X_k$, one uses

$$\frac{X_n}{n} = \frac{S_n}{n} - \frac{(n-1)}{n}\,\frac{S_{n-1}}{n-1} \to 0 \text{ a.s.}$$

to get $\mathbf{P}(|X_n| > n \text{ i. o.}) = 0$, whence by the Borel — Cantelli lemma

$$\sum \mathbf{P}(|X_n| > n) = \sum \mathbf{P}(|X| > n) < \infty,$$

which says $\mathbf{E}|X| < \infty$. See also Fréchet [24], Lévy [57].

Kolmogorov's strong law of large numbers is certainly one of the most important results — and arguably **the** most important result — in the entire subject of probability theory. The theory outlined above has been refined, streamlined and incorporated into all the classic textbooks of the subject, as exemplified for instance by Doob [14, III, §§ 2, 3, 5], Loève [58, § 16], Chung [11, Ch. 5].

So far as methods are concerned, the classical approach is to follow Kolmogorov, using his two inequalities and giving a combined treatment of the strong law and the convergence of random series: Loève's book contains a particularly efficient treatment of this kind. It became clear that Kolmogorov's second (and harder) inequality could be dispensed with (cf. the preface to Chung [11]). It was long believed that Kolmogorov's first inequality was indispensable, but Etemadi [16] obtained a short proof without it — and incidentally weakened independence to pairwise independence, complementing Chung's result for the weak law. Etemadi's proof has already entered the textbook literature; see Grimmett & Stirzaker [26].

5. Generalisations. Kolmogorov's strong law may be generalised well eyond the classical setting of sums of independent random variables; see for instance the monograph of Révèsz [80]. The first such generalisation (which actually preceded the completion of Kolmogorov's result) was Birkhoff's pointwise ergodic theorem (Birkhoff [6]). This was extended to full generality by Khinchin [39]. The pointwise ergodic theorem is thus known also as the Birkhoff — Khinchin theorem, particularly in the Russian literature, its proof was simplified in Kolmogorov [49]. For a full account of such matters, see Krengel [52], Stout [85, § 3.5].

Although the strong law may be obtained as a special case of the ergodic theorem, the results are of different natures, as may be seen in their converses. As noted in § 4, the strong law has a full converse: existence of the mean is necessary as well as sufficient. By contrast, the converse of the ergodic theorem fails: one may have a.s. convergence of the «ergodic averages»

$$\frac{1}{n}\sum_{1}^{n} X(T^k\omega)$$

with T measure-preserving and ergodic, even though

$$\mathbf{E}\,|X| = \infty;$$

for an example, see e.g. Halmos [28, p. 32].

The other main setting in which the strong law arises as a special case is that of the martingale convergence theorem. For this result see Doob [14, VII.4], and for a short deduction of the strong law, VII.6.

Doob raises here the possibility of a common treatment of the pointwise ergodic and martingale convergence theorems. Such an approach has been attempted by several authors, including Rota, Neveu, and A. & C. Ionescu Tulcea; for a survey, see Rao [79], Krengel [52, § 9.4.3].

We have already pointed out that Kolmogorov's strong law is the special case $p = 1$ of the strong law for L_p of Marcinkiewicz & Zygmund [64]: for $X, X_1, \ldots$ iid and $0 < p < 2$, there exists a constant c with

$$\frac{1}{n^{1/p}}\left(\sum_{1}^{n} X_k - nc\right) \to 0 \text{ a. s.}$$

if and only if

$$\mathbf{E}\,(|X|^p) < \infty,$$

and then c may be taken as $\mathbf{E}X$ if $1 \leqslant p < 2$, 0 if $0 < p < 1$. (Observe that the restriction $p < 2$ is forced here, because of the central limit theorem.) This strong law is closely connected with the Marcinkiewicz — Zygmund inequalities proved in the same paper (see e. g. Chow & Teicher [10], §§ 5.2, 10.3 for a modern treatment).

The interconnections between these results are perhaps best seen by considering their extensions to more general settings. Kolmogorov's strong law extends to the Banach — valued case for arbitrary Banach spaces, as was shown by Mourier [69]. More complicated results may generalise to some but not all Banach spaces, thus singling out some geometrical property of Banach spaces as characterising their domain of validity. An early instance is the result of Beck, showing that a type of convexity condition is necessary and sufficient for the validity of a form of the strong law. The martingale convergence theorem holds if and only if the Banach space has the Radon — Nikodým property (Chatterji (1968); cf. Diestel & Uhl [13]). The Marcinkiewicz — Zygmund inequalities for L_p suggest a notion of spaces of type p; the Marcinkiewicz — Zygmund strong law for L_p holds on B if and only if B has type p (de Acosta [1]).

Necessary and sufficient conditions for the strong law of large numbers n the independent, non-identically distributed, real-valued case are known: see Prokhorov [77], Nagaev [70], Stout [86, §§ 3.4, 5.2]. In the Banach-valued case matters are more complicated; see Kuelbs & Zinn [54]. In particular, in the zero-mean case, the «Kolmogorov condition».

$$\sum \mathbf{E}\,(\| X_n \|^2)/n^2 < \infty$$

need not imply the strong law, but does reduce the strong law to the weak law. For recent work on Banach spaces with smooth norm, see Heinkel [30].

Rate-of-convergence estimates may be obtained for the strong law (see e.g. Petrov [75, IX.4, 5], Bingham [4]); this subject is closely linked with the law of the iterated logarithm, considered below. For extensions to the dependent case, see e. g. Peligrad [74].

6. The law of the iterated logarithm. Following the Borel normal number theorem of 1909, much interest attached to the frequency behaviour of the number of 1s (say) in the first n places of the dyadic expansion of a «typical» real number in [0, 1]. The problem was considered by Hausdorff, Hardy & Littlewood, and finally by Khinchin [34]. He showed that, if X_n is the indicator of the event that the nth place is a 1 (by properties of dyadic expansions, the X_n are independent, Bernoulli distributed with parameter $^1/_2$), one has

$$\limsup_{n\to\infty} \frac{\sum_1^n X_k - n/2}{{}^1/_2\sqrt{2n \log\log n}} = +1 \text{ a. s.}$$

and the a.s. liminf is -1. This «law of the iterated logarithm» (LIL) completes, with the law of large numbers (LLN) and central limit theorem (CLT) the trilogy of classical limit theorems. More accurately, one may distinguish the two strong (measure-theoretic) limit theorems, the strong law and the LIL, from the two weak (distributional) limit theorems, the weak law and the CLT. The LIL was extended to the case of not necessarily identically distributed Bernoulli trials by Khinchin [36].

In his second major contribution to the theory of strong limit theorems, Kolmogorov [46] extended Khinchin's LIL to the case of general independent random variables X_n. Write S_n for $\sum_1^n X_k$, B_n for the variance of S_n, and assume $B_n \to \infty$. Kolmogorov assumed the order restriction

$$X_n = o\,((B_n/\log\log B_n)^{1/2}) \text{ a. s.,}$$

and showed that

$$\limsup S_n/(2B_n \log\log B_n)^{1/2} = 1 \text{ a. s.}$$

His method of proof involves the Kolmogorov exponential bounds, followed by separate treatment of the «upper» and «lower» parts:

(a) For arbitrary $\varepsilon, \delta > 0$ there is an n such that for all $p \geqslant 0$,

$$\mathbf{P}\Big(\bigcup_{k=n}^{n+p} \{S_k > (1+\delta)\,(2B_k \log\log B_k)^{1/2}\}\Big) < \varepsilon,$$

(b) For arbitrary $\varepsilon, \delta > 0$ and m there is a p such that

$$\mathbf{P}\Big(\bigcap_{k=m}^{m+p} \{S_k < (1-\delta)\,(2B_k \log\log B_k)^{1/2}\}\Big) < \varepsilon.$$

For textbook accounts, see e.g. Loève [58], § 18, Stout [86], § 5.2.

It is interesting to note that Kolmogorov's order restriction is essentially best-possible: if it is weakened to

$$| X_n | = O\left((B_n/\log\log B_n)^{1/2}\right) \text{ a. s.}$$

the result is false (Marcinkiewicz & Zygmund [65]).

Kolmogorov's LIL led to the definitive result on the iid case, the LIL of Hartman & Wintner [29]: if $X_1, X_2, \ldots$ are iid with mean μ and variance σ^2,

$$\limsup \frac{S_n - n\mu}{\sigma\sqrt{2n \operatorname{loglog} n}} = 1 \text{ a. s.}$$

and the liminf is -1 (in fact, the a.s. limit point set is $[-1, 1]$ as follows easily).

The most important subsequent result on the LIL has been that of Strassen [87], who obtained his well-known functional version. This gives, in particular (Th. 4) information on how thickly the a.s. limit points are distributed over subintervals of $[-1, 1]$.

Kolmogorov's LIL has inspired many subsequent developments. Among these are theorems of Stout [85] on the martingale case, Tomkins [88], which allows X_n unbounded, Moricz [68] on rates of convergence, Major [60], extending Strassen's theorem to the non-identically distributed case, and Kuelbs [53] on the Banach-valued case. Th «LIL problem» for Banach spaces has been completely solved, very recently, by Ledoux & Talagrand [55].

Returning to the two halves (a) and (b) above of Kolmogorov's LIL, one says (following Khinchin [40]) that

$$\Phi^{\varepsilon}(n) := (1 + \varepsilon)(2B_n \log\log B_n)^{1/2}$$

is an upper function for S_n:

$$\mathbf{P}(S_n > \Phi^{\varepsilon}(n) \text{ i. o.}) = 0,$$

while

$$\Phi_{\varepsilon}(n) := (1 - \varepsilon)(2B_n \log\log B_n)^{1/2}$$

is a lower function:

$$\mathbf{P}(S_n > \Phi_{\varepsilon}(n) \text{ i. o}) = 1.$$

Following unpublished work by Kolmogorov in the early 1930s, Petrowsky [76] showed that for Brownian motion, a non-decreasing function $\Phi(\cdot)$ is upper [lower] (definitions as above) if

$$\int^{\infty} \Phi(t)\, t^{-1} \exp\left(-\frac{1}{2}\Phi^2(t)\right) dt < \infty \; [= \infty]$$

(cf. Itô & McKean [31, p. 161—164]). There are similar integral tests for random walks by Erdös [15], Feller [18], [21] and by Jain, Jogdeo & Stout [33, § 5] for martingales.

For a full account of the LIL up to 1974 see Chapter 5 of Stout [86], and for a survey of stubsequent developments see Bingham [5].

7. Randomness and computational complexity. In classical probability theory, of which Kolmogorov is the principal creator, one specifies the mechanism generating the randomness (to fix ideas, say an infinite sequence of independent tosses of a fair coin) and — in the strong limit theorems to which this article is devoted — deduces properties which outcomes will possess with probability one. Conversely, statistics allows one to test an observed outcome for consistency with such a meshanism.

Such an approach was rejected by R. von Mises (1883 — 1953), who attempted to «define probability», or «define a random sequence»; see e. g. von Mises [67]. Furthermore, the development of «random number generators» for use on modern computers has provided a powerful stimulus to approaches to randomness other than the classical one; see e. g. Knuth [43, Ch. 3].

Beginning in the early 1960s, Kolmogorov developed an approach to randomness based on ideas of information theory and the theory of algorithms; see the volume 'IA' (Theory of Information and the Theory of Algorithms (in Russian), Nauka, Moscow, 1987, particularly papers 6, 9 and 13) and the survey by Kolmogorov & Uspensky [5] based on an address given at Tashkent in 1986. Here it is proposed to call random those sequences in the class of «chaotic» sequences — loosely, those of maximal complexity; by the Levin — Schnorr theorem of 1973, this coincides with the class of «typical» sequences, considered in earlier work of Martin-Löf. (For background, see the survey of Zvonkin & Levin [91] and the books of Schnorr [82], Chaitin [9, § 7.2].) The class R of random sequences so defined is a subset of the class KS of «stochastic sequences in Kolmogorov's sense». These are the sequences possessing a «limiting frequency», and such that all suitably defined subsequences possess the same limiting frequency, thereby encapsulating (in precise form) the idea of a «collective», which is the crux of the von Mises approach. Incidentally, it is unknown whether or not the inclusion $\mathbf{R} \subset KS$ is proper. The von Mises programme can be carried out in full if (and only if) equality holds; in any case, the theory above may be regarded as essentially bringing von Mises' approach to fruition.

The power of this theory may be seen in the paper of Vovk [90] immediately following Kolmogorov & Uspensky [51]. Here a random, or chaotic, sequence is shown to obey the strong law of large numbers and the law of the iterated logarithm (of course, without the qualification «a.s.», characteristic of the classical theory but absent in the present «pointwise» approach). Furthermore, sharp results are obtained showing how far from randomness a sequence may deviate and still satisfy LLN and LIL (in terms of a concept of «$f(n)$-chaotic sequences», measuring how far below the maximum the algorithmic complexity of the initial segment $(x_1, \ldots, x_n)$ may fall). These results provide definitive analogues of the strong limit theorems of classical probability theory.

REFERENCES

1. *Acosta A. de* Inequalities for B-valued random vectors with applications to the strong law of large numbers.— Ann. Probab., 1981, v. 9, p. 157—161.
2. *Barone J., Novikoff A.* A history of the axiomatic formulation of probability from Borel to Kolmogorov. Part I.— Arch. Hist. Exact. Sci., 1978, v. 18, p. 123—190.
3. *Beck A.* On the law of large numbers.— In: Ergodic Theory. / Ed. by F. B. Wright. N.-Y.: Academic Press, p. 21—53.
4. *Bingham N. H.* Extensions of the strong law.— Adv. Appl. Probab. Supplement, 1986, p. 27—36.

5. *Bingham N. H.* Variants on the law of the iterated logarithm.— Bull. London Math. Soc., 1986, v. 18, p. 433—467.
6. *Birkhoff G. D.* Proof of the ergodic theorem.— Proc. Nat. Acad. Sci. U. S. A., 1931, v. 17, p. 656—660.
7. *Borel E.* Sur les probabilités dénombrables et leurs applications arithmétiques.— Rend. Circ. Mat. Palermo, 1909, t. 26, p. 247—261.
8. *Cantelli F. R.* Sulla probabilitá come limite della frequenza.— Rend. Accad. Lincei, 1917, v. 26, p. 39—45.
9. *Chaitin G. J.* Algorithmic information theory. Cambridge: Cambridge Univ. Press, 1987.
10. *Chow Y.-S.*, *Teicher H.* Probability Theory. Independence, Interchangeability, Martingales. Berlin — Heidelberg etc.: Springer, 1978.
11. *Chung K.-L.* A Course in Probability Theory. 2nd ed. N.-Y. etc.: Academic Press, 1974(1968).
12. *David F. N.* Games, Gods and Gambling. The Origins and History of Probability and Statistical Ideas from the Earliest Times to the Newtonian Era. N.-Y.: Hafner, 1962.
13. *Diestel J.*, *Uhl J. J.* Vector measures. Providence: Amer. Math. Soc., 1977.
14. *Doob J. L.* Stochastic Processes. N.— Y. etc.: Wiley, 1953.— Русск. перев.— *Дуб Дж. Л.* Вероятностные процессы. М.: ИЛ, 1956, 605 с.
15. *Erdös P.* On the law of the iterated logarithm.— Ann. Math., 1942, v. 43, p. 419—436.
16. *Etemadi N.* An elementary proof of the strong law of large numbers.— Z. Wahrscheinlichkeitstheor. verw. Geb., 1981, B. 55, S. 119—122.
17. *Feller W.* Über das Gesetz der grossen Zahlen.— Acta Litt. Sci. Szeged, 1937, B. 8, S. 191—201.
18. *Feller W.* The general form of the so-called law of the iterated logarithm.— Trans. Amer. Math. Soc., 1943, v. 54, p. 373—402.
19. *Feller W.* The fundamental limit theorems of probability theory.— Bull. Amer. Math. Soc., 1945, v. 51, p. 800—832.
20. *Feller W.* Note on the law of large numbers and «fair» games.— Ann. Math. Statist., 1945, v. 16, p. 301—304.
21. *Feller W.* The law of the iterated logarithm for identically distributed random variables.— Ann. Math., 1946, v. 47, p. 631—638.
22. *Feller W.* An Introduction to Probability Theory and Its Applications. V. I. 3rd ed., rev. N.— Y. etc.: Wiley, 1968.
23. *Feller W.* An Introduction to Probability Theory and Its Applications. V. II. 2nd ed., rev. N.— Y. etc.: Wiley, 1971.— Русск. перев.— *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 2. М.: Мир, 1984, 752 с.
24. *Fréchet M.* Generalités sur les probabilités. Variables aléatoires. Paris: Gauthier — Villars, 1937.
25. *Garber D.*, *Zabell S.* On the emergence of probability.— Arch. Hist. Exact Sci., 1979, v. 21, p. 33—53.
26. *Grimmet G. R.*, *Stirzaker D.* Probability and Random Processes. Oxford: Oxford Univ. Press, 1982.
27. *Hacking I.* The Emergence of Probability. A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference. Cambridge: Cambridge Univ. Press, 1975.
28. *Halmos P. R.* Lectures on Ergodic Theory. N.— Y.: Chelsea, 1960.
29. *Hartman P.*, *Wintner A.* On the law of the iterated logarithm.— Amer. J. Math., 1941, v. 63, p. 169—176.
30. *Heinkel B.* Une extension de la loi des grandes nombres de Prohorov.— Z. Wahrscheinlichkeitstheor. verw. Geb., 1984, B. 67, S. 349—362.
31. *Itô K.*, *McKean H. P.* Diffusion Processes and Their Sample Paths. Berlin — Heidelberg etc.: Springer, 1965.— Русск. перев.— *Ито К.*, *Маккин Г.* Диффузионные процессы и их траектории. М.: Мир, 1968, 394 с.
32. *Itô K.*, *Nisio M.* On the convergence of sums of independent Banach space valued random variables.— Osaka J. Math., 1968, v. 5, p. 35—48.

33. *Jain N. C., Jogdeo K., Stout W. F.* Upper and lower functions for martingales and mixing processes.— Ann. Probab., 1975, v. 3, p. 119—145.
34. *Khinchin A. Ya.* Über einen Satz der Wahrscheinlichkeitsrechnung.— Fundam. Math., 1924, B. 6, S. 9—20.
35. *Хинчин А. Я.* О Петербургской игре.— Матем. сб., 1925, т. 32, с. 330—341.
36. *Khinchin A. Ya.* Über das Gesetz der grossen Zahlen.— Math. Ann., 1926, B. 96, S. 152—158.
37. *Khinchin A. Ya.* Sur la loi forte des grandes nombres.— C. r. Acad. Sci. Paris, 1928, t. 186, p. 285—287.
38. *Khinchin A. Ya.* Sur la loi des grandes nombres.— C. r. Acad. Sci. Paris, 1929, t. 188, p. 477—479.
39. *Khinchin A. Ya.* Zu Birkhoffs Lösung des Ergodenproblems.— Math. Ann., 1932, B. 107, S. 485—488.
40. *Khinchin A. Ya.* Asymptotische Gesetze der Wahrscheinlichkeitsrechnung.— Ergebnisse Math., 1933, B. 2, H. 4.
41. *Khinchin A. Ya., Kolmogorov A. N.* Über Konvergenz von Reihen, deren Glieder durch den Zufall bestimmt werden.— Матем. сб., 1925, т. 32, № 4, с. 668—677.
42. *Knopp K.* Infinite Sequences and Series. N.— Y.: Dover, 1956.
43. *Knuth D. E.* The Art of Computer Programming. V. 2: Semi-Numerical Algorithms. 2nd ed. N.— Y.: Addison—Wesley, 1981 (1969).
44. *Kolmogorov A. N.* Über die Summen durch den Zufall bestimmter unabhängiger Grossen.— Math. Ann., 1928, B. 99, S. 309—319.
45. *Kolmogorov A. N.* Bemerkungen zu meiner Arbeit «Über die Summen durch den Zufall bestimmter unabhängiger Grossen».— Math. Ann., 1929, B. 102, S. 484—488.
46. *Kolmogorov A. N.* Über das Gesetz des iterierten Logarithmus.— Math. Ann., 1929, B. 101, S. 126—135.
47. *Kolmogorov A. N.* Sur la loi forte des grandes nombres.— C. r. Acad. Sci. Paris, 1930, t. 191, p. 910—912.
48. *Kolmogorov A. N.* Grundbegriffe der Wahrscheinlichkeitsrechnung. Berlin — Heidelberg: Springer, 1933.— Русск. изд.— Колмогоров А. Н. Основные понятия теории вероятностей. М.: ОНТИ, 1936, 120 с.— 2-е изд.— М.: Наука, 1974, 120 с. *Kolmogorov A. N.* Foundations of the Theory of Probability. N. — Y.: Chelsea, 1950.
49. *Kolmogorov A. N.* Ein vereinfachter Beweis des Birkhoff — Khintchine'schen Ergodensatzes.— Матем. сб., 1937, т. 2, с. 367—368.
50. *Колмогоров А. Н.* Теория информации и теория алгоритмов. М.: Наука, 1987, 304 с.
51. *Колмогоров А. Н., Успенский В. А.* Алгоритмы и случайность.— Теория вероятн. и ее примен., 1987, т. XXXII, в. 3, с. 425—455.
52. *Krengel U.* Ergodic Theorems. Berlin — N.-Y.: Walter de Gruyter, 1985.
53. *Kuelbs J.* Kolmogorov's law of iterated logarithm for Banach space valued random variables.— Illinois J. Math., v. 21, p. 784—800.
54. *Kuelbs J., Zinn J.* Some stability results for vector valued random variables.— Ann. Probab., 1979, v. 7, p. 75—84.
55. *Ledoux M., Talagrand M.* Characterization of the law of the iterated logarithm in Banach spaces.— Ann. Probab., 1988, v. 16, p. 1242—1264.
56. *Lévy P.* Calcul des probabilités. Paris: Gauthier — Villars, 1925.
57. *Lévy P.* Théorie de l'addition des variables aléatoires. Paris: Gauthier — Villars, 1937.— 2nd. ed. 1954.
58. *Loève M.* Probability Theory. 3rd ed. Princeton: Van Nostrand, 1963.— Русск. перев.— *Лоэв М.* Теория вероятностей. М.: ИЛ, 1962, 719 с.
59. *Maistrov L. E.* Probability Theory: An Historical Sketch. N.— Y.: Academic Press, 1973, 281 p.— Русск. изд.— *Майстров Л. Е.* Развитие понятия вероятности. М.: Наука, 1980, 269 с.
60. *Major P.* A note on Kolmogorov's law of the iterated logarithm.— Studia Sci. Math. Hungarica, 1977, v. XII, p. 161—167.
61. *Maller R. A.* Relative stability and the strong law of large numbers.— Z. Wahrscheinlichkeitstheor. verw. Geb., 1978, B. 43, S. 141—148.

62. *Maller R. A.* Relative stability, characteristic functions and stochastic compactness.— J. Austral. Math. Soc., Ser. A, 1978, v. 28, p. 499—509.
63. *Marcinkiewicz J.* Sur les fonctions indépendantes.— I—III. Fundam. Math., 1938, v. 30, p. 202—214, 349—364; v. 31, p. 86—102.— Reprinted in: Collected Papers. Warsaw: PWN, 1964.
64. *Marcinkiewicz J., Zygmund A.* Sur les fonctions indépendantes.— Fundam. Math., 1937, v. 29, p. 60—90.
65. *Marcinkiewicz J., Zygmund A.* Remarque sur la loi du logarithm itéré.— Fundam. Math., 1937, v. 29, p. 215—222.
66. *Maurey B., Pisier G.* Séries des variables aléatoires vectorielles indépendantes et propriétés géometriques des espaces de Banach.— Studia Math., 1976, v. 58, p. 49—90.
67. *von Mises R.* Wahrscheinlichkeit, Statistik und Wahrheit. Berlin: Springer, 1928.— 2nd ed. 1951.— Engl. transl.: Probability, Statistics and Truth. 2nd ed. N.— Y.: Dover, 1981.— Русск. перев.— *Мизес Р.* Вероятность и статистика.— М.— Л.: ГИЗ, 1930.
68. *Moricz F.* Probability inequalities of exponential type and laws of the iterated logarithm.— Acta Sci. Math. Szeged, 1976, v. 38, p. 325—341.
69. *Mourier E.* Éléments aléatoires dans un espace de Banach,— Ann. Inst. H. Poincaré, Ser. B, t. 13, 1953, p. 159—244.
70. *Нагаев С. В.* О необходимых и достаточных условиях для усиленного закона больших чисел.— Теория вероятн. и ее примен., 1972, т. XVII, в. 4, с. 609—618.
71. *Novikoff A., Barone J.* The Borel law of normal numbers, the zero — one law, and the work of Van Vleck.— Historia Math., 1977, v. 4, p. 43—65.
72. The Correspondence Between A. A. Markov and A. A. Chuprov. On the Theory of Probability and Mathematical Statistics. / Ed. by K. O. Ondar. An Introduction by J. Neyman. Transl. from the Russian by C. Stein, M. Stein. Berlin — Heidelberg etc.: Springer, 181 S.— Русск. изд.— Переписка А. А. Маркова и А. А. Чупрова. О теории вероятностей и математической статистике. М.: Наука, 1978.
73. *Ore O.* Cardano, the gambling scholar. Princeton: Princeton Univ. Press, 1953.
74. *Peligrad M.* Convergence rates of the strong law for stationary mixing sequences.— Z. Wahrscheinlichkeitstheor. verw. Geb., 1985, B. 70, S. 307—314.
75. *Петров В. В.* Суммы независимых случайных величин. М.: Наука, 1972, 416 с.
76. *Petrowsky I. G.* Zur ersten Randwertaufgabe der Warmleitungsgleichung.— Comp. Math., 1935, B. 1, S. 383—419.
77. *Прохоров Ю. В.* Несколько замечаний к усиленному закону больших чисел.— Теория вероятн. и ее примен., 1959, т. IV, в. 2, с. 215—220.
78. *Прохоров Ю. В.* Закон больших чисел и закон повторного логарифма.— Успехи матем. наук, 1983, т. 38, № 4. с. 281—287.
79. *Rao M. M.* Abstract martingales and ergodic theory.— In: Multivariate Analysis III. / Ed. by P. R. Krishnaiah. N.— Y.: Academic Press, 1973, p. 45—60.
80. *Révész P.* The Laws of Large Numbers. N.— Y.: Academic Press, 1968.
81. *Рогозин Б. А.* Относительно устойчивые блуждания.— Теория вероятн. и ее прнмен., 1976, т. XXI, в. 2, с. 383—387.
82. *Schnorr C.-P.* Zufälligkeit und Wahrscheinlichkeit. Eine algoritmische Begrundung der Wahrscheinlichkeitstheorie.— Lect. Notes Math., 1971, B. 218.
83. *Sheynin O. B.* On the early history of the law of large numbers.— Biometrika, 1968, v. 55, p. 459—467. (Paper 16 in: Studies in the History of Statistics and Probability.—/Ed. by E. S. Pearson, M. G. Kendall. V. I. London: Griffin, 1970.)
84. *Sheynin O. B.* On the prehistory of the theory of probability.— Arch. Hist. Exact Sciences, 1975, v. 12, p. 97—141.
85. *Stout W. F.* A martingale analogue of Kolmogorov's law of the iterated logarithm.— Z. Wahrscheinlichkeitstheor. verw. Geb., 1970, B. 15, S. 279—290.
86. *Stout W. F.* Almost sure convergence. N.— Y.: Academic Press, 1974, 381 p.
87. *Strassen V.* An invariance principle for the law of the iterated logarithm.— Z. Wahrscheinlichkeitstheor. verw. Geb., 1964, B. 3, S. 211—226.
88. *Tomkins R. J.* A generalisation of Kolmogorov's law of the iterated logarithm.— Proc. Amer. Math. Soc., 1972, v. 32, p. 268—274.

89. *van Vleck E.* On non-measurable sets of points, with examples.— Trans Amer. Math. Soc., 1908, v. 9, p. 237—244.
90. *Вовк В. Г.* Закон повторного логарифма для случайных колмогоровских или хаотических последовательностей.— Теория вероятн. и ее примен., 1987, т. XXXII, в. 2, с. 413—425.
91. *Звонкин А. К., Левин Л. А.* Сложность конечных объектов и развитие понятия информации и случайности при помощи теории алгоритмов.— Матем. сб., 1970, т. 25, № 6, с. 83—124.

Поступила в редакцию
12.X.1988