APPRECIATION OF KHINCHIN

J. L. DOOB
UNIVERSITY OF ILLINOIS

Probability students can hardly visualize the state of probability theory in the nineteen twenties. There were numerous texts on the subject, almost all of which could as well have been written a century earlier, since they were largely collections of elementary combinatorial problems without a unified point of view, together with perhaps some unorganized material on "continuous probabilities." Their authors did not treat probability theory as a mathematical subject, but as the analysis of certain more or less practical problems. One of the better books contained a "proof" that no such theorem as what is now called the strong law of large numbers could be valid for independent random variables with a common distribution! Reputable statisticians were not sure of the relation between independence and orthogonality of random variables, in particular whether orthogonality implies independence. The place of probability theory was so low that one prominent statistician remarked that he supposed it was possible to teach probability apart from statistics, but that doing so would be a tour de force in which he could see no point.

The mathematical background of probability was confused even in the periodical literature. It was not yet clear that there was a distinction between probability as a mathematical theory and as a theory of real events. This confusion enlivened meetings with heated controversies, unhappily absent now that probability has lost its youthful charm and vagueness.

A change was imminent, however. Borel's discussion of "denumerable probabilities" in 1909 had attracted attention to a new class of problems, those involving complete additivity, and Lebesgue's measure theory had already provided the needed mathematical background. It was becoming clear that the connection between probability and measure theory was at least very close. In certain special cases at least, for example in Wiener's discussion of Brownian motion in 1923, studies of infinite collections of random variables were carried out by representing the random variables as measurable functions on a measure space.

The turning point was the appearance of Kolmogorov's monograph in 1933, which laid the basis for probability in terms of measure theory. The Russian mathematical group was at a great advantage in that they were the heirs of a strong tradition, going back to the work of Chebycheff and A. A. Markov and continuing with Bernstein. Unfortunately the significance of the Russian work was misunderstood or ignored until about the middle thirties. One of the leaders of the new Russian school, whose span of mathematical activity covered the final establishment of probability as a mathematical subject, and who contributed greatly to this establishment, was Khinchin.

Alexander Iacovlevich Khinchin was born in 1894 in the village of Kondrovo, about one hundred miles southwest of Moscow. He died November 18, 1959, a professor at Moscow University, one of the founders and still one of the leaders of the Moscow probability group. Some of his work is discussed below.

Khinchin was a student at Moscow University, where he was strongly influenced by Luzin. His papers were mostly on various topics connected with measure theory until 1924, when he wrote his first probability paper, on the law of iterated logarithm for a Bernoulli sequence of random variables. This initiated the work of the Moscow probability school, and Khinchin's own work thereafter was almost entirely devoted to probability theory and its applications.

Besides his research papers, Khinchin wrote extremely useful and influential monographs on limit distributions (1927, 1933, 1938), on statistical mechanics (1943, 1950, 1951), on mass service (1955). His papers on information theory, although less original, were also stimulating and influential.

Khinchin's writing was distinguished by its elegance and its concentration on the essentials of the ideas involved. These qualities of his writing made him ideally suited to clarify the ergodic theorem and to expound statistical mechanics and information theory without encumbering these subjects with the usual irrelevancies.

In the following, a few results typical of Khinchin's research in probability theory will be discussed.

In 1925 Khinchin and Kolmogorov initiated the systematic study of the convergence theory of infinite series whose terms are independent random variables. In this paper, Khinchin proved that, for countably valued random variables, convergence of means and variances guarantees probability one convergence. It is interesting historically to observe that Khinchin considered it necessary to construct his random variables as functions on the interval [0, 1] with Lebesgue measure. It is of course no longer necessary to go through such construction procedures.

Let $x_1, x_2, \cdots$ be mutually independent random variables with a common distribution. Khinchin proved in 1929 that $\sum_1^n x_j/n \to E\{x_1\}$ stochastically, if the indicated expectation is finite. This important step has since been overshadowed by the fact that a few years later it was shown that there is even convergence with probability 1 under the same hypotheses.

A distribution is called infinitely divisible if, for each positive integer n, the distribution is the nth convolution of some distribution with itself. For large n, the latter distribution will necessarily be concentrated near 0. Conversely, Khinchin, and independently Feller, proved in 1937 that a distribution is infinitely divisible if it is the convolution of distributions concentrated arbitrarily near 0. This implies that the distribution of an increment of a continuous parameter stochastic process with independent increments (under hypotheses which only eliminate trivial irrelevancies) are infinitely divisible.

Khinchin was one of the leaders in the study of limit distributions of sums of independent random variables. Only a few of his many results in this area will be singled out here.

A class of distributions is called a "distribution type" if (a) any two distributions in the class, corresponding say to distribution functions F, G, are related by an equation of the form $F(x) = G(ax + b)$, where a and b are constants with $a > 0$, and if (b) any two distributions related in this way are either both or neither in the class. The type is called improper if its distributions are all concentrated at single points.

Khinchin proved in 1937 that a sequence of distributions taken from an assigned sequence of types can converge only to a distribution in a single proper type, that is, a sequence of types can converge to only one proper type.

A type is called stable if the convolution of two distributions of this type is of the same type. The members of a stable type are called stable distributions. In 1936, completing earlier work by Lévy, Lévy and Khinchin jointly found the characteristic functions of the stable distributions.

If $x_1, x_2, \cdots$ are independent random variables with a common distribution, and if y_n is defined by

$$y_n = B_n^{-1} \sum_1^n x_j - A_n, \tag{1}$$

where A_n and B_n are constants, the sequence of distributions of $y_1, y_2, \cdots$ can have only stable distributions as limit distributions. Khinchin proved in 1937, however, that, allowing limit distributions of y_n subsequences, every infinitely divisible distribution can be obtained as such a limit.

The class of distributions is a semigroup under the convolution operation considered as multiplication. The distribution concentrated at 0 is the identity element. Khinchin proved in 1937 that, with the obvious interpretation of the algebraic language, every distribution with no prime factor is infinitely divisible. Conversely, however, there is an infinitely divisible distribution with a prime factor. Furthermore, every distribution is the product of two, the first of which contains no prime factor and the second of which is either the identity or the product of a countable number of prime factors. This decomposition may not be unique.

In 1931, Birkhoff proved the (pointwise limit) ergodic theorem, in an obscurely written paper whose mathematics carried more than a vestigial trace of the physical background of the theorem. In 1933 Khinchin proved Birkhoff's theorem, using Birkhoff's method but cleaning out the irrelevancies and thereby clarifying the mathematical significance of the theorem. Khinchin understood at once the relation between this theorem and the law of large numbers, namely that the ergodic theorem in Birkhoff's form is simply the strong law of large numbers for strictly stationary processes. In 1934, Khinchin returned to this subject, in an important and historically influential study of what are now called wide sense stationary processes (continuous parameter case). In this paper, Khinchin found the spectral form of the covariance function and the law of large numbers appropriate to these processes. Although it is now known that these results can be deduced from general Hilbert space theorems known at the time (in fact are identical with them) this fact was not clear then and does not nullify the historical significance of Khinchin's work.

Khinchin had been interested in the theory of mass servicing for some years before he published his monograph on the subject in 1955. In fact he had been consulted in 1930 by engineers of the Moscow telephone system, and he published his first paper on the subject in that year. In his 1955 monograph he gives an elegant treatment of stationary streams of events, setting up differential coefficients $\phi_k(t)$ giving the probability that k events will occur in an interval of length t if one has occurred at the beginning of the interval. This monograph is typical of the clarity of his analysis in illuminating an already well-developed field.