ASYMPTOTIC NORMALITY FOR SUMS OF DEPENDENT RANDOM VARIABLES

ARYEH DVORETZKY
THE HEBREW UNIVERSITY OF JERUSALEM
AND COLUMBIA UNIVERSITY, NEW YORK

1. Introduction

1.1. Limiting distributions of sums of independent random variables have been exhaustively studied and there is a satisfactory general theory of the subject (see the monograph of B. Gnedenko and A. Kolmogorov [6], or advanced text books on probability theory such as that of M. Loève [10]). Our knowledge of the corresponding theory for dependent random variables is much more meagre. Although a great number of papers have been published on the subject, not many general results are known. In recent years the author has shown ([4], [5]) that the necessary and sufficient conditions for convergence in distribution to any specified, infinitely divisible law remain sufficient also in the most general dependent case, provided that quantities such as means and the like, are replaced by conditional means, and the like, the conditioning being relative to the preceding sum. (The necessity of the conditions requires, in general, further assumptions.)

In the present paper we are concerned almost exclusively with asymptotic normality. Though our general results about asymptotic normality can be obtained by direct specialization of the results mentioned above, we preferred to develop them here independently. We hope that the greater accessibility of the present proofs will compensate for this sacrifice of brevity.

After establishing the general results we give a few applications. It would be quite easy to extend the list of applications indefinitely by going through various results in the literature and seeing how they can be improved by using our general theorems.

1.2. We consider random variables arranged in a double array

$$
\begin{array}{c}
X_{1,1}, X_{1,2}, \cdots, X_{1,k_1} \\
X_{2,1}, X_{2,2}, \cdots, X_{2,k_2} \\
\vdots \\
X_{n,1}, X_{n,2}, \cdots, X_{n,k_n} \\
\vdots
\end{array}
\tag{1.1}
$$

and, putting

$$S_{n,k} = \sum_{j=1}^{k} X_{n,j} \qquad \text{for} \quad k = 0, 1, \cdots, k_n, \tag{1.2}$$

we establish conditions which imply that the row sums in (1.1) are asymptotically standard normal, that is that

$$\lim_{n=\infty} P(S_{n,k_n} \leqq u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-v^2/2}\, dv \tag{1.3}$$

for all real u.

Let $\mathscr{F}_{n,k} = \mathscr{B}(S_{n,k})$. Throughout the paper $\mathscr{B}(\cdot)$ denotes the σ-field generated by the indicated random variable or variables. ($\mathscr{F}_{n,0}$ is the trivial field). Most of our results assume that the random variables in (1.1) have finite second moments; then the conditional means

$$\mu_{n,k} = E(X_{n,k} | \mathscr{F}_{n,k-1}) \tag{1.4}$$

and the conditional variances

$$\sigma_{n,k}^2 = E(X_{n,k}^2 | \mathscr{F}_{n,k-1}) - \mu_{n,k}^2 \tag{1.5}$$

exist almost surely.

Our main results are given in Section 1. Theorem 2.1 asserts (1.3) under the assumptions that the random variables are centered at their conditional expectations, that the conditional variances in each row add up to 1 and that the Lindeberg condition,

$$\lim_{n=\infty} \sum_{k=1}^{k_n} E[X_{n,k}^2 I(|X_{n,k}| > \varepsilon)] = 0 \qquad \text{for every } \varepsilon > 0, \tag{1.6}$$

holds (the symbol $I(\cdot)$ denotes the indicator function of the set within the parentheses). Equation (1.6) is the ordinary Lindeberg condition; it is stated for the variables in (1.1) themselves, not for those obtained from them through conditioning.

The crucial new feature of this result is that no assumptions are made about the conditional variances being nearly constant in some sense. Most results in the literature (for example, S. Bernstein [1], Loève [10], [11], J. L. Doob [3], B. Rosén [14], and W. Philipp [13]) are similar to Corollary 2.1 in making some near constancy assumptions about the conditional variances or related quantities. A most remarkable exception is furnished by P. Lévy ([9], p. 243, and earlier work quoted in [9]), and Theorem 2.1 may be considered as a further development of Lévy's ideas.

Theorem 2.2 relaxes the conditions of Theorem 2.1. Instead of assuming the conditional means to be zero, it is assumed that their sum in each row tends to zero in probability. About the conditional variances it is only assumed that their sum in each row tends in probability to one; equation (1.6) is also similarly relaxed. Corollary 2.2 gives a simple extension and Theorem 2.3 gives a more interesting extension to the case when no moment conditions are imposed on the random variables (1.1). Theorem 2.4 is much easier than the previous ones. It suffers from a "near constancy" assumption and is indeed not new (see for example Loève [11], p. 375), but we need it to illustrate some applications.

1.3. In order not to interrupt the proofs of the results of Section 2, we establish four lemmas in Section 3. Lemma 3.1 is of independent interest. It allows us to assume without loss of generality that the partial sums (1.2) in each row form a Markov sequence. Remark 3.4 raises a warning about disregarding measurability conditions (see also Section 6.1). We also draw attention to Lemma 3.3, especially to its simplest case.

1.4. Basing our argument on the preceding lemmas, we prove in Section 4 our main results.

1.5. Before giving some illustrative applications of the general results, we study in Section 5.1 a measure of dependence (Definition 5.1) used in connection with the study of asymptotic normality of stationary sequences by M. Rosenblatt [15], I. A. Ibragimov [8], W. Philipp [13] and others. The lemmas proved here are used in Sections 5.2 and 5.3. Theorem 5.1 may be regarded as a generalization of some results of Bernstein [1] and their descendants. The reduction of this theorem to those of Section 2 follows the method introduced by Bernstein and used by all his followers. Theorem 5.2 and Corollary 5.1 improve the results of Ibragimov [8] and Philipp [13]. In Section 5.4 we show how an improved form of a result of R. J. Serfling [16] follows from Theorem 2.2. Theorem 5.4 shows how to reduce to our general theorems the study of random variables (1.1) in which the conditional mean of each random variable depends nearly linearly on the preceding sum. The results of Rosén [14] follow from this theorem.

1.6. The last section contains a counter example and some remarks on further applications as well as generalizations.

2. Main results

2.1. We start with a special case patterned after a classical version of Lindeberg's theorem on sums of independent random variables. It exhibits many features of the more general results and subsumes many theorems on sums of dependent random variables.

THEOREM 2.1. *If the random variables* (1.1) *satisfy*

$$\mu_{n,k} = 0, \tag{2.1}$$

$$\sum_{k=1}^{k_n} \sigma_{n,k}^2 = 1, \tag{2.2}$$

and (1.6), *then* (1.3) *holds.*

We already discussed the main feature of this theorem in Section 1.2. Here we note that the conditioning in our theorems is relative to the preceding row sum, not relative to the richer σ-field generated by all preceding random variables in the same row, which, in cases like (2.1) or (2.2), would constitute a more stringent requirement.

The one on the right side of (2.2) could, of course, be replaced by any *constant*, provided a corresponding change of scale is made in (1.3). This condition cannot, however, be weakened to the condition that the sum in (2.2) be a random variable whose distribution is the same for all n. The weaker condition does not even ensure the convergence in distribution of the row sums of (1.1) (see 6.1).

2.2. The following generalization of Theorem 2.1 is perhaps the basic result of the present paper. Here and in the sequel $\xrightarrow{P}$ denotes convergence in probability.

THEOREM 2.2. *Let the random variables* (1.1) *satisfy*

$$\sum_{k=1}^{k_n} \mu_{n,k} \xrightarrow[n\to\infty]{P} 0, \tag{2.3}$$

$$\sum_{k=1}^{k_n} \sigma_{n,k}^2 \xrightarrow[n\to\infty]{P} 1 \tag{2.4}$$

and

$$\sum_{k=1}^{k_n} E[X_{n,k}^2 I(|X_{n,k}| > \varepsilon) | \mathscr{F}_{n,k-1}] \xrightarrow[n\to\infty]{P} 0 \qquad \text{for every} \quad \varepsilon > 0. \tag{2.5}$$

Then (1.3) *holds.*

We note that (2.5), a conditioned Lindeberg condition, is indeed a weaker assumption than (1.6), the ordinary Lindeberg condition. In fact, the sum in (2.5) is a nonnegative (possibly generalized) random variable, whose expectation is the sum in (1.6). Thus (1.6) implies not only (2.5), but even the convergence to zero in L_1 norm. It is trivial to construct examples where (2.5) holds but (1.6) is not satisfied.

Specializing to near constant conditional means and variances, we have the following useful result.

COROLLARY 2.1. *Let $a_{n,k}$ and $b_{n,k}$, for $n = 1, 2, \cdots$ and $k = 1, \cdots, k_n$, be numbers satisfying*

$$\lim_{n=\infty} \sum_{k=1}^{k_n} a_{n,k} = 0 \tag{2.6}$$

and

$$\lim_{n=\infty} \sum_{k=1}^{k_n} b_{n,k}^2 = 1. \tag{2.7}$$

If the random variables (1.1) *satisfy*

$$\lim_{n=\infty} \sum_{k=1}^{k_n} E|\mu_{n,k} - a_{n,k}| = 0, \tag{2.8}$$

$$\lim_{n=\infty} \sum_{k=1}^{k_n} E|\sigma_{n,k}^2 - b_{n,k}^2| = 0 \tag{2.9}$$

and (1.6), *then* (1.3) *holds.*

Indeed, (2.6) and (2.8) imply (2.3), while (2.4) is implied by (2.7) and (2.9).

For future reference we make the following remarks.

REMARK 2.1. Equation (2.3) is implied by

$$\sum_{k=1}^{k_n} |\mu_{n,k}| \xrightarrow[n\to\infty]{P} 0, \tag{2.10}$$

which implies also

$$\sum_{k=1}^{k_n} \mu_{n,k}^2 \xrightarrow[n\to\infty]{P} 0. \tag{2.11}$$

When (2.11) holds, then (2.4) is equivalent to

$$\sum_{k=1}^{k_n} E(X_{n,k}^2 | \mathscr{F}_{n,k-1}) \xrightarrow[n\to\infty]{P} 1. \tag{2.12}$$

2.3. We state next two versions of the asymptotic normality result which do not assume any moment conditions. The first is an immediate consequence of Theorem 2.2.

COROLLARY 2.2. *Let* $H_{n,k}$, for $n = 1, 2, \cdots$ *and* $k = 1, \cdots, k_n$, *be positive numbers, and put* $\bar{X}_{n,k} = X_{n,k} I(|X_{n,k}| \leqq H_{n,k})$. *If*

$$\lim_{n=\infty} P\left[\bigcup_{k=1}^{k_n} (|X_{n,k}| > H_{n,k})\right] = 0 \tag{2.13}$$

and the $\bar{X}_{n,k}$ *satisfy the conditions imposed on* $X_{n,k}$ *in Theorem* 2.2., *then* (1.3) *holds.*

Note that if $H_{n,k} \leqq H$, for $n = 1, 2, \cdots$ and $k = 1, \cdots, k_n$, then the ordinary Lindeberg condition can be expressed more simply. Indeed, if the random variables (1.1) are uniformly bounded, then (1.6) is equivalent to

$$\lim_{n=\infty} \sum_{k=1}^{k_n} P(|X_{n,k}| > \varepsilon) = 0 \qquad \text{for every} \quad \varepsilon > 0, \tag{2.14}$$

Our next extension of Theorem 2.2 is more interesting.

THEOREM 2.3. *Let the random variables* (1.1) *satisfy*

$$\sum_{k=1}^{k_n} P(|X_{n,k}| > \varepsilon | \mathscr{F}_{n,k-1}) \xrightarrow[n\to\infty]{P} 0 \qquad \text{for every} \quad \varepsilon > 0, \tag{2.15}$$

$$\sum_{k=1}^{k_n} E\left(\frac{X_{n,k}}{1 + X_{n,k}^2} | \mathscr{F}_{n,k-1}\right) \xrightarrow[n\to\infty]{P} 0, \tag{2.16}$$

$$\sum_{k=1}^{k_n} \left[E\left(\frac{X_{n,k}}{1 + X_{n,k}^2} | \mathscr{F}_{n,k-1}\right)\right]^2 \xrightarrow[n\to\infty]{P} 0, \tag{2.17}$$

and

$$\sum_{k=1}^{k_n} E\left(\frac{X_{n,k}^2}{1 + X_{n,k}^2} | \mathscr{F}_{n,k-1}\right) \xrightarrow[n\to\infty]{P} 1; \tag{2.18}$$

then (1.3) *holds.*

We note that the expectations in (2.16) and (2.17) exist. Notice also that (2.15) is implied by (2.14) and that (2.15) is equivalent to (2.5) when the random variables (1.1) are uniformly bounded.

It is worth commenting that the theorem no longer remains valid if (2.18) and (2.17) are replaced by the single condition that the difference of their left sides tend to 1.

2.4. After these general results we state a much simpler one. It involves a near constancy assumption about the characteristic functions of the random variables (1.1) and, as stated in Section 1.2, it is not new.

THEOREM 2.4. *Let* $X^*_{n,k}$, *for* $n = 1, 2, \cdots$ and $k = 1, \cdots, k_n$, *have the same distribution functions as the random variables* $X_{n,k}$ *of* (1.1) *and let the* $X^*_{n,k}$ *in each row be independent. If*

$$\lim_{n=\infty} \sum_{k=1}^{k_n} E\left|E\left(e^{itX_{n,k}} \mid \mathscr{F}_{n,k-1}\right) - Ee^{itX_{n,k}}\right| = 0. \tag{2.19}$$

then the sequence S_{n,k_n} *converges in distribution if and only if the sequence*

$$S^*_{n,k} = \sum_{k=1}^{k_n} X^*_{n,k_n} \tag{2.20}$$

converges in distribution. In the case of convergence the limits are the same.

For the applications in this paper we need only the sufficiency part of the following corollaries. We stated the theorem because it follows immediately from one of the lemmas in the sequel and illustrates one of the ways to derive necessary and sufficient conditions about convergence in distribution of sums of certain classes of dependent random variables.

COROLLARY 2.3. *Let the random variables* (1.1) *satisfy*

$$EX_{n,k} = 0, \tag{2.21}$$

$$\lim_{n=\infty} \sum_{k=1}^{k_n} EX^2_{n,k} = 1, \tag{2.22}$$

and (2.19). *Then* (1.3) *holds if and only if* (1.6) *is satisfied.*

COROLLARY 2.4. *Let the random variables* (1.1) *satisfy* (2.19), (2.21) *and*

$$\lim_{n=\infty} \sum_{k=1}^{k_n} EX^2_{n,k} = 0, \tag{2.23}$$

then S_{n,k_n} *converges in probability to* 0.

These two corollaries follow at once from the preceding theorem and from classical results on independent random variables.

3. Some lemmas

3.1. The first lemma is very simple but most useful as it permits replacing the conditioning by $\mathscr{B}(S_{n,k-1})$ in the preceding theorems by a finer conditioning relative to an increasing sequence of σ-fields.

LEMMA 3.1. *Let $S_1, S_2, \cdots, S_n$ be random variables defined on some probability space $(\Omega, \mathscr{A}, P)$. Then there exists a probability space $(\tilde{\Omega}, \tilde{\mathscr{A}}, \tilde{P})$ and random variables $\tilde{S}_1, \tilde{S}_2, \cdots, \tilde{S}_n$ defined on it such that*

$$\tilde{P}(\tilde{S}_k \leqq u, \tilde{S}_{k-1} \leqq v) = P(S_k \leqq u, S_{k-1} \leqq v), \qquad k = 2, \cdots, n, \tag{3.1}$$

for all real u, v, and

$$\tilde{P}(\tilde{S}_k | \mathscr{F}_{k-1}) = \tilde{P}(\tilde{S}_k | \mathscr{G}_{k-1}), \qquad k = 2, \cdots, n, \tag{3.2}$$

where $\mathscr{F}_k = \mathscr{B}(\tilde{S}_1, \cdots, \tilde{S}_k)$ and $\mathscr{G}_k = \mathscr{B}(\tilde{S}_k)$.

Condition (3.1) is equivalent to $\tilde{P}(\tilde{S}_1 \leqq u) = P(S_1 \leqq u)$ and

$$\tilde{P}(\tilde{S}_k \leqq u | \tilde{S}_{k-1} = v) = P(S_k \leqq u | S_{k-1} = v), \quad k = 2, \cdots, n, \tag{3.3}$$

while (3.2) asserts that the conditional distribution of $\tilde{S}_k$ given $\tilde{S}_{k-1}$ is unaffected by specifying the values assumed by $\tilde{S}_i$ with $i < k - 1$. This illustrates clearly the Markovian nature of the lemma.

PROOF. There is nothing to prove for $n = 1, 2$. Now let $n > 2$ and assume the lemma proved for $n - 1$. Denote by $S'_1, \cdots, S'_{n-1}$ and $(\Omega', \mathscr{A}', P')$ respectively the random variables and probability space whose existence is asserted by the lemma for $n - 1$ applied to $S_1, \cdots, S_{n-1}$ and $(\Omega, \mathscr{A}, P)$. Let R be the real line and $\mathscr{B}$ the Borel σ-field on R, and put $\tilde{\Omega} = \Omega' \times R$ and $\tilde{\mathscr{A}} = \mathscr{A}' \times \mathscr{B}$. Let $f(v, B)$, for $B \in \mathscr{B}$ and real v, be a regular version of the conditional probability $P(S_n \in B | S_{n-1} = v)$.

Define $\tilde{S}_k(\omega', r) = S'_k(\omega')$ for $k = 1, \cdots, n - 1$ and $\tilde{S}_n(\omega', r) = r$. Furthermore, let $\tilde{P}$ be defined by

$$\tilde{P}(A' \times R) = P'(A) \tag{3.4}$$

for every $A' \in \mathscr{A}'$ and

$$\tilde{P}(\tilde{S}_n \in B | \tilde{S}_1 = v_1, \cdots, \tilde{S}_{n-2} = v_{n-2}, \tilde{S}_{n-1} = v_{n-1}) = f(v_{n-1}, B) \tag{3.5}$$

for all $B \in \mathscr{B}$ and real $v_1, \cdots, v_{n-1}$.

It is easy to check that $\tilde{S}_1, \cdots, \tilde{S}_n$, $(\tilde{\Omega}, \tilde{\mathscr{A}}, \tilde{P})$ satisfy the conditions of the lemma. Indeed, we have only to verify (3.1) and (3.2) for $k = n$. Since $v_1, \cdots, v_{n-2}$ do not appear on the right side of (3.5), we at once have (3.2). Again by (3.5), $\tilde{P}(\tilde{S}_n \in B | \tilde{S}_{n-1} = v_{n-1}) = P(S_n \in B | S_{n-1} = v_{n-1})$. Taking $B = (-\infty, u]$ and integrating on v_{n-1} from $-\infty$ to v, we obtain (3.1) since by the induction assumption, $\tilde{P}(\tilde{S}_{n-1} \leqq v) = P'(S'_{n-1} \leqq v) = P(S_{n-1} \leqq v)$ for all v. This completes the proof.

REMARK 3.1. The lemma with the same proof holds also for infinite sequences of random variables and even for more general directed systems. Also $\mathscr{F}_k$ can be replaced by still larger σ-fields, for example by applying the lemma to random variables $T_1, S_1, \cdots, T_n, S_n$ and then disregarding $T_1, \cdots, T_n$.

REMARK 3.2. Let $X_1, \cdots, X_n$ be random variables; we shall usually apply the lemma to the sequence of partial sums $S_k = \Sigma_{j=1}^k X_j$ for $k = 1, \cdots, n$. Putting $\tilde{X}_k = \tilde{S}_k - \tilde{S}_{k-1}$ with $\tilde{S}_0 = 0$, conditions (3.1) and (3.2) are equivalent to the conditions obtained from them by replacing S_k and $\tilde{S}_k$ by X_k and $\tilde{X}_k$ (but keeping S_{k-1} and $\tilde{S}_{k-1}$ in (3.2)). Note also that $\mathscr{F}_k = \mathscr{B}(X_1, \cdots, X_k)$.

REMARK 3.3. For technical reasons it is very convenient to work with an increasing sequence of σ-fields, and this is precisely what Lemma 3.1 lets us do. When studying the distribution of a sum of random variables, $X_1 + \cdots + X_n$, it is clear that this distribution is determined by the conditional distributions of the summands relative to the sum of the preceding terms. The purport of our lemma is to remark that these conditional distributions can be preserved while introducing a Markovian structure. Thus, in particular, if the conditional expectations of X_k given S_{k-1} are almost surely zero, it may be assumed, when convenient, that $X_1, \cdots, X_n$ is a martingale difference sequence.

3.2. In the next lemma it is important that the σ-fields be increasing. It should be noted that $\mathscr{F}_0$ need not be the trivial field (though in the applications given in the present paper it will be). We denote by $N(\cdot, \cdot)$ the normal distribution with the indicated mean and variance.

LEMMA 3.2. *Let $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \cdots \subset \mathscr{F}_n$ be σ-fields in a probability space. Let $\sigma_1, \cdots, \sigma_n$ be nonnegative random variables with σ_k measurable $\mathscr{F}_{k-1}$, and let Y_k, for $k = 1, \cdots, n$ be $N(0, 1)$ mutually independent and independent of $\mathscr{F}_n$. Then if*

$$\sigma^2 = \sigma_1^2 + \cdots + \sigma_n^2 \tag{3.6}$$

is $\mathscr{F}_0$ measurable, we have

$$\sum_{j=k}^{n} \sigma_j Y_j = \left(\sum_{j=k}^{n} \sigma_j^2\right)^{1/2} Z_k, \qquad k = 1, \cdots, n, \tag{3.7}$$

with Z_k also $N(0, 1)$ and independent of $\mathscr{F}_n \times \mathscr{B}(Y_1, \cdots, Y_{k-1})$.

In particular if σ^2 is almost surely a constant, then $\Sigma_{j=1}^n \sigma_j Y_j$ is $N(0, \sigma^2)$.

PROOF. We proceed by induction on k from $k = n$ to $k = 1$. For $k = n$ there is nothing to prove. Let $k < n$ and assume the validity of (3.7) for $k + 1$. Then

$$\sum_{i=k}^{n} \sigma_i Y_i = \sigma_k Y_k + \sum_{i=k+1}^{n} \sigma_i Y_i = \sigma_k Y_k + \left(\sigma^2 - \sum_{i=1}^{k} \sigma_i^2\right)^{1/2} Z_{k+1} \tag{3.8}$$

and, by the $\mathscr{F}_0$ measurability of σ^2, both σ_k^2 and $\sigma^2 - (\sigma_1^2 + \cdots + \sigma_k^2)$ are $\mathscr{F}_{k-1}$ measurable. Also, given $\mathscr{F}_n$, the random variable Y_k is $N(0, 1)$ and Z_{k+1} is $N(0, 1)$ and independent of it. This yields (3.7) with Z_k independent of $\mathscr{F}_n$ as required. Formally,

$$\begin{aligned} & P\left[\sigma_k Y_k + \left(\sum_{j=k+1}^{n} \sigma_j^2\right)^{1/2} Z_{k+1} \leqq u \,\middle|\, \mathscr{F}_n\right] \\ & = P\left[\sigma_k Y_k + \left(\sum_{j=k+1}^{n} \sigma_j^2\right)^{1/2} Z_{k+1} \leqq u \,\middle|\, \sigma_k^2, \sum_{j=k+1}^{n} \sigma_j^2\right] \\ & = P\left[\left(\sum_{j=k}^{n} \sigma_j^2\right)^{1/2} Z \leqq u\right] \end{aligned} \tag{3.9}$$

with Z standard normal.

REMARK 3.4. The condition that σ^2 given by (3.7) be $\mathscr{F}_0$ measurable is essential. Otherwise the lemma fails entirely; indeed, the data do not determine the distribution of $\sigma_k Y_k + \cdots + \sigma_n Y_n$ for $k < n$.

A very simple example is constructed as follows: Let $n = 2$, $\mathscr{F}_0$ be the trivial field, $\sigma_1 = 1$, let $\mathscr{F}_1$ be the σ-field generated by the standard normal variable Y_1, and let us consider two choices of σ_2: (a) $\sigma_2 = 1$ if $Y_1 \geqq 0$ and 0 otherwise, and (b) $\sigma_2 = 0$ if $Y_1 \geqq 0$ and 1 otherwise. In both cases, (a) and (b), we have $P(\sigma^2 = 1) = P(\sigma^2 = 2) = 1/2$. But it is easily checked that in case (a) we have $P(\sigma_1 Y_1 + \sigma_2 Y_2 > 0) < 1/2$, while in case (b) we have $P(\sigma_1 Y_1 + \sigma_2 Y_2 > 0) > 1/2$.

3.3. The next lemma is somewhat curious.

LEMMA 3.3. *Let X be an integrable random variable and $\mathscr{F}$ a σ-field in the probability space and let $\mu = E(X|\mathscr{F})$. Then we have for every $\varepsilon > 0$*

$$4E[X^2 I(|X| > \varepsilon)] \geqq E[(X - \mu)^2 I(|X - \mu| > 2\varepsilon)]. \tag{3.10}$$

In particular

$$4E[X^2 I(|X| > \varepsilon)] \geqq E[(X - EX)^2 I(|X - EX| > 2\varepsilon)], \tag{3.11}$$

and the constant 2 is best possible.

PROOF. We first prove (3.11). This is equivalent to proving that if $EY = 0$ and c is any constant, then

$$4E[(Y + c)^2 I(|Y + c| > \varepsilon)] \geqq E[Y^2 I(|Y| > 2\varepsilon)]. \tag{3.12}$$

Equation (3.12) is easily checked for random variables Y satisfying $P(Y = q) = p$ and $P(Y = -p) = q$, where $0 \leqq p \leqq 1$ and $p + q = 1$. Also the remark about the constant 2 is verified for, say, the case $p = q$.

Since the distribution function of any random variable Y with $EY = 0$ can be approximated by $w_1 F_1 + \cdots + w_n F_n$ where $w_j > 0$ and $w_1 + \cdots w_n = 1$ and F_j is the distribution function of the random variable assuming the values $a_j q_j$ and $-a_j p_j$ with probabilities p_j and q_j respectively ($0 < p_j < 1, p_j + q_j = 1$), the validity of (3.12), in general, follows from the special case.

To complete the proof of (3.10) it is enough to establish it for the case when F is generated by a finite number of atoms A_j. Let $\mu_j = E(X|A_j)$; applying (3.12) to $XI(A_j)$ we have

$$4E[X^2 I(|X| > \varepsilon) I(A_j)] \geqq E[(X - \mu_j)^2 I(|X - \mu_j| > 2\varepsilon) I(A_j)]. \tag{3.13}$$

Adding up for all j, we obtain the required result.

3.4. Our next lemma is on characteristic functions. It can be traced at least as far back as Lindeberg and may be essentially found in [11] or [13] for example.

LEMMA 3.4. *Let $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \cdots \subset \mathscr{F}_n$ be σ-fields in a probability space. Let the random variables X_k, for $k = 1, \cdots, n$, be $\mathscr{F}_k$ measurable. Let the random variables Z_k, for $k = 1, \cdots, n$, be such that $Z_{k+1} + \cdots + Z_n$ is measurable relative to a σ-field $\mathscr{F}_{k-1} \times \mathscr{H}_{k-1}$, with $\mathscr{H}_{k-1}$ and $\mathscr{F}_{k-1} \times \mathscr{B}(Z_k)$ independent* *for* $k = 1, \cdots, n$. *Then*

$$(3.14)\qquad \left|E \exp\left\{it \sum_{k=1}^{n} X_k\right\} - E \exp\left\{it \sum_{k=1}^{n} Z_k\right\}\right| \leqq \sum_{k=1}^{n} E|E([\exp\{itX_k\} - \exp\{itZ_k\}]|\mathscr{F}_{k-1})|$$

for all real t.

PROOF. Putting

$$(3.15)\qquad S_k = \sum_{j=1}^{k} X_j, \qquad R_k = \sum_{j=k+1}^{n} Z_j,$$

we have

$$(3.16)\qquad \begin{aligned}\exp\{itS_n\} - \exp\{itR_0\} &= \sum_{k=1}^{n} [\exp\{it(S_k + R_k)\} - \exp\{it(S_{k-1} + R_{k-1})\}] \\ &= \sum_{k=1}^{n} \exp\{it(S_{k-1} + R_k)\}(\exp\{itX_k\} - \exp\{itZ_k\})\end{aligned}$$

and, therefore,

$$(3.17)\qquad \begin{aligned}&|E \exp\{itS_n\} - E \exp\{itR_0\}| \\ &= \sum_{k=1}^{n} E(\exp\{it(S_{k-1} + R_k)\} E(\exp\{itX_k\} - \exp\{itZ_k\}|\mathscr{F}_{k-1} \times \mathscr{H}_{k-1})) \\ &\leqq \sum_{k=1}^{n} E|E(\exp\{itX_k\} - \exp\{itZ_k\}|\mathscr{F}_{k-1} \times \mathscr{H}_{k-1})|\end{aligned}$$

and since X_k and Z_k are independent of $\mathscr{H}_{k-1}$, we obtain (3.14) on noticing that the inner expectations in (3.14) and (3.17) are equal.

3.5. We need the next result only in the proof of Theorem 2.3.

LEMMA 3.5. *Let* $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \cdots \subset \mathscr{F}_n$ *be σ-fields in a probability space, and let* $A_k \in \mathscr{F}_k$. *Then we have for every* $\varepsilon > 0$

$$(3.18)\qquad P\left(\bigcup_{k=1}^{n} A_k\right) \leqq \varepsilon + P\left[\sum_{k=1}^{n} P(A_k|\mathscr{F}_{k-1}) > \varepsilon\right].$$

PROOF. We shall prove the somewhat stronger result

$$(3.19)\qquad P\left(\bigcup_{k=1}^{n} A_k|\mathscr{F}_0\right) \leqq \varepsilon + P\left[\sum_{k=1}^{n} P\left(A_k|\mathscr{F}_{k-1}\right) > \varepsilon|\mathscr{F}_0\right]$$

where ε is any nonnegative $\mathscr{F}_0$ measurable function.

For $n = 1$ the assertion becomes

$$(3.20)\qquad P(A_1|\mathscr{F}_0) \leqq \varepsilon + P[P(A_1|\mathscr{F}_0) > \varepsilon|\mathscr{F}_0],$$

which is obvious, since if $P(A_1|\mathscr{F}_0) > \varepsilon$ the second summand on the right is 1.

Next we assume $n > 1$ and the lemma established for $n - 1$ sets. Clearly,

$$P\left(\bigcup_{k=1}^{n} A_k | \mathscr{F}_0\right) \leqq P(A_1|\mathscr{F}_0) + E\left[P\left(\bigcup_{k=2}^{n} A_k|\mathscr{F}_1\right)|\mathscr{F}_0\right]. \tag{3.21}$$

By the induction assumption

$$P\left(\bigcup_{k=2}^{n} A_k | \mathscr{F}_1\right) \leqq \delta + P\left[\sum_{k=2}^{n} P(A_k|\mathscr{F}_{k-1}) > \delta | \mathscr{F}_1\right] \tag{3.22}$$

where δ is any nonnegative $\mathscr{F}_1$ measurable function.

Note that, as in the case of (3.20), relation (3.19) is obvious when $P(A_1|\mathscr{F}_0) > \varepsilon$. Now let $\delta = \varepsilon - P(A_1|\mathscr{F}_0)$ if $P(A_1|\mathscr{F}_0) \leqq \varepsilon$, and $\delta = 0$ otherwise. In view of the remark just made it is enough to consider the first possibility. Then the right side of (3.22) becomes

$$\varepsilon - P(A_1|\mathscr{F}_0) + P\left[\sum_{k=1}^{n} P(A_k|\mathscr{F}_{k-1}) > \varepsilon|\mathscr{F}_1\right], \tag{3.23}$$

and using this estimate in (3.21) we obtain (3.19).

4. Proofs of the main results

Because of Lemma 3.1 we may assume that the σ-fields defined by (1.2) satisfy

$$\mathscr{F}_{n,1} \subset \cdots \subset \mathscr{F}_{n,k-1} \subset \mathscr{F}_{n,k} \subset \cdots \subset \mathscr{F}_{n,k_n}. \tag{4.1}$$

We denote by $\mathscr{F}_{n,0}$ the trivial field. ($X_{n,k}$ is assumed to be $\mathscr{F}_{n,k}$ measurable, but $\mathscr{B}$ $(X_{n,1}, \cdots, X_{n,k})$ may be a proper subfield of $\mathscr{F}_{n,k}$).

4.1. PROOF OF THEOREM 2.1. Let $Y_{n,k}$ be $N(0, 1)$ independent and independent of $\mathscr{F}_{n,k_n}$. (No loss of generality is involved in the assumption that there exist such $Y_{n,k}$, since it is always possible to imbed the probability spaces in larger ones having the required properties. This applies whenever similar assumptions are made in the sequel.) Then, by Lemma 3.2,

$$\sum_{j=k+1}^{k_n} \sigma_{n,j} Y_{n,j} \tag{4.2}$$

are measurable $\mathscr{F}_{n,k-1} \times \mathscr{B}(Y_{n,k+1}, \cdots, Y_{n,k_n})$ and hence we can apply Lemma 3.4 to $X_{n,k}$ with Z_k replaced by $\sigma_{n,k} Y_{n,k}$. Since, again by Lemma 3.2, (4.2) is $N(0, 1)$ for $k = 0$, we have from (3.14):

$$\begin{aligned} &|E \exp\{itS_{n,k_n}\} - \exp\{-\tfrac{1}{2}t^2\}| \\ &\quad \leqq \sum_{k=1}^{k_n} E|E(\exp\{itX_{n,k}\} - \exp\{it\sigma_{n,k}Y_{n,k}\}|\mathscr{F}_{n,k-1})|. \end{aligned} \tag{4.3}$$

It remains only to show that for every real t the sum in (3.14) tends to 0 as $n \to \infty$; but this is quite easy. Using standard inequalities we have, by (2.1),

$$(4.4)\quad \begin{aligned}&|E[\exp\{itX_{n,k}\}\,|\,\mathscr{F}_{n,k-1}] - 1 - \tfrac{1}{2}t^2\sigma_{n,k}^2|\\ &\leqq \tfrac{1}{6}|t|^3E(|X_{n,k}|^3 I(|X_{n,k}| \leqq \varepsilon)|\mathscr{F}_{n,k-1}) + t^2E(X_{n,k}^2 I(|X_{n,k}| > \varepsilon)|\mathscr{F}_{n,k-1})\\ &\leqq \tfrac{1}{6}\varepsilon t^2\sigma_{n,k}^2 + t^2E[X_{n,k}^2 I(|X_{n,k}| > \varepsilon)|\mathscr{F}_{n,k-1}]\end{aligned}$$

and thus, by (2.2),

$$(4.5)\quad \sum_{k=1}^{k_n} |E(\exp\{itX_{n,k}\} - 1 - \tfrac{1}{2}t^2\sigma_{n,k}^2|\mathscr{F}_{n,k-1})| \leqq \tfrac{1}{6}\varepsilon t^2 + t^2 \sum_{k=1}^{k_n} E[X_{n,k}^2 I(|X_{n,k}| > \varepsilon)].$$

Since $\varepsilon > 0$ is arbitrary, it follows from (1.6) that the left side of (4.5) approaches 0 as $n \to \infty$. A similar estimate holds when $X_{n,k}$ is replaced by $\sigma_{n,k}Y_{n,k}$, or we can use the better estimate, following from the normality of $Y_{n,k}$ and its independence of $\mathscr{F}_{n,k-1}$,

$$(4.6)\quad \sum_{k=1}^{k_n} E|E(\exp\{it\sigma_{n,k}Y_{n,k}\} - 1 - \tfrac{1}{2}t^2\sigma_{n,k}^2|\mathscr{F}_{n,k-1})| \leqq 3t^4 \sum_{k=1}^{k_n} E\sigma_{n,k}^4 \leqq 3t^4E(\max_{1\leqq k\leqq k_n} \sigma_{n,k}^2)$$

and the fact that the last expectation tends to zero by (1.6). Since the sum on the right in (4.3) is less than the sum of the left sides of (4.5) and (4.6), this completes the proof.

4.2. PROOF OF THEOREM 2.2. First we remark that in the proof of Theorem 2.1 we could have used the weaker condition (2.5) instead of (1.6). Indeed, (1.6) was used only to show that the left sides of (4.5) and (4.6) approach 0 as $n \to \infty$. But (4.5) and (4.6) remain valid when the expectation on the right is replaced by $E(\cdot|\mathscr{F}_{n,k-1})$ and the approach to zero is already entailed by (2.5).

Next we retain (2.1) and show that Theorem 2.1 still holds when (2.2) is replaced by (2.4).

Let t_n be the stopping time defined by

$$(4.7)\quad t_n = \min\left[k_n, \max\left(k : \sum_{j=1}^{k} \sigma_{n,j}^2 \leqq 1\right)\right],$$

thus $t_n = k_n$ when (2.2) holds. Put

$$(4.8)\quad \begin{aligned}\tilde{X}_{n,k} &= X_{n,k} &&\text{if } k \leqq t_n\\ \tilde{X}_{n,k} &= 0 &&\text{otherwise,}\end{aligned}$$

and

$$(4.9)\quad \tilde{X}_{n,k_n+1} = \left(1 - \sum_{j=1}^{t_n} \sigma_{n,j}^2\right)Y$$

where Y is $N(0, 1)$ and independent of $\mathscr{F}_{n,k_n}$. The important thing about Y is not its normality but that

$$(4.10)\quad E(\tilde{X}_{n,k_n+1}|\mathscr{F}_{n,k_n}) = 0, \qquad E(\tilde{X}_{n,k_n+1}^2|\mathscr{F}_{n,k_n}) = 1 - \sum_{j=1}^{t_n} \sigma_{n,j}^2.$$

The $\tilde{X}_{n,k}$, for $n = 1, 2, \cdots$ and $k = 1, \cdots, k_n + 1$, obviously satisfy the conditions of Theorem 1.1 about the conditional means and variances. We proceed to show that they also satisfy the remaining condition, that is the Lindeberg condition. Since by (4.8) $|\tilde{X}_{n,k}| \leqq |X_{n,k}|$ for $k \leqq k_n$, it suffices to show that

$$\lim_{n=\infty} E\tilde{X}^2_{n,k_n+1} = 0. \tag{4.11}$$

But, by (4.10),

$$E\tilde{X}^2_{n,k_n+1} = 1 - E\sum_{j=1}^{t_n} \sigma^2_{n,j} \tag{4.12}$$

and (4.11) would follow from

$$\sum_{j=1}^{t_n} \sigma^2_{n,j} \xrightarrow[n\to\infty]{P} 1, \tag{4.13}$$

since the left side of (4.13) is $\leqq 1$ by (4.7). This, again by (4.7), is implied by

$$\sigma^2_{n,t_n+1} \xrightarrow[n\to\infty]{P} 0 \tag{4.14}$$

(where the left side is taken as 0 when $t_n = k_n$). But (4.14) follows at once from (1.6). Now

$$S_{n,k_n} = \sum_{j=1}^{k_n+1} \tilde{X}_{n,j} + \sum_{j=t_n+1}^{k_n} X_{n,j} - \tilde{X}_{n,k_n+1}, \tag{4.15}$$

and, by Theorem 2.1, the first sum on the right is asymptotically $N(0, 1)$.

Let t_n' be defined similarly to t_n except for replacing the $\leqq 1$ at the end of (4.8) by $\leqq 2$. Then

$$\sum_{j=t_n+1}^{k_n} X_{n,j} = \sum_{j=t_n+1}^{t_n'} X_{n,j} + \sum_{j=t_n'+1}^{k_n} X_{n,j}. \tag{4.16}$$

We have

$$E\left(\sum_{j=t_n+1}^{t_n'} X_{n,j}\right)^2 = E\sum_{j=t_n+1}^{t_n'} \sigma^2_{n,j} \tag{4.17}$$

and the right side tends to 0 as $n \to \infty$ by (2.4), (4.13), and the boundedness of the sum. Therefore, the first summand on the right in (4.16) converges in probability to 0 as $n \to \infty$. The same is true of the last summand, since $P(t_n' \neq k_n) \to 0$ as $n \to \infty$ by (2.4). Therefore the left side of (4.16) also converges in probability to zero as $n \to \infty$. Since, by (4.11), the same is true of $\tilde{X}_{n,k_n+1}$, it follows from (4.15) that S_{n,k_n} is asymptotically $N(0, 1)$. This completes the proof of Theorem 2.2 with (2.3) replaced by (2.1).

To complete the proof of the theorem it remains to show that, under the assumptions of Theorem 2.2, the random variables $\bar{X}_{n,k} = X_{n,k} - \mu_{n,k}$ satisfy the conditions of the case which we have already proved. The only point that requires proof is that (2.5) still holds when $X_{n,k}$ is replaced by $\bar{X}_{n,k}$; but this is an immediate consequence of Lemma 3.3.

4.3. PROOF OF THEOREM 2.3. Let

$$\bar{X}_{n,k} = X_{n,k} I(|X_{n,k}| \leqq 1), \qquad \bar{S}_{n,k} = \sum_{j=1}^{k} \bar{X}_{n,j}. \tag{4.18}$$

By (2.15) and Lemma 3.5, $P(S_{n,k_n} \neq \bar{S}_{n,k_n})$ approaches 0 as $n \to \infty$. Thus it suffices to prove that the $\bar{X}_{n,k}$ satisfy the conditions imposed on $X_{n,k}$ in Theorem 2.2.

Let $\bar{\mu}_{n,k} = E(\bar{X}_{n,k} | \mathscr{F}_{n,k-1})$, then

$$\left|\bar{\mu}_{n,k} - E\left(\frac{X_{n,k}}{1 + X_{n,k}^2} | \mathscr{F}_{n,k-1}\right)\right| \tag{4.19}$$

$$= \left|E\left(\frac{X_{n,k}^3 I(|X_{n,k}| \leqq 1)}{1 + X_{n,k}^2} | \mathscr{F}_{n,k-1}\right) + E\left(\frac{X_{n,k} I(|X_{n,k}| > 1)}{1 + X_{n,k}^2} | \mathscr{F}_{n,k-1}\right)\right|$$

$$\leqq \varepsilon E\left(\frac{X_{n,k}^2}{1 + X_{n,k}^2} | \mathscr{F}_{n,k-1}\right) + P(|X_{n,k}| > \varepsilon | \mathscr{F}_{n,k-1})$$

for every $\varepsilon > 0$. From (2.15) and (2.18) it follows then that the $\bar{X}_{n,k}$ satisfy the first condition of Theorem 2.3.

In the same way we obtain a similar estimate for

$$\left|E(\bar{X}_{n,k}^2 | \mathscr{F}_{n,k-1}) - E\left(\frac{X_{n,k}^2}{1 + X_{n,k}^2} | \mathscr{F}_{n,k-1}\right)\right| \tag{4.20}$$

from which we deduce

$$\sum_{k=1}^{n_k} E(\bar{X}_{n,k}^2 | \mathscr{F}_{n,k-1}) \xrightarrow[n \to \infty]{P} 1. \tag{4.21}$$

Also

$$\left|\bar{\mu}_{n,k}^2 - \left[E\left(\frac{X_{n,k}}{1 + X_{n,k}^2} | \mathscr{F}_{n,k-1}\right)\right]^2\right| \tag{4.22}$$

is at most twice the left side of (4.19) (since $|\bar{\mu}_{n,k}| \leqq 1$ and similarly for the other term). From this, (2.17) and (4.21) we conclude that the second condition of Theorem 2.2 is satisfied (see Remark 2.1).

Finally, the last condition is obtained from the boundedness of $\bar{X}_{n,k}$, Equation (2.14) and Lemma 2.3.

4.4. PROOF OF THEOREM 2.4. This follows immediately from Lemma 3.4 on taking, as Z_k, random variables having the same distributions as $X_{n,k}$ but independent of $\mathscr{F}_{n,k_n}$ and of one another.

5. Some applications

5.1. Our next results will deal with sums of random variables whose dependence diminishes as their distance increases. Such results can be formulated for any suitable measure of dependence. We shall formulate our results in Sections 5.2 and 5.3 in terms of the quantity α given by the following definition.

DEFINITION 5.1. *The α dependence of two fields $\mathscr{F}$ and $\mathscr{G}$ in a probability space is defined by*

$$\alpha(\mathscr{F}, \mathscr{G}) = \sup |P(F \cap G) - P(F)P(G)|, \tag{5.1}$$

the supremum being taken over all sets $F \in \mathscr{F}$ and $G \in \mathscr{G}$.

We refer to [8] and [13] for other notions of dependence and the relations among them. In general it is less restrictive for the α dependence to be small than for most other measures of dependence which were studied.

We need some lemmas on $\alpha(\mathscr{F}, \mathscr{G})$, but first we show

LEMMA 5.1. *Let X, Y be random variables satisfying $|X| \leqq 1$, $|Y| \leqq 1$ and let*

$$\Delta = \sup_B |P(X \in B) - P(Y \in B)|, \tag{5.2}$$

the supremum being taken over all Borel sets B. Then

$$|EX - EY| \leqq 2\Delta \tag{5.3}$$

and the constant 2 is exact.

PROOF. Let $\nu(B) = P(X \in B) - P(Y \in B)$ for all Borel sets B. Then ν is a signed measure and

$$|EX - EY| = \left|\int t\nu(dt)\right| \leqq \int |t|\,|\nu|\,(dt) \leqq \int |\nu|\,(dt) \tag{5.4}$$

where $|\nu|\,(B)$ is the total variation of ν on B. Let B^+ and B^- be a Jordan decomposition of $[-1, 1]$ corresponding to ν; then $|\nu|\,([-1, 1]) = 2\nu(B^+)$ with $\nu(B^+) = \Delta$ as defined by (5.2). Substituting in (5.4) we obtain (5.3).

Using this lemma we obtain

LEMMA 5.2. *Let X be a random variable satisfying $|X| \leqq 1$, let $\mathscr{F} = \mathscr{B}(X)$ and $\mathscr{G}$ be any σ-field in the probability space, then*

$$E|E(X|\mathscr{G}) - EX| \leqq 4\alpha(\mathscr{F}, \mathscr{G}) \tag{5.5}$$

and the constant 4 is exact.

PROOF. Let G denote the set where $E(X|\mathscr{G}) \geqq EX$, and G' its complement; then $G \in \mathscr{G}$ and we have

$$\begin{aligned} 0 &= E[E(X|\mathscr{G}) - EX] \\ &= E[E(X|\mathscr{G}) - EX|G]P(G) + E[E(X|\mathscr{G}) - EX|G']P(G'), \end{aligned} \tag{5.6}$$

$$\begin{aligned} &E|E(X|\mathscr{G}) - EX| \\ &= E[E(X|\mathscr{G}) - EX|G]P(G) - E[E(X|\mathscr{G}) - EX|G']P(G'). \end{aligned} \tag{5.7}$$

Combining these two equations we obtain

$$(5.8)\qquad E|E(X|\mathscr{G}) - EX| = 2E[E(X|\mathscr{G}) - EX|G]P(G).$$

Now let $\tilde{X}$ be a random variable with the same distribtuion as X and independent of $\mathscr{G}$ (if need be the probability space can be enlarged in order to carry such a random variable). Then

$$(5.9)\qquad E[E(X|\mathscr{G}) - EX|G] = E(X|G) - E(\tilde{X}|G) \leqq 2 \sup_B |P(X \in B|G) - P(\tilde{X} \in B|G)|$$

by Lemma 5.1, the sup being taken over all Borel sets. Writing the conditional probabilities in (5.8) in full, recalling that $P[(\tilde{X} \in B) \cap G] = P(\tilde{X} \in B)P(G) = P(X \in B)P(G)$, and substituting in (5.9), we obtain (5.5).

That 4 is exact is seen, for example, by letting X assume the values ± 1 only and $\mathscr{G} = \{\varnothing, G, G', \Omega\}$.

We need a similar lemma for complex valued random variables.

LEMMA 5.3. *Let ξ be a complex valued random variable satisfying $|\xi| \leqq 1$, let $\mathscr{F} = \mathscr{B}(\xi)$ and $\mathscr{G}$ be a σ-field in the probability space. Then*

$$(5.10)\qquad E|E(\xi|\mathscr{G}) - E\xi| \leqq 2\pi\alpha(\mathscr{F}, \mathscr{G}).$$

PROOF. Put $\eta = E(\xi|\mathscr{G}) - E\xi$; then we have for every real u,

$$(5.11)\qquad E|\eta| = E\left[\frac{1}{2}\int_0^\pi |\mathrm{Re}(\eta e^{iu})|\, du\right] = \frac{1}{2}\int_0^\pi E|\mathrm{Re}(\eta e^{iu})|\, du \leqq \frac{\pi}{2} \sup_u |\mathrm{Re}(\eta e^{iu})|$$

by Fubini's theorem and a trivial estimation of the integral. Applying (5.5) with $X = \mathrm{Re}(\xi e^{iu})$ and substituting in (5.11) we obtain (5.10).

We also need the following consequence of Lemma 5.2.

LEMMA 5.4. *Let X and Y be random variables with $|X| \leqq c$ and $E(X) = 0$. Let $\mathscr{F} = \mathscr{B}(X)$ and $\mathscr{G} = \mathscr{B}(Y)$; then*

$$(5.12)\qquad |E(XY)| \leqq 4cE|Y|\alpha(\mathscr{F}, \mathscr{G}).$$

Indeed, by (5.5),

$$(5.13)\qquad |E(XY)| = |E[YE(X|\mathscr{G})]| \leqq E[4c\alpha(\mathscr{F}, \mathscr{G})|Y|].$$

5.2. The next theorem has many antecedents starting with Bernstein [1]. It is a little clumsy, but it can be easily specialized to give various results on m dependence or stationarity (see [2], [7], [8], [12], [13], [14], [15]).

THEOREM 5.1. *Let the random variables* 1.1 *satisfy $EX_{n,k} = 0$ and put*

$$(5.14)\qquad \alpha_n(m) = \sup_{1 \leqq k < k_n - m} \alpha(\mathscr{F}_{n,k}, \mathscr{G}_{n,k+m+1})$$

where $\mathscr{F}_{n,k} = \mathscr{B}(X_{n,1}, \cdots, X_{n,k})$ and *$\mathscr{G}_{n,k} = \mathscr{B}(X_{n,k}, \cdots, X_{n,k_n})$. Assume there* *exist integers*

(5.15) $$0 = j_n(0) < j_n(1) < \cdots < j_n(r_n) = k_n$$

such that, putting

(5.16) $$Y_{n,i} = \sum_{k=j_n(i-1)+1}^{j_n(i)} X_{n,k},$$

we have

(5.17) $$\lim_{n=\infty} \sum_{i\,\text{even}} EY_{n,i}^2 = 0,$$

(5.18) $$\lim_{n=\infty} \sum_{i\,\text{odd}} EY_{n,i}^2 = 1,$$

and

(5.19) $$\lim_{n=\infty} \sum_{i=1}^{r_n} E[Y_{n,i}^2 I(|Y_{n,i}| > \varepsilon)] = 0 \qquad \text{for all} \quad \varepsilon > 0.$$

Then

(5.20) $$\lim_{n=\infty} r_n \alpha_n(m_n) = 0,$$

where $m_n = \min_{1<i<k_n} [j_n(i) - j_n(i-1)]$ *implies that* (1.3) *holds.*

PROOF. Because of equation (5.20) and Lemma 5.3 we have

(5.21) $$\sum_{k\,\text{even}} E|E(\exp\{itY_{n,k}\} | \mathscr{F}_{n,j_n(k-1)}) - E\exp\{itY_{n,k}\}| < 2\pi r_n \alpha_n(m_n).$$

Therefore, by (5.20) and (5.17), Corollary 2.4 applies to $\Sigma\, Y_{n,2i}$ and this sum converges in probability to zero. But (5.21) holds also when the summation is done on the odd k. Therefore, by (5.18) and (5.19), Corollary 2.3 applies to $\Sigma\, Y_{n,2i-1}$. Combining these two results we obtain the assertion of the theorem.

5.3. Here we shall be more specific and obtain a more easily applicable result. Instead of dealing with many dependence functions $\alpha_n(m)$ we shall, for simplicity, operate with a single one $\alpha(m)$, satisfying

(5.22) $$\alpha(m) \geqq \alpha(m+1), \qquad m = 1, 2, \cdots,$$

and

(5.23) $$\alpha(m) \geqq \sup_n \alpha_n(m), \qquad m = 1, 2, \cdots,$$

where $\alpha_n(m)$ is given by (5.14) and the sup is taken over all n for which $k_n > m$.

THEOREM 5.2. *Let the random variables* (1.1) *satisfy*

(5.24) $$|X_{n,k}| \leqq c_n, \qquad \lim_{n=\infty} c_n = 0,$$

(5.25) $$EX_{n,k} = 0, \qquad \lim_{n=\infty} ES_{n,k_n}^2 = 1.$$

Assume there exist positive integers m_n, $n = 1, 2, \cdots$, *satisfying*

(5.26) $$\lim_{n=\infty} m_n = \infty, \qquad \lim_{n=\infty} m_n c_n = 0$$

to which there correspond sequences (5.15) *with*

$$(5.27)\qquad j_n(2i) - j_n(2i-1) = m_n, \qquad EY^2_{n,2i-1} \geqq \frac{1}{m_n},$$

for which (5.19) *holds. Then the condition*

$$(5.28)\qquad \sum_{m=1}^{\infty} \alpha(m) < \infty$$

implies (1.3).

W. Philipp [13] has a similar result, however, with (5.28) replaced by the stronger condition $\Sigma\,[\alpha(m)]^{1/2} < \infty$. For a discussion of various specializations, see [13].

PROOF. All we have to do is to check that conditions (5.17), (5.18) and (5.20) of Theorem 5.1 are satisfied.

First we estimate

$$(5.29)\qquad \begin{aligned} EY^2_{n,2i} &= \sum_{k=j_n(2i-1)+1}^{j_n(2i)} EX^2_{n,k} + 2 \sum_{k=j_n(2i-1)+1}^{j_n(2i)} E\left(X_{n,k} \sum_{p=k+1}^{j_n(2i)} X_{n,p}\right) \\ &\leqq m_n c_n^2 + 8 m_n c_n^2 \sum_{m=1}^{m_n} \alpha(m) < K_1 m_n c_n^2 \end{aligned}$$

by Lemma 5.4 and (5.28). (The $K_1, K_2, \cdots$ denote finite positive constants.)

The same computation also gives

$$(5.30)\qquad |\mathrm{E}(Y_{n,2i} \sum_{k>1} Y_{n,2k})| < K_1 m_n c_n^2.$$

Thus, writing

$$(5.31)\qquad Y_n' = \sum Y_{n,2i-1}, \qquad Y_n'' = \sum Y_{n,2i}$$

we have

$$(5.32)\qquad E(Y_n'')^2 < K_2 r_n m_n c_n^2$$

with r_n defined in (5.15).

We also have

$$(5.33)\qquad \begin{aligned} |E(Y_{n,2i-1} \sum_{k>i} Y_{n,2k-1})| &\leqq \sum_{k>j_n(2i)} |E(Y_{n,2i-1} X_{n,k})| \\ &\leqq 8 c_n E|Y_{n,2i-1}| \sum_{m=1}^{\infty} \alpha(m) < K_3 c_n (EY^2_{n,2i-1})^{1/2}, \end{aligned}$$

again by Lemma 5.4. The last member is $o(EY^2_{n,2i-1})$ as $n \to \infty$, uniformly in i, by (5.26) and (5.27). Therefore

$$(5.34)\qquad E(Y_n')^2 = [1 + o(1)] \sum EY^2_{n,2i-1}.$$

On the other hand we have from (5.29) and (5.30), again on account of (5.26) and (5.27), that

$$(5.35)\qquad E(Y_n'')^2 = o(\sum EY^2_{n,2i-1}).$$

From (5.34), (5.35), and (5.25) we obtain (5.18). From (5.18) and (5.27) we have $r_n < K_4 m_n$, which by (5.29) yields (5.17). Finally, (5.20) follows from $r_n < K_4 m_n$, (5.22), (5.28), and (5.26).

5.4. As a direct application of Theorem 2.2 we bring the following result which improves a result of Serfling [16] under somewhat weaker conditions.

THEOREM 5.3. *Let X_n, for $n = 1, 2, \cdots$, be a sequence of random variables. Put $S_n = X_1 + \cdots + X_n$ and, for nonnegative integers a, let*

$$T_a(m) = \frac{1}{\sqrt{m}}(X_{a+1} + \cdots + X_{a+m}) \tag{5.36}$$

and $\mathscr{F}_a = \mathscr{B}(S_a)$. If the following four conditions

$$\lim_{m=\infty} ET_a(m)^2 = 1, \tag{5.37}$$

$$\lim_{m=\infty} m^{-\gamma} E|T_a(m)|^{2+\beta} = 0, \tag{5.38}$$

$$\lim_{m=\infty} m^{\theta} E|E(T_a(m)|\mathscr{F}_a)| = 0, \tag{5.39}$$

and

$$\lim_{m=\infty} E|E[T_a(m)^2|\mathscr{F}_a] - ET_a(m)^2| = 0 \tag{5.40}$$

hold uniformly in a, then $S_n/\sqrt{n}$ is asymptotically $N(0, 1)$, provided

$$\gamma \leqq \beta\theta. \tag{5.41}$$

PROOF. We put $X_{n,k} = n^{-1/2}X_k$ for $k = 1, \cdots, n = k_n$ and apply Theorem 2.2. However, we apply it not to the $X_{n,k}$ but to $Y_{n,j}$ where $Y_{n,1}$ is the sum of the first $m_n = [n^{\delta}]$ random variables $X_{n,k}$; the $Y_{n,2}$ is the sum of the next m_n terms $X_{n,k}$, and so forth (the last $Y_{n,j}$ is perhaps a sum of fewer summands). The number of $Y_{n,j}$ is roughly $n^{1-\delta}$. Also each $Y_{n,j}$ is equal to some $(m_n/n)^{1/2}T_a(m_n)$. We now check the conditions of the theorem.

First, the sum in (2.10) is estimated by $(n/m_n)o[(m/n)^{1/2}m_n^{-\theta}]$ in view of (5.38). Thus, putting $m_n = [n^{\delta}]$ we see that equation (2.3) holds provided

$$\frac{1-\delta}{2} \leqq \theta\delta. \tag{5.42}$$

Since $\Sigma\, EY_{n,i}^2 = (n/m_n)(m_n/n)[1 + o(1)] \to 1$ by (5.37), it follows from (5.40) that

$$\sum_i E(Y_{n,i}^2|\mathscr{F}_{n,im_n-1}) \xrightarrow[n\to\infty]{P} 1. \tag{5.43}$$

In view of (2.10) the above implies (2.4). Thus it remains to check the Lindeberg condition. This would certainly be implied by the Liapounoff condition $\Sigma\, E|Y_{n,i}|^{2+\beta} \to 0$, but by (5.38) this sum is $(n/m_n)o(m_n^{\gamma}(m_n/n)^{1+\beta/2})$ and thus tends to zero provided

$$\frac{1-\delta}{2}\beta \geqq \gamma\delta. \tag{5.44}$$

Equation (5.41) ensures the possibility of finding δ in (0, 1) satisfying (5.42) and (5.44) simultaneously.

5.5. B. Rosén gives interesting central limit theorems and many useful applications [14]. It is not difficult to deduce generalizations of his results from Theorem 2.2 and the following Theorem 5.4. This theorem, motivated by Rosén's results, considers the case when the mean of the summands depends, roughly speaking, linearly on the sum of the previous terms. (In several places in [14] conditional second moments should be replaced by conditional variances).

THEOREM 5.4. *Let the random variables* (1.1) *and the constants* $a_{n,k}$, *for* $n = 1, 2, \cdots$ *and* $k = 1, \cdots, k_n$, *satisfy the following conditions:*

$$\sum_{k=1}^{k_n} |\mu_{n,k} - a_{n,k}S_{n,k-1}| \xrightarrow[n\to\infty]{P} 0, \tag{5.45}$$

$$\sum_{k=1}^{k_n} \sigma_{n,k}^2 \prod_{j=k+1}^{n} (1 + a_{n,j})^2 \xrightarrow[n\to\infty]{P} 1, \tag{5.46}$$

$$\limsup_{n=\infty} \max_{1\leqq k\leqq k_n} \prod_{j=k+1}^{k_n} |1 + a_{n,j}| < \infty. \tag{5.47}$$

Then the condition

$$\lim_{n=\infty} \sum_{k=1}^{k_n} E[(X_{n,k} - \mu_{n,k})^2 I(|X_{n,k} - \mu_{n,k}| > \varepsilon)] = 0 \tag{5.48}$$

for every $\varepsilon > 0$, *implies* (1.3).

REMARK 5.1. By Lemma 3.3 condition (5.48) is less restrictive than the Lindeberg condition (1.6) or the condition

$$\lim_{n=\infty} E[(X_{n,k} - a_{n,k}S_{n,k-1})^2 I(|X_{n,k} - a_{n,k}S_{n,k-1}| > \varepsilon)] = 0. \tag{5.49}$$

(We can, of course, replace (5.48) by weaker 'conditional type' conditions.)

Note also that (5.47) is satisfied, for example, if the sums of $|a_{n,k}|$ in each row ($n = 1, 2, \cdots$) are bounded.

PROOF. We may, and do, retain assumption (4.1). Putting

$$Y_{n,k} = X_{n,k} - a_{n,k}S_{n,k-1}, \tag{5.50}$$

we have

$$S_{n,k} = (1 + a_{n,k})S_{n,k-1} + Y_{n,k}. \tag{5.51}$$

Applying (5.50) with $k = k_n$, then with $k = k_n - 1$, and so on, we obtain

$$S_{n,k_n} = \sum_{k=1}^{k_n} \prod_{j=k+1}^{k_n} (1 + a_{n,j})Y_{n,k}. \tag{5.52}$$

Or, putting,

$$Z_{n,k} = \prod_{j=k+1}^{k_n} (1 + a_{n,j})\, Y_{n,k}, \tag{5.53}$$

we have

$$S_{n,k_n} = \sum_{k=1}^{k_n} Z_{n,k}. \tag{5.54}$$

Let M be a bound on all products appearing in (5.47). Then, by (5.50) and (5.53),

$$|E(Z_{n,k}|F_{n,k-1})| \leqq M|\mu_{n,k} - a_{n,k}S_{n,k-1}| \tag{5.55}$$

and hence, by (5.54) and (5.45), to prove (1.3) we have to show that

$$\sum_{k=1}^{k_n} [Z_{n,k} - E(Z_{n,k}|\mathscr{F}_{n,k-1})] \tag{5.56}$$

is asymptotically $N(0, 1)$. But it is immediately verified that the random variables under the summation sign in (5.56) satisfy the conditions of Theorem 2.2. Indeed, condition (2.3) holds because they are centered at the conditional expectations; condition (2.4) holds since the summands in (5.46) are precisely the conditional variances; while condition (2.5) follows from (5.48) and $|Z_{n,k} - E(Z_{n,k}|\mathscr{F}_{n,k-1})| \leqq M|X_{n,k} - \mu_{n,k}|$.

6. Remarks

6.1. Occasionally there are attempts to establish results on asymptotic weighted normality. For instance, one may try to show that if (2.1) and (1.6) hold, and the sum on the left in (2.2) is a random variable σ_n^2 with $P(\sigma_n^2 \leqq u) \to W(u)$, then S_{n,k_n} approaches in distribution the weighted normal $N(0, W)$, (this is simply the average of ordinary normal distributions $N(0, \cdot)$ relative to the weight $W(\cdot)$).

However, no such result can hold, not even if $P(\sigma_n^2 \leqq u)$ is independent of n. The distribution of S_{n,k_n} need not tend to a limit, and if it does, the limit need not be of the form described above. Imposing a martingale structure or strong boundedness conditions does not change the situation.

The following furnishes a simple example (see Remark 3.4). Consider (1.1) with $k_n = 2n + 1$. Let ρ_n be irrational numbers tending to zero and let the $X_{n,k}$ with $k \leqq n$ be independent and $P(X_{n,k} = 1/[n^{1/2}]) = P(X_{n,k} = -1/[n^{1/2}]) = 1/2$. Let B_n be any set for which $P(S_{n,n} \in B_n)$ has a limit, say p, and let $X_{n,n+1} = 0$ if $S_{n,n} \notin B$ and assume the values ρ_n or $-\rho_n$, each with probability 1/2, otherwise. Similarly, for $k > n + 1$ let $X_{n,k} = 0$ if $S_{n,k-1}$ is rational and assume the values $\pm 1/[n]^{1/2}$, each with probability 1/2, otherwise. Then the left side of (2.2) assumes two values tending to 1 and 2 with probabilities $1 - p$ and p respectively. To finish the construction of the example we have merely to specify B_n. If $B_n = (0, \infty)$, then S_{n,k_n} has a limit distribution which is non-symmetric. If $B_n = (-\infty, 0)$, it tends to another such distribution. If u_n is such that $P(|S_n| < u_n) \to 1/2$, then choosing $B_n = (-u_n, u_n)$, S_{n,k_n} tends in distribution to a symmetric limit; taking $B_n = (-\infty, -u_n) \cup (u_n, \infty)$ it tends to another such limit.

6.2. There are many applications of quite a different nature from those considered in Section 5. The most interesting ones are obtained by combining the results of the present paper with stopping rules. We mention the following result, related to the three series theorem for dependent random variables.

Let $|X_n| \leqq c$, $\mathscr{F}_n = \mathscr{B}(X_1, \cdots, X_n)$, let $\mu_n = E(X_n | \mathscr{F}_{n-1})$, and let $\sigma_n^2 = E(X_n^2 | \mathscr{F}_{n-1}) - \mu_n^2$. If $|\mu_n| < K\sigma_n^2$ for some $K < \infty$, then the convergence sets of the series ΣX_n and $\Sigma \sigma_n^2$ coincide (except for null sets).

Recently N. Langberg (Ph.D. thesis, Jerusalem) applied results of the present paper to the study of asymptotic normality for stochastic approximation procedures.

Lévy has applied his results to such problems as the law of the iterated logarithm for dependent random variables [9]. Advances since his book was written permit various improvements, but it is still an invaluable source of ideas.

6.3. It is not difficult to obtain results by our methods on the rate of convergence to normality. Extensions to several dimensions also present no problems.

It is also possible, following a line going back to [1] and [9], to study correlation functions and limiting processes. An invariance principle related to the questions studied here has recently been established by R. Drogin (Ph.D. Thesis, Berkeley).

Our results can be reformulated so as to exhibit not only sufficiency but also necessity, but this requires additional assumptions in order to avoid cases such as $X_{n,k} = X_n I(A_{n,k})$ with X_n asymptotically $N(0, 1)$ and $A_{n,k}$ a partition of the probability space.

REFERENCES

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

[2] P. Billingsley, "The Lindeberg–Lévy theorem for martingales," *Proc. Amer. Math. Soc.*, Vol. 12 (1961), pp. 788–792.

[3] J. L. Doob, *Stochastic Processes*, New York, Wiley, 1953.

[4] A. Dvoretzky, "Central limit theorems for dependent random variables and some applications," 1969 Rietz lecture, Institute of Mathematical Statistics.

[5] ——— "The central limit problem for dependent random variables," *Proceedings of the International Congress of Mathematicians*, Nice, 1970.

[6] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, Reading, Addison-Wesley, 1954.

[7] W. Hoeffding and H. Robbins, "The central limit theorem for dependent random variables," *Duke Math. J.*, Vol. 15 (1948), pp. 773–780.

[8] I. A. Ibragimov, "Some limit theorems for stationary processes," *Teor. Verojatnost. i Premenen.*, Vol. 7 (1962), pp. 361–392.

[9] P. Lévy, *Theorie de l'Addition des Variables Aléatoires*, Paris, Gauthier-Villars, 1937.

[10] M. Loève, "Étude asymptotique des sommes de variables aléatoires liées," *J. Math. Pures Appl.*, Vol. 24 (1945), pp. 249–318.

[11] ———, *Probability Theory*, Princeton, Van Nostrand, 1960.

[12] S. OREY, "A central limit theorem for m-dependent random variables," *Duke Math. J.*, Vol. 25 (1958), pp. 543–546.

[13] W. PHILIPP, "The central limit problem for mixing sequences of random variables," *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, Vol. 12 (1969), pp. 155–171.

[14] B. ROSÉN, "On the central limit theorem for sums of dependent random variables," *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, Vol. 7 (1967), pp. 48–82.

[15] M. ROSENBLATT, "A central limit theorem and a mixing condition," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 42 (1956), pp. 43–47.

[16] R. J. SERFLING, "Central limit theory for dependent variables," *Ann. Math. Statist.*, Vol. 39 (1968), pp. 1158–1175.