THE CENTRAL LIMIT THEOREM FOR m-DEPENDENT VARIABLES

By P. H. DIANANDA

Communicated by M. S. Bartlett

Received 9 February 1954

1. *Introduction and summary.* In a previous paper (4) central limit theorems were obtained for sequences of m-dependent random variables (r.v.'s) asymptotically stationary to second order, the sufficient conditions being akin to the Lindeberg condition (3). In this paper similar theorems are obtained for sequences of m-dependent r.v.'s with bounded variances and with the property that $n = O(s_n'^2)$ for large n, where s_n' is the standard deviation of the nth partial sum of the sequence. The same basic ideas as in (4) are used, but the proofs have been simplified. The results of this paper are examined in relation to earlier ones of Hoeffding and Robbins (5) and of the author (4). The cases of identically distributed r.v.'s and of vector r.v.'s are mentioned.

2. *Preliminaries.* We deal with sequences of r.v.'s

$$\{X_i\} \quad (i = 1, 2, 3, \ldots), \tag{1}$$

and define G_i, G, S_n, s_n, C_p, $S(A)$, H_i as in (4). Unless otherwise stated, limits are those as $n \to \infty$, and the summation suffix i runs through the values $1, 2, \ldots, n$. When (1) is a sequence of scalar r.v.'s with zero means,

$$s_n' = \sqrt{\mathscr{E}(S_n^2)}. \tag{2}$$

We refer to Theorem 1 of (3) as Theorem 1 and to Lemma 1 of (4) as Lemma 1. We note, however, that conclusion (i) of this lemma holds even if ϵ and m are functions of n. The proof of (i) given in (4) holds in this wider case too, without amendment. For the purpose of conclusion (ii), however, ϵ and m are constants. In this paper we use the fact that in (i) ϵ can be a function of n, thus simplifying our proofs.

We use also the following central limit theorem (1).

Theorem 2. *Let* (1) *be a sequence of independent scalar r.v.'s with means zero and finite variances. Suppose that, for every fixed* $\epsilon > 0$,

$$\frac{1}{s_n^2} \sum_i \int_{|x| > \epsilon s_n} x^2 \, dG_i \to 0. \tag{3}$$

Then the distribution function (d.f.) of $S_n/s_n \to$ *the standardized normal d.f.*

3. *Scalar r.v.'s.* We prove for scalar r.v.'s the following

Theorem 3. *Let* (1) *be a sequence of m-dependent scalar r.v.'s with means zero and variances bounded. Suppose that*

$$\liminf (s_n'^2/n) > 0, \tag{4}$$

and that, for every positive ϵ,

$$\frac{1}{n} \sum_i \int_{|x| > \epsilon \sqrt{n}} x^2 \, dG_i \to 0. \tag{5}$$

Then the d.f. of S_n/s_n' *tends to the standardized normal d.f.*

https://doi.org/10.1017/S0305004100029959 Published online by Cambridge University Press

We note that $$E(X_i^2) < b^2 \quad (i = 1, 2, 3, \ldots) \tag{6}$$

for some $b^2 > 0$, and so $$s_n^2/n < b^2 \tag{7}$$

and $$s_n'^2/n \leqslant (2m+1)\, s_n^2/n < (2m+1)\, b^2, \tag{8}$$

the X_i being m-dependent. From (4) it follows that

$$s_n'^2/n > a^2 \quad (n \to \infty) \tag{9}$$

for some $a^2 > 0$. From (8) and (9) we get

$$s_n^2/n > a^2/(2m+1) \quad (n \to \infty). \tag{10}$$

(8)–(10) give us that $$s_n^2 \underset{\frown}{\cup} s_n'^2 \underset{\frown}{\cup} n \quad (n \to \infty) \tag{11}$$

so that we have the

COROLLARY. *In Theorem* 3, (3) *can replace* (5) *and so can* (3) *with* s_n' *in place of* s_n.

Let h, k, r be as in Theorem 1,

$$\left.\begin{aligned} Y_i &= X_{(i-1)k+1} + X_{(i-1)k+2} + \ldots + X_{ik-m} \\ Z_i &= X_{ik-m+1} + X_{ik-m+2} + \ldots + X_{ik} \end{aligned}\right\} \quad (i = 1, 2, \ldots, h), \tag{12}$$

$$\left.\begin{aligned} &s_{n,k} = \sqrt{\mathscr{E}\,(Y_1^2 + Y_2^2 + \ldots + Y_h^2)}, \quad U_{n,k} = (Y_1 + Y_2 + \ldots + Y_h)/s_{n,k}, \\ &V_{n,k;1} = (Z_1 + Z_2 + \ldots + Z_h)/s_n', \quad V_{n,k;2} = (S_n - S_{hk})/s_n', \\ &V_{n,k;3} = (Y_1 + Y_2 + \ldots + Y_h)\left(\frac{1}{s_n'} - \frac{1}{s_{n,k}}\right), \end{aligned}\right\} \tag{13}$$

$$V_{n,k} = V_{n,k;1} + V_{n,k;2} + V_{n,k;3}. \tag{14}$$

Then $$U_{n,k} + V_{n,k} = S_n/s_n' \equiv W_n. \tag{15}$$

LEMMA 2. *For large* k *and* $\epsilon > 0$, *as* $h \to \infty$,

$$L \equiv \frac{1}{s_{n,k}^2} \sum_{i=1}^{h} \int_{|y| > \epsilon s_{n,k}} y^2\, dH_i \to 0, \tag{16}$$

where $$y = x_1 + x_2 + \ldots + x_{k-m}. \tag{17}$$

Proof. We have

$$|y| \leqslant |x_1| + |x_2| + \ldots + |x_{k-m}| \leqslant k \max_{1 \leqslant j \leqslant k-m} |x_j|, \tag{18}$$

so that $$S(|y| > \epsilon s_{n,k}) \subset S \equiv S\left(\max_{1 \leqslant j \leqslant k-m} |x_j| > \frac{\epsilon}{k} s_{n,k}\right). \tag{19}$$

Hence

$$\begin{aligned} L &\leqslant \frac{1}{s_{n,k}^2} \sum_{i=1}^{h} \int_S (x_1 + x_2 + \ldots + x_{k-m})^2\, dH_i \\ &\leqslant \frac{k}{s_{n,k}^2} \sum_{i=1}^{h} \int_S (x_1^2 + x_2^2 + \ldots + x_{k-m}^2)\, dH_i \\ &\leqslant \frac{k}{s_{n,k}^2} \sum_{i=1}^{hk} \int_{\max_{1 \leqslant j \leqslant hk} |x_j| > \frac{\epsilon}{k} s_{n,k}} x_i^2\, dG \\ &\leqslant \frac{k}{s_{n,k}^2} \left(2m + 1 + \frac{k^2 s_{hk}^2}{\epsilon^2 s_{n,k}^2}\right) \sum_{i=1}^{hk} \int_{|x| > \frac{\epsilon}{k} s_{n,k}} x^2\, dG_i, \end{aligned} \tag{20}$$

by Lemma 1. Now $$s_{n,k}^2 = \mathscr{E}\{S_{hk} - (Z_1 + Z_2 + \ldots + Z_h)\}^2. \tag{21}$$

https://doi.org/10.1017/S0305004100029959 Published online by Cambridge University Press

The Z_i are independent if $k \geqslant 2m$ and, by (6),

$$\mathscr{E}(Z_i^2) < m^2b^2, \tag{22}$$

so that, θ being a general notation for a number with modulus $\leqslant 1$,

$$\begin{aligned} s_{n,k}^2 &= s_{hk}'^2 + 2\theta s_{hk}' mb\sqrt{h} + \theta m^2b^2h \\ &= s_{hk}'^2\left(1 + \frac{2\theta mb\sqrt{h}}{a\sqrt{(hk)}} + \frac{\theta m^2b^2h}{a^2hk}\right), \end{aligned} \tag{23}$$

by (9), for large hk. Hence

$$s_{n,k}^2/s_{hk}'^2 \to 1 \tag{24}$$

uniformly in h as $k \to \infty$. Hence, for large fixed k,

$$s_{n,k}^2 > \tfrac{1}{2}s_{hk}'^2 > \tfrac{1}{2}hka^2, \tag{25}$$

by (9). Hence, from (20), we get that for large k,

$$L \leqslant \frac{k}{\frac{1}{2}hka^2}\left(2m + 1 + \frac{b^2hk^3}{\frac{1}{2}hka^2\epsilon^2}\right)\sum_{i=1}^{hk} \int_{|x| > \frac{\epsilon}{k}\sqrt{(\frac{1}{2}hka^2)}} x^2\, dG_i \to 0 \tag{26}$$

as $h \to \infty$, by (5), and the lemma follows.

LEMMA 3. *For large k, the d.f. of $U_{n,k}$ tends to the standardized normal d.f. ($F_k(x) = F(x)$, say) as $h \to \infty$.*

This follows immediately from (13), Theorem 2 and Lemma 2, since the Y_i are independent.

LEMMA 4. *$V_{n,k}$ converges in probability to zero uniformly in r as $h, k \to \infty$.*

Proof. We have from (6), (9) and (13),

$$\mathscr{E}(V_{n,k;1}^2) \leqslant hm^2b^2/na^2 \leqslant m^2b^2/ka^2 \to 0 \tag{27}$$

and

$$\mathscr{E}(V_{n,k;2}^2) \leqslant (2m+1)rb^2/na^2 \leqslant (2m+1)b^2/ha^2 \to 0, \tag{28}$$

uniformly in r as $h, k \to \infty$; also

$$\mathscr{E}(V_{n,k;3}^2) = s_{n,k}^2\left(\frac{1}{s_n'} - \frac{1}{s_{n,k}}\right)^2 = \left(\frac{s_{n,k}}{s_{hk}'}\frac{s_{hk}'}{s_n'} - 1\right)^2. \tag{29}$$

Now, from (2) and (13),

$$\begin{aligned} \frac{s_{hk}'^2}{s_n'^2} &= \mathscr{E}\left(\frac{S_{hk}^2}{s_n'^2}\right) = \mathscr{E}\left\{\left(\frac{S_n}{s_n'} - V_{n,k;2}\right)^2\right\} \\ &= \mathscr{E}\left(\frac{S_n^2}{s_n'^2} - 2\frac{S_n}{s_n'}V_{n,k;2} + V_{n,k;2}^2\right) \to 1, \end{aligned} \tag{30}$$

uniformly in r as $h, k \to \infty$. Using (24) and (30) in (29), we get

$$\mathscr{E}(V_{n,k;3}^2) \to 0, \tag{31}$$

uniformly in r as $h, k \to \infty$. From (14), (27), (28) and (31) the lemma now follows.

We can now complete the proof of Theorem 3, noting (15) and using Lemmas 3, 4 and Theorem 1, which can now be applied to W_n. The conditions of Theorem 1 are satisfied if we take $V = 0$.

4. *The Hoeffding-Robbins conditions.* Hoeffding and Robbins (5) prove central limit theorems for sequences (1) of m-dependent r.v.'s, with zero means and bounded third absolute moments, which are such that

$$\lim_{p\to\infty}\frac{1}{p}\mathscr{E}\{(X_{i+1} + X_{i+2} + \ldots + X_{i+p+m})^2 - (X_{i+1} + X_{i+2} + \ldots + X_{i+m})^2\} = A, \tag{32}$$

https://doi.org/10.1017/S0305004100029959 Published online by Cambridge University Press

uniformly for all $i = 0, 1, 2, \ldots$. We show that these conditions imply those of Theorem 3 if $A > 0$. The case $A = 0$ is trivial. The boundedness of the third absolute moments implies that the variances are bounded. (32) implies that

$$\lim_{p\to\infty} \frac{1}{p} \mathscr{E}\{S^2_{p+m} - S^2_m\} = A, \tag{33}$$

which implies that

$$\lim \frac{1}{n} \mathscr{E}(S^2_n) = A, \tag{34}$$

so that (4) holds if $A > 0$. Also since the third absolute moments are bounded (3) and, (11) being true, (5) hold (1). Thus the Hoeffding-Robbins conditions imply those of Theorem 3 if $A > 0$. We note that in these conditions (32) can be replaced by (34) or by (4), and the boundedness of third absolute moments by (*a*) that of absolute moments of any particular order higher than the second or (*b*) that of the second moments, together with any condition of Liapounoff type (1).

5. *Identically distributed r.v.'s.* In this case (5) is automatically satisfied and the variances are bounded if only they exist finitely. Thus we have the following extension of the Lindeberg-Lévy theorem (2).

THEOREM 4. *Let* (1) *be a sequence of m-dependent identically distributed scalar r.v.'s with means zero and variances finite. Suppose that* (4) *holds. Then the d.f. of* S_n/s'_n *tends to the standardized normal d.f.*

6. *Further remarks.* The results given in (4) are particular cases of those given in this paper, except once again in a trivial case. When the r.v.'s of (1) are asymptotically stationary to second order, the variances are bounded. Also

$$s'^2_n/n \to \sum_{-m}^{m} C_p. \tag{35}$$

Thus (4) is satisfied unless the right-hand side of (35) is zero. Of course, in the previous paper too, the condition (5) was imposed.

In the case of vector r.v.'s, Theorems 3 and 4 can be extended in the usual way (see, for example, (6)) to give us analogous results. For this a vector analogue of Theorem 2 has to be used in place of Theorem 2.

REFERENCES

(1) CRAMÉR, H. *Random variables and probability distributions* (Cambridge, 1937), pp. 57, 60.
(2) CRAMÉR, H. *Mathematical methods of statistics* (Princeton, 1946), p. 215.
(3) DIANANDA, P. H. Some probability limit theorems with statistical applications. *Proc. Camb. phil. Soc.* 49 (1953), 239–46.
(4) DIANANDA, P. H. The central limit theorem for *m*-dependent variables asymptotically stationary to second order. *Proc. Camb. phil. Soc.* 50 (1954), 287–92.
(5) HOEFFDING, W. and ROBBINS, H. The central limit theorem for dependent variables. *Duke Math. J.* 15 (1948), 773–80.
(6) KHINTCHINE, A. *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung* (Berlin, 1933).

INSTITUTE OF STATISTICS
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA*

* Present addres: Department of Mathematics, University of Malaya, Singapore.

https://doi.org/10.1017/S0305004100029959 Published online by Cambridge University Press