FUNDAMENTAL LIMIT THEOREMS OF PROBABILITY THEORY[1]

By M. Loève[2]

University of California, Berkeley

no sooner is Proteus caught
than he changes his shape

1. Introduction. The fundamental limit theorems of Probability theory may be classified into two groups. One group deals with the problem of *limit laws* of sequences of sums of random variables, the other deals with the problem of *limits of random variables*, in the sense of almost sure convergence, of such sequences. These problems will be labelled, respectively, the Central Limit Problem (CLP) and the Strong Central Limit Problem (SCLP). Like all mathematical problems, the CLP and SCLP are not static; as answers to old queries are discovered they experience the usual development and new problems arise. The development consists in (i) simplifying proofs and forging general tools out of the special ones (ii) sharpening and strengthening results (iii) finding general notions behind the results obtained and extending their domains of validity. *Analysis of this growth will put in relief the role and the interconnections of the fundamental limit theorems.*

Summary. The growth of the CLP for independent summands can be divided into three (overlapping) periods. The first covers the Bernoulli case and the corresponding limit theorems of Bernoulli, de Moivre and Poisson. The first two theorems gave rise to the notions—from which the classical CLP stems—of the Law of Large Numbers (LLN) and of Normal Convergence (NC). Poisson's approach belongs to the set-up of the modern CLP.

The second period extends over two centuries and is devoted to the extension of the domains of validity of LLN and NC. This is the classical CLP period. Lyapunov's crucial work, submitted to the above treatment, led to the discovery of the natural boundaries of these domains by Lindeberg, Kolmogorov, Feller and P. Lévy.

However, the LLN and NC problems are but two particular cases of the general problem of limit laws of sequences of sums of independent random variables. The coming into sight and the solution of this problem—the third period of the CLP—covers less than ten years. The tools forged for the classical CLP proved to be powerful enough and the final solution is due to P. Lévy, Khintchine, Gnedenko and Doeblin.

[1] This paper was presented to the New York meeting of the Institute of Mathematical Statistics on December 27, 1949.

Editor's Note: The Institute of Mathematical Statistics has formed a Committee on Special Invited Papers to invite lecturers to deliver expository addresses to the Institute with the understanding that the Special Invited Papers are to be published in the *Annals of Mathematical Statistics*. This paper is the first one invited by the Committee.

[2] This work is supported in part by the Office of Naval Research.

The CLP for dependent variables started with so called Markoff chains. The study of their limit properties is due essentially to Markov, S. Bernstein and Doeblin. For more general forms of dependence the LLN and NC problems were investigated by P. Lévy and Loève after the crucial work of S. Bernstein. The modern CLP was considered only recently (Loève).

The SCLP stems from the strengthening by Borel of the Bernoulli theorem and the sharpening of Borel's result by Khintchine. They gave rise to the notions of Strong Law of Large Numbers (SLLN) and of the Law of the Iterated Logarithm (LIT).[3] The domains of validity were extended to their boundaries by Kolmogorov, P. Lévy and Feller. In the case of dependence, results are due to G. D. Birkhoff, P. Lévy, W. Doeblin, and Loève. However, the SCLP has not attained, at present, the harmonious development of the CLP.

Notations. Let $\mathfrak{L}(X)$ be the *law* of a (real) random variable (r.v.) X. The law is defined by the *distribution function* (d.f.) $F(x) = P(X < x)$. As is well known $\mathfrak{L}(X)$ is determined by the *characteristic function* (ch. f.)

$$f(u) = \int_{-\infty}^{+\infty} e^{iux}\, dF(x), \qquad -\infty < u < +\infty.$$

When a r.v. possesses subscripts, the same subscripts will be used for its d.f. and ch.f. EX will denote the *expectation* of X:

$$EX = \int_{-\infty}^{-\infty} x\, dF(x),$$

and $\sigma^2(X)$ will denote the *variance* of X:

$$\sigma^2(X) = E(X - EX)^2.$$

With a random event A we associate a r.v., to be called *indicator* of the event A, which takes values 1 and 0 respectively, according as A occurs or does not occur. If X is the indicator of an event A of probability p, then $EX = p$ and $\sigma^2(X) = pq$, where $q = 1 - p$. To avoid trivialities we shall assume that $pq \neq 0$.

Two laws $\mathfrak{L}(X_1)$ and $\mathfrak{L}(X_2)$ will be said to belong to the same *complete type* if there exist two numbers $a \neq 0$ and b such that $P\{X_1 \leqq x\} = P\{aX_2 + b \leqq x\}$. If values of a are restricted to positive values, then the two laws are said to belong to the same *type*. If two independent r.v.'s obey $\mathfrak{L}$ and their sum belongs to the type of $\mathfrak{L}$, then $\mathfrak{L}$ and its type are said to be *stable*. Three classes of laws play an essential role in the CLP: the normal and the degenerate types and the Poisson complete types.

$\mathfrak{N}(m, \sigma)$ is a *normal* law if it is defined by

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(1/2\sigma^2)(t-m)^2}\, dt \qquad (\sigma > 0).$$

[3] For a very thorough and deep analysis of the NC and LIT problems and their solutions see FELLER, *Bull. Am. Math. Soc.*, Vol. 51 (1945), pp. 800–832, under the same title as that of the present paper.

$\mathfrak{L}(m)$ is a law *degenerate* at m, if it attaches probability 1 to the value m. $\mathcal{P}(\lambda;\, a,\, b)$ is a *Poisson* law if

$$P(X = ak + b) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad (\lambda > 0), \qquad k = 0, 1, 2, \cdots;$$

the familiar Poisson law is $\mathcal{P}(\lambda;\, 1,\, 0)$.

A law $\mathfrak{L}(X_n)$ is said to converge to the law $\mathfrak{L}(X)$ as $n \to \infty$, if $F_n(x)$ converges to $F(x)$ at the continuity points of the latter. In this paper, all limits will be considered for $n \to \infty$, if not otherwise stated.

The structure of sequences of r.v.'s whose limit properties are investigated will be called the *limiting process* of the problem. The limiting process of *sequences of sums* is that of sequences of the form $S_{n,\nu_n} = \sum_{k=1}^{\nu_n} X_{n,k}$, where $\nu_n \to \infty$. The limiting process of *normed sums* is that of sequences of the form $\frac{S_n}{a_n} - b_n$ with $S_n = \sum_{k=1}^{n} X_k$, where $a_n > 0$ and b_n are real numbers. Normed sums are a special form of sequences of sums: take $\nu_n = n$, $X_{n,k} = \frac{X_k}{a_n} - \frac{b_n}{n}$, then $S_{n,\nu_n} = \frac{S_n}{a_n} - b_n$.

To avoid repetitions we shall note, once and for all, that limit types rather than limit laws appear in the case of normed sums, because, if $\mathfrak{L}(X)$ is their limit law, then any law of its type is obtainable as a limit law by a convenient change of origin b_n and of scale a_n, independent of n. The importance of the notion of type is due, primarily, to this property. In fact, even more is true: *if $\mathfrak{L}(X_n)$ converges to $\mathfrak{L}(X)$ and $\mathfrak{L}(a_nX_n + b_n)$ converges to $\mathfrak{L}(Y)$, then $\mathfrak{L}(X)$ and $\mathfrak{L}(Y)$ belong to the same type, provided neither is degenerate* (Khintchine [20]).

I. Central Limit Problem

2. Origin of the CLP: Binomial case. Three limit theorems are at the origin of the CLP; the first, due to Bernoulli ([2], 1713), laid the ground. *Let S_n be the number of occurrences of an event A of probability p in n identical and independent trials. Then, for every $\epsilon > 0$,*

$$P\left\{\left|\frac{S_n}{n} - p\right| > \epsilon\right\} \to 0.$$

Bernoulli found this result by a direct, but cumbersome, analysis of the behaviour of the binomial probabilities

$$P\{S_n = k\} = C_n^k p^k q^{n-k}, \qquad k = 0, 1, \cdots, n.$$

Sharpening this analysis, de Moivre ([7], 1730) obtained the second limit theorem of probability theory which, in the form given to it by Laplace, states that: *For every x*

$$P\left\{\frac{S_n - np}{\sqrt{npq}} < x\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt.$$

Suppose now, with Poisson ([36], 1837), that the probability $p = p_n$ depends upon the number n of trials and, more precisely, that $p_n = \frac{\lambda}{n}$, where λ is a positive constant. Write then $S_{n,n}$, instead of S_n, for the number of occurrences of the considered event in a group of n trials. By a direct analysis of the binomial probabilities, much easier to carry out than the preceding ones, it follows that for $k = 0, 1, \cdots$,

$$P\{S_{n,n} = k\} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}.$$

Let X_k be the indicator of the event A in the k-th trial. The number of occurrences S_n is the sum $\sum_{k=1}^{n} X_k$ of n of these independent and identically distributed indicators. The first two limit theorems mean that

$$\mathfrak{L}\left(\frac{S_n - ES_n}{n}\right) \rightarrow \mathfrak{L}(0) \quad \text{and} \quad \mathfrak{L}\left(\frac{S_n - ES_n}{\sigma S_n}\right) \rightarrow \mathfrak{N}(0, 1).$$

Thus we have two limiting processes, (both special and completely specified forms of normed sums), and two limit laws (more precisely two limit types, see introduction), a degenerate and a normal one.

Poisson's limiting process is utterly different. $S_{n,n}$ is still a sum $\sum_{k=1}^{n} X_{n,k}$ of independent and identically distributed indicators but, as n varies, *all* $X_{n,k}$ change, $P(X_{n,k} = 1) = \frac{\lambda}{n}$ and

$$\mathfrak{L}(S_{n,n}) \rightarrow \mathcal{P}(\lambda; 1, 0).$$

While the two first theorems with their special limiting processes and limit laws played a central role in the development of Probability theory, Poisson's result stood isolated and ignored until about fifteen years ago.[4] We shall see further that there was a deep reason for its isolation and also that, surprisingly enough, Poisson laws are, in a sense, more fundamental for the CLP, than the normal law.

3. The classical CLP and its extension. From the time of Laplace until 1935, research in the domain of limit laws was centered about the extension to summands other than indicators of the validity of the two first limit theorems. This is the period of the classical CLP: *Let* $S_n = \sum_{k=1}^{n} X_k$ *be sums of independent r.v.'s. Find necessary and sufficient conditions for the LLN and for NC*, i.e., conditions under which, respectively,

$$\text{LLN: } \mathfrak{L}\left(\frac{S_n - ES_n}{n}\right) \rightarrow \mathfrak{L}(0),$$

$$\text{NC: } \mathfrak{L}\left(\frac{S_n - ES_n}{\sigma(S_n)}\right) \rightarrow \mathfrak{N}(0, 1).$$

[4] In Uspensky's textbook (1937!) Poisson's law is mentioned once—in an exercise.

It is assumed that EX_k's and EX_k^2's exist. The d.f. not being completely specified as in the Bernoulli case, the direct Bernoulli-de Moivre approach is of no avail and general methods are necessary. The first to appear was the method of moments relative to bounds of d.f. in terms of their moments (Tchebicheff [40], Markov [37]). The relation

$$P\left\{\frac{|S_n - ES_n|}{n} > \epsilon\right\} \leqq \frac{\sigma^2(S_n)}{\epsilon^2 n^2}, \qquad \epsilon > 0,$$

together with

$$\sigma^2(S_n) = \sum_{k=1}^{n} \sigma^2(X_k),$$

entails at once a LLN theorem (Tchebicheff-Markov): *If*

$$\frac{1}{n^2} \sum_{k=1}^{n} \sigma^2(X_k) \to 0,$$

then the LLN holds.

This result can be easily improved (bringing it into closer analogy with Lyapunov's theorem): *If there exists a constant* $\delta > 0$ *such that*

$$\frac{1}{n^{1+\delta}} \sum_{k=1}^{n} E\,|X_k - EX_k|^{1+\delta} \to 0$$

then the LLN holds.

It contains then a Markov's LLN condition: LLN holds if $E\,|X_k - EX_k|^{1+\delta} \leqq C$ where C is independent of k.

In a much more elaborate form the method of moments gives also a NC theorem (Tchebicheff-Markov): *If* $EY_n^k \to EZ^k$ *for* $k = 1, 2, \cdots$, *and* $\mathfrak{L}(Z) = \mathfrak{N}(0, 1)$, *then* $\mathfrak{L}(Y_n) \to \mathfrak{N}(0, 1)$.

This theorem has been extended to more general limit laws. However the inherent defects of the method of moments remain. Even if moments of all orders exist, they do not necessarily determine a unique d.f. A definitive result in this direction is the Fréchet-Shohat theorem: *If* $EY_n^k \to m^{(k)}$ *for all* k, *there exists a subsequence* $\mathfrak{L}(Y_{n_\nu})$ *which converges to a limit law* $\mathfrak{L}$ *with moments* $m^{(k)}$. *Moreover, if the moment problem is determined, i.e., if the* $m^{(k)}$ *determine a unique law, then the whole sequence* $\mathfrak{L}(Y_n)$ *converges to* $\mathfrak{L}$.

To apply the convergence theorem to the NC part of the classical CLP, one has to assume existence of moments of all orders. In particular, it does not seem suitable for proving Lyapunov's theorem. Yet, the simple *truncation* idea (Markov) not only overcomes this seemingly insurmountable obstacle, but also provides a method per se. It associates with the summands X_k "truncated" r.v.'s X_k'; for $k \leqq n$ and c_n conveniently chosen real numbers,

$$X_k' = X_k \text{ if } |X_k| \leqq c_n,$$

$$X_k' = 0 \quad \text{if } |X_k| > c_n.$$

Nevertheless, the method of moments is too cumbersome and was soon to be discarded in favor of that of ch.f.'s.

The turning point for the entire CLP is Lyapunov's introduction of the *method of ch.f.'s*. The ch.f.'s were well known and used already by Laplace. However, the first convergence property, proved but not stated, is due to Lyapunov [28]: *If the ch.f.'s* $g_n(u)$ *of* $\mathfrak{L}(Y_n)$ *converge to the ch.f.* $e^{-u^2/2}$ *of* $\mathfrak{N}(0, 1)$, *then* $\mathfrak{L}(Y_n) \to \mathfrak{N}(0, 1)$. From it he deduced the first general NC theorem [28, 29]: *If there exists a number* $\delta > 0$ *such that*

$$\frac{1}{\sigma^{2+\delta}(S_n)} \sum_{k=1}^{n} E \mid X_k - EX_k \mid^{2+\delta} \to 0,$$

then NC holds.

The ch.f. became, in the hands of P. Lévy [21], a general tool, instrumental in the subsequent tremendous growth of the CLP, with the so called

CONTINUITY THEOREM. *If the ch.f.'s* $g_n(u)$ *converge to a function* $g(u)$ *continuous at* $u = 0$, *then* $\mathfrak{L}(Y_n)$ *converge to a limit law* $\mathfrak{L}$ *and* $g(u)$ *is its ch.f.; and conversely.*

The methods of ch.f. and of truncation dominate at present the limit problems of Probability theory.

In spite of the generality of the above conditions for LLN and NC, they are not *necessary* conditions. In fact they are not sharp enough since they assume the existence of moments of higher order than those which figure in the classical CLP. However the tools forged proved powerful enough to get its complete solution. The truncation method yielded to Kolmogorov ([16, 1928) the complete answer to the LLN problem. A "smoothing" device, due to Lyapunov, provided Lindeberg ([20], 1922) with adequately sharp sufficient conditions; using ch.f.'s P. Lévy ([22], 1922) proved Lindeberg's result and Feller ([11], 1935) showed that, under a natural restriction, these conditions are also necessary.

Solution of the classical CLP.

1. *LLN holds if, and only if,*

$$\sum_{k=1}^{n} \int_{|x|>n} dF_k(x + EX_k) \to 0 \quad \textit{and} \quad \sum_{k=1}^{n} \frac{1}{n^r} \int_{|x|<n} x^r\, dF_k(x + EX_k) \to 0 \quad \textit{for } r = 1, 2.$$

2. *NC holds and* $\max_{k \leq n} \dfrac{\sigma(X_k)}{\sigma(S_n)} \to 0$ *if, and only if, for every* $\epsilon > 0$,

$$\sum_{k=1}^{n} \frac{1}{\sigma^2(S_n)} \int_{|x|>\epsilon\sigma(S_n)} x^2\, dF_k(x + EX_k) \to 0.$$

An unsatisfactory feature of the classical CLP is the assumption, made at the start, of existence of certain moments. They are used to avoid, as $n \to \infty$, the shift, towards infinite values, of the probability spread by changing the origin and the scale of values of S_n. However there is no specific reason for these special choices of norming quantities a_n and b_n except that, historically, they appeared as a straightforward extension of Bernoulli and de Moivre ones. Moreover, even if these moments do not exist, there is no reason not to try to find norming quantities. (Take X_k's to be independent and identically distributed as follows: to $\pm\sqrt{m}$ where $m = 1, 2, \cdots$, attach probabilities $\frac{3}{\pi^2 m^2}$. The second moments are infinite; yet norming S_n by $c\sqrt{n \log n}$, we have NC.) Thus the CLP becomes the problem of the LLN and NC for general normed sums $\frac{S_n}{a_n} - b_n$.

The extended classical NC problem was solved, masterfully and independently, by Feller ([10], 1935) using ch.f.'s and by P. Levy ([25], 1935) who applied the method of truncation. The extension of the results to the more general set-up of the following section is trivial and will be given there. Feller also solved ([11], 1937) the extended LLN problem.

In this new set-up a question arises at once. *Given the r.v.'s X_k, do there exist numbers which will produce the desired convergence? If so, how can they be found?* This problem is perhaps more difficult than the previous one and is specifically linked with the limiting process of normed sums. We shall give here a criterion, due to Feller ([10], 1935), which solves entirely the NC problem.[5] Take as origin of values of the summands their medians and let $c_n(\epsilon)$ be the g.l.b. of the x's for which $\sum_{k=1}^{n} P(|X_k| > x) \leqq \epsilon$. Then *norming quantities a_n and b_n such that $\mathfrak{L}\left(\frac{S_n}{a_n} - b_n\right) \to \mathfrak{N}(0, 1)$ and $\max_{k \leqq n} P\left\{\left|\frac{X_k}{a_n} - \frac{b_n}{n}\right| > \epsilon\right\} \to 0$ exist if, and only if, for every $\epsilon > 0$,*

$$\frac{1}{c_n^2(\epsilon)} \int_{|x| < c_n(\epsilon)} x^2 \, dF_k(x) \to \infty.$$

4. Modern CLP. At the same time that the classical CLP neared its happy end, a new and much wider problem of limit laws appeared and, because the necessary tools were at hand, was solved almost at once. Various particular problems, of which the classical CLP is one, contributed to its set-up.

Since the discovery, in the Bernoulli case, of the LLN and NC, the problem of limit laws has been centered about extensions of their domains of validity for more and more general normed sums. A similar query about the Poisson convergence would have provided us with a new problem. As soon as we drop the restriction that in $S_{n,\nu_n} = \sum_{k=1}^{n} X_{n,k}$ the r.v.'s $X_{n,k}$ are indicators, we are led to the problem of finding conditions under which laws of sums of independent r.v.'s will converge to a Poisson law. We have here not only a different limit law than in the CLP but also a more general limiting process. An utterly different problem, stated and solved by P. Lévy [21], is the following: *find the*

[5] As for the LLN, norming numbers, such that the LLN holds always exist whatever be the r.v.'s X_k. Hence, from the point of view of limit types of normed sums, the degenerate type is to be considered as a degenerate form of every limit type.

possible limit laws of normed sums of independent and identically distributed r.v.'s (the answer is that *they are the stable laws*). For the first time one does not inquire about a completely specified limit law but about the class of *all* limit laws for a fairly general limiting process. Thus, starting with limit theorems with completely specified limiting processes and limit laws, after two centuries of struggle Probability theory got rid of initial restrictions.

The general set-up is now visible. The limiting process is that of sequences of sums of independent r.v.'s. The queries are about the classes of possible limit laws and conditions of convergence. However, so general a limit problem is without content. In fact, the limiting process is that of arbitrary sequences of r.v.'s: let $\{Y_n\}$ be any sequence of r.v.'s and take $X_{n,1} = Y_n$, $\mathfrak{L}(X_{n,k}) = \mathfrak{L}(0)$ for $k > 1$. Any law $\mathfrak{L}$ belongs to the class of limit laws: take $\mathfrak{L}(Y_n) \equiv \mathfrak{L}$. Hence some restriction is needed. To find a "natural" restriction consider the previous problems. Their common feature is that the limiting process is that of sequences of *sums* of independent r.v.'s, *the number of summands increasing indefinitely*. If we wish to emphasize this feature, a relatively small number of summands ought not to have a preponderant role in the determination of the limit laws. A "natural" restriction is then a requirement of *uniform asymptotic negligibility* (uan) *of the summands*, i.e., for every $\epsilon > 0$, $P\{\mid X_{nk} \mid > \epsilon\} \to 0$ uniformly in k. We come thus to the *Modern CLP. Let* $S_{n,\nu_n} = \sum_{k=1}^{\nu_n} X_{n,k}$, $\nu_n \to \infty$, *be sums of r.v.'s* $X_{n,k}$, *mutually independent for every fixed n, and such that*

$$\max_k P\{X_{n,k} \mid > \epsilon\} \to 0;$$

characterize the class $\{D\}$ *of limit laws of the* S_{n,ν_n} *and find necessary and sufficient conditions for convergence to any element of this class.*

The solution of this problem is essentially due to the results of investigation of random functions $X(t)$ with independent increments. Let $X(0) = 0$, divide the interval $(0, t)$ into ν_n subintervals (t_{k-1}, t_k) with $t_0 = 0$, and denote by X_{nk} the increment $X(t_k) - X(t_{k-1})$. Then $X(t) = \sum_{k=1}^{\nu_n} X_{nk}$ where X_{nk} are independent r.v.'s. If, moreover, $X(t)$ is continuous in probability for every t, i.e., if $\mathfrak{L}\{X(t+h) - X(t)\} \to \mathfrak{L}(0)$ as $h \to 0$, then the $X_{n,k}$ can be chosen to obey the uan restriction as $\nu_n \to \infty$. Hence $\mathfrak{L}\{X(t)\}$ might be expected to belong to $\{D\}$.

The particular case of the modern CLP for summands and limit laws with the finite second moments was solved by Bawly [1], using Kolmogorov's characterization of $X(t)$'s with finite second moments [7]. The general problem, thanks to a much more general result by P. Lévy ([24], 1934), was solved by P. Lévy, Khintchine ([20], 1937), Gnedenko ([14], [15], 1938, 1939) and Doeblin ([8], 1938–1939). The method used throughout was that of ch.f.'s. (except in the case of Doblin who used also the P. Lévy "dispersion" function).

One can avoid an explicit introduction of the considered random function $X(t)$, limiting oneself to the corresponding (infinitely divisible) laws. For a very large n, S_{n,ν_n} is, roughly speaking, a very large number ν_n of very small (in probability) independent summands. This leads at once to the consideration of laws which possess such a property for any ν_n and, first, the *infinitely divisible* (i.d.) laws. A law is i.d. if it is a law of sums of an arbitrarily large number of independent and identically distributed r.v.'s. In other words, *$f(u)$ is the ch.f. of an i.d. law if $[f(u)]^{1/n}$ is a ch.f. for every positive integer n.* One might expect i.d. laws to belong to $\{D\}$ and, surprisingly enough, it turns out that, because of the uan, $\{D\}$ contains *only* i.d. laws.

We can now state the solution of the modern CLP, in three parts. Let $\mathop{\int\!\!\!\!-}_{-a}^{+a} = \int_{-a}^{-0} + \int_{+0}^{+a}$, let $\phi(x)$ be any function, defined and non-decreasing in $(-\infty, -0)$ and $(+0, +\infty)$, with $\phi(-\infty) = \phi(+\infty) = 0$ and $\mathop{\int\!\!\!\!-}_{-\epsilon}^{+\epsilon} x^2\, d\phi(x) < \infty$, and let α and β be real numbers.

I. *The function $f(u)$ is the ch.f. of an i.d. law if, and only if,*

$$\log f(u) = i\alpha u - \frac{\beta^2}{2} u^2 + \mathop{\int\!\!\!\!-}_{-\infty}^{+\infty} \left(e^{iux} - 1 - \frac{iux}{1 + x^2}\right) d\phi(x),$$

and $f(u)$ determines uniquely α, β and $\phi(x)$ at all the continuity points of the latter. (P. Lévy).

Normal laws are obtained for $\phi(x) \equiv 0$ and Poisson laws correspond to the $\phi(x)$ with one point of increase ($x \neq 0$) only. The fundamental role of Poisson laws appears clearly since, roughly speaking, an i.d. law is the convolution of a normal law and a continuum of Poisson ones. This role is further emphasized by the following theorem (Khintchine [20]): *A law is i.d. if, and only if, it is the limit law of sequences of sums of independent Poisson r.v.'s.* In other words, the class of i.d. laws is the closure of laws of finite sums of independent Poisson r.v.'s.

II. *The class $\{D\}$ of limit laws of the modern CLP coincides with that of i.d. laws* (P. Lévy-Khintchine).

Together with I this result characterizes in an explicit manner the class $\{D\}$. An immediate question arises (Khintchine). What about the limit laws of normed sums? The answer is the following (P. Lévy [27]). Let $y = \log |x|$, $\psi_1(y) = -\phi(x)$ for $x < 0$, $\psi_2(y) = \phi(x)$ for $x > 0$ where $y = \log |x|$. *The limit laws of normed sums, under uan, are the i.d. laws with convex $\psi_k(y)$, $k = 1, 2$.*

In particular a Poisson law does not belong to this subclass $\{D_N\}$ of $\{D\}$, hence cannot be obtained as a limit law of normed sums. This brings out the deep reason for the isolation in which the Poisson law remained as long as the limiting process was restricted to that of normed sums.[6] II shows that, with respect to the possible limit laws, the limiting process of the modern CLP is definitely wider than that of the classical CLP and of its extension. However the entire class $\{D\}$ can be obtained with normed sums, provided we consider

[6] A problem, specific for normed sums, arises: given r.v.'s X_k, find necessary and sufficient conditions for existence of norming numbers such that the laws of normed sums would converge to a given element of $\{D_N\}$ and, if they exist, find them. Feller's NC criterion solves a particular case of this problem.

not only limit laws but also "accumulation" laws (P. Lévy-Khintchine): *A law is i.d. if, and only if, it is the limit law of a subsequence of normed sums of independent and identically distributed r.v.'s.*

I and II provided Gnedenko and, independently, Doeblin with the properties which allowed them to find conditions of convergence, thus completing the solution of the modern CLP. Let

$$\sigma_\epsilon^2(X) = \int_{|x|<\epsilon} x^2\, dF(x) - \left[\int_{|x|<\epsilon} x\, dF(x)\right]^2$$

denote a "truncated" variance of X.

III. *Under uan, $\mathfrak{L}(S_{n,\nu_n} - b_n)$ converges, necessarily to an i.d. law "for a convenient choice of b_n", if, and only if,*

$$\text{(i)} \quad \sum_{k=1}^{\nu_n} F_{nk}(x) \to \phi(x) \text{ for } x < 0, \qquad \sum_{k=1}^{\nu_n} [1 - F_{nk}(x)] \to -\phi(x) \text{ for } x < 0$$

at the continuity points of $\phi(x)$, and

$$\text{(ii)} \qquad \lim_{\epsilon\to 0} \liminf_{n} \sum_{k=1}^{\nu_n} \sigma_\epsilon^2(X_{n,k}) = \beta^2 .$$

In particular, since normal laws correspond to $\phi(x) \equiv 0$, the NC conditions of Feller and P. Lévy follow: *$\mathfrak{L}(S_{n,\nu_n} - b_n)$ converges to $\mathfrak{N}(0, 1)$ for a convenient choice of b_n and uan holds if, and only if, for every $\epsilon > 0$,*

$$\text{(i)} \quad \sum_{k=1}^{\nu_n} \int_{|x|>\epsilon} dF_{nk}(x) \to 0 \qquad \textit{and} \qquad \text{(ii)} \quad \sum_{k=1}^{\nu_n} \sigma_\epsilon^2(X_{nk}) \to 1.$$

The first condition shows that among all limit laws under uan, limit normality corresponds to a sufficiently strong asymptotic negligibility of the summands, and, more precisely, to

$$\sum_{k=1}^{\nu_n} P(|X_{nk}| > \epsilon) \to 0,$$

or, equivalently, to

$$P(\max_k |X_{nk}) > \epsilon) \to 0.$$

Another illuminating characterization of NC (Raikov [39]) follows also from III. Take for origin of values of summands the truncated first moments $\int_{|x|<1} x\, dF_{nk}(x)$. Then *$\mathfrak{L}(S_{n,\nu_n} - b_n) \to \mathfrak{N}(0, 1)$ for a convenient choice of b_n if, and only if, $\mathfrak{L}(\sum_{k=1}^{\nu_n} X_{nk}^2) \to \mathfrak{L}(1)$.*

5. CLP in the case of dependence. Limit problems for sums of dependent r.v.'s. were considered for the first time by Markov [37], less than fifty years ago. He extended the first two limit theorems of probability theory to the case of events linked in *chain*, i.e., such that $P(A_k \mid A_1, \cdots A_{k-1}) = P(A_k \mid A_{k-1})$.

However the crucial work in this field is the celebrated memoir by S. Bernstein ([3], 1927) which has the same historical importance for the dependence case as that of Lyapunov has for the classical CLP.

Let $\{X_k\}$ be a sequence of r.v.'s. $E'X_k$ will denote the conditional expectation of X_k, given $X_1, \cdots X_{k-1}$. Consider the sequence of sums $S_n = \sum_{k=1}^{n} X_k$, with $EX_k \equiv 0$ and let $\sigma_n = \sqrt{\sum_{k=1}^{n} \sigma^2(X_n)}$.

BERNSTEIN'S NC THEOREM. *If*

$$\text{(i)} \quad \frac{1}{\sigma_n} \sum_{k=1}^{n} \sup | E'X_k | \to 0, \qquad \text{(ii)} \quad \frac{1}{\sigma_n^2} \sum_{k=1}^{n} \sup | E'X_k^2 - EX_k^2 | \to 0,$$

and

$$\text{(iii)} \quad \frac{1}{\sigma_n^3} \sum_{k=1}^{n} \sup E' \, | X_k |^3 \to 0,$$

then

$$\mathcal{L}\left(\frac{S_n}{\sigma_n}\right) \to \mathfrak{N}(0, 1).$$

Obviously, if the X_k's are independent, this theorem reduces to Lyapunov's with $\delta = 1$. The method used is still that of ch.f.'s. From this result Bernstein deduces various particular NC cases and, applying them to Markov chains, extends the latter's results.

The unpleasant feature of the above theorem is the use of suprema of conditional expectations and, except when the r.v.'s X_k are bounded, one cannot exfect these suprema to be finite. On the other hand, the conditional expectations are r.v.'s and it would be natural to associate their values with the corresponding probabilities. This can be done and Bernstein's theorem can be improved in various directions simultaneously. First it may be stated for sequences of sums S_{n,ν_n}—this is trivial; next it extends to $\delta > 0$ instead of $\delta = 1$—this contains completely Lyapunov's result but is of secondary interest. Then NC can be replaced by *asymptotic normality*, i.e., by the existence of a sequence of normal laws $\mathfrak{N}(0, \sigma_n)$ such that the "distance" between $\mathcal{L}(S_{n,\nu_n})$ and $\mathfrak{N}(0, \sigma_n)$ would approach zero as $n \to \infty$—this is quite simple to get. However, significant improvements are obtained on replacing suprema by expectations. Let $F_n(x)$ be the d.f. of S_{n,ν_n} and $G_n(x)$ be that of $\mathfrak{N}(0, \sigma_n^2)$. Then, taking $EX_{nk} \equiv 0$, we have the following

NC THEOREM. *If* (i) $\sum_k E \, | E'X_{nk} | \to 0$, (ii) $\sum_k E \, | E'X_{nk}^2 - EX_{nk}^2 | \to 0$ *and* (iii) *there exists a constant* $\delta > 0$ *such that* $\sum_k | X_{nk} |^{2+\delta} \to 0$, *then* $F_n(x) - G_n(x) \to 0$.

This theorem shows that, so far as moments of order higher then the second are concerned, the NC condition is the same as in the case of independence. In this last case the theorem is a slight improvement of that of Lyapunov. In 1941 conditions for LLN and NC were given (Loève [31], [32]) in the frame of the modern CLP, without assuming the existence of moments; when independence is assumed, they reduce to those given by Feller. Conditions for NC which in the case of independence, reduce to Lindeberg's, were then deduced in the particular case of finite second moments and special cases of NC, including those considered by S. Bernstein, were obtained.

The whole modern CLP had not been considered until lately (Loève, [33–35]). It appeared useful to extend the CLP to an "Asymptotic Central Problem" (ACP); primarily, to the behavior of $\mathfrak{L}(S_{n,\nu_n})$ as $n \to \infty$. This in turn, led to the introduction of laws "in a wide sense," i.e., with possible positive probabilities for infinite values. To the sequence $\{\mathfrak{L}(S_{n,\nu_n})\}$ is associated another conveniently chosen sequence $\mathfrak{L}_n$ of laws of sums; if $\mathfrak{L}_n \to \mathfrak{L}$ or $\mathfrak{L}_n \equiv \mathfrak{L}$ then the ACP reduce to the CLP. The investigation uses an extension of the P. Lévy convergence theorem for ch.f.'s and the modern CLP solutions are obtained as particular cases. The case of sums of a random number of r.v.'s,[7] as well as the multidimensional case, are easily treated by the same methods [35].

Many new problems arise in ACP. The foremost corresponds to possible relaxations of the uan condition. For instance, in the case of independence, the relaxed condition

$$\max_k P\{|X_{nk} - Y_k| > \epsilon\} \to 0, \qquad \text{for every } \epsilon > 0,$$

where $Y_1, Y_2, \cdots$ are independent, does not change, essentially, the nature of the ACP. Yet, as soon as dependence is introduced, the whole outlook changes and it would be interesting to investigate various new possibilities which thus arise. On the other hand, stricter than uan conditions are of special interest when independence is not assumed. The one which seems natural is the following:

$$\max_k \sup P'\{|X_{nk}| > \epsilon\} \to 0, \qquad \text{for every } \epsilon > 0,$$

where $P'(A_{nk})$ denotes the conditional probability of the event $A_{n,k}$, given $X_{n,1}, \cdots, X_{n,k-1}$. An immediate problem is whether this or an analogous restriction enables us to find, not only sufficient, but also necessary conditions for various convergences and various cases of dependence.

II. The Strong Central Limit Problem

6. The Bernoulli case and its extension. A sequence $\{X_n\}$ such that the corresponding sequence of laws converges does not, in general, determine a r.v. X which might be considered, in some sense, as the limit of X_n. However, if we define two r.v.'s X and X' such that $P(X \neq X') = 0$ as equivalent, then, whenever $\mathfrak{L}(X_m - X_n) \to \mathfrak{L}(0)$ as $\dfrac{1}{m} + \dfrac{1}{n} \to 0$, the sequence $\{X_n\}$ determines a

[7] H. Robbins (*Bull. Am. Math. Soc.*, Vol. 54 (1948), pp. 1151–1161. studied in detail the case of independent and identically distributed X_k's with $EX_k^2 < \infty$ and ν_n, independent of X_k's, with $E_n^2 < \infty$.

unique r.v. X (up to an equivalence)—for which $P\{| X_n - X | > \epsilon\} \to 0$ for every $\epsilon > 0$. This X is the limit *in probability* of X_n.

Yet, an observed sequence of values of $\{X_n\}$ need not converge to the observed value of X. For instance, let Y be a r.v. uniformly distributed over (0, 1). Consider the sequence $\{D_n\}$ of partitions of (0, 1) into n equal subintervals and to the k-th subinterval of D_n attach the indicator $X_{n,k}$ of the event when Y falls within this subinterval. The sequence $X_{1,1}$; $X_{2,1}$, $X_{2,2}$; $X_{3,1}$, $X_{3,2}$, $X_{3,3}$; $\cdots$ converges in probability to zero since $P(X_{nk} \neq 0) = \frac{1}{n}$, for $k = 1, 2, \cdots, n$, approaches zero as $n \to \infty$. On the other hand, observed values of X_{nk}'s, for $k = 1, 2, \cdots, n$, will contain $n - 1$ zeros and a one, except in cases of total probability zero. Hence, except in these cases, any observed sequence will contain infinitely many zeros and infinitely many ones and will not converge.

The Bernoulli theorem means only that $f_n = \frac{S_n}{n}$ converges in probability to zero. Borel showed, in a fundamental memoir ([5], 1909), that Bernoulli's statement is too weak, and, in fact, that observed values of f_n converge to zero, except in cases of total probability zero. Borel's proof is based upon a direct analysis of the de Moivre-Laplace approach to NC. Thus a new domain in probability theory was opened to exploration.

First Strong Limit Theorem. *In the Bernoulli case*

$$P\{\lim_n f_n = p\} = 1.$$

This leads to the introduction in probability theory of the notion of almost sure (a.s.) convergence:

$$X_n \xrightarrow{\text{a.s.}} X \text{ if } P\{\lim_n X_n = X\} = 1,$$

or, equivalently, if for every $\epsilon > 0$,

$$P\{ | X_{n+k} - X | > \epsilon \text{ for } k = 1, \quad \text{or 2 or } \cdots \text{ ad inf.}\} \to 0 \text{ as } n \to \infty.$$

If we denote by A_n the event $| X_n - X | > \epsilon$, we see that we are concerned here with

$P = P$ (realization of infinitely many events A_n) $= \lim_{n\to\infty} \lim_{\nu\to\infty} P(A_{n+1} \cup \cdots \cup A_{n+\nu})$.[8] From Boole's inequality

$$P(A_{n+1} \cup \cdots \cup A_{n+\nu}) \leqq \sum_{k=n+1}^{n+\nu} P(A_k)$$

follows, at once, the fundamental Borel-Cantelli Lemma. *If* $\sum_n P(A_n) < \infty$ *then* $P = 0$. This lemma can be extended, using sharper inequalities (Loève [32]).

[8] Already Poincaré considered such probabilities in his investigation of "recurrence", and this, before the notion of completely additive measures was born.

Now apply the Tchebicheff-Markov inequality

$$P\{|X_n - X| > \epsilon\} \leqq \frac{E|X_n - X|^r}{\epsilon^r}, \qquad r > 0,$$

and the Cantelli criterion follows: *if for some* $r > 0$, $\sum E|X_n - X|^r < \infty$ *then* $X_n \xrightarrow{\text{a.s.}} X$.

Applying it, with $r = 4$, to the Bernoulli case, Cantelli [6] obtained an almost immediate proof of Borel's result. An even simpler proof is as follows: $\sum_n E|f_{n^2} - p|^2 < \infty$ since $E(f_n - p)^2 = \frac{pq}{n}$, hence $f_{n^2} - p \xrightarrow{\text{a.s.}} 0$. Moreover, $|f_\nu - f_{n^2}| \leqq \frac{2}{n}$ for $0 \leqq \nu - n^2 \leqq 2n$, hence $f_\nu - f_{n^2} \to 0$ in the usual sense, uniformly in ν, and the theorem is proved. This last method applies as well to sequences of dependent events $\{B_n\}$, which constitute a natural extension of the Bernoulli case. Let

$$p_1(n) = \frac{1}{n}\sum_{k=1}^{n} P(B_k), \qquad p_2(n) = \frac{1}{C_n^2}\sum_{1 \leq k < l \leq n} P(B_k B_l),$$

$\delta_n = p_2(n) - p_1^2(n)$ (in the Bernoulli case $\delta_n = 0$!). It is very easy to show that $f_n - p_1(n) \to 0$ in probability if, and only if, $\delta_n \to 0$; this extends the Bernoulli theorem. Moreover, if $n|\delta_n| \leqq C < \infty$ then $f_n - p_1(n) \xrightarrow{\text{a.s.}} 0$ (Loève [31]), and Dvoretzky [10] proved that it is enough to have $\Sigma \frac{|\delta_n|}{n} < \infty$. Thus we have a simple extension of Borel's result.

The method used by Borel, while uselessly complicated in view of the result obtained, is very powerful and, by sharpening it, *the law of the iterated logarithm* (Khintchine [18]) follows.

SECOND STRONG LIMIT THEOREM. *In the Bernoulli case*

$$P\left\{\limsup_n \frac{S_n - ES_n}{\sigma_n(2\log\log\sigma_n)^{1/2}} = 1\right\} = 1.$$

where $\sigma_n = \sigma(S_n)$.

Let us use the following terminology (P. Lévy [26]). A non-decreasing sequence $\{\phi_n\}$ of positive numbers belongs to the *lower class L*, if the probability that $S_n \leqq \phi_n$, from some n onwards, is 1, and it belongs to the upper class U if this probability is 0. The following criterion (Kolmogorov) applies: *In the Bernoulli case* $\{\phi_n\}$ *belongs to L or U, respectively, according as* $\sum_n \frac{1}{\sigma_n^2}\phi_n e^{-1/2\phi_n^2} = \infty$ *or* $< \infty$. Clearly this result contains the Khintchine's LIT.

7. The general case. The question of domains of validity of the obtained results arises immediately and thus the SCLP appears in its present form. Let $S_n = \sum_{k=1}^n X_k$ be sums of r.v.'s X_k, independent or not. Find conditions for 1° a.s. convergence of $\frac{S_n}{n}$ or, more generally [31] of $\frac{S_n}{a_n}$, $a_n \uparrow \infty$ (SLLN). 2° the law of the iterated logarithm (LIT) and, more generally, criteria for classifying sequences $\{\phi_n\}$.

The second problem, in the case of independent summands possesses almost complete solutions due, respectively, to Kolmogorov [17] and to Feller [13].

a. *If* $\sup |X_k| = o(\sigma_n/(\log\log\sigma_n)^{-1/2})$ *for* $k \leqq n$, *then LIT holds.*

b. *If* $\sup |X_k| = O(\sigma_n/(\log\log\sigma_n)^{-3/2})$ *for* $k \leqq n$, *then the criterion for the Bernoulli case continues to hold.* (Feller also gave sharper criteria).

In the case of dependent summands general results were obtained by P. Lévy [26] and for Markov chains by Doeblin [7]. The problem belongs (at present) to the domain of NC; it is complicated and pries deeply into the behavior of probabilities as $n \to \infty$. Yet, in the case of independence, the dichotomy into classes L and U is more general as shown by the following property (P. Levy [26]). *If* $\{S_n\}$ *is a sequence of consecutive sums of independent r.v.'s, and cannot be reduced by adding constants to an a.s. convergent sequence, then, for any given sequence* $\{c_n\}$ *of sure numbers,* $P(S_n > c_n$ *for an infinity of values of* $n) = 0$ *or* 1.

The SLLN problem *seems* easier. Nevertheless it is far from being solved; we don't even know necessary *and* sufficient conditions for the SLLN in the case of independent summands in terms of individual d.f's.[9] The essential tools are, besides the fundamental Borel-Cantelli lemma, 1° the truncation method together with the convergence in r-mean: $X_n \xrightarrow{r} X$ if $E|X_n - X|^r \to 0 (r > 0)$, 2° the Kronecker lemma: *If* $\sum_n x_k/a_k$ *is convergent, then* $\frac{1}{a_n}\sum_{k=1}^n x_k \to 0$ $(a_n \uparrow \infty)$. It provides a possibility of transforming problems about the SLLN into those of a.s. convergence of series of r.v.'s, at least when sufficient conditions are sought for.

In the case of independent summands one can start with the following property of series (Lévy [23]): *a.s. convergence of* $\sum_1^n X_k$ *is equivalent to convergence in probability.* (It can be shown that this property holds also for certain classes of dependent summands.) On the other hand, convergence in q.m. $(r = 2)$ entails convergence in probability. Hence, when $EX_k^2 < \infty$, taking EX_k as the origin of values of X_k, it follows that *If* $\sum_n \sigma^2(X_n) < \infty$, *then* S_n *a.s. converges.* Kolmogorov proved this result using his celebrated inequality which considerably strengthens that of Tchebicheff:

$$P\{\max_{k\leqq n} |S_n| > \epsilon\} \leqq \frac{\sigma^2(S_n)}{\epsilon^2}.$$

This inequality has been extended by P. Lévy [26], and by Loève [32] to dependent summands and conditions for a.s. convergence were deduced from it. If the EX_k^2 are not finite, the truncation method is applied. Put $X_k' = X_k$, if $|X_k| \leqq 1$ and $= 0$ if $|X_k| > 1$. Then (Khintchine-Kolmogorov) $\sum_n X_n$,

[9] A first step in this direction is due to U. V. Prokhorov, "On the strong law of large numbers" (in Russian), *Dokl. Ak. Nauk.* Vol. 69 (1949), pp. 607–610. See also a paper by K. L. Chung to appear in the *Proceedings of the Second Berkeley Symposium.*

where X_n are independent r.v.'s, is a.s. convergent if, and only if, $\sum_n P(X_n \neq X'_n)$, $\sum_n \sigma^2(X'_n)$, $\sum_n (X'_n)$ *converge.*

It is not difficult to obtain conditions for series of dependent summands. Let $q_n(t) = P\{ | X_n | > t\}$, $\xi_n = \int_{-\epsilon}^{+\epsilon} x\, dF'_n(x)$, where $F'_n(n)$ is the conditional d.f. of X_n, given $X_1, \cdots X_{n-1}$. *If* $\sum_n \int_0^\epsilon tq_n(t)\, dt < \infty$ *for an* $\epsilon > 0$, *then* $\sum_n (X_n - \xi_n)$ *a.s. converges.*

By using Kronecker's lemma the results above yield immediately sufficient conditions for the SLLN. Those which come from the last one would in turn yield without difficulty the following: *Let* $a_n \uparrow \infty$ *and* $\eta_n = \int_{-\epsilon a_n}^{+\epsilon a_n} x\, dF'_n(x)$. *If* $\sum_n q_n(a_n t) \leqq q(t)$ *and* $\int_0^\epsilon tq(t)\, dt < \infty$, *then* $\frac{1}{a_n} \sum_{k=1}^{n} (X_k - \eta_k) \to 0$.

Take now the particular case: $a_n = n^r$; and X_k's *independent and identically distributed.* From the stated result follows:

1. *If* $EX_k = m$ *exist, then* $\frac{1}{n} \sum_{k=1}^{n} X_k \xrightarrow{\text{a.s.}} m$ *and conversely* (Kolmogorov).

2. *If* $0 < r < 2$, $r \neq 1$, $E | X_k |^r < \infty$ *and* $\lim_{a \to \infty} \int_{-a}^{+a} x\, dF_n(x) = 0$, *then*

$$\frac{1}{n^{1/r}} \sum_{k=1}^{n} X_k \xrightarrow{\text{a.s.}} 0 \text{ (Marcinkiewicz).}$$

Other conditions for SLLN, in the case of dependence, are known (Lévy [27], Loève [32]).

The above result of Kolmogorov is a particular case of the celebrated ergodic theorem (Birkhoff [3]) which can be considered as a SLLN for a special case of dependence. Let A_n be an event defined on the set $\{X_{k_1}, \cdots X_{k_n}\}$ and let $A_n^{(m)}$ be an event defined in the same manner on the translated set $\{X_{k_1+m}, \cdots, X_{k_n+m}\}$. The sequence $\{X_k\}$ is called *stationary* if $P(A_n^{(m)}) = P(A_n)$ for every finite set $\{k_1, \cdots, k_n\}$ and every finite m. The ergodic theorem states that *If the sequence* $\{X_k\}$ *is stationary and* $E\, |X_k| < \infty$, *then* $\frac{1}{n} \sum_{k=1}^{n} X_k$ *converges a.s.*[10]

However an unsatisfactory feature of Birkhoff's theorem (and of its extensions) is that the conditions are not asymptotic—they have to be satisfied for every n and not for $n \to \infty$—while the conclusion is an asymptotic one. Let us only mention that more satisfactory ones, at least from this point of view, which contain the previous ones, can be found.

[10] For about fifteen years Khintchine, Kolmogorov, Wiener, Yosida and Kakutani, F. Riesz, worked to simplify the proof of this theorem. It is only lately that its domain of validity has been extended by Hurewicz, by Halmos, and by Dunford and Miller. See also a forthcoming paper by the author in the *Proceedings of the Second Berkeley Symposium.*

The bird's-eye view above of the SCLP shows that this problem is only in a tentative stage, perhaps because no adequately powerful methods or no adequately general approach to the problem had been found until now.

REFERENCES

[1] G. BAWLY, Über einige Verallgemeinerungen der Grenzwertsätze der Wahrscheinlichkeitsrechnung," *Rec. Math. Moscow*, Vol. 43 (1936), pp. 917–929.

[2] J. BERNOULLI, *Ars conjectandi*, 1713.

[3] S. BERNSTEIN, "Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes," *Math. Ann.*, Vol. 97 (1927), pp. 1–59.

[4] G. D. BIRKHOFF, "Proof of the ergodic theorem," *Proc. Nat. Acad. Sc.* Vol. 17 (1931), pp. 650–665.

[5] E. BOREL, "Sur les probabilités dénombrables et leurs applications arithmétiques," *Circ. Mat. d. Palermo*, Vol. 26 (1909), pp. 247–271.

[6] F. P. CANTELLI, "Sulla probabilita come limite della frequenza," *Rend. Accad. d. Lincei*, Vol. 26 (1917), pp. 39–45.

[7] A. DE MOIVRE, *Miscellanea analytica*, 1730.

[8] W. DOEBLIN, "Sur les propriétés asymptotiques . . . de chaînes simples," *Bull. Math. Soc. Roum. Sc.* (1937), pp. 1–120.

[9] W. DOEBLIN, "Sur les sommes d'un grand nombre de variables aléatoires indépendantes," *Bull. Sc. Math.*, Vol. 63 (1939), pp. 23–64.

[10] A. DVORETZKY, "On the strong stability of a sequence of events, *Annals of Math. Stat.*, Vol. 20 (1949), pp. 296–299.

[11] W. FELLER, "Ueber den Zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung," *Math. Zeit.*, Vol. 40 (1935), pp. 521–539; Vol. 42 (1937), pp. 301–312.

[12] W. FELLER, "Ueber das Gesetz der grossen Zahlen," *Acta. Litt. Ac. Scient.* Vol. 8 (1937), pp. 191–201.

[13] W. FELLER, "The general form of the so-called law of the iterated logarithm," *Trans. Am. Math. Soc.*, Vol. 54 (1943), pp. 373–402.

[14] B. GNEDENKO, "On convergence of distribution laws of independent random variables" (in Russian), *Dokl. Ak. Nauk*, Vol. 18 (1938), pp. 231–234.

[15] B. GNEDENKO, "On the theory of limit theorems for sums of independent random variables" (in Russian), *Bull. Ac. Sci. URSS* (1939), pp. 181–232; pp. 643–647.

[16] A. KOLMOGOROFF, "Ueber die Summen durch den Zufall bestimmter unabhängiger Grössen," *Math. Ann.*, Vol. 99 (1928), pp. 309–319; 102 (1929), pp. 484–489.

[17] A. KOLMOGOROFF, "Ueber das Gesetz des iterierten Logarithms," *Math. Ann.*, Vol. 101 (1929), pp. 126–135.

[18] A. KOLMOGOROFF, "Sulla forma generale di un processo stocastico omogeneo," *Rend. Accad. d. Lincei*, Vol. 15 (1932), pp. 805–806; pp. 866–869.

[19] A. KHINTCHINE, "Ueber einen Satz der Wahrscheinlichkeitsrechnung," *Fund. Math.*, Vol. 6 (1924), pp. 9–20.

[20] A. KHINTCHINE, "Zur Theorie der unbeschrankt teilbaren Verteilungsgesetze," *Rec. Math. Moscow*, Vol. 44 (1937), pp. 79–119.

[21] P. S. LAPLACE, *Traité des Probabilities*, 1812.

[22] P. LÉVY, *Calcul des Probabilités*, Gauthier-Villars, 1925.

[23] P. LÉVY, "Sur les séries dont les termes sont des variables éventuelles indépendantes," *Studia Math.* Vol. 3 (1931), pp. 117–155.

[24] P. LÉVY, "Sur les integrales dont les éléments sont des variables aléatoires indépendantes," *Ann. d. Scuola Nor. di Pisa*, B (1934), pp. 337–366.

[25] P. LÉVY, "Propriétés asymptotiques des sommes de variables aléatoires indépendantes on enchaînées," *Jour. Math. Pures Appl.*, Vol. 14 (1935), pp. 347–402.

[26] P. Lévy, "La loi forte des grands nombres pour les variables aléatoires enchaînées," *Jour. Math. Pures Appl.*, Vol. 15 (1936), pp. 11-24.

[27] P. Lévy, *Théorie de l'addition des variables aléatoires*, Gauthier-Villers, 1937.

[28] A. Liapounoff, "Sur une proposition de la théorie des probabilités," *Bull. Acad. Sci. St. Petersbourg*, (5); Vol. 13 (1900), pp. 359-386.

[29] A. Liapounoff, "Nouvelle forme du théorème sur la limite de probabilités," *Mem. Ac. St. Petersbourg*, (8): Vol. 12 (1901).

[30] J. Lindeberg, "Eine nene Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung," Vol. 15 (1922), pp. 211-225.

[31] M. Loève, "La tendance centrale pour des variables aléatoires liées," *C. R. Ac. Sc. Paris*, Vol. 212 (1940).

[32] M. Loève, "Etude ásymptotique des sommes de variables aléatoires liées," *Jour. Math. Pures Appl.*, Vol. 24 (1945), pp. 249-318.

[33] M. Loève, "Sur l'équivalence asymptotique des lois," *C. R. Ac. Sc. Paris*, Vol. 227 (1948), pp. 1335-1337.

[34] M. Loève, "On the Central Probability problem," *Proc. Nat. Acad. Sci.*, Vol. 35 (1949), pp. 328-332.

[35] M. Loève, "On sets of Probability laws and their limit elements," *Statistical Series*, VoL 1, pp. 53-88, University of California Press, Berkeley, California.

[36] J. Marcinkiewicz, "Sur les fonctions indépendantes," *Fund. Math.*, Vol. 30 (1938), pp. 349-364.

[37] A. Markoff, *Calculus of probabilities* (Russian), 1913.

[38] Poisson, *Recherches sur la Probabilité des Jugements* 1837, pp. 205-207.

[39] D. Raïkoff, "On connection between the central limit theorem of probability theory and the law of large numbers," *Bull. Ac. Sci. URSS*, Ser. Math., (1938), pp. 323-338.

[40] P. L. Tchebicheff, *Oeuvres*, Vol. I, pp. 687-694; Vol. II, pp. 481-492.