Actes, Congrès intern. Math., 1970. Tome 2, p. 565 à 570.

CENTRAL LIMIT THEOREMS FOR DEPENDENT RANDOM VARIABLES

by Aryeh DVORETZKY

1. – Limiting distributions of sums of 'small' independent random variables have been extensively studied and there is a satisfactory general theory of the subject (see e.g. the monograph of B.V. Gnedenko and A.N. Kolmogorov [2]). These results are conveniently formulated for double arrays $X_{n,k}$ ($k = 1, \ldots, k_n$; $n = 1, 2, \ldots$) of random variables where the $X_{n,k}$ ($k = 1, \ldots, k_n$), the random variables in the n-th row, are assumed mutually independent for every n. The smallness assumption is that, as $n \to \infty$, the random variables $X_{n,k}$ converge in probability to 0, uniformly in k. This assumption implies that the limit distributions of $X_{n,1} + \ldots + X_{n,k_n}$ are infinitely divisible ones, and necessary and sufficient conditions on the distributions of the summands are known for convergence to any specific infinitely divisible distribution.

Our knowledge of the corresponding theory for sums of dependent random variables is much less satisfactory. Though a great number of papers have been published on the subject, not many general results are known. A notable exception to this statement is provided by some pioneering work of P. Lévy (see [3] and the references there). Recently [1] we have shown that the classical necessary and sufficient conditions mentioned above for the independent case, remain sufficient for the most general dependent case, provided we replace in their formulation the distributions of the summands $X_{n,k}$ by their conditional distributions. In *2* we present some of these results, in *3* we state three lemmas used in proving the results of *2* while in *4* we note some applications and make some comments.

2. – The typical results which we present here are patterned after familiar results for independent random variables.

Let $(\Omega_n, \mathcal{a}_n, P_n)$, ($n = 1, 2, \ldots$) be a sequence of probability spaces, not necessarily distinct, and let $X_{n,k}$ ($k = 1, 2, \ldots, k_n$) be any random variables defined on $(\Omega_n, \mathcal{a}_n, P_n)$. Put $S_{n,k} = \sum_{j=1}^{k} X_{n,j}$ ($k = 0, \ldots, k_n$) and let $\mathcal{F}_{n,k}$ be the σ-field generated by $S_{n,k}$. Conditional expectation and conditional probability relative to $\mathcal{F}_{n,k}$ are denoted by $E_{n,k}$ and $P_{n,k}$ respectively. $I\{.\}$ is the indicator of the set within the braces. $\mathcal{L}(.)$ is the distribution of the random variable within the brackets ; $\mathcal{N}(0, 1)$ is the standard normal law. Convergence $\to$ always refers to $n \to \infty$ and $\xrightarrow{P}$ denotes convergence in probability. We abbreviate $\sum_{k=1}^{k_n}$ to $\sum_k$ and $\max_{1 \leqslant k \leqslant k_n}$ to $\max_k$.

The first result we quote concerns asymptotic normality.

THEOREM 1. – *The conditions*

$$(1) \qquad \sum_k E_{n,k-1} X_{n,k} \xrightarrow{P} 0,$$

$$(2) \qquad \sum_k E_{n,k-1} (X_{n,k} - E_{n,k-1} X_{n,k})^2 \xrightarrow{P} 1$$

and

$$(3) \qquad \sum_k E_{n,k-1} (X_{n,k}^2 I\{|X_{n,k}| > \epsilon\}) \xrightarrow{P} 0 \quad \text{for every} \quad \epsilon > 0,$$

imply

$$(4) \qquad \mathcal{L}(S_{n,k_n}) \to \mathcal{N}(0, 1).$$

We note that (3) is a weaker condition than the standard, non-conditioned, Lindeberg condition

$$(3') \qquad \sum_k E(X_{n,k}^2 I\{|X_{n,k}| > \epsilon\}) \to 0 \quad \text{for every} \quad \epsilon > 0.$$

We also remark that (1) is automatically satisfied when $X_{n,k}$ ($k = 1, \ldots, k_n$) are, for each n, martingale differences. The classical result about asymptotic normality (4) is, of course, a further specialization. Theorem 1 already exhibits the recurring feature of our generalizations : Replacing, where appropriate in the classical limit theorems, expectations by conditional ones relative to the preceding row sum we obtain sufficient conditions for convergence to a given limit law.

A certain generality is gained by formulating the conditions in terms of convergence in probability. However, the most important generality feature of Theorem 1 relative to the usual results on asymptotic normality for sums of dependent random variables is that no assumption beyond (2) is made about the conditional variances. Usually they are assumed to be nearly constant in some sense.

Similar results can be stated for convergence to a Poisson law. Instead of giving further special results we state one of moderate generality.

THEOREM 2. – *Assume that* $E_{n-1} X_{n,k} = 0$ *and that* $\sum_k E_{n,k-1} X_{n,k}^2$ ($n = 1, 2, \ldots$) *are uniformly bounded in probability and let* $K(.)$ *be a bounded monotone function on the real line. Then the conditions*

$$(5) \qquad \sum_k E_{n,k-1} (X_{n,k}^2 I\{X_{n,k} < x\}) \xrightarrow{P} K(x)$$

at every continuity point x *of* K *and*

$$(6) \qquad \sum_k (E_{n,k-1} X_{n,k}^2)^2 \xrightarrow{P} 0$$

imply that $\mathcal{L}(S_{n,k_n}) \to$ *the infinitely divisible law whose characteristic function is given by* $\exp \int_{-\infty}^{\infty} \frac{e^{itx} - 1 - itx}{x^2} \, dK(x)$.

The conditioning requirements become more stringent if the conditioning σ-fields are replaced by finer ones. Thus if we denote by $E^*_{n,k}$ expectations relative to $\mathfrak{F}^*_{n,k}$, the σ-field generated by $X_{n,1}, \ldots, X_{n,k}$, then the conditions (5*) and (6*) obtained on replacing E by E^* in (5) and (6), respectively, certainly imply the conclusion. We note that, given the other assumptions, the condition (6*) is implied by

$$\max_k E^*_{n,k-1} X_{n,k}^2 \xrightarrow{P} 0 , \tag{7}$$

or by

$$\max_k P^*_{n,k-1} \{ |X_{n,k-1}| > \epsilon \} \xrightarrow{P} 0 \quad \text{for every} \quad \epsilon > 0 \tag{8}$$

($P^*_{n,k}$ is the conditional probability relative to $\mathfrak{F}^*_{n,k}$).

Conditions (7) and (8) are smallness conditions similar to those used in the independent case. They are, however, expressed through finer conditionings than those relative to $\mathfrak{F}_{n,k}$. But it is possible, through Lemma 1 below, to replace them by a condition which is expressed in terms of the conditional distributions of $X_{n,k}$ relative to $\mathfrak{F}_{n,k-1}$ only (see (15)).

The previous theorem assumed first and second moment conditions. The following is an example of a result which dispenses entirely with such assumptions.

THEOREM 3. – *Let a be a real number and K be a bounded monotone function on the real line. Put $a_{n,k} = E_{n,k-1} (X_{n,k} \, I\{|X_{n,k}| < 1\})$ and $Y_{n,k} = X_{n,k} - a_{n,k}$. Then the conditions*

$$\sum_k \left(a_{n,k} + E_{n,k-1} \frac{Y_{n,k}}{1 + Y_{n,k}^2} \right) \xrightarrow{P} 0 , \tag{9}$$

$$\sum_k E_{n,k-1} \left(\frac{Y_{n,k}^2}{1 + Y_{n,k}^2} \, I\{|Y_{n,k}| < x\} \right) \xrightarrow{P} K(x) \tag{10}$$

for every continuity point x of K,

$$\sum_k E_{n,k-1} \left(\frac{Y_{n,k}^2}{1 + Y_{n,k}^2} \right) \xrightarrow{P} K(\infty) - K(-\infty) \tag{11}$$

where $K(\infty) = \lim_{x = \infty} K(x)$, $K(-\infty) = \lim_{x = -\infty} K(x)$, *and*

$$\sum_k \left(E_{n,k-1} \left(\frac{Y_{n,k-1}^2}{1 + Y_{n,k}^2} \right) \right)^2 \xrightarrow{P} 0 \tag{12}$$

imply that $\mathcal{L}(S_{n,k_n}) \to$ *the infinitely divisible law whose characteristic function is given by* $\exp\left(ita + \int_{-\infty}^{\infty}\left(e^{itx} - 1 - \frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}\, d\,K(x)\right)$.

Condition (12) is again implied by the smallness condition (15). (The theorem obtained from Theorem 3 on substituting, throughout its statement, E^* by E is, of course, weaker than Theorem 3. Then (12*), the condition obtained from (12), is implied by the smallness condition (8)).

3. – Technically the fact that the σ-fields $\mathcal{F}_{n,k}$ $(k = 1, \dots, k_n)$ are not necessarily increasing is very cumbersome. This is overcome by the following simple but useful result.

LEMMA 1. – *Let* $S_1, S_2, \dots S_n$ *be random variables on a probability space* $(\Omega, \mathcal{A}, P)$. *Then there exists a probability space* $(\widetilde{\Omega}, \widetilde{\mathcal{A}}, \widetilde{P})$ *and random variables* $\widetilde{S}_1, \widetilde{S}_2, \dots, \widetilde{S}_n$ *on it such that*

$$(13)\quad \widetilde{P}(\widetilde{S}_k \leqslant u\ ,\ \widetilde{S}_{k-1} \leqslant v) = P(S_k \leqslant u\ ,\ S_{k-1} \leqslant v)\ ,\quad (k = 2, \dots, n)$$

for all real u, v *and*

$$(14)\quad \widetilde{P}(\widetilde{S}_k \mid \widetilde{\mathcal{F}}_{k-1}) = \widetilde{P}(\widetilde{S}_k \mid \widetilde{\mathcal{F}}_{k-1}^{*})\ ,\quad (k = 2, \dots, n)$$

where $\widetilde{\mathcal{F}}_k$ *and* $\widetilde{\mathcal{F}}_k^*$ *are the* σ*-fields generated by* $\widetilde{S}_k$ *and by* $\widetilde{S}_1, \dots, \widetilde{S}_k$, *respectively.*

Condition (13) asserts that the distributions of the $\widetilde{S}_k$ are the same as those of the S_k and that, moreover, the conditional distributions given that $S_{k-1} = v$, respectively $\widetilde{S}_{k-1} = v$, are also the same. The point of (14) is that conditioning by $\widetilde{\mathcal{F}}_k$ is like conditioning by an increasing sequence of σ-fields. Since joint distributions of $\widetilde{S}_1, \dots, \widetilde{S}_n$ can be obtained from the conditional distributions of the $\widetilde{S}_k$ given $\widetilde{S}_{k-1}$, i.e. of the S_k given S_{k-1}, it follows that, after obvious adaptation of notation, the condition

$$(15)\qquad \max_k \widetilde{P}_{n,k-1}\{|\widetilde{X}_{n,k}| > \epsilon\} \xrightarrow{P} 0 \quad \text{for every} \quad \epsilon > 0$$

is a smallness condition expressed in terms of the conditional distributions of the $X_{n,k}$ relative to $\mathcal{F}_{n,k-1}$. (15) is precisely the smallness condition which can replace (8) or (12).

The next two lemmas rely heavily on the fact that the conditioning σ-fields are increasing.

LEMMA 2. – *Let* $\mathcal{G}_0 \subset \mathcal{G}_1 \subset \dots \subset \mathcal{G}_n$ *be* σ*-fields in a probability space and let* $\eta_1, \dots, \eta_n$ *be bounded, complex-valued random variables with* η_k *measurable* $\mathcal{G}_k (k = 1, \dots, n)$. *Put* $\varphi_k = E(\eta_k \mid \mathcal{G}_{k-1})$ *and* $\psi = \prod_{k=1}^{n} \varphi_k$. *If* ψ *is measurable* $\mathcal{G}_0$ *and* $\psi \neq 0$ *almost surely, then* $\psi = E\left(\prod_{k=1}^{n} \eta_k \mid \mathcal{G}_0\right)$.

LEMMA 3. – *Let* $\mathcal{G}_0 \subset \mathcal{G}_1 \subset \ldots \subset \mathcal{G}_n$ *be* σ*-fields in a probability space and let* $\eta_1, \ldots, \eta_n, \zeta_1, \ldots, \zeta_n$ *be bounded, complex-valued random variables with* η_k *and* ζ_k *measurable* $\mathcal{G}_k$ $(k = 1, \ldots, n)$. *Let the random variables* $\prod_{j=1}^{k-1} \zeta_j \prod_{j=k+1}^{n} \eta_j$, $(k = 1, \ldots, n)$ *be bounded by the constant c, then*

$$(16) \qquad \left| E \prod_{k=1}^{n} \zeta_k - E \prod_{k=1}^{n} \eta_k \right| \leqslant c \sum_{k=1}^{n} E \left| E(\zeta_k - \eta_k \mid \mathcal{G}_{k-1}) \right| .$$

Lemma 2 is proved by backward induction. Lemma 3 is easy and its antecedents can be traced to Lindeberg at least. When applying the above lemmas to deduce the theorems of *2* we have $|\zeta_k| \leqslant 1$ for all k and $|\eta_k| \leqslant 1$ except for the last one which is estimated, e.g. in case of Theorem 2, via the differences between the random variables on the left and the expressions on the right of (5) and (6). Through (16) we can obtain an explicit bound for the difference between the characteristic function of S_{n,k_n} and that of the limit law. With the aid of Berry-Esseen and similar estimates we can thus obtain explicit results on the rate of convergence.

4. – We refer to [1] for a number of specializations of Theorem 1 which improve various known results on asymptotic normality for dependent random variables. This can also be applied to derive a version of the three series theorem for dependent random variables. Its other applications include results on stochastic approximation, on optimal stopping and on convergence to the Brownian motion process.

The general theory can also be applied to study domains of attraction and similar questions. The generalization to m-dimensional random variables presents no difficulties.

Recently V.M. Zolotarev and others (see [4] and the references there) succeeded in developing a theory of limit laws for sums of independent random variables without assuming a 'smallness' condition. It would be interesting to extend this theory to the dependent case.

In some special cases we can show that our sufficient conditions are also necessary, but our results in this direction are either fragmentary or involve very cumbersome conditions. Further study of these problems should be of interest.

REFERENCES

[1] DVORETZKY A. — " Asymptotic normality for sums of dependent random variables ", *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability,* Berkeley 1970 (to appear). Also the 1969 Reitz lecture of the Institute of Mathematical Statistics.

[2] GNEDENKO B.V. and KOLMOGOROV A.N. — *Limit Distributions of Sums of Independent Random Variables,* Reading, Addison-Wesley, 1954.

[3] LÉVY P. — *Théorie de l'Addition des Variables Aléatoires,* Paris, Gauthier-Villars, 1937.
[4] ZOLOTAREV V.M. — Théorèmes limites généraux pour les sommes de variables aléatoires indépendantes, *C. R. Acad. Sc. Paris,* 270, 1970, p. 899-902.

Hebrew University
Dept. of Mathematics,
Jerusalem
Israël