PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 25, Fasc. 2 (2005), pp. 317–329

A NOTE ON THE ALMOST SURE CONVERGENCE OF CENTRAL ORDER STATISTICS

BY

SIEGFRIED HÖRMANN (GRAZ)

Abstract. We prove almost sure versions of distributional limit theorems for central order statistics. We develop a new method which not only gives a simplified proof of existing results in the literature, but also extends them for general summation methods, leading to considerably sharper results.

2000 Mathematics Subject Classification: 60F05, 60F15, 40G99.

Key words and phrases: Central order statistics, almost sure limit theorem, summation methods.

1. INTRODUCTION AND MAIN RESULTS

Let $X_1, X_2, \ldots$ be i.i.d. r.v.'s and denote their order statistics by

$$X_{1:n} \leqslant X_{2:n} \leqslant \ldots \leqslant X_{n:n}.$$

If $r \geqslant 0$ is some fixed integer, then $X_{n-r:n}$ is called an *extreme order statistic*. It is well known that if

$$a_n(X_{n-r:n}-b_n) \xrightarrow{\mathscr{L}} G \tag{1.1}$$

for some non-degenerate distribution function G, then G belongs to one of three classes of distribution functions, the so-called *extremal distributions* (cf. Galambos [12]). If $r_n \in \{0, \ldots, n-1\}$ satisfies

$$\min\{r_n, n-r_n\} \to \infty, \tag{1.2}$$

$X_{n-r_n:n}$ is called a *central order statistic*. It is also well known that under weak conditions on the underlying distribution function, central order statistics are asymptotically normally distributed (cf. Reiss [19]), i.e. for some numerical sequences (a_n) and (b_n)

$$a_n(X_{n-r_n:n}-b_n) \xrightarrow{\mathscr{L}} \mathscr{N}_{0,1}. \tag{1.3}$$

Recently several authors dealt with almost sure versions of the extremal limit theorems (1.1) and (1.3). In the case $r = 0$, Cheng et al. [8] and Fahrner and Stadtmüller [10] proved that the weak convergence relation (1.1) implies

$$(1.4) \qquad \frac{1}{\log N}\sum_{n=1}^{N}\frac{1}{n}I\{a_n(X_{n:n}-b_n)\leqslant x\}\to G(x) \text{ a.s.} \quad \text{for all } x\in C_G,$$

where C_G denotes the set of all continuity points of G. Relation (1.4) is called "a.s. max-limit theorem", and it is one of the natural extensions of the almost sure central limit theorem (ASCLT), a remarkable pathwise form of the classical (weak) CLT investigated intensively in the past two decades. In its simplest form the ASCLT states that if $X_1, X_2, \dots$ are i.i.d. r.v.'s with $EX_1 = 0$, $EX_1^2 = 1$, then

$$(1.5) \qquad \frac{1}{\log N}\sum_{n=1}^{N}\frac{1}{n}I\{n^{-1/2}(X_1+\ldots+X_n)\leqslant x\}\to\Phi(x) \text{ a.s.} \quad \text{for all } x\in\boldsymbol{R},$$

where Φ denotes the standard normal distribution function. Relation (1.5) was proved by Brosamler [6] and Schatte [21] under more restrictive moment conditions and by Lacey and Philipp [16] and Fisher [11] under finite second moments. Later the ASCLT has been generalized in many directions. The main focus was to extend (1.5) for dependent or not identically distributed r.v.'s $X_1, X_2, \dots$ and studying refinements such as the corresponding CLT and LIL and a.s. invariance principles. We do not go into detail here, but refer to Atlagh and Weber [2] and Berkes [3] for surveys.

The papers of Fahrner and Stadtmüller and Cheng et al. cited above were the first examples for the a.s. version of a "nonlinear" limit theorem, i.e. a weak limit theorem for nonlinear functionals of independent random variables. Later, a.s. versions of other nonlinear limit theorems have been found, and Berkes and Csáki [4] showed that *every* weak limit theorem of a certain generic form and subject to minor technical conditions has an almost sure version. Such a result is known as "universal ASCLT". For a precise formulation and several examples we refer to [4].

In this paper we concentrate on almost sure limit theorems for central order statistics. Let us first review the existing results in the field. Stadtmüller [22] proved that if for some numerical sequences (a_n) and (b_n) we have

$$(1.6) \qquad a_n(X_{n-r_n:n}-b_n)\xrightarrow{\mathscr{L}}G,$$

with some non-degenerate distribution function G, and

$$(1.7) \qquad r_n/n = q+O\left((n\log^{\varepsilon}n)^{-1/2}\right) \quad (\varepsilon>0),$$

then the a.s. analogue

$$(1.8)\qquad \frac{1}{\log N}\sum_{n=1}^{N}\frac{1}{n}I\{a_n(X_{n-r_n:n}-b_n)\leqslant x\}\to G(x)\ \text{a.s.}$$

holds for any $x\in \boldsymbol{R}$ with $G(x)=\Phi(x)$. He showed that (1.6) implies (1.8) also if

$$(1.9)\qquad r_n = O\big((\log n)^{1-\varepsilon}\big)\quad (\varepsilon>0).$$

Note that the last condition covers extreme order statistics. There is a gap between (1.7) and (1.9) which was filled by Peng and Qi [18], who proved the following result.

THEOREM A. *Let $X_1, X_2, \ldots$ be i.i.d. r.v.'s and assume that for some non-degenerate distribution function G there exist constants $a_n>0$ and b_n such that* (1.6) *holds. Then under the condition* (1.2)

$$\frac{1}{\log N}\sum_{n=1}^{N}\frac{1}{n}I\{a_n(X_{n-r_n:n}-b_n)\leqslant x\}\to G(x)\ \textit{a.s.}\quad \textit{for all } x\in C_G.$$

In particular, if $X_1, X_2, \ldots$ are i.i.d. r.v.'s which are uniformly distributed over the interval (0, 1), the limit distribution G is normal and we can choose

$$(1.10)\qquad a_n=\left(\frac{n^3}{r_n(n-r_n)}\right)^{1/2}\quad \text{and}\quad b_n=1-\frac{r_n}{n}.$$

The proof of Theorem A uses the classical method of covariance estimates (see [16], [21]), but such estimates are not easy to get and the argument of Peng and Qi [18] is very technical. In this paper we develop a new approach to the problem which not only yields a quick proof of Theorem A, but enables us to extend the theorem for a large class of summation procedures, leading to considerably sharper results. Before formulating our results, we make some preliminary remarks on summation methods.

Given a positive sequence $\boldsymbol{D}=(d_k)$ with $D_n=\sum_{k=1}^{n}d_k\to\infty$, we say that a sequence (x_n) is ***D**-summable* to x if

$$\lim_{N\to\infty} D_N^{-1}\sum_{n=1}^{N} d_n x_n = x.$$

By a result of Hardy (see [7], p. 35), if $\boldsymbol{D}$ and $\boldsymbol{D}^*$ are summation procedures with $D_N^*=O(D_N)$, then under minor technical assumptions, the summation $\boldsymbol{D}^*$ is stronger than $\boldsymbol{D}$, i.e. if a sequence (x_n) is $\boldsymbol{D}$-summable to x, then it is also $\boldsymbol{D}^*$-summable to x. Moreover, by a result of Zygmund (see [7], p. 35), if $D_N^\alpha\leqslant D_N^*\leqslant D_N^\beta$ $(N\geqslant N_0)$ for some $\alpha>0$, $\beta>0$, then $\boldsymbol{D}$ and $\boldsymbol{D}^*$ are equivalent, and if $D_N^*=O(D_N^\varepsilon)$ for any $\varepsilon>0$, then $\boldsymbol{D}^*$ is strictly stronger than $\boldsymbol{D}$. For example, logarithmic summation defined by $d_n=1/n$ is stronger than ordinary

(Cesàro) summation defined by $d_n = 1$ and weaker than loglog summation defined by $d_n = 1/(n \log n)$. On the other hand, all summation methods defined by

$$d_n = (\log n)^\alpha / n, \qquad \alpha > -1,$$

are equivalent to logarithmic summation, and all summation methods defined by

$$d_n = n^\alpha, \qquad \alpha > -1,$$

are equivalent to Cesàro summation. The characteristic feature of a.s. central limit theory is logarithmic summation, but even in the simplest case when X_n are i.i.d. r.v.'s with mean 0 and variance 1, there exists a large class of weight sequences (d_n), other than $d_n = 1/n$, such that

$$(1.11) \quad \frac{1}{D_N} \sum_{n=1}^{N} d_n I\{n^{-1/2}(X_1 + \ldots + X_n) \leqslant x\} \to \Phi(x) \text{ a.s.} \quad \text{for all } x \in \mathbf{R},$$

where $D_N = \sum_{k=1} d_k$. For example, (1.11) holds for all $d_n \leqslant 1/n$ with $\sum d_n = \infty$ and also for many sequences $d_n \geqslant 1/n$. Moreover, in the case of independent, not identically distributed r.v.'s, the weights $d_n = 1/n$ are generally not suitable, and one should use different summation methods, see Atlagh [1] and Ibragimov and Lifshits [15]. The same holds for nonlinear limit theorems: for example, the a.s. versions of the Darling–Erdős theorem require loglog summation, see Berkes and Csáki [4]. By Hardy's theorem mentioned above, the larger weight sequence (d_n) we choose in (1.11), the stronger the result becomes, and thus the strongest, optimal form of the ASCLT is obtained for the maximal weight sequence (d_n). This optimal weight sequence was determined, up to an unknown constant, in our recent paper Hörmann [14]. In this paper we will investigate the analogous problem for central order statistics and we will prove the following results.

THEOREM 1. *Let $X_1, X_2, \ldots$ be i.i.d. r.v.'s and assume that for some non-degenerate distribution function G there exist constants $a_n > 0$ and b_n such that*

$$a_n(X_{n-r_n:n} - b_n) \xrightarrow{\mathcal{L}} G.$$

Assume that (1.2) *holds, that*

$$(1.12) \quad \liminf_n nd_n > 0 \quad \text{and} \quad d_n n^\alpha \text{ is non-increasing for some } 0 < \alpha < 1,$$

and that for some $\varrho > 0$

$$(1.13) \qquad d_n = O\left(\frac{D_n}{n(\log D_n)^\varrho}\right).$$

Then we have

$$(1.14)\qquad \frac{1}{D_N}\sum_{n=1}^{N} d_n I\{a_n(X_{n-r_n:n}-b_n)\leqslant x\}\to G(x)\ \text{a.s.}\quad \text{for any } x\in C_G.$$

As noted above, the larger the sequence (d_n) is, the stronger the statement of Theorem 1 becomes. The second relation of (1.12) implies that $d_n = O(n^{-\alpha})$ for some $0 < \alpha < 1$, which puts no restriction on the growth speed of (d_n) since, as our next theorem will show, the conclusion of Theorem 1 already fails for $d_n = n^{-\alpha}$, which determines a summation equivalent to Cesàro summation. The first relation of (1.12) is also a natural one, since the theorem holds for $d_n = 1/n$, and thus by a similar argument to that given in [4] it follows for smaller sequences as well. The crucial restriction on (d_n) is (1.13) which is an asymptotic negligibility condition resembling Kolmogorov's condition for the LIL, except the factor n in the denominator of (1.13), which is characteristic for a.s. limit theory. Of course, condition (1.13) fails in the Cesàro case $D_n = n$, but it permits

$$D_n = \exp\big((\log n)^{\alpha}\big),\qquad 0<\alpha<1,$$

which borders on the Cesàro case $\alpha = 1$, and thus we see the surprising fact that the optimal weight sequence in Theorem 1 is in some sense closer to Cesàro summation than to logarithmic summation.

Our next theorem states the fact, already mentioned above, that the statement of Theorem 1 becomes false for Cesàro summation. This is a usual feature in this circle of problems; note, however, that its proof presents substantial difficulties in the present case.

THEOREM 2. *Let $U_1, U_2, \ldots$ be i.i.d. r.v.'s, where U_1 is uniformly distributed over* (0, 1), *and assume that* (1.2) *holds. Let (a_n) and (b_n) be the same as in* (1.10). *Assume further that we have positive constants $\alpha_1, \alpha_2, C_1, C_2$ such that*

$$(1.15)\qquad C_1\left(\frac{k}{l}\right)^{\alpha_1}\leqslant\sqrt{\frac{(k-r_k)\,r_l}{r_k(l-r_l)}}\leqslant C_2\left(\frac{l}{k}\right)^{\alpha_2}.$$

Then for any $x\in\mathbf{R}$

$$\frac{1}{N}\sum_{n=1}^{N} I\{a_n(U_{n-r_n:n}-b_n)\leqslant x\}\to\Phi(x)$$

does not hold almost surely or in probability.

It is likely that Theorem 2 holds without (1.15) but this remains open. However, (1.15) contains most cases of interest. For example, it is easily checked that if $r_k = qk+o(k)$, $q\in(0, 1)$, or if r_k is non-decreasing and r_k/k is non-increasing, then (1.15) is satisfied.

2. PROOFS

Our first lemma is the extension of the ASCLT for general summation methods, proved in Hörmann [13].

LEMMA 1. *Let $X_1, X_2, \ldots$ be i.i.d. r.v.'s with $EX_1 = 0$ and $EX_1^2 = 1$. Assume that (D_N) defines a summation method such that* (1.12) *and* (1.13) *hold. Then for any $x \in \boldsymbol{R}$*

$$\frac{1}{D_N} \sum_{n=1}^{N} d_n I\{n^{-1/2}(X_1 + \ldots + X_n) \leqslant x\} \to \Phi(x) \text{ a.s.}$$

This result will be crucial in the sequel. Note, however, that for technical reasons we need Lemma 1 for triangular arrays as well; this is given by the next lemma.

LEMMA 2. *Let $\{\xi_{n,i}\colon 1 \leqslant i \leqslant n;\ n \geqslant 1\}$ be a triangular array of r.v.'s satisfying*

(2.1) $$E\xi_{n,i} = 0 \quad \text{for each } (n, i),$$

(2.2) *the sequences $(\xi_{n,1})_{n \geqslant 1}, (\xi_{n,2})_{n \geqslant 2}, \ldots$ are mutually independent,*

(2.3) *there is some C such that $E\xi_{n,i}^2 \leqslant C/n$ for each (n, i),*

(2.4) $$\lim_{n \to \infty} \sum_{i=1}^{n} E(\xi_{n,i}^2) = 1,$$

(2.5) $$\lim_{n \to \infty} \sum_{i=1}^{n} E(\xi_{n,i}^2;\ |\xi_{n,i}| > \varepsilon) = 0 \quad \text{for all } \varepsilon > 0.$$

Assume that (D_N) defines a summation method such that (1.12) *and* (1.13) *hold. Then for any $x \in \boldsymbol{R}$*

$$\frac{1}{D_N} \sum_{n=1}^{N} d_n I\{\xi_{n,1} + \ldots + \xi_{n,n} \leqslant x\} \to \Phi(x) \text{ a.s.}$$

Lemma 2 is a common generalization of Lemma 1 and a version of the ASCLT for triangular arrays due to Lesigne [17], which states Lemma 2 in the case $d_n = 1/n$. As Lesigne observed, the standard proof of Lacey and Philipp applies also in the triangular array case, and in fact the proof of Lemma 2 is essentially the same as that of Lemma 1.

Proof of Theorem 1. As noted by Peng and Qi ([18], proof of Theorem 2), Theorem A can be reduced to the case of i.i.d. uniform r.v.'s by a simple quantile-transformation argument, and for the same reason it suffices to prove Theorem 1 for this special case. Our proof is based on the following useful and easily verified duality relation: For any sequence of random variables

$X_1, X_2, \ldots$ we have

$$\{X_{r:n} \leqslant x\} = \{\sum_{i=1}^{n} I\{X_i \leqslant x\} \geqslant r\}. \tag{2.6}$$

As our random variables X_n are uniform, we will denote them by U_n $(n = 1, 2, \ldots)$, and $U_{i:n}$ $(1 \leqslant i \leqslant n,\ n \geqslant 1)$ will denote the corresponding order statistics. Using the values of a_n and b_n in (1.10), we infer from (2.6) and some simple algebra that

$$\{a_n(U_{n-r_n:n}-b_n) \leqslant x\} = \left\{-\frac{a_n}{n}\sum_{i=1}^{n}\left(I\{U_i \leqslant x/a_n+b_n\}-(x/a_n+b_n)\right) \leqslant x\right\}. \tag{2.7}$$

By (1.10) and (1.2) we have

$$a_n b_n = \sqrt{\frac{n(n-r_n)}{r_n}} \to \infty \quad \text{and} \quad a_n(1-b_n) = \sqrt{\frac{nr_n}{n-r_n}} \to \infty. \tag{2.8}$$

Thus for any fixed $x \in \boldsymbol{R}$ we have $0 < x/a_n+b_n < 1$ if n is large enough, which we assume from now on. Define

$$\xi_{n,i} = -\frac{a_n}{n}\left(I\{U_i \leqslant x/a_n+b_n\}-(x/a_n+b_n)\right) \quad (1 \leqslant i \leqslant n). \tag{2.9}$$

Using (1.10), by easy calculations we obtain

$$E\xi_{n,i}^2 = \frac{1}{n}+\frac{x}{n}\left(\frac{r_n}{n(n-r_n)}\right)^{1/2}-\frac{x}{n}\left(\frac{n-r_n}{nr_n}\right)^{1/2}-\frac{x^2}{n^2}.$$

Clearly,

$$\frac{r_n}{n(n-r_n)} \leqslant \frac{1}{n-r_n} \quad \text{and} \quad \frac{n-r_n}{nr_n} \leqslant \frac{1}{r_n},$$

and hence by (1.2) we have $E\xi_{n,i}^2 \sim 1/n$. Now it is easily checked that the triangular array $\{\xi_{n,i}\colon 1 \leqslant i \leqslant n;\ n \geqslant 1\}$ satisfies the conditions of Lemma 2, and the proof of Theorem 1 is completed by (2.7).

Proof of Theorem 2. It suffices to show that

$$\liminf_{N\to\infty}\frac{1}{N^2}\operatorname{Var}\sum_{n=1}^{N} I\{a_n(U_{n-r_n:n}-b_n) \leqslant x\} > 0. \tag{2.10}$$

For this purpose we define $\xi_{n,i}$ as in (2.9) and again set $S_n = \xi_{n,1}+\ldots+\xi_{n,n}$.

LEMMA 3. *Under the conditions of Theorem 2 there exists a $K > 0$ such that for $\kappa_1 = \frac{1}{2}\min\{C_1, C_2^{-1}\}$ and $\gamma = \frac{1}{2}+\max\{\alpha_1, \alpha_2\}$ we have*

$$\operatorname{Cov}(S_k, S_l) \geqslant \kappa_1 (k, l)^{\gamma}, \quad K \leqslant k \leqslant l.$$

Proof. Since the $\xi_{n,1}, \ldots, \xi_{n,n}$ are i.i.d. and the condition (2.2) is satisfied, we have for $k \leqslant l$

$$\operatorname{Cov}(S_k, S_l) = \operatorname{Cov}\Big(S_k, \sum_{i=1}^{k} \xi_{l,i}\Big) = k\operatorname{Cov}(\xi_{k,1}, \xi_{l,1})$$

$$= a_k \frac{a_l}{l}\big(\min\{xa_k^{-1}+b_k,\, xa_l^{-1}+b_l\} - (xa_k^{-1}+b_k)(xa_l^{-1}+b_l)\big)$$

$$= \frac{1}{l}(x+a_k b_k)\big(a_l(1-b_l)-x\big),$$

where we assumed first that $xa_k^{-1}+b_k \leqslant xa_l^{-1}+b_l$. Now we use (2.8) and conclude by (1.15) that

$$\operatorname{Cov}(S_k, S_l) = \big(1+o(1)\big)\left(\frac{k}{l}\right)^{1/2}\sqrt{\frac{(k-r_k)r_l}{r_k(l-r_l)}} \geqslant \frac{1}{2}C_1\left(\frac{k}{l}\right)^{1/2+\alpha_1}$$

if $k \geqslant K$. (Here $o(1)$ is meant for $\min\{k, l\} \to \infty$.) Similarly one can show that if $xa_k^{-1}+b_k > xa_l^{-1}+b_l$, then

$$\operatorname{Cov}(S_k, S_l) = \big(1+o(1)\big)\left(\frac{k}{l}\right)^{1/2}\sqrt{\frac{(l-r_l)r_k}{r_l(k-r_k)}} \geqslant \frac{1}{2}\frac{1}{C_2}\left(\frac{k}{l}\right)^{1/2+\alpha_2}$$

if $k \geqslant K$.

LEMMA 4. *Let (T_k, T_l) be a 2-dimensional Gaussian vector with zero expectation and the same covariance matrix as (S_k, S_l). Let further $\varphi_{k,l}(s, t) = E\big(\exp(isS_k+itS_l)\big)$ be the characteristic function of (S_k, S_l) and $\psi_{k,l}(s, t)$ the characteristic function of (T_k, T_l). Then for any $(s, t) \in \mathbf{R}^2$*

(2.11) $$|\varphi_{k,l}(s, t) - \psi_{k,l}(s, t)| \to 0 \quad \textit{if} \ \min\{k, l\} \to \infty.$$

Proof. In the following we assume that $k \leqslant l$. Let $\sigma_{kl} := \operatorname{Cov}(S_k, S_l)$. Then clearly

(2.12) $$\psi_{k,l}(s, t) = \exp\big(-\tfrac{1}{2}(\sigma_{kk}s^2 + \sigma_{ll}t^2 + 2\sigma_{kl}st)\big)$$

and observe also that

(2.13) $$\sigma_{kl} = kE\xi_{k,1}\xi_{l,1}.$$

Since $\xi_{l,1}, \ldots, \xi_{l,l}$ is an i.i.d. sequence, we have

(2.14) $$\varphi_{k,l}(s, t) = \big(E\exp(is\xi_{k,1}+it\xi_{l,1})\big)^k \big(E\exp(it\xi_{l,1})\big)^{l-k}.$$

Using

$$\left|e^{ix} - \sum_{k=0}^{n}\frac{(ix)^k}{k!}\right| \leqslant \frac{|x|^{n+1}}{(n+1)!},$$

we derive easily

$$\left|E\exp(is\xi_{k,1}+it\xi_{l,1})-\left(1-\tfrac{1}{2}E(s\xi_{k,1}+t\xi_{l,1})^2\right)\right| \leqslant C\cdot\max\{|s|,|t|\}^3\,[E|\xi_{k,1}|^3+E|\xi_{l,1}|^3].$$

Relations (1.2) and (1.10) imply $a_n = o(n)$; further from the definition of $\xi_{k,1}$ and $E\xi_{k,1}^2 \sim k^{-1}$ it follows that

$$E|\xi_{k,1}|^3 \leqslant \frac{2a_k}{k}E\xi_{k,1}^2 = o(k^{-1}).$$

Thus,

$$E\exp(is\xi_{k,1}+it\xi_{l,1}) = 1-\tfrac{1}{2}E(s\xi_{k,1}+t\xi_{l,1})^2+\varrho(s,t,k,l) \tag{2.15}$$

with $|k\varrho(s,t,k,l)| = o(1)$ for $k\to\infty$. Next we observe that by (2.13)

$$E(s\xi_{k,1}+t\xi_{l,1})^2 = s^2\frac{\sigma_{kk}}{k}+t^2\frac{\sigma_{ll}}{l}+2st\frac{\sigma_{kl}}{k}.$$

Some simple analysis shows that for any $r>0$ and $0\leqslant t\leqslant 1$

$$|(1-t)^r-e^{-rt}| \leqslant re^{-rt+t}\left(e^{-t}-(1-t)\right) \leqslant \frac{rt^2}{2}e^{-rt+t} \leqslant \frac{t}{2}.$$

Hence from $E(s\xi_{k,1}+t\xi_{l,1})^2\to 0$ for $k\to\infty$ it follows that

$$\left|\left(1-\frac{1}{2}E(s\xi_{k,1}+t\xi_{l,1})^2\right)^k-\exp\left(-\frac{1}{2}\left(s^2\sigma_{kk}+\frac{k}{l}\sigma_{ll}t^2+2st\sigma_{kl}\right)\right)\right|\to 0 \quad (k\to\infty).$$

Further we have by (2.15)

$$\begin{aligned}(2.16)\quad &\left|\left(E\exp(is\xi_{k,1}+it\xi_{l,1})\right)^k-\exp\left(-\frac{1}{2}\left(s^2\sigma_{kk}+\frac{k}{l}\sigma_{ll}t^2+2st\sigma_{kl}\right)\right)\right|\\ &\leqslant\left|\left(1-\frac{1}{2}E(s\xi_{k,1}+t\xi_{l,1})^2+\varrho(s,t,k,l)\right)^k-\left(1-\frac{1}{2}E(s\xi_{k,1}+t\xi_{l,1})^2\right)^k\right|\\ &\quad+\left|\left(1-\frac{1}{2}E(s\xi_{k,1}+t\xi_{l,1})^2\right)^k-\exp\left(-\frac{1}{2}\left(s^2\sigma_{kk}+\frac{k}{l}\sigma_{ll}t^2+2st\sigma_{kl}\right)\right)\right|.\end{aligned}$$

From the fact that $|1-\frac{1}{2}E(s\xi_{k,1}+t\xi_{l,1})^2+\varrho(s,t,k,l)|\leqslant 1$ (since by (2.15) it is a characteristic function) and $|1-\frac{1}{2}E(s\xi_{k,1}+t\xi_{l,1})^2|\leqslant 1$ for k large enough (since $E(s\xi_{k,1}+t\xi_{l,1})^2\to 0$) we infer by the mean value theorem that (2.16) is less than or equal to $|k\varrho(s,t,k,l)|$ which tends to zero for $k\to\infty$. This proves that

$$\left|\left(E\exp(is\xi_{k,1}+it\xi_{l,1})\right)^k-\exp\left(-\frac{1}{2}\left(s^2\sigma_{kk}+\frac{k}{l}\sigma_{ll}t^2+2st\sigma_{kl}\right)\right)\right|\to 0 \quad (k\to\infty). \tag{2.17}$$

Similarly one can show that

$$(2.18)\qquad \left|\left(E\exp(it\xi_{l,1})\right)^{l-k}-\exp\left(-\tfrac{1}{2}(1-k/l)\,\sigma_{ll}\,t^2\right)\right|\to 0 \qquad (k\to\infty).$$

Combining (2.12), (2.14), (2.17) and (2.18), we obtain (2.11).

LEMMA 5. *Define κ_1 and γ as in Lemma 3, let $f=1_{(-\infty,x]}$ for some $x\in\boldsymbol{R}$, and let $\varepsilon>0$. Then there exist an $A=A(\varepsilon)>0$ and positive constants κ_2, μ such that*

$$\operatorname{Cov}\left(f(S_k),f(S_l)\right)\geqslant\kappa_2\,(k/l)^{\mu}\qquad(k\leqslant l)$$

if $k\geqslant A$ and $\kappa_1(k/l)^{\gamma}\geqslant\varepsilon$.

It is needless to say that Lemma 5 does not hold in the trivial cases $x=\pm\infty$. In the sequel $c_1, c_2, \ldots$ denote positive constants.

Proof. Again we assume that $k\leqslant l$ and denote by $P_{k,l}$ and $Q_{k,l}$ the probability measures belonging to (S_k, S_l) and (T_k, T_l), respectively, defined in Lemma 4. Since the difference of the corresponding characteristic functions $|\varphi_{k,l}(s,t)-\psi_{k,l}(s,t)|$ tends to zero for $k\to\infty$, we see that the Prokhorov distance $\varepsilon_{k,l}:=\pi(P_{k,l},Q_{k,l})\to 0$ for $k\to\infty$ (see, e.g., Lemma 2.2 in Berkes and Philipp [5]). By a special case of the Strassen–Dudley theorem (cf. Dudley [9], Theorem 11.6.2), there exist for every (k,l) a probability space $(\Omega_{kl},\mathscr{F}_{kl},\mathscr{P}_{kl})$ and random vectors (S_k^*, S_l^*) and (T_k', T_l') defined on it, with respective distributions P_{kl} and Q_{kl} such that

$$(2.19)\qquad \mathscr{P}_{kl}\left(\|(S_k^*,S_l^*)-(T_k',T_l')\|>\varepsilon_{k,l}\right)\leqslant\varepsilon_{k,l},$$

where $\|\cdot\|$ denotes the Euclidean distance. Setting

$$(T_k^*,T_l^*)=\left(\frac{T_k'}{\sqrt{\sigma_{kk}}},\frac{T_l'}{\sqrt{\sigma_{ll}}}\right)$$

we get from (2.19) and $\sigma_{kk}\to 1$ that for every $\delta>0$ there is a $k(\delta)$ such that for $k>k(\delta)$

$$(2.20)\qquad \mathscr{P}_{kl}\left(\|(S_k^*,S_l^*)-(T_k^*,T_l^*)\|>\delta\right)<\delta.$$

Define

$$c_{k,l}:=\operatorname{Cov}(S_k^*,S_l^*),\qquad d_{k,l}:=\operatorname{Cov}\left(f(S_k^*),f(S_l^*)\right),$$
$$c_{k,l}^*:=\operatorname{Cov}(T_k^*,T_l^*),\qquad d_{k,l}^*:=\operatorname{Cov}\left(f(T_k^*),f(T_l^*)\right).$$

We note here that the sequences (S_k^*) and (T_k^*) are uniformly integrable. This is clear for T_k^* and can be easily verified for S_k^*, e.g. by showing that $E(S_k^*)^4\leqslant M$, where M is some constant which does not depend on k. Thus (2.20) implies that

$$|c_{k,l}-c_{k,l}^*|\to 0\qquad(k\to\infty),$$

and hence if $k\geqslant A_1(\varepsilon)$, we have $|c_{k,l}-c_{k,l}^*|\leqslant\varepsilon^2$. Without loss of generality we may assume $\varepsilon\leqslant 1/2$ and $A_1\geqslant K$, where K stems from Lemma 3. By Lemma 3 and $\kappa_1 (k/l)^\gamma \geqslant \varepsilon$ we get $|1-c^*_{k,l}/c_{k,l}| \leqslant \varepsilon$, whence

$$c^*_{k,l} \geqslant (1-\varepsilon)\, c_{k,l} \geqslant \tfrac{1}{2} c_{k,l}. \tag{2.21}$$

Since the vector (T_k^*, T_l^*) is Gaussian with standard normally distributed components, we get

$$d^*_{k,l} = \sum_{\nu=0}^{\infty} \frac{(c^*_{k,l})^\nu}{\nu!} \alpha_\nu^2,$$

where α_ν are the coefficients in the Hermite expansion of $f - Ef(\mathcal{N}_{0,1})$. (See e.g. the proof of Lemma 10.2 in Rozanov [20].) Since f is non-constant, there is some $\nu_0 \geqslant 0$ such that $|\alpha_{\nu_0}| > 0$. This shows that

$$d^*_{k,l} \geqslant c_1 (c^*_{k,l})^{\nu_0}. \tag{2.22}$$

Clearly, Lemma 3 and (2.21)–(2.22) imply that if $k \geqslant A_1$ and $\kappa_1 (k/l)^\gamma \geqslant \varepsilon$, then

$$d^*_{k,l} \geqslant c_2 (k/l)^{\gamma\nu_0}, \tag{2.23}$$

and thus $d^*_{k,l} \geqslant c_3 \varepsilon^{\nu_0}$. Remember that $f = 1_{(-\infty,x]}$, and thus for any $\delta > 0$ we have

$$\mathscr{P}_{kl}\big(\|(f(S_k^*),\, f(S_l^*)) - (f(T_k^*),\, f(T_l^*))\| > 0\big)$$
$$\leqslant \mathscr{P}_{kl}\big(T_k^* \in U_x(\delta)\big) + \mathscr{P}_{kl}\big(T_l^* \in U_x(\delta)\big) + \mathscr{P}_{kl}\big(\|(S_k^*, S_l^*) - (T_k^*, T_l^*)\| > \delta\big),$$

where $U_\delta(x) = (x-\delta, x+\delta)$. Since $\delta > 0$ is arbitrary and T_k^*, T_l^* are standard normal r.v.'s, (2.20) implies that

$$\big(f(S_k^*),\, f(S_l^*)\big) - \big(f(T_k^*),\, f(T_l^*)\big) \xrightarrow{P} 0,$$

and thus $|d_{k,l} - d^*_{k,l}| \to 0$ as $k \to \infty$. Now we choose $A_2(\varepsilon)$ such that $|d^*_{k,l} - d_{k,l}| \leqslant c_3 \varepsilon^{\nu_0+1}$ for $k \geqslant A_2$. Since $d^*_{k,l} \geqslant c_3 \varepsilon^{\nu_0}$, this yields, by the same argument as above,

$$d_{k,l} \geqslant (1-\varepsilon)\, d^*_{k,l} \geqslant \tfrac{1}{2} d^*_{k,l}.$$

We set $A = \max\{A_1, A_2\}$ and the lemma is proved by (2.23).

LEMMA 6. *Let $f = 1_{(-\infty,x]}$, $x \in \mathbf{R}$. Then there is some $L > 0$ such that for all $N \geqslant N_0$*

$$\operatorname{Var}\Big(\sum_{k=1}^{N} f(S_k)\Big) \geqslant LN^2.$$

Proof. Define A, κ_1 and κ_2 as before. Then from $|f| \leqslant 1$ we get

$$\operatorname{Var}\Big(\sum_{k=1}^{N} f(S_k)\Big) = \sum_{k=1}^{N} \operatorname{Var} f(S_k) + 2 \sum_{1 \leqslant k < l \leqslant N} \operatorname{Cov}\big(f(S_k),\, f(S_l)\big)$$

$$= O(N)+2 \sum_{\substack{1\leqslant k<l\leqslant N\\ \kappa_1(k/l)^\gamma<\varepsilon}} \operatorname{Cov}\big(f(S_k), f(S_l)\big)+2 \sum_{\substack{A\leqslant k<l\leqslant N\\ \kappa_1(k/l)^\gamma\geqslant\varepsilon}} \operatorname{Cov}\big(f(S_k), f(S_l)\big)$$

$$=: O(N)+2S^{(1)}+2S^{(2)}.$$

Trivially,

$$|S^{(1)}| \leqslant \sum_{l=1}^{N} \sum_{1\leqslant k\leqslant l(\varepsilon/\kappa_1)^{1/\gamma}} 2 \leqslant c_4\, \varepsilon^{1/\gamma} N^2. \tag{2.24}$$

By Lemma 5 we get

$$S^{(2)} \geqslant \sum_{A\leqslant k<l\leqslant N} \kappa_2 \left(\frac{k}{l}\right)^{\mu} - \sum_{\substack{A\leqslant k<l\leqslant N\\ \kappa_1(k/l)^\gamma<\varepsilon}} \operatorname{Cov}\big(f(S_k), f(S_l)\big). \tag{2.25}$$

It is easily seen that

$$\sum_{A\leqslant k<l\leqslant N} \kappa_2 \left(\frac{k}{l}\right)^{\mu} \sim \frac{\kappa_2}{2(\mu+1)} N^2 \qquad (N\to\infty),$$

and thus using the same argument as in (2.24) to estimate the second sum in (2.25) we can always achieve that for sufficiently large N

$$S^{(2)} \geqslant \left(\frac{\kappa_2}{3(\mu+1)} - c_4\, \varepsilon^{1/\gamma}\right) N^2.$$

Summing up we get for N large enough

$$\operatorname{Var}\Big(\sum_{k=1}^{N} f(S_k)\Big) \geqslant \left(\frac{\kappa_2}{3(\mu+1)} - 2c_4\, \varepsilon^{1/\gamma} + o(1)\right) N^2,$$

and the term in brackets is greater than or equal to $\kappa_2/(4(\mu+1))$ if ε is small and N is large enough. This proves Lemma 6.

Using again the duality (2.7) we get immediately (2.10).

Acknowledgement. I am grateful to Professor István Berkes for careful reading the manuscript. Especially I appreciate his valuable comments to Lemma 5 closing a gap in the first version.

REFERENCES

[1] M. Atlagh, *Théorème central limite presque sûr et loi du logarithme itéré pour des sommes de variables aléatoires indépendantes*, C. R. Acad. Sci. Paris Sér. I. 316 (1993), pp. 929–933.

[2] M. Atlagh and M. Weber, *Le théorème central limite presque sûr*, Expo. Math. 18 (2000), pp. 97–126.

[3] I. Berkes, *Results and problems related to the pointwise central limit theorem*, in: *Asymptotic Methods in Probability and Statistics* (*Ottawa, ON, 1997*), North-Holland, Amsterdam 1998, pp. 59–96.

[4] I. Berkes and E. Csáki, *A universal result in almost sure central limit theory*, Stochastic Process. Appl. 94 (2001), pp. 105–134.
[5] I. Berkes and W. Philipp, *Approximation theorems for independent and weakly dependent random vectors*, Ann. Probab. 7 (1979), pp. 29–54.
[6] G. A. Brosamler, *An almost everywhere central limit theorem*, Math. Proc. Cambridge Philos. Soc. 104 (1988), pp. 561–574.
[7] K. Chandrasekharan and S. Minakshisundaram, *Typical Means*, Oxford University Press, 1952.
[8] S. Cheng, L. Peng and Y. Qi, *Almost sure convergence in extreme value theory*, Math. Nachr. 190 (1998), pp. 43–50.
[9] R. M. Dudley, *Real Analysis and Probability*, Cambridge Stud. Adv. Math., Vol. 74, Cambridge University Press, Cambridge 2002.
[10] I. Fahrner and U. Stadtmüller, *On almost sure max-limit theorems*, Statist. Probab. Lett. 37 (1998), pp. 229–236.
[11] A. Fisher, *A pathwise central limit theorem for random walks*, preprint, 1989.
[12] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, Wiley Ser. Probab. Math. Statist., New York–Chichester–Brisbane 1978.
[13] S. Hörmann, *An extension of the almost sure central limit theorem*, Statist. Probab. Lett. 76 (2006), pp. 191–202.
[14] S. Hörmann, *Critical behavior in almost sure central limit theory*, preprint.
[15] I. A. Ibragimov and M. A. Lifshits, *On limit theorems of "almost sure" type*, Theory Probab. Appl. 44 (2000), pp. 254–272.
[16] M. T. Lacey and W. Philipp, *A note on the almost sure central limit theorem*, Statist. Probab. Lett. 9 (1990), pp. 201–205.
[17] E. Lesigne, *Almost sure central limit theorem for strictly stationary processes*, Proc. Amer. Math. Soc. 128 (2000), pp. 1751–1759.
[18] L. Peng and Y. Qi, *Almost sure convergence of the distributional limit theorem for order statistics*, Probab. Math. Statist. 23 (2003), pp. 217–228.
[19] R.-D. Reiss, *Approximate Distributions of Order Statistics*, Springer, New York 1989.
[20] Y. A. Rozanov, *Stationary Random Processes*, Holden-Day, 1967.
[21] P. Schatte, *On strong versions of the central limit theorem*, Math. Nachr. 137 (1988), pp. 249–256.
[22] U. Stadtmüller, *Almost sure versions of distributional limit theorems for certain order statistics*, Statist. Probab. Lett. 58 (2002), pp. 413–426.

Institute of Statistics
Graz University of Technology
Steyrergasse 17/IV
8010 Graz, Austria
E-mail: shoermann@tugraz.at

Received on 21.11.2005