ON ALMOST SURE CONVERGENCE

MICHEL LOÈVE
UNIVERSITY OF CALIFORNIA

1. Introduction

Since the discovery by Borel[1] (1907) of the strong law of large numbers in the Bernoulli case, there has been much investigation of the problem of almost sure convergence and almost sure summability of series of random variables. So far most of the results concern series of independent random variables. In the case of dependent random variables, the first general result is the celebrated Birkhoff ergodic theorem [1], or the strong law of large numbers for a stationary sequence with a finite first moment. This theorem contains the Kolmogoroff strong law of large numbers as a particular case. P. Lévy studied series that are the same as those of independent random variables in their properties of second order as described by the first and second conditional moments. The author [8] investigated series which behave asymptotically as those of P. Lévy. There are also properties of martingales, due essentially to Doob [2], P. Lévy and Ville [10].

We shall proceed to a systematic investigation of almost sure convergence of sequences of random variables, emphasizing the methods and assuming as little as possible about the stochastic structure of the sequences. The known results will appear as various particular cases of a few propositions, and the necessary known tools will be established. In that respect, this paper is self contained.

Part I is devoted to definitions, notations, and general criteria. In part II the truncation and centering ideas are expounded. Part III is concerned with the use of "determining" set functions and with two propositions. The particular cases of one of them contain the martingale properties and of the other the ergodic theorem and its known extensions.

PART I. BASIC CONCEPTS AND ELEMENTARY INEQUALITIES

2. Fundamental notions

Let (F, P) represent a *probability field* defined on a space Ω of "points" ω.

F is a σ-field of *events* A in Ω, that is a family of sets in Ω such that (i) the *sure event* Ω belongs to F, (ii) if $A \in F$, then $A' = (\Omega - A) \in F$, (iii) if $A_n \in F$ for $n = 1, 2, \ldots$, then $\bigcap_n A_n \in F$. It follows then, from $(\bigcup_n A_n)' = \bigcap_n A_n'$, that (iii)′ if $A_n \in F$ for $n = 1, 2, \ldots$, then $\bigcup_n A_n \in F$ and, in general, if a sequence of events has a limit, the limit is also an event. When the events are disjoint we shall repre-

This work was supported in part by the Office of Naval Research.

[1] References, denoted by numbers in square brackets, are listed at the end of the paper except for those of the known results which can be found in P. Lévy, *Théorie de l'addition des variables aléatoires*, Gauthier-Villars (1937).

sent their union by $\sum$ instead of $\cup$. A basic role is played throughout this paper by the familiar transformation of $\cup$ into $\sum$ given by $A_1 + A_1'A_2 + \ldots$, to be written

(1) $$\bigcup_n A_n = \sum_n A_1'A_2' \ldots A_{n-1}'A_n .$$

P is a *probability measure* on F, that is (i) $P\Omega = 1$, (ii) $PA \geqq 0$ if $A \in F$, and (iii) $P\left(\sum_n A_n\right) = \sum_n PA_n$ if $A_n \in F$ for all n. It follows that P is a continuous function of events, that is $P(\lim_n A_n) = \lim_n P(A_n)$ if the sequence of events A_n has a limit.

Real *random variables* (r.v.) on (F, P) will be represented by X, Y, Z, U, V, W, with or without subscripts. A r.v., say U, is a real, single valued, finite function of $\omega \in \Omega$ such that for all real u, the sets of the form $[U < a]$ of the ω's are events in Ω. In general, the inverse image by a countable set of r.v. $\{U_\lambda\}$ where $\lambda \in \Lambda$, of the Borel sets on the range space R_Λ is to be a σ-field $F\{U_\lambda\}$ contained in F, or the σ-field *defined by* $\{U_\lambda\}$. Given a sequence $\{U_n\}$, we shall have to consider later the σ-field $F\{U_m, \ldots, U_n\}$ of events defined on the segment $\{U_m, \ldots, U_n\}$, the *tail* σ-field $F\{TU_n\} = \bigcap_n F\{U_n, U_{n+1}, \ldots\}$ of events defined on the tail of $\{U_n\}$ and the field $F(U_1, U_2, \ldots) = \bigcup_n F\{U_1, \ldots, U_n\}$ of events defined on a finite number of U's. $F(U_1, U_2, \ldots)$ is not, in general, a σ-field and $F\{U_n\}$ is the smallest σ-field containing it. Two r.v. X and Y are *equivalent*, or almost surely (a.s.) equal, if $P[X \neq Y] = 0$. Since equivalence is symmetric, reflexive and transitive, this relationship will be written $X = Y$.

Throughout this paper, unless otherwise stated, a passage to the limit will be for $n \to \infty$, ϵ will be a positive number and the integer k (or l) will run from m to n $(m \leqq k \leqq n)$.

3. Almost sure convergence

Given a sequence $\{U_n\}$, the events

$$A_{mn}(\epsilon) = [\sup_{k,l} |U_k - U_l| \leqq \epsilon]$$

clearly have the property that

(2) $$A_{mn}(\epsilon) \subset A_{m'n'}(\epsilon') \text{ for } m \leqq m' < n' \leqq n \text{ and } \epsilon \leqq \epsilon' .$$

Accordingly, the limits

$$C(\epsilon) = \lim_m \lim_n A_{mn}(\epsilon) \text{ and } C = \lim_{\epsilon \to 0} C(\epsilon)$$

exist (*and are tail events*). Moreover,

(3) $$C \subset C(\epsilon) \subset C(\epsilon') \quad \text{for} \quad \epsilon \leqq \epsilon' .$$

By the Cauchy mutual convergence criterion C is the event $[U_n$ converges], consequently $P[U_n$ converges] is well defined (Kolmogoroff) and, when this probability is one, the sequence $\{U_n\}$ is *almost surely* (*a.s.*) *convergent*. (Clearly the limit r.v. is determined up to an equivalence.) The same event C is obtained if we use

$$B_{mn}(\epsilon) = [\sup_{k,l} | U_k - U_l | \geq \epsilon],$$

since

$$(4) \qquad A_{mn}(\epsilon) \subset B_{mn}(\epsilon) \subset A_{mn}(2\epsilon).$$

If we start with

$$C_{mn}(\epsilon) = [\sup_k | U_k - U | \geq \epsilon],$$

C becomes $[U_n \to U]$. In any case, the relation (3) holds and we have the well known

LEMMA 3.1. *U_n converges a.s. if, and only if, $P[C(\epsilon)] = 1$ for every ϵ.*

$P[C(\epsilon)]$ is the probability of occurrence of an infinity of events $[|U_i - U_j| > \epsilon]$ defined on $(U_1, U_2, \ldots)$ or of events $[|U_n - U| > \epsilon]$. In general, given a sequence $\{A_n\}$ of events

$$(5) \qquad P_\infty = P[\infty \text{ of } A_n\text{'s occurs}] = \lim_m \lim_n P(\bigcup_k A_k).$$

From the inequality (Boole)

$$(6) \qquad P(\bigcup_k A_k) = P\left(\sum_k A'_m \ldots A'_{k-1} A_k\right) \leqq \sum_k PA_k,$$

there follows (Borel-Cantelli)

LEMMA 3.2. *If $\sum_n PA_n < \infty$, then $P_\infty = 0$.*

This with lemma 3.1 gives at once (Cantelli)

LEMMA 3.3. *If, for every $\epsilon > 0$, $\sum_n P[|U_n - U| > \epsilon] < \infty$, then $U_n \xrightarrow{\text{a.s.}} U$.*

And, from

$$(7) \qquad P[\bigcup_{n=m}^{\infty} (|U_n - U_m| > \epsilon)] \leqq \sum_{n=m}^{\infty} P[|U_n - U_m| > \epsilon],$$

there follows

LEMMA 3.3′. *If, for every $\epsilon > 0$,*

$$\liminf_m \sum_{n=m}^{\infty} P[|U_n - U_m| > \epsilon] = 0,$$

then U_n converges a.s.

Also, the event $B_m = \bigcap_{n=m}^{\infty} [|U_{n+1} - U_n| \leqq \epsilon_n]$, where $\sum_n \epsilon_n < \infty$, entails

$$(8) \qquad |U_n - U_m| \leqq \sum_{k=m}^{n} |U_{k+1} - U_k| \leqq \sum_{k=m}^{n} \epsilon_k \to 0 \text{ as } \frac{1}{m} + \frac{1}{n} \to 0,$$

that is, $B_m \subset [U_n \text{ converges}]$. Hence applying Boole's inequality,

$$(9) \quad P[U_n \text{ converges}] \geqq PB_m = 1 - PB'_m \geqq 1 - \sum_{k=m}^{\infty} P[|U_{n+1} - U_n| > \epsilon_n].$$

Thus,

LEMMA 3.3″. *If* $\sum_n P[|U_{n+1} - U_n| > \epsilon_n] < \infty$, *where* $\sum_n \epsilon_n < \infty$, *then* U_n *converges a.s.*

By the Tchebicheff-Markoff inequality,

$$P[|U_i - U_j| > \epsilon] \leqq \frac{1}{\epsilon^r} E|U_i - U_j|^r, \qquad r > 0. \tag{10}$$

Lemmas 3.3, 3.3′ and 3.3″ yield

LEMMA 3.4. *If, for an* $r > 0$, $\sum_n E|U_n - U|^r < \infty$, *then* $U_n \xrightarrow{\text{a.s.}} U$. (Cantelli).

LEMMA 3.4′. *If, for an* $r > 0$, $\liminf_m \sum_{n=m}^{\infty} E|U_n - U_m|^r = 0$, *then* U_n *converges a.s.*

LEMMA 3.4″. *If, for an* $r > 0$ *and* $\sum_n \epsilon_n < \infty$, $\sum_n \frac{1}{\epsilon_n^r} E|U_{n+1} - U_n|^r < \infty$, *then* U_n *converges a.s.*

Thus Boole's inequality provides us with relatively simple sufficient conditions for a.s. convergence. Relatively simple necessary conditions can be obtained as follows. Let

$$p_{mk} = P(A_k; A'_m, \ldots, A'_{k-1}), \qquad p_{mm} = PA_m.$$

We have

$$P(\bigcup_k A_k) = 1 - P(\bigcap_k A'_k) = 1 - P(A'_m)P(A'_{m+1}; A'_m) \ldots \times P(A'_n; A'_m, \ldots, A'_{n-1}), \tag{11}$$

hence

$$P(\bigcup_k A_k) = 1 - \prod_k (1 - p_{mk}) \geqq 1 - e^{-\sum_k p_{mk}}, \tag{12}$$

or

$$0 \leqq \sum_k p_{mk} \leqq |\log[1 - P(\bigcup_k A_k)]|. \tag{13}$$

Letting $n \to \infty$, then $m \to \infty$, there follows

LEMMA 3.5. *If* $P_\infty = 0$, *then* $\lim_m \sum_{n=m}^{\infty} p_{mn} = 0$.

In the particular case of independent events A_n, $p_{mn} = PA_m$ and the lemma above becomes a converse to lemma 3.2. Moreover, the relation (12) becomes

$$P(\bigcup_k A_k) = 1 - \prod_k (1 - PA_k) \tag{14}$$

and, if $\sum_n PA = \infty$, the product on the right hand side diverges to zero as $n \to \infty$, hence $P_\infty = 1$. Taking this with lemma 3.2, we have the first obtained law of 0 or 1 (Borel). Namely,

LEMMA 3.5′. *If the events* A_n *are independent, then* $P = 0$ *or* $= 1$, *according as* $\sum_n PA < \infty$ *or* $= \infty$.

Applied to the events $[|U_n - U| > \epsilon]$, this becomes

LEMMA 3.5″. *If the r.v.* $U_n - U$ *are independent, then* $U_n - U$ *converges or diverges a.s. to zero, according as* $\sum_n P[|U_n - U| > \epsilon] < \infty$ *or* $\sum_n P[|U_n - U| > \epsilon] = \infty$.

Sharpened inequalities. At present, Boole's inequality is basic in the search for conditions for a.s. convergence in terms of expectations. It is also very simple in that it uses only probabilities of the events taken one by one. However, for this very reason, it is a rough tool since it does not take into account the structure of the *set* of events. In fact, the domain of validity of theorem A (next section) and its particular cases can be extended, and it would be interesting to do so, using sharper tools such as the following ones.

Let

$$C_{n-m}^{r} p_{mn}(r) = \sum_{k_1, \ldots, k_r} P(A_{k_1} \cup \ldots \cup A_{k_r})$$

and

$$p^*(r) = \liminf_m \liminf_n \frac{n}{r} p_{mn}(r), \qquad p_*(r) = \limsup_m \limsup_n p_{mn}(r).$$

We have, Loève [8],

$$(15) \quad p_*(1) \leqq \ldots \leqq p_*(r-1) \leqq p_*(r) \leqq \ldots \leqq p_\infty \leqq p^*(r) \leqq p^*(r-1) \leqq \ldots \leqq p^*(1).$$

Thus two scales of conditions for P_∞ vanishing are at our disposal, one provides us with necessary and the other with sufficient conditions. As r increases, they become sharper. The Borel-Cantelli lemma corresponds to $p^*(1) = 0$ and is the roughest, that is, we may have $p^*(1) > 0$ while $p^*(r) = 0$ for some $r > 0$ but the converse is not necessarily true. It would be of some interest to examine the structure of sets $\{A_n\}$ for which the condition $p^*(r) = 0$ for a given r is not only sufficient but also necessary; Borel's law of 0 or 1 answers this question for $r = 1$.

A.s. absolute convergence. The study of a.s. absolute convergence of series of r.v. or, which comes to the same, of monotonic sequences of r.v. is particularly simple. Let $U_n = \sum_{i=1}^{n} |X_i|$. U_n does not decrease as $n \rightarrow \infty$ and, consequently, has always a limit. However, this limit can be infinite with positive probability. Yet if the law of U_n converges, then such a possibility is excluded and conversely and, changing if necessary $\lim_n U_n$ on the event $[U_n = +\infty]$ of probability zero, this a.s. limit is a r.v. Thus

LEMMA 3.6. *A series of r.v. converges absolutely a.s. if, and only if, its law converges.*

The possibility above is also excluded if, for an $r > 0$, $\lim_n E|U_n|^r < \infty$. Thus

LEMMA 3.6′. *If, for an* $r > 0$, $E\left(\sum_n |X_n|\right)^r < \infty$ *then* $\sum_n |X_n| < \infty$ *a.s.*

Moreover, we have

(16) $$E^{r'}\sum_{i=1}^{n}|X_i|^r \leqq \sum_{i=1}^{n} E^{r'}|X_i|^r,$$

where $r' = 1$ for $0 < r \leqq 1$ and $r' = \frac{1}{r}$ for $r > 1$, and by lemma 3.6′, there follows

LEMMA 3.6″. *If, for an $r > 0$, $\sum_n E^{r'}|X_n|^r < \infty$, then $\sum_n |X_n| < \infty$ a.s.*

In particular, if X_n is an indicator IA_n, that is, $= 1$ if the event A occurs and $= 0$ otherwise, then $\sum_n IA_n$ equals the number of occurrences of the A_n and lemma 3.6″ reduces to the Borel-Cantelli lemma 3.2.

PART II. TRUNCATION AND CENTERING

4. Equivalent sequences

The primary purpose of truncation (introduced by Markoff for the study of normal convergence and fully exploited by Khintchine, Kolmogoroff and P. Lévy) is to replace a sequence of r.v. by a sequence of *bounded* r.v. equivalent to the original in the properties that interest us. For the investigation of a.s. convergence these equivalences can be of two kinds, one preserves all the limit properties and the other only the convergence. $\{U_n\}$ is *tail equivalent* to $\{U'_n\}$ if $\{U_n\}$ and $\{U'_n\}$ differ only by a finite number of terms. $A_n = [U_n \neq U'_n]$, the definition corresponds to $P_\infty = 0$. Similarly, writing $U_n = \sum_{i=1}^{n} X_i$, the series $\sum_n X_n$ and $\sum_n X'_n$ are *convergence equivalent* if they differ only by a finite number of terms, that is, writing $A_n = [X_n \neq X'_n]$, if $P_\infty = 0$.

The relation (15) provides us with scales of necessary and of sufficient conditions for these equivalences. In particular, lemma 3.2 ($r = 1$) yields

LEMMA 4.1. (i) *If $\sum_n P[U_n \neq U'_n] < \infty$, then $\{U_n\}$ and $\{U'_n\}$ are tail equivalent.*

(ii) *If $\sum_n P[X_n \neq X'_n] < \infty$, then $\sum_n X_n$ and $\sum_n X'_n$ are convergence equivalent* (Khintchine).

Now, given $\{U_n\}$, we can always find a sufficiently fast increasing sequence $\{L_n\}$ such that $\sum_n P[|U_n| > L_n] < \infty$. Taking U_n *truncated at L_n*, that is $U'_n = U_n$ if $|U_n| \leqq L_n$ and $= 0$ otherwise, the sequence $\{U'_n\}$ is tail equivalent to $\{U_n\}$. Similarly, with $\sum_n P[|X_n| > L_n] < \infty$, the series $\sum_n X_n$ and $\sum_n X'_n$ of the X_n truncated at L_n are convergence equivalent.

One of the purposes of truncation is to introduce moments. In general, one has at the same time to conveniently *center* r.v.: a r.v. X is centered at ξ if it is replaced by $X - \xi$. One of the purposes of centering is to avoid the shift of probability spreads toward infinite values.

5. Series of expectations and a.s. convergence

Let $p_n(t) = P[|X_n| < t]$, $q_n(t) = P[|X_n| \geqq t]$ and $p_n'(t)$, $q_n'(t)$ be the conditional probabilities of the corresponding events, given $X_1, \ldots, X_{n-1}$. E' will be the symbol of the conditional expectation of a r.v. (truncated or not) in a sequence, given the preceding terms and $\xi_n = E'X_n$, $\xi_n' = E'X_n'$, where X_n' is X_n truncated at L_n. We shall denote by $f_n(t)$ functions of $t \geqq 0$ such that $f_n(t) \geqq ct^2$ for $t < L_n$ and $\geqq c_1 > 0$ for $t \geqq L_n > 0$. First, we shall give the following extension due to P. Lévy of a Kolmogoroff inequality, for further extensions, see Loève [8]. Let $X_i = U_i - U_{i-1}$, $i = 1, 2, \ldots \nu$.

LEMMA 5.1. *If* $E'X_i \equiv 0$, *then* $P[\sup_i |U_i| > \epsilon] \leqq \frac{1}{\epsilon^2} EU_\nu^2$.

Let

$$A_i = [|U_i| > \epsilon] \quad \text{and} \quad B_i = A_1' \ldots A_{i-1}'A_i\,.$$

Then, by (1), section 2,

$$A = \bigcup_i A_i = [\sup_i |U_i| > \epsilon] = \sum_i B_i\,. \tag{17}$$

The hypothesis yields

$$E(U_\nu^2; B_i) = E(U_i^2; B_i) + E\{(U_\nu - U_i)^2; B_i\} \geqq \epsilon^2\,. \tag{18}$$

Consequently, the inequality follows from

$$EU_\nu^2 \geqq \sum_i PB_i E(U_\nu^2; B_i) \geqq \epsilon^2 \sum_i PB_i = \epsilon^2 PA\,. \tag{19}$$

LEMMA 5.2. *If* (i) $\sum_n q_n(L_n) < \infty$ *and* (ii) $\sum_n \int_0^{L_n} f_n(t)dp_n(t) < \infty$, *then* $\sum_n (X_n - \xi_n')$ *converges a.s.*

From (i) and lemma 4.1 above it follows that $\sum_n (X_n - \xi_n')$ and $\sum_n (X_n' - \xi_n')$ are convergence equivalent and we have only to prove that the last series is a.s. convergent to establish lemma 5.2. By definition of ξ_n' we have $E'(X_n - \xi_n') = 0$; consequently

$$\sigma^2(X_n') = E(X_n' - \xi_n')^2 = EE'(X_n'^2 - \xi_n'^2) \leqq EE'X_n'^2 = EX'^2 = \int_0^{L_n} t^2 dp_n(t)\,. \tag{20}$$

From the hypothesis (ii) and the first property of $f_n(t)$, there follows then that

$$\sum_n \sigma^2(X_n') \leqq \sum_n \int_0^{L_n} t^2 dp_n(t) \leqq \frac{1}{c} \sum_n \int_0^{L_n} f_n(t)\, dp_n(t) < \infty\,. \tag{21}$$

But, for $m \leqq k \leqq n$, we have applying lemma 5.1

$$P\left[\sup_k \left|\sum_{i=1}^{k} (X_i' - \xi_i')\right| > \epsilon\right] \leqq \frac{1}{\epsilon^2} \sum_k \sigma^2(X_k')\,. \tag{22}$$

Consequently, by lemma 3.1, the series $\sum_n (X'_n - \xi'_n)$ converges a.s. and the proof is completed.

From the lemma follows

LEMMA 5.2′. *If* $\sum_n Ef_n(|X_n|) < \infty$, *then* $\sum_n (X_n - \xi'_n)$ *converges a.s.*

For

$$\sum_n \int_0^{L_n} f_n(t)\,dp_n(t) \leqq \sum_n Ef_n(|X_n|) < \infty \tag{23}$$

and

$$\sum_n q_n(L_n) \leqq \frac{1}{c'}\int_{L_n}^{\infty} f_n(t)\,dp_n(t) \leqq \frac{1}{c'}\sum_n Ef_n(|X_n|) < \infty . \tag{24}$$

Assuming that $f_n(t)$ are also *continuous and nondecreasing for* $t \leqq L_n$ *and* $= 0$ *at* $t = 0$, lemma 5.2 yields, by integration by parts,

LEMMA 5.2″. *If* $\sum_n \int_0^{L_n} q_n(t)df_n(t) < \infty$, *then* $\sum_n (X_n - \xi'_n)$ *converges a.s.*

Let a_n be positive constants and $\eta_n = E'X''_n$, where X''_n is X_n truncated at a_nL_n. The two lemmas above, together with the Kronecker lemma that, if $a_n \uparrow \infty$ and $\sum_n \frac{X_n}{a_n}$ converges, then $\frac{1}{a_n}\sum_{i=1}^n X_i \to 0$, yield at once the first part of

THEOREM A. *If* (i) $\sum_n Ef_n(|X_n|/a_n) < \infty$ *or* (ii) $\sum_n \int_0^{L_n} q_n(a_nt)df_n(t) < \infty$, *then* $\sum_n \frac{X_n - \eta_n}{a_n}$ *converges a.s. and, when* $a_n \uparrow \infty$, $\frac{1}{a_n}\sum_{i=1}^n (X_i - \eta_i) \xrightarrow{\text{a.s.}} 0$.

Moreover, under (i), η_n *can be replaced by zero or* ξ_n *according as* (iii′) $f_n(t)/t \geqq c''$ *for* $t \leqq L_n$ *or* (iii″) $f_n(t)/t \geqq c''$ *for* $t \geqq L_n$.

Let $\sum'$ and $\sum''$ denote summations over subscripts n for which (iii′) and (iii″) hold, respectively. From

$$\begin{aligned}
\sum{}' \frac{E|\eta_n|}{a_n} + \sum{}'' \frac{E|\xi_n - \eta_n|}{a_n}
&= \sum{}' \frac{1}{a_n}\int_0^{a_nL_n} t\,dp_n(t) + \sum{}'' \frac{1}{a_n}\int_{a_nL_n}^{\infty} t\,dp_n(t) \\
&\leqq \frac{1}{c''}\sum{}' \int_0^{a_nL_n} \frac{f_n(t)}{a_n}\,dp_n(t) + \frac{1}{c''}\sum{}'' \int_{a_nL_n}^{\infty} \frac{f_n(t)}{a_n}\,dp_n(t) \\
&\leqq \frac{1}{c''}\sum_n Ef_n\left(\frac{|X_n|}{a_n}\right) \\
&< \infty
\end{aligned} \tag{25}$$

follows by lemma 3.6″ that the series

$$\sum{}' \frac{\eta_n}{a_n} + \sum{}'' \frac{\xi_n - \eta_n}{a_n}$$

is a.s. absolutely convergent and this entails the second part of the theorem.

Particular cases. 1°. Let $f_n(t) = t^{r_n}$ with $0 < r_n \leqq 2$; the conditions imposed upon the functions f_n are satisfied and the criterion (i), for instance, yields

If $a_n \uparrow \infty$ *and* $\sum_n E|X_n|^{r_n}/a_n^{r_n} < \infty$, *where* $0 < r_n \leqq 2$, *then* $\frac{1}{a_n}\sum_{i=1}^n X_i \xrightarrow{\text{a.s.}} 0$ *or* $\frac{1}{a_n}\sum_{i=1}^n (X_i - \xi_i) \xrightarrow{\text{a.s.}} 0$ *according as, for almost all* n, $r_n \leqq 1$ *or* $r_n \geqq 1$.

For $r_n \equiv 2$, $a_n = n$, X_n independent and $EX_n \equiv 0$, this becomes the Kolmogoroff criterion for the SLLN.

2°. Let $f_n(t) = t^2$ in $[0, L]$. The criterion (ii) becomes $\sum_n \int_0^L tq_n(a_n t)dt < \infty$. When, moreover, $q_n(t) \leqq q(t) = P[|X| \geqq t]$, it becomes $\int_0^L t\left[\sum_n q(a_n t)\right]dt < \infty$. To specialize further we shall establish

LEMMA 5.3. (i) $E|X|^r < \infty$ *for an* $r > 0$ *if, and only if* (ii) $t^r \sum_n q(n^{1/r}t) \leqq c < \infty$ *where* c *is independent of* t; *and the last relation holds for* $t > 0$ *if it holds for* $t = 1$.

In fact,

$$E|X|^r = -\int_0^\infty u^r dq(u) = -\sum_n \int_{(n-1)^{1/r}t}^{n^{1/r}t} u^r dq(u) \tag{26}$$

and

$$\begin{aligned}(n-1)t^r[q(n^{1/r}t) - q\{(n-1)^{1/r}t\}] &\leqq \int_{(n-1)^{1/r}t}^{n^{1/r}t} u^r dq(u) \\ &\leqq nt^r[q(n^{1/r}t) - q\{(n-1)^{1/r}t\}].\end{aligned} \tag{27}$$

Summing over n in (26) and rearranging the terms, the first part of the lemma follows. Now, (26) and (27) hold with $t = 1$ and the last part follows at once.

We take now $a_n = n^{1/r}$. By lemma 5.3, the criterion becomes $\int_0^L \frac{dt}{t^{r-1}}\,dt < \infty$, that is $r < 2$. On the other hand, if $r = 1$ and $\frac{1}{n}\sum_{i=1}^n \eta_i \xrightarrow{\text{a.s.}} \eta$, then the conclusion becomes $\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{\text{a.s.}} \eta$.

If $r < 1$, then $\sum_n (\eta_n/n^{1/r})$ converges a.s. because lemma 3.6″ applies:

$$\begin{aligned}\sum_n \frac{E|\eta_n|}{n^{1/r}} = -\sum_n \int_0^L t\,dq(n^{1/r}t) &\leqq \int_0^L \left[\sum_n q(n^{1/r}t)\right]dt \\ &\leqq c\int_0^L \frac{dt}{t^r} < \infty\end{aligned} \tag{28}$$

and we can suppress the η_n in the conclusion.

If $r > 1$, then $\int_0^L \frac{dt}{t^r} < \infty$ and, applying the same lemma 3.6″, we can replace

η_n by ξ_n in the conclusion. Thus

If $q_n(t) < P[|X| \geqq t]$ *and* $E|X|^r < \infty$, *then*

(i) $$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{a.s.}} \eta\,, \quad \text{if } \frac{1}{n}\sum_{i=1}^{n} \eta_i \xrightarrow{\text{a.s.}} \eta\,,$$

(ii) $$\sum_n \frac{X_n}{n^{1/r}} \text{ converges a.s. and } \frac{1}{n^{1/r}}\sum_{i=1}^{n} X_k \xrightarrow{\text{a.s.}} 0\,,$$

if $0 < r < 1$,

(iii) $$\sum_n \frac{(X_n - \xi_n)}{n^{1/r}} \text{ converges a.s. and } \frac{1}{n^{1/r}}\sum_{i=1}^{n} (X_i - \xi_i) \xrightarrow{\text{a.s.}} 0\,,$$

if $1 < r < 2$.

Let $r = 1$ and $q_n'(t) \leqq q(t)$ a.s. for $t \geqq n$. Then a.s.

(29) $$|\xi_n - \eta_n| \leqq q(n) - \int_0^{\infty} t\,dq(t) \to 0\,,$$

hence, *a fortiori*,

(30) $$\frac{1}{n}\sum_{i=1}^{n} (X_i - \xi_i) \xrightarrow{\text{a.s.}} 0$$

and, taking $\xi_n \equiv 0$, we get a P. Lévy SLLN.

Now, let the X_n be independent and identically distributed r.v. Then, (i) becomes the Kolmogoroff SLLN and (ii) and (iii) a Marcinkiewicz [9] result. Moreover, the converse is true, that is,

If the X_n are independent and identically distributed r.v., then $E|X|^r < \infty$ *if*

(i) $$r = 1 \text{ and } \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{a.s.}} 0 \text{ (Kolmogoroff)},$$

(ii) $$0 < r < 1 \text{ and } \frac{1}{n^{1/r}}\sum_{i=1}^{n} X_i \xrightarrow{\text{a.s.}} 0\,,$$

(iii) $$1 < r < 2 \text{ and } \frac{1}{n^{1/r}}\sum_{i=1}^{n} (X_i - EX_i) \xrightarrow{\text{a.s.}} 0\,.$$

In fact, in case (ii) for instance, there follows that $X_n/n^{1/r} \xrightarrow{\text{a.s.}} 0$ and by lemma 3.5″,

(31) $$\sum_n P[|X_n| > \epsilon n^{1/r}] = \sum_n q(\epsilon n^{1/r}) < \infty\,.$$

Hence, by lemma 5.3, $E|X|^r < \infty$.

Convergent subsequences and truncation. Given a sequence $\{U_n\}$, *if* (i) *there exists* $U_{n_i} \xrightarrow{\text{a.s.}} U$ *as* $n_i \to \infty$ *and* (ii) *for* $n_i \leqq k < n_{i+1}$, $V_i = \sup_k |U_k - U_{n_i}| \xrightarrow{\text{a.s.}} 0$, *then* $U_n \xrightarrow{\text{a.s.}} U$.

This is shown by selecting n_i such that $n_i < n < n_{i+1}$. Then, as $n_i \to \infty$,

$n \to \infty$ and

(32) $$|U_n - U| \leqq |U_n - U_{n_i}| + |U_{n_i} - U| \leqq V_i + |U_{n_i} - U| \to 0 .$$

In the following section, we shall see how this can be achieved, by a convenient centering, for sequences which converge in probability. There the special form of $\{n_i\}$ will be immaterial. However if the n_i and the X_n increase relatively slowly, this can be achieved by truncation. In fact,

Let $U_n = \sum_{i=1}^{n} X/n^\gamma$. *If* (i) *for every* $\epsilon > 0$, $\sum_n P[|U_n - U| > \epsilon]/n^\alpha < \infty$ *with* $0 < \alpha \leqq 1$, *and* (ii) $\sum_n P[|X_n| > cn^\beta] < \infty$ *with* $\beta > 0$, *then* $U_n \xrightarrow{\text{a.s.}} U$ *for* $\gamma \geqq \alpha + \beta$.

First the following slight extension of a property of series (Dvoretsky [4], for $\alpha = 1$) is easy to prove: if $\sum_n |p_n|/n^\alpha < \infty$ with $0 < \alpha \leqq 1$, then $\sum_i |p_{n_i}| < \infty$ with $n_{i+1} - n_i = 0(n_i^\alpha)$. Thus, (i) implies that $\sum_n P[|U_{n_i} - U| > \epsilon] < \infty$ and, by lemma 3.3, $U_{n_i} \xrightarrow{\text{a.s.}} U$. Now, let X'_n be X_n truncated at cn^β and $U'_n = \sum_{i=1}^{n} X'_i/n^\gamma$. From (ii), there follows by lemma 4.1 that $U_n - U'_n \xrightarrow{\text{a.s.}} 0$, so that $U'_{n_i} \xrightarrow{\text{a.s.}} U$. On the other hand, we have

(33) $$|U'_k - U'_{n_i}| = \left| \frac{k^\gamma - n_i^\gamma}{k^\gamma} U'_{n_i} + \frac{X'_{n_i+1} + \ldots + X'_k}{k} \right| ,$$

hence, because of the structure of $\{n_i\}$

(34) $$V_i \leqq |U'_{n_i}| \, o\,(n_i^{\alpha-1}) + o\,(n_i^{\alpha+\beta-\gamma}) \to 0 ,$$

and the proposition is proved.

It is easily seen that in the hypothesis (i) we can replace $P[|U_n - U| > \epsilon]$ by $\bar{\omega}_n(\epsilon) = \sup_m P[|U_{n+m} - U_n|]$. Then, not only do we not have to know U but also $\{n_i\}$ can be explicitly determined. In fact, elementary computations, using the relation $\bar{\omega}_{n+m}(2\epsilon) \leqq \bar{\omega}_n(\epsilon)$, show that

If $\alpha < 1$, *then* $\sum_n \bar{\omega}_n(\epsilon)/n^\alpha < \infty$ *for every* $\epsilon > 0$ *is equivalent to* $\sum_n \bar{\omega}_{[n^{1/(1-\alpha)}]}(\epsilon) < \infty$ *for every* $\epsilon > 0$ *and if* $\alpha = 1$ *then it is equivalent to* $\sum_n \bar{\omega}_{[q^n]}(\epsilon) < \infty$ *for every* $\epsilon > 0$ *and every* $q > 1$.

Clearly the proposition proved above holds *a fortiori* if $P[|U_n - U| > \epsilon]$ is replaced by $E|U_n - U|^r$ or $\sup_m E|U_{n+m} - U_n|^r$, $r > 0$. Then, for $\alpha = 1$, $\beta = 0$, $\gamma = 1$, $r = 2$ and $U = 0$, it reduces to a result of Dvoretzky [4] which generalized a result of the author [8] which, in turn, contains Borel's SLLN in the Bernoulli case and the SLLN in the Poisson case.

With $\bar{\omega}_n(\epsilon)$ and $\alpha = 1, \beta = 0, \gamma = 1$, the proposition reduces to a SLLN (Khintchine) for sequences of events which contain sequences of exchangeable events (de Finetti).

6. Centering of sequences convergent in probability

A sequence $\{U_n\}$ which converges in probability to U ($U_n \xrightarrow{P} U$) contains, as is well known, an a.s. convergent subsequence $U_{n_i} \xrightarrow{\text{a.s.}} U$ since for every ϵ, because of $P[|U_n - U| > \epsilon] \to 0$, there exists for a given $\sum_i \epsilon_i < \infty$ a sequence of $\{n_i\}$ such that $P[|U_{n_i} - U| > \epsilon_i] < \epsilon_i$, hence $\sum_i P[|U_{n_i} - U| > \epsilon_i] < \infty$ and, by lemma 3.3, $U_{n_i} \xrightarrow{\text{a.s.}} U$. Similarly, because of $P[|U_m - U_n| > \epsilon] \to 0$ for every ϵ, as $\frac{1}{m} + \frac{1}{n} \to 0$, there exists a sequence $n_i \uparrow \infty$ (there will be no possible confusion between the two sequences $\{n_i\}$) such that $\sum_i P[|U_{n_{i+1}} - U_{n_i}| > \epsilon_i] < \infty$ and, by lemma 3.3″, U_{n_i} converges a.s.; this implies that U_{n_i} converges in probability, whence to U and, finally $U_{n_i} \xrightarrow{\text{a.s.}} U$. In any case, condition (i) stated at the beginning of the last part of the previous section is fulfilled and we shall now fulfill condition (ii) by centering upon convenient conditional quantiles of order p, $0 < p \leqq 1/2$.

Let $\{U_k\}$ and $\{V_k\}$, $k = m, m + 1, \ldots, n$, be two sets of r.v. and $\{u_k\}$ and $\{v_k\}$, $v_k = u_k$, two sets of nonnegative sure or random quantities. Let

$$A_k^+ = [U_k > u_k], \quad A_k^- = [U_k < -u_k], \quad A_k = A_k^+ + A_k^- = [|U_k| > u_k], \quad A = \bigcup_k A_k,$$

$$B_k^+ = [V_k > v_k], \quad B_k^- = [V_k < -v_k], \quad B_k = B_k^+ + B_k^- = [|V_k| > v_k], \quad B = \bigcup_k B_k,$$

$$C_k^+ = [V_k - U_k \geqq v_k - u_k], \quad C_k^- = [V_k - U_k \leqq -v_k + u_k].$$

LEMMA 6.1. *If*

$$\text{(i)} \quad P(C_k^+; A_k^+ A'_{k+1} \ldots A'_n) \geqq p \leqq P(C_k^-; A_k^- A'_{k+1} \ldots A'_n)$$

or

$$\text{(i')} \quad P(C_k^+; A'_m \ldots A'_{k-1} A_k^+) \geqq p \leqq P(C_k^-; A'_m \ldots A'_{k-1} A_k^-)$$

then

$$pP(A) \leqq P(B).$$

In fact, by (1),

$$(35) \quad A = \bigcup_k A_k = \sum_k A_k A'_{k+1} \ldots A'_n = \sum_k A_k^+ A'_{k+1} \ldots A'_n + \sum_k A_k^- A'_{k+1} \ldots A'_n,$$

and, since

$$(36) \qquad C_k^{\pm}(A_k^{\pm} A'_{k+1} \ldots A'_n) \subset C_k^{\pm} A_k^{\pm} \subset B_k^{\pm} \subset B,$$

we have

$$(37) \quad PB \geqq \sum_k P(A_k^+ A'_{k+1} \ldots A'_n) P(C_k^+; A_k^+ A'_{k+1} \ldots A'_n) + \sum_k P(A_k^- A'_{k+1} \ldots A')P(C_k^-; A_k^- A'_{k+1} \ldots A'_n),$$

hence

$$PB \geqq p \sum_k P(A_k A'_{k+1} \dots A'_n) = pPA$$

and the conclusion is reached; similarly for (i′). Clearly, if some of the *a priori* probabilities are null, the corresponding terms in (37) disappear and the conclusion still holds

LEMMA 6.2. *If*

(i) $P(V_k - U_k \geqq 0;\ U_k, \dots, U_n) \geqq p \leqq P(V_k - U_k \leqq 0;\ U_k, \dots, U_n)$

or

(i′) $P(V_k - U_k \geqq 0;\ U_m, \dots, U_k) \geqq p \leqq P(V_k - U_k \leqq 0;\ U_m, \dots, U_k)$,

then, for every $\epsilon > 0$,

$$pP[\sup_k |U_k| > \epsilon] \leqq P[\sup_k |V_k| > \epsilon].$$

Let $D \in F\{U_k, \dots, U_n\}$. From (i) and the definition of conditional expectation (see section 7), it follows that

$$P(D)P(V_k - U_k \geqq 0;\ D) = \int_D P(V_k - U_k \geqq 0;\ U_k, \dots, U_n)\, dP \geqq pPD, \tag{38}$$

and similarly for $P(V_k - U_k \leqq 0;\ D)$. The same follows from (i′) for $D \in F\{U_m, \dots, U_k\}$. Applying lemma 6.2 above with $y_k \equiv z_k \equiv \epsilon$, the proposition is proved.

Fixing p, let $\mu(X; G)$ denote conditional quantiles of order p on the σ-field $G \subset F$ (see section 7) defined by

$$P[X - \mu(X;G) \geqq 0; G] \geqq p \leqq P[X - \mu(X;G) \leqq 0; G]. \tag{39}$$

The conditions under which the conclusion of the lemma above holds can be fulfilled by centering $V_k - U_k$ upon corresponding μ's. For instance, (i) *is true if we replace every* U_k *by* $U'_k = U_k - \mu(U_k - V_k;\ U_k, \dots, U_n)$. In fact, by the definition (39) of μ's, we have for example

$$P\{V_k - U_k - \mu(V_k - U_k;\ U_k, \dots, U_n) \geqq 0;\ U_k, \dots, U_n\} \geqq p, \tag{40}$$

and it implies

$$P(V_k - U'_k \geqq 0;\ U'_k, \dots, U'_n) \geqq p \tag{41}$$

because, for $i \geqq k$, U'_i is a function of $\{U_i, U_{i+1}, \dots, U_n\}$ only.

Particular cases. 1°. With $V_k \equiv U_n$, the lemma becomes *if* $\mu(U_n - U_k;\ U_k, \dots, U_n) = 0$ *or* $\mu(U_n - U_k;\ U_m, \dots, U_k) = 0$, *then*

$$pP[\sup_k |U_k| > \epsilon] = P[|U_n| > \epsilon].$$

2°. With U_k replaced by $U_n - U_k$ and $V_k \equiv U_n$ the conclusion becomes

$$pP[\sup_k |U_n - U_k| > \epsilon] \leqq P[|U_n| > \epsilon]. \tag{42}$$

But

$$P[\sup_k |U_k| > \epsilon] \leqq P\left[\sup_k |U_k - U_n| > \frac{\epsilon}{2}\right] + P\left[|U_n| > \frac{\epsilon}{2}\right] \tag{43}$$

and we have

If

$$\mu\,(U_k;\ U_n - U_k, \ldots, U_n - U_{n-1}) = 0\,,$$

then

$$pP\,[\sup_k |\,U_k| > \epsilon] \leqq (p+1)P\Big[\,|\,U_n| > \frac{\epsilon}{2}\Big]\,.$$

3°. Because of the Tchebicheff-Markoff inequality, the conclusions above hold *a fortiori* when $P[|U_n| > \epsilon']$ is replaced by $\frac{1}{\epsilon'^r} E|U_n|^r$, $r > 0$. They are then similar to the Kolmogoroff-P. Lévy inequality (for $r = 2$), but the centering is now upon conditional quantiles of order p instead of conditional expectations which might not exist. Clearly, if $U_k = \sum_{i=1}^{k} X_i$ where $X_1, \ldots, X_n$ are independent, the conditional quantiles of order p degenerate into *a priori* quantiles of order p. The inequalities thus obtained can be used to deduce criteria similar to theorem A.

4°. We shall need below the following form of lemma 6.2. Let $V_k \equiv U_{n_{i+1}} - U_{n_i}$, U_k be replaced by $U_{k_{i+1}} - U_{k_i}$ and $m = n_i + 1$, $n = n_{i+1} - 1$. Then we have, using (i) for instance,

If $P(U_k - U_{n_i} \geqq 0;\ U_{n_{i+1}} - U_k, \ldots, U_{n_{i+1}} - U_{n_{i+1}-1}) \geqq p$

$$\leqq P(U_k - U_{n_i} \leqq 0;\ U_{n_{i+1}} - U_k, \ldots, U_{n_{i+1}} - U_{n_{i+1}-1})$$

then

$$p\ P\,[\sup_k |\,U_{n_{i+1}} - U_k| > \epsilon] = P\,[\,|\,U_{n_{i+1}} - U_{n_i}| > \epsilon]\,,$$

and, writing $X_n = U_n - U_{n-1}$, *the knowledge of* $U_{n_{i+1}} - U_{n_{i+1}-1} = X_{n_{i+1}}$, $U_{n_{i+1}} - U_{n_{i+1}-2} = X_{n_{i+1}} + X_{n_{i+1}}$, $\ldots$, $U_{n_{i+1}} - U_k = X_{n_{i+1}} + \ldots + X_{k+1}$, *is equivalent to that of* $X_{n_{i+1}}, X_{n_{i+1}-1}, \ldots, X_{k+1}$.

We can now get back to $U_n \xrightarrow{P} U$, the sequences $\{n_i\}$ which correspond to a given $\sum_i \epsilon_i < \infty$ chosen, according to the beginning of this section, such that $\sum_i P[|U_{n_{i+1}} - U_{n_i}| > \epsilon_i] < \infty$. Here let $k = n_i + 1, \ldots, n_{i+1} - 1$.

THEOREM B. *If* $U_n = \sum_{i=1}^{n} X_i \xrightarrow{P} U$ *and*

(i) $P\,(X_{n_i+1} + \ldots + X_k \geqq 0;\ X_{k+1}, \ldots, X_{n_{i+1}}) \geqq p$

$$\leqq P\,(X_{n_i+1} + \ldots + X_k \leqq 0;\ X_{k+1}, \ldots, X_{n_{i+1}})$$

or

(i′) $P\,(U_k - U \geqq 0;\ X_{k+1}, \ldots, X_{n_{i+1}}) \geqq p \leqq P\,(U_k - U \leqq 0;$ $X_{k+1}, \ldots, X_{n_{i+1}})$

then $U_n \xrightarrow{\text{a.s.}} U$.

In fact, letting $W_i = \sup_k |X_{n_i+1} + \ldots + X_k|$, it follows from (i) and 4° that

$$P\,(W_i > \epsilon_i) \leqq \frac{1}{p} P\,(X_{n_i+1} + \ldots + X_{n_{i+1}}) \leqq \frac{\epsilon_i}{p}\,. \tag{44}$$

Consequently $\sum_i P(W_i > \epsilon_i) < \infty$ and, by 3.3, $W_i \xrightarrow{\text{a.s.}} 0$. Similarly, in case (i′) or others in which lemma 6.2 applies. The conclusion holds *a fortiori* if $U_n - U$ are centered about $\mu(U_n - U; X_{n+1}, \ldots)$.

THEOREM B′. *If $U_n \xrightarrow{P} U$, then there exists a sequence of zeros and conditional quantiles of order p, $\mu_n \xrightarrow{P} 0$ such that $U_n - \mu_n \xrightarrow{\text{a.s.}} U$.*

In fact, taking $\mu_k = \mu(X_{n_i+1} + \ldots + X_k; X_{k+1}, \ldots, X_{n_{i+1}})$ for $k = n_i + 1, \ldots, n_{i+1} - 1$, and $\mu_{n_i} \equiv 0$, the proof above applies to $U_n - \mu_n$ which thus converges a.s. But $U_{n_i} - \mu_{n_i} = U_{n_i} \xrightarrow{\text{a.s.}} U$, *hence* $U_n - \mu_n \xrightarrow{\text{a.s.}} U$. Consequently $U_n - \mu_n \xrightarrow{P} U$ and since $U_n \xrightarrow{P} U$, it follows that $\mu_n \xrightarrow{P} 0$.

If the μ_n are such that $\mu_n \xrightarrow{P} 0$ is equivalent to $\mu_n \xrightarrow{\text{a.s.}} 0$ then, clearly, $U_n \xrightarrow{P} U$ is equivalent to $U_n \xrightarrow{\text{a.s.}} U$. This is certainly true when the μ_n are degenerate and, in particular, when the X_n are independent. Thus (P. Lévy) for a series of independent r.v., convergence in probability and a.s. convergence are equivalent.

When $U_n = (X_1 + \ldots + X_n)/n$ we can, centering conveniently, reduce convergence of $\{U_n\}$ to that of a sequence of $\{V_i\}$ whose terms do not have common X's. In fact, let $q_i = [q^i]$ with $q > 1$ and $V_i = (X_{q_i+1} + \ldots + X_{q_{i+1}})/q_i$ with $i = 0, 1, 2, \ldots$. Since $q_{i+1} - q_i \to \infty$ as $i \to \infty$, we can assume, neglecting perhaps a finite number of X's, that $q_{i+1} - q_i > 1$ for $i = 0, 1, 2, \ldots$. To begin with

$$U_{q_i} \xrightarrow{\text{a.s.}} U \text{ if and only if } V_i \xrightarrow{\text{a.s.}} (q-1)U.$$

This follows at once from

$$(45)\qquad V_i = \frac{1}{q_i}(q_{i+1}U_{q_{i+1}} - q_i U_{q_i}) \qquad \text{and} \qquad U_{q_i} = \sum_{h=0}^{i-1} \frac{1}{q_i}(X_1 + q_h V_h).$$

Replacing U_n by $U_n - U$ and V_i by $V_i - (q-1)U$, that is, replacing X_n by $X_n - U$, we have $U_{q_i} \xrightarrow{\text{a.s.}} 0$ if and only if $V_i \xrightarrow{\text{a.s.}} 0$.

But, for $q_i < k < q_{i+1}$,

$$(46)\qquad U_k = \frac{q_i}{k} U_{q_i} + \frac{1}{k}(X_{q_{i+1}} + \ldots + X_k).$$

Hence,

$$(47)\qquad |U_k| \leqq q\,|U_{q_i}| + \frac{1}{q_i}|X_{q_i+1} + \ldots + X_k|.$$

Once the sums $X_{q_i+1} + \ldots + X_k$ are centered in order that lemma 6.2 applies, for instance upon $\mu(X_{q_i+1} + \ldots + X_k; X_{k+1}, \ldots, X_{q_{i+1}})$, it gives

$$(48)\qquad P\left[\sup_k \frac{1}{q_i}|X_{q_i+1} + \ldots + X_k| > \epsilon\right] = \frac{1}{p} P[\,|V_i| > \epsilon].$$

If $\sum_i P[|V_i| > \epsilon] < \infty$, then, by (47) and (48), $U_n \xrightarrow{\text{a.s.}} 0$. Thus

If $\sum_i P[|V_i| > \epsilon] < \infty$ *and* $\mu(X_{q_i+1} + \ldots + X_k;\ X_{k+1}, \ldots, X_{q_{i+1}}) = 0$ *for* $i = 0, 1, 2, \ldots$ *and* $k = q_i + 1, \ldots, q_{i+1} - 1$, *then* $U_n \xrightarrow{\text{a.s.}} 0$.

In the case of independent X_n's, the μ are degenerate, the V_i are independent and, by lemma 3.6, $V_i \xrightarrow{\text{a.s.}} 0$ if and only if, for every $\epsilon > 0$, $\sum_i P[|V_i| > \epsilon] < \infty$. With $q = 2$, we have thus the necessary and sufficient condition for the SLLN, due to Prokhoroff. (See K. L. Chung in these Proceedings.)

PART III. DETERMINING SET FUNCTIONS

7. Determining set functions and conditional expectations

Let (G, P_G) be a contraction of the probability field (F, P), that is G is a σ-field containing Ω and contained in F and $P_G(B) = P(B)$ for every $B \in G$. Let ϕ, with or without subscripts, denote a finite, σ-additive and P-continuous function defined on F and ϕ_G denote its contraction on G, that is, $\phi_G(B) = \phi(B)$ for every $B \in G$. A r.v. X with $E|X| < \infty$, determines uniquely its indefinite integral

$$\phi(A) = \int_A X dP, \qquad A \in F. \tag{49}$$

Conversely, by the Radon-Nikodym theorem a function ϕ defined on F determines (up to an equivalence) a r.v. $X = \frac{d\phi}{dP}$ or derivative of ϕ on (F, P) given by (49); ϕ will be called the *determining set function* (d.s.f.) of X. By the same theorem ϕ_G determines a r.v. $Y = \frac{d\phi_G}{dP_G}$ or derivative of ϕ on (G, P_G), by the relation

$$\phi(B) = \int_B Y dP, \qquad B \in G. \tag{50}$$

Y is a smoothed X, roughly a set of partial mean values of X (for instance, if B is an atom of G but not of F, then Y is constant on B and is a P-weighted mean of X on B). It will be the *conditional expectation of* X *on* G and written $E(X; G)$. When $X = I(A)$, then $E\{I(A); G\}$ is written $P(A; G)$ and is the conditional probability of A on G. If G is the σ-field determined in F by the set $\{U_\lambda\}$, $\lambda \in \Lambda$ then $E(X; G)$ will be denoted by $E(X; \{U_\lambda\})$—the conditional expectation of X on $\{U_\lambda\}$. Conditional expectations obey the rules of differential formalism. The conditional expectation $E(X; A)$ of X on $A \in F$ is defined by

$$PAE(X; A) = \phi(A) \tag{51}$$

and the indeterminacy of $E(X; A)$ when $PA = 0$ can be removed by making the convention, for a countable number of A's, $E(X; A) = 0$ if $PA = 0$.

We shall first briefly examine a few details indirectly related to the study of a.s. convergence.

1°. Because the Radon-Nikodym theorem applies to σ-finite, σ-additive and P-continuous set functions the definition of conditional expectation on G (such that ϕ_G is σ-finite) extends at once to a r.v. X whose indefinite integral exists but is not necessarily finite. It would be of some interest to adapt the results of the following sections, where *we assume once for all that the r.v. possess a finite expectation*, to this more general case.

2^0. Consideration of d.s.f.'s leads, naturally, to a kind of convergence different from those used at present in probability theory because, if ϕ_n converges on G to a finite set function, then this limit function is d.s.f. of a r.v. X. Consequently, if ϕ_n is d.s.f. of X_n, then $E(X_n; B)$ converges to $E(X; B)$, $B \in G$, and we can say that X_n converges to X *conditionally on* G ($X_n \xrightarrow{G} X$). Clearly $X_n \xrightarrow{G} X$ implies $EX_n \to EX$.

Convergence in probability, a.s. or ordinary, does not imply conditional convergence on F and conversely. The first three do not even imply convergence of EX_n to EX; examples are abundant. For the converse, take $X_n = \cos 2\pi n\omega$ where ω is uniformly distributed on (0, 1) and apply the Riemann-Lebesgue theorem.

Conditional convergence on F will be *uniform* if, given ϵ, there exists an n_ϵ such that $|\phi_n(A) - \phi(A)| < \epsilon$ for $n > n_\epsilon$ and all $A \in F$ and it follows at once from this definition that *uniform conditional convergence on F is equivalent to convergence in the mean.* We can define a stronger form of uniform convergence—uniform convergence of conditional expectations: $|E(X_n; A) - E(X; A)| < \epsilon$, that is, $|\phi_n(A) - \phi(A)| < \epsilon PA$, for $n > n_\epsilon$ and all $A \in F$. We shall see that it is related to various criteria for a.s. convergence.

3^0. *The zero one law.* One might say that two events A_1 and A_2 are a.s. independent on G if $P(A_1A_2; G) = P(A_1; G)\cdot P(A_2; G)$. If $A_1 = A_2 = A$, this relation becomes $P(A; G) = P^2(A; G)$. Hence the r.v. $P(A; G)$ has (up to an equivalence) only two possible values 0 and 1 and one can say that *A obeys the zero one law on G.* It is almost immediate that *an event A obeys the zero one law on G if and only if A belongs to G a.s.* (up to an event of probability zero). Thus *A obeys the zero one law*, that is, obeys it on all $G \subset F$ if and only if it coincides a.s. with an event belonging to all G, hence either with Ω or with $\emptyset$. Consequently, $PA = 0$ or $PA = 1$ and conversely.

Let $\{U_n\}$ be a sequence of r.v. and $C \in F\{TU_n\}$. C obeys the zero one law on $F\{TU_n\}$, hence on $F\{U_n\} \supset F\{TU_n\}$, but in general $PC \neq (PC)^2$. However, if $C = \lim_m \lim_n B_{mn}$ with $B_{mn} \in F\{U_m, \ldots, U_n\}$ and B_{mn} and $B_{m'n'}$ become asymptotically independent as the *disjoint* segments on which they are defined with the distance $|n - m'| \to \infty$, then C obeys the zero one law. More precisely,

If (i) $C = \lim_m \lim_n B_{mn}$ *and* (ii) $P(B_{mn}B_{m'n'}) - P(B_{mn})\, P(B_{m'n'}) \to 0$ *as* $n - m$, $n' - m'$, m, m' *and* $|n - m'| \to \infty$, *then C obeys the zero one law* because, passing to the limit in (ii), we get by (i) $PC - (PC)^2 = 0$. In the particular case of independent U_n, we have $P(B_{mn}B_{m'n'}) - P(B_{mn})P(B_{m'n'}) = 0$. Hence $F\{TU_n\} = (\Omega, \emptyset)$ up to events of probability zero. Similarly, when $U_n = \sum_{i=1}^{n} X_i$ with independent X_i and the events B_{mn} are defined in terms of differences $U_k - U_m$ ($m \leq k \leq n$). Specializing to $C = [U_n \text{ converges}]$, we have (P. Lévy): a sequence or a series of independent r.v. a.s. converges or a.s. diverges and we see that this conclusion extends to asymptotically independent U_n or X_n in the sense (ii) [or (ii′) below] with $B_{mn} = \sup_k |U_k - U_n| > \epsilon$, ϵ arbitrary. *The hypothesis* (ii) *can be replaced by* (ii′),

$$P(B_{m'n'};\, U_m, \ldots, U_n) \xrightarrow{P} PC .$$

In fact,

$$(52) \quad \int_{B_{mn}} P(B_{m'n'};\, U_m, \ldots, U_n)\, dP = \int_{B_{mn}} I(B_{m'n'})\, dP = P(B_{mn}B_{m'n'}) \longrightarrow PC ,$$

while the first integral, which can be written $\int_\Omega I(B_{mn})P(B_{m'n'};\, U_m, \ldots, U_n)dP$ converges, by (ii′) and the Lebesgue convergence theorem, to $\int_\Omega I(C)P(C)dP = (PC)^2$. *The conclusion holds a fortiori with* $P(C;\, U_1, \ldots, U_n) \xrightarrow{P} PC$, and the particular case $P(C;\, U_1, \ldots, U_n) = PC$ for every n becomes the Kolmogoroff zero one law (for tail events). Also, (ii) and (ii′) hold if $P(B_{mn}C) - P(B_{mn})PC \to 0$ and $P(B_{mn};\, F\{TU_n\}) \xrightarrow{P} PC$.

8. D.s.f. and a.s. convergence of sequences

Let $\{\phi_n\}$ be the sequence of d.s.f. which corresponds to $\{U_n\}$, $G = F\{U_n\}$ and $H = F\{TU_n\}$. Events A_{ij} are disjoint in j if A_{ij_1} and A_{ij_2} are disjoint for $j_1 \neq j_2$. ψ' will denote the P-continuous part of ψ and, as usual, $k = m, m+1, \ldots, n$.

THEOREM C. *If, for all disjoint in* k *events* A_{mk} *defined on* $\{U_m, \ldots, U_k\}$ *such that* $\lim_m \lim_n \sum_k A_{mk} = A$ *exists, and all events* B *defined on a finite number of* U_n, ψ *defined by* $\lim_m \lim_n \sum_k \phi_k(A_{mk}B) = \psi(AB)$ *exists and is bounded and* σ*-additive, then* $U \xrightarrow{\text{a.s.}} U = \frac{d\psi'_H}{dP_H}$; *if* $\{U_m, \ldots, U_k\}$ *is replaced by* $\{U_1, \ldots, U_k\}$, *then* $U = \frac{d\psi'_G}{dP_G}$.

To prove the existence of a limit r.v., we shall show that, writing U_* for $\liminf_n U_n$ and U^* for $\limsup_n U_n$, (i) $P[U_* \neq U^*] = 0$ and (ii) $P[U_* = -\infty] = P[U^* = +\infty] = 0$. To begin with, (i) is equivalent to (i′) $PC_{ab} = 0$ for any couple $a < b$ of rationals, $C_{ab} = [U_* < a < b < U^*]$ because

$$(53) \qquad C_0 = [U_* \neq U^*] = \bigcup_{a,b} C_{ab} .$$

Now, selecting $A_{mk} = [U_m \geqq a, \ldots, U_{k-1} \geqq a,\ U_k < a]$, hence $A = [U_* \leqq a]$, we have

$$(54) \quad \sum_k \phi_k(A_{mk}B) = \sum_k \int_{A_{mk}B} U_k dP \leqq a \sum_k P(A_{mk}B) = aP\Big(\sum_k A_{mk}B\Big).$$

From the hypothesis, there follows that as $n \to \infty$ and $m \to \infty$,

$$(55) \qquad \psi(AB) \leqq a\, PAB .$$

Since ψ has a unique extension (to be denoted by the same letter) on G, this relation holds when B is replaced by any event in G and, in particular, by $C_{ab} (\supset A)$. Therefore

$$\psi(C_{ab}) \leqq aPC_{ab}\,. \tag{56}$$

Similarly, starting with $A_{mk} = [U_m \leqq b, \ldots, U_{k-1} \leqq b, U_k > b]$, we get

$$\psi(C_{ab}) \geqq bPC_{ab}\,. \tag{57}$$

Consequently, since $a < b$, it follows that $PC_{ab} = 0$ and (i′), hence (i), is proved.

Now let $B = \Omega$. Being finite and σ-additive, the function ψ is bounded by a constant $c < \infty$ and, for $a < 0$, (55) gives

$$P[U_* < a] \leqq \frac{c}{|a|} \to 0 \quad \text{as} \quad a \to -\infty\,. \tag{58}$$

Similarly, for $b > 0$

$$P[U^* > b] \leqq \frac{c}{b} \to 0 \quad \text{as} \quad b \to +\infty\,, \tag{59}$$

and (ii) is proved. Thus, taking for U the r.v. equal to $U_* = U^*$ on C' and to zero on $C = C_0 \cup [U^* = +\infty] \cup [U_* = -\infty]$ we have proved that $U_n \xrightarrow{\text{a.s.}} U$.

There remains to show that $U = \dfrac{d\psi'_H}{dP_H}$. Let $A \in H$ and

$$U_\lambda = h\lambda \quad \text{on} \quad C_h = [h\lambda \leqq U < (h+1)\lambda] \tag{60}$$

where $\lambda > 0$, $h = \ldots, -1, 0, +1, \ldots$, and

$$U_\lambda = 0 \quad \text{on}\ C\,. \tag{61}$$

From the hypothesis, there follows as for (i) that

$$h\lambda P(AC_hC') \leqq \psi(AC_hC') \leqq (h+1)\lambda P(AC_hC')\,. \tag{62}$$

Summing over h we get, because of $U_\lambda \leqq U < U + \lambda$,

$$\psi(AC') - \lambda \leqq \int_{AC'} U_\lambda dP \leqq \int_{AC'} U dP \leqq \int_{AC'} U_\lambda dP + \lambda \tag{63}$$

$$\leqq \psi(AC') + \lambda\,.$$

Thus, letting $\lambda \to 0$,

$$\psi(AC') = \int_{AC'} U dP \tag{64}$$

or, since $PC = 0$, $U = \dfrac{d\psi'_H}{dP_H}$.

If $A_{mk} \in F\{U_1, \ldots, U_k\} \supset F\{U_m, \ldots, U_k\}$, $U_n \xrightarrow{\text{a.s.}} U$ *a fortiori* and the last part of the proof applies with $A \in G$ instead of H, hence $U = \dfrac{d\psi'_G}{dP_G}$.

Also A_{mk} can be replaced by $A_{kn} = [U_k < a, U_{k+1} \geqq a, \ldots, U_n \geqq a]$ with $B \in H$ instead of $F(U_1, U_2, \ldots)$ and then $U = \dfrac{d\psi'_H}{dP_H}$.

The theorem above has a partial converse and, perhaps, it would be possible to improve it so as to get a necessary and sufficient condition for a.s. convergence in terms of d.s.f.

THEOREM C′. *If* $U_n \xrightarrow{\text{a.s.}} U$ *and* $|U_n| \leqq V (EV < \infty)$, *then* $\lim_m \lim_n \sum_k \phi_k(A_k) = \phi(A)$ *where* $\lim_m \lim_k \sum A_k = A$, A_k *varying or not with m and n.*

In fact, given $\epsilon > 0$, there exists an m_ϵ such that for $m > m_\epsilon$,

$$(65)\qquad PC_\epsilon < \epsilon \quad \text{for} \quad C_\epsilon = [\sup_k |U_n - U| > \epsilon] .$$

On the other hand,

$$(66)\qquad \sum_k \phi_k(A_k) = \sum_k \int_{A_k} U dP + \sum_k \int_{A_k} (U_k - U)\, dP$$

or

$$(67)\qquad \sum_k \phi_k(A_k) = \phi\Big(\sum_k A_k\Big) + \sum_k \int_{A_k C'_\epsilon} (U_k - U)\, dP - \int_{\sum_k A_k C_\epsilon} U dP + \sum_k \int_{A_k C_\epsilon} U_k dP .$$

As $\epsilon \to 0$ the first term on the right hand side converges to $\phi(A)$ and the third to zero; the second term is bounded by $\epsilon P\Big(\sum_k A_k C'_\epsilon\Big) \leqq \epsilon \to 0$. The fourth is the only term for which a restriction seems essential. If $|U_k| \leqq V$, $(EV < \infty)$ then

$$(68)\qquad \Big|\sum_k \int_{A_k C_\epsilon} U_k dP\Big| \leqq \sum_k \int_{A_k C_\epsilon} |U_k|\, dP \leqq \int_{\sum_k A_k C_\epsilon} V dP \leqq \int_{C_\epsilon} V dP \to 0 \quad \text{as} \quad \epsilon \to 0$$

and the theorem is proved.

An obvious weaker restriction under which the conclusion still holds is

$$(69)\qquad \sum_k \phi_k(A_k C_\epsilon) \to 0 \quad \text{as} \quad n \to \infty , \quad m \to \infty , \quad \text{when} \quad P(C_\epsilon) \to 0 .$$

Particular cases. 1^0. (i) *If* $E(U_n; A) \to E(V; A)$ *uniformly on* $A \in F\{U_1, U_2, \ldots, U_n\}$, *then* $U_n \xrightarrow{\text{a.s.}} U = E(V; G)$ *and* (ii) *if* $A \in F\{U_n, U_{n+1}, \ldots\}$, *then* $U = E(V; H)$.

In the first case, given $\epsilon > 0$ and $B \in F(U_1, U_2, \ldots)$, there exists an n_ϵ independent of $A \in F\{U_1, \ldots, U_n\}$ such that

$$(70)\qquad |\phi_n(AB) - \phi_V(AB)| < \epsilon\, PAB ,$$

and, consequently, with $A_{mk} \in F\{U_1, \ldots, U_n\}$,

$$(71)\qquad \Big|\sum_k \phi(A_{mk}B) - \phi_V\Big(\sum_k A_{mk}B\Big)\Big| < \epsilon P\Big(\sum_k A_{mk}B\Big) \leqq \epsilon .$$

Thus, passing to the limit, theorem C applies with $U = E(V; G)$. In the second case, we apply the theorem with A_{kn} instead of A_{mk}.

(iii) *If* $E(U_m;\ A) - E(U_n;\ A) \to 0$ *as* $\frac{1}{m}+\frac{1}{n} \to 0$ *uniformly in* $A \in F\{U_1, \ldots, U_n\}$ *and* $E|U_n| < c < \infty$, *then* U_n *converges a.s.*

In fact, by Cauchy convergence criterion $E(X_n; A)$ converges to a set function $e(A)$ uniformly in $A \in F\{U_1, \ldots, U_{n_0}\}$ for an arbitrarily fixed n_0. Consequently, given $\epsilon > 0$, there exists an n_ϵ such that

$$(72) \qquad |\phi_n(A) - \psi(A)| < \epsilon PA \quad \text{with} \quad A \in F\{U_1, \ldots, U_{n_0}\}$$
$$\text{and} \quad \psi(A) = e(A)\,PA\,.$$

It follows that $\psi(A)$ is bounded by c and is σ-additive on $F\{U_1, \ldots, U_{n_0}\}$ for every n_0. Hence, by a trivial extension of Kolmogoroff's theorem on probability measures in R_∞, the set function ψ, which so far is defined only on $F(U_1, U_2, \ldots)$, can be extended to a bounded and σ-additive set function ψ on G.

Now, for $m > n_\epsilon$ sufficiently large

$$(73) \qquad \left|\sum_k \phi_k(A_{mk}) - \psi\left(\sum_k A_{mk}B\right)\right| < \epsilon P\left(\sum_k A_{mk}B\right) \leqq \epsilon\,.$$

Thus, $\psi\left(\sum_k A_{mk}\right)$ approaches $\psi(AB)$ as $\sum_k A_{mk}$ approaches $A \in G$, and theorem C applies with $U = \frac{d\psi'_G}{dP_G}$. Specializing further we have

2⁰. (i′) *If* $U_n - E(V;\ U_1, \ldots, U_n) \to 0$, *uniformly in* $\{U_1, \ldots, U_n\}$, *then* $U_n \xrightarrow{\text{a.s.}} U = E(V; G)$. (ii′) *If* $U_1, \ldots, U_n$ *are replaced by* $U_n, U_{n+1}, \ldots$, *then* $U_n \xrightarrow{\text{a.s.}} U = E(V; H)$. (iii′) *If* V *is replaced by* U_{n+1} *and* $E|U_n| < c < \infty$, *then* U_n *converges a.s.*

These propositions follow from those above in the same fashion. For instance, case (i) follows from the fact that for $n > n_\epsilon$ sufficiently large,

$$(74) \quad |E(U_n; A) - E(V; A)| \leqq \frac{1}{PA}\int_A |U_n - E(V;\ U_1, \ldots, U_n)|\,dP < \epsilon.$$

Taking the particular cases (i″) $U_n = E(V;\ U_1, \ldots, U_n)$, (ii″) $U_n = E(V;\ U_n, U_{n+1}, \ldots)$, (iii″) $U_n = E(U_{n+1};\ U_1, \ldots, U_n)$ the above propositions reduce to martingale properties due to Doob [2], P. Lévy and Ville [10].

3⁰. In the case of exchangeable r.v. $\{X_n\}$ it is known that if the second common moment is finite, the SLLN holds. It is easy to show that finiteness of the first common moment suffices and also to find the limit r.v. In fact, let $S_n = \sum_{i=1}^{n} X_i$ and $H_S = F\{TS_n\}$. We have, for every fixed i,

$$(75) \qquad E(X_i;\ S_n, S_{n+1}, \ldots) \xrightarrow{\text{a.s.}} E(X_i;\ H_S)\,.$$

But the X_i being exchangeable, for all $i \leqq n$,

$$(76) \qquad E(X_i;\ S_n, S_{n+1}, \ldots) = U_n = E(X_1;\ S_n, S_{n+1}, \ldots)\,,$$

and

$$U_n \xrightarrow{\text{a.s.}} E(X_1;\ H_S)\,. \tag{77}$$

Consequently

$$\frac{S_n}{n} = E\left(\frac{S_n}{n}; S_n, S_{n+1}, \ldots\right) = \frac{1}{n}\sum_{i=1}^{n} (X_i, S_n, S_{n+1}, \ldots) = U_n \xrightarrow{\text{a.s.}} E(X_1;\ H_S)\,. \tag{78}$$

Thus, *if the* X_i *are exchangeable (and* $E|X_1| < \infty$*), then* $S_n/n \xrightarrow{\text{a.s.}} E(X_1;\ H_S)$. In particular, if the X_n are independent and identically distributed, this reduces to the Kolmogoroff SLLN and the proof specializes to that of Doob.

9. D.s.f. and a.s. convergence of ratios of sums

The method used in the preceding section but adapted to sequences of ratios of sums of r.v. by the Fr. Riesz lemma below will yield a proposition containing as particular cases the celebrated Birkhoff ergodic theorem as well as its known extensions obtained by various methods by Hopf [6], Hurewicz [7], Halmos [5] and Dunford and Miller [3].

Let $a_1, a_2, \ldots, a_\nu$ be real numbers, $m < \nu$, $l = 0, 1, 2, \ldots, m - 1$. The term a_i is said to be "favorable" if $\sup_l (a_i + a_{i+1} + \ldots + a_{i+l}) > 0$. There need not be any favorable terms but if there are any, then,

LEMMA 9.1. *The sum of favorable terms in* $\{a_1, \ldots, a_\nu\}$ *is positive.*

In fact, let a_{k_1} be the first favorable term and $a_{k_1} + a_{k_1+1} + \ldots + a_{k_1+l_1}$ ($l_1 < m$) the *shortest* positive sum beginning with a_{k_1}. All terms of this sum are favorable because if $a_{k_1+l_0}$ $(0 \leqq l_0 \leqq l_1)$ were not favorable, we would have $a_{k_1+l_0} + \ldots + a_{k_1+l_1} \leqq 0$ hence $a_{k_1} + \ldots + a_{k_1+l_0-1} > 0$ which contradicts the assumption made. Thus the successive favorable terms form stretches of positive sums and the lemma is proved.

We are now going to give conditions under which $\lim_n U_n$ exists a.s. with $U_n = S_n/T_n$, $S_n = \sum_{i=1}^{n} X_i$, $T_n = \sum_{i=1}^{n} Y_i$, ϕ_i and ψ_i d.s.f. of X_i and Y_i, respectively. We assume that *the r.v.* Y_i *are positive* and that

$$\liminf_n \frac{1}{n}\sum_{i=1}^{n} \psi_i(C) > 0 \tag{79}$$

for all events C of positive probability defined on $\{X_n, Y_n\}$ which remain invariant under translations on Y's.

All events below will be defined on $F\{X_n, Y_n\}$, A_{mi} will represent an event $A_m = A_{m1}$ translated by $i - 1$, that is obtained by adding $i - 1$ to the subscripts of the r.v. X_n, Y_n which figure in the definition of A_m.

THEOREM D. *If, for every* $A_m \downarrow \phi$ *as* $m \to \infty$, (i) $\lim_m \limsup_n \frac{1}{n}\left|\sum_{i=1}^{n} \phi_i(A_{mi})\right| = 0$, (i′) $\lim_m \limsup_n \frac{1}{n}\sum_{i=1}^{n} \psi_i(A_{mi}) = 0$, *and* (ii) $\lim_n \frac{E|X_n|}{n} = 0$, (ii′) $\lim_n \frac{EY_n}{n} = 0$,

then $\lim_n U_n$ *exists a.s.*

We want to prove that $P[U_* = U^*] = 1$. Thus, as in the proof of theorem C, we have to show that

$$(80) \qquad PC_{ab} = 0, \qquad C_{ab} = [U_* < a < b < U^*];$$

to simplify the writing we shall drop the subscripts a and b in C_{ab}. Let $B_m = [\inf_{i \leq m} U_i < a]$. Clearly

$$(81) \qquad B_m \uparrow B = [\inf_n U_n \leqq a] \supset C.$$

Thus, $B_m C \uparrow C$ and we can write

$$(82) \qquad C = B_m C + A_m \qquad \text{with} \qquad A_m \downarrow \emptyset \qquad \text{as} \qquad m \to \infty.$$

Now we apply the lemma above with $\nu = n + m$ to the values of $aY_i - X_i$, $i = 1, 2, \ldots, n + m$. If B_m^i represents the event $[aY_i - X_i$ favorable$]$, that is,

$$B_m^i = \left[\inf_{l<m} \frac{X_i + \ldots + X_{i+l}}{Y_i + \ldots + Y_{i+l}} < a\right], \qquad B_m^1 = B_m,$$

then

$$(83) \qquad Z_{n+m} = \sum_{i=1}^{n+m} (aY_i - X_i)\, I(B_m^i) \geqq 0$$

and, *a fortiori*,

$$(84) \qquad Z_{n+m} I(C) = \sum_{i=1}^{n+m} (aY_i - X_i)\, I(B_m^i C) \geqq 0.$$

But, for $i \leqq n$, we have $B_m^i = B_{mi}$ hence, applying to (82) a translation $i - 1$, we get

$$(85) \qquad C = B_{mi} C + A_{mi}, \qquad i \leqq n.$$

Thus, *a fortiori*,

$$(86) \qquad \sum_{i=1}^{n} (aY_i - X_i)[I(C) - I(A_{mi})] + \sum_{i=n+1}^{n+m} |aY_i - X_i| > 0.$$

Integrating with respect to P and dividing by n we get, *a fortiori*

$$(87) \qquad \int_C \left(a\frac{T_n}{n} - \frac{S_n}{n}\right) dp + \frac{1}{n}\left|\sum_{i=1}^{n} \phi_i(A_{mi})\right| + \frac{|a|}{n}\sum_{i=1}^{n} \psi_i(A_{mi})$$

$$+ \frac{1}{n}\sum_{i=n+1}^{n+m} E|X_i| + \frac{|a|}{n}\sum_{i=1}^{n} EY_i \geqq 0.$$

Now we let $n \to \infty$, then $m \to \infty$. From the hypotheses there follows that

$$(88) \qquad \liminf_n \int_C \left(a\frac{T_n}{n} - \frac{S_n}{n}\right) dP \geqq 0.$$

Similarly, starting with $B_m = [\sup_{i \leq m} U_i > b]$, we get

$$(89) \qquad \liminf_n \int_C \left(\frac{S_n}{n} - b\frac{T_n}{n}\right) dP \geqq 0 .$$

Together with (88) this implies that

$$(90) \qquad \liminf_n \int_C (a-b)\frac{T_n}{n}\, dP \geqq 0$$

and, since $a < b$, that

$$(91) \qquad \liminf_n \frac{1}{n}\sum_{i=1}^{n} \psi_i(C) = 0$$

which, because of property (79) of $\{Y_n\}$ implies that $PC = 0$, and the theorem follows.

Particular cases. 1⁰. *If $PA_{mi} \leqq cPA_m$, then $U_n = S_n/n$ converges a.s. to a r.v.*

In fact, $Y_n = 1$, $|\phi_i|(A_{mi}) \leqq c|\phi_1|(A_m)$ and the conditions are satisfied since (i) $\lim_m |\phi_1|(A_m) = 0$, (i′) $\lim_m PA_m = 0$, (ii) $\lim_n |\phi_1|(\Omega)/n = 0$, and (ii′) $\lim_n 1/n = 0$. Moreover $E|S_n|/n \leqq cE|X_1| < \infty$ hence, by the Fatou lemma $\lim_n U_n$ is a r.v. (up to an equivalence).

Specializing to $PA_{mi} = PA_m$, we have the Birkhoff theorem.

2⁰. *If $\frac{1}{n}\sum_{i=1}^{n} PA_{mi} \leqq cPA_m$, then $U_n = S_n/n$ converges a.s. to a r.v.*

In fact, it follows that

$$(92) \qquad \frac{1}{n}\sum_{i=1}^{n} |\phi_i|(A_{mi}) \leqq c|\phi_1|(A_m)$$

and, in particular, for every h

$$(93) \qquad \frac{1}{n}\sum_{i=1}^{n} |X_{h+i}E| \leqq cE|X_h|$$

hence $E|X_n|/n \to 0$. The conditions of the theorem are satisfied, the Fatou lemma applies and this extension of Birkhoff's theorem (Dunford-Miller) is proved.

3⁰. *If the d.s.f. of X_i and Y_i are defined by $\phi_i(A_{mi}) = \phi_1(A_m)$ and $\psi_i(A_{mi}) = \psi_1(A_m)$, then $U_n = S_n/T_n$ converges a.s. to a r.v.*

This corresponds to Hurewicz's extension of Birkhoff's theorem and follows as above from our theorem. It contains Hopf's extension which corresponds to $P(A_{mi}) = PA_m$ for $A_m \in F\{X_n, Y_n\}$ with $Y_n = 1$.

4⁰. From $\frac{1}{n}\left|\sum_{i=1}^{n} \phi_i(A_{mi})\right| \leqq c\phi(A_m)$, $\frac{1}{n}\sum_{i=1}^{n} \psi_i(A_{mi}) \leqq c\psi(A_m)$ (ϕ and ψ P-continuous), (i) and (i′) follow at once. The corresponding particular case would represent an extension of the ergodic theorem containing Hurewicz's and Dunford-Miller's theorems.

REFERENCES

[1] G. D. Birkhoff, "Proof of a recurrence theorem for strongly transitive systems," *Proc. Nat. Acad. Sci.*, Vol. 17 (1931), pp. 650–660.
[2] J. L. Doob, "Regularity properties of certain families of chance variables," *Trans. Amer. Math. Soc.*, Vol. 47 (1940), pp. 455–486.
[3] N. Dunford and D. S. Miller, "On the ergodic theorem," *Trans. Amer. Math. Soc.*, Vol. 60 (1946), pp. 538–549.
[4] A. Dvoretzky, "On the strong stability of a sequence of events," *Annals of Math. Stat.*, Vol. 20 (1949), pp. 296–299.
[5] P. R. Halmos, "An ergodic theorem," *Proc. Nat. Acad. Sci.*, Vol. 32 (1946), pp. 156–161.
[6] E. Hopf, *Ergodentheorie*, Springer, Berlin, 1937.
[7] W. Hurewicz, "Ergodic theorem without invariant measure," *Annals of Math.*, Ser. 2, Vol. 45 (1944), pp. 196–206.
[8] M. Loève, "Étude asymptotique des sommes de variables aléatoires liées," *Jour. de Math.*, Vol. 24 (1945), pp. 249–318.
[9] J. Marcinkiewicz, "Sur les fonctions indépendantes, I, II," *Fund. Math.*, Vol. 30 (1938), pp. 202–214, 349–364.
[10] J. Ville, *Étude Critique de la Notion de Collectif*, Gauthier-Villars, Paris, 1939.