YOKOHAMA MATHEMATICAL
JOURNAL VOL. 32, 1984

AN EXACT RATE OF CONVERGENCE IN THE INVARIANCE PRINCIPLE FOR MARTINGALE DIFFERENCE ARRAYS

By

SHÛYA KANAGAWA

(Received March 22, 1984)

ABSTRACT. A convergence rate in the invariance principle for martingale difference arrays is obtained. The rate obtained depends essentially on the asymptotic behavior of the conditional variances. An example which shows the rate obtained is exact is shown. It will be also remarked, by an example, that the convergence rate in the central limit theorem is not always faster than the one in the invariance principle.

1. Introduction and results. Let $X_n=(X_{n,1}, X_{n,2}, \cdots, X_{n,n})$, $n\geqq 1$, be a triangular array of real valued and square integrable random variables satisfying

$$E(X_{n,i}|\mathscr{F}_{n,i-1})=0 \text{ a.s.}, \quad 1\leqq i\leqq n, \tag{1.1}$$

where $\mathscr{F}_{n,i}$ is the σ-field generated by $X_{n,1}, X_{n,2}, \cdots, X_{n,i}$ and $\mathscr{F}_{n,0}$ is the trivial σ-field. We write

$$S_{n,k}=\sum_{i=1}^{k} X_{n,i}, \quad 1\leqq k\leqq n, \quad V_n^2=\sum_{i=1}^{n} E(X_{n,i}^2|\mathscr{F}_{n,i-1}).$$

Under some assumptions for X_n (see e.g. Theorem 4.1 in Hall and Heyde [1]), it follows from Donsker's invariance principle that, for each $\lambda\geqq 0$,

$$\lim_{n\to\infty} P\left(\max_{1\leq k\leq n}|S_{n,k}|\leqq\lambda\right)=T(\lambda), \tag{1.2}$$

$$T(\lambda)=P\left(\max_{0\leq t\leq 1}|\boldsymbol{B}(t)|\leqq\lambda\right)=\frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}\exp\left\{-\frac{(2k+1)^2\pi^2}{8\lambda^2}\right\}, \tag{1.3}$$

$\{\boldsymbol{B}(t), t\geqq 0\}$ being a standard Wiener process.

On the other hand, as a special case of the result due to Hall and Heyde (Theorem 4.6 in [1]), we have the following result: If we assume that

$$\max_{1\leq k\leq n}|X_{n,k}|\leqq K_1 n^{-1/2} \text{ a.s.}, \quad n\geqq 1, \tag{1.4}$$

$$P(|V_n^2-1|\geqq 20K_1^2 n^{-1/2}(\log n)^2)=O(n^{-1/4}(\log n)^{3/2}), \quad n\to\infty, \tag{1.5}$$

then it holds that

$$\Delta_n \equiv \sup_{\lambda} \left| P\left(\max_{1 \leqq k \leqq n} |S_{n,k}| \leqq \lambda \right) - T(\lambda) \right| = O(n^{-1/4}(\log n)^{3/2}) , \quad n \to \infty . \tag{1.6}$$

K_1, K_2, $\cdots$ denote some positive constants independent of n throughout the paper.

We shall in this paper obtain a better rate $O(n^{-1/2} \log n)$ in (1.6) under the assumption (1.4) and a condition which is weaker than (1.5) as a special case of the following theorem.

Theorem 1. *Let* $\{X_n\}_{n \geqq 1}$ *be a triangular array of martingale difference sequences satisfying* (1.1). *Let* $g(\cdot)$ *be a positive function such that*

$$\liminf_{n \to \infty} g(n) n^{1/2} (\log n)^{-1} \geqq 1 . \tag{1.7}$$

If we assume (1.4) *and that*

$$P(|V_n^2 - 1| \geqq g(n)) = O(g(n)) , \quad n \to \infty , \tag{1.8}$$

then we have

$$\Delta_n = O(g(n)) , \quad n \to \infty . \tag{1.9}$$

Our rate is sharp in the sense that there exists an example of a martingale difference array (see the example in the last section) which satisfies the condition (1.4) and

$$|V_n^2 - 1| = g(n)/2 + O(n^{-1}) \text{ a.s.} , \quad n \to \infty , \tag{1.10}$$

and hence $\Delta_n \leqq K_2 g(n)$ by Theorem 1, while

$$\Delta_n \geqq K_3 g(n) , \tag{1.11}$$

for some constant $K_3 > 0$.

We now point out about the relationship between our result and the rate of convergence in the central limit theorem. It is plausible that for independent random variables the ordinary rate of convergence to the normal law of single sums is much faster than that for maximums of sums (see Sawyer [3]). However this situation does not keep true for the case of martingale difference arrays. Actually, in an example of Hall and Heyde which satisfies all assumptions of Theorem 1 with $g(n) = n^{-1/2}(\log n)^2$,

$$\sup_{-\infty < \lambda < \infty} \left| P(S_{n,n} \leqq \lambda) - (2\pi)^{-1/2} \int_{-\infty}^{\lambda} e^{-t^2/2} dt \right| \geqq K_4 n^{-1/4} \log n \tag{1.12}$$

holds (see Remarks in [1], p. 84), but by our present result the convergence rate of (1.2) is given by

$$\Delta_n = O(n^{-1/2}(\log n)^2) , \quad n \to \infty ,$$

which is faster than (1.12).

2. Proof of Theorem 1. In proving the theorem we shall use the following lemma which is due to Skorokhod (Lemma 1 in [4], p. 150) and Kato (Lemma 2 in [2]).

Lemma 1. *Suppose that* $\{\mathscr{M}_1, \mathscr{M}_2, \cdots, \mathscr{M}_n\}$ *is a nondecreasing sequence of* σ*-fields and that* $\eta_1, \eta_2, \cdots, \eta_n$ *are random variables such that* η_i *is* $\mathscr{M}_i$*-measurable,* $E(\eta_i|\mathscr{M}_{i-1})=0$ a.s., *and*

$$E(\eta_i^2|\mathscr{M}_{i-1})\leqq K_5 n^{-1} \text{ a.s.}, \quad E(\eta_i^4)\leqq K_6 n^{-2}, \tag{2.1}$$

for each i. *Then we have*

$$P\left(\left|\sum_{i=1}^{n}\eta_i\right|\geqq \log n\right)\leqq n^{-1}\{2\exp(K_5(1+e^2/2))+K_6\}. \tag{2.2}$$

Write

$$F_n(\lambda)=P\left(\max_{1\leq k\leq n}|S_{n,k}|\leqq\lambda\right).$$

By the Skorokhod embedding theorem due to Strassen [5] applied to the martingale difference sequence $\boldsymbol{X}_n$, there is a standard Wiener process $\{\boldsymbol{B}(t), t\geqq 0\}$ together with a sequence of nonnegative random variables $\tau_1, \tau_2, \cdots, \tau_n$ on a new probability space such that

$$\{S_{n,1}, S_{n,2}, \cdots, S_{n,n}\}\overset{\mathrm{d}}{=}\left\{\boldsymbol{B}(\tau_1), \boldsymbol{B}(\tau_1+\tau_2), \cdots, \boldsymbol{B}\left(\sum_{i=1}^{n}\tau_i\right)\right\} \tag{2.3}$$

("$\overset{\mathrm{d}}{=}$" means the equivalence in joint distributions),

$$E(\tau_i|\mathscr{B}_{n,i-1})=E(Y_{n,i}^2|\mathscr{B}_{n,i-1})=E(Y_{n,i}^2|\mathscr{G}_{n,i-1}) \text{ a.s.}, \tag{2.4}$$

$$E(\tau_i^k|\mathscr{B}_{n,i-1})\leqq K_7 E(Y_{n,i}^{2k}|\mathscr{B}_{n,i-1})=K_7 E(Y_{n,i}^{2k}|\mathscr{G}_{n,i-1}) \text{ a.s.}, \tag{2.5}$$

for each positive integer k and $1\leqq i\leqq n$, where

$$Y_{n,1}\equiv \boldsymbol{B}\left(\sum_{k=1}^{i}\tau_k\right)-\boldsymbol{B}\left(\sum_{k=1}^{i-1}\tau_k\right)$$

and $\mathscr{G}_{n,i}$ (respectively $\mathscr{B}_{n,i}$) is the σ-field generated by $Y_{n,1}, Y_{n,2}, \cdots, Y_{n,i}$ (respectively $Y_{n,1}, Y_{n,2}, \cdots, Y_{n,i}$ and $\boldsymbol{B}(t)$ for $0\leqq t\leqq \sum_{k=1}^{i}\tau_k$) for each $1\leqq i\leqq n$. From the assumption (1.4) and the construction of $\{\tau_i\}$, we easily find that

$$\max_{0\leq k\leq n-1}\max_{0\leq t\leq \tau_{k+1}}\left|\boldsymbol{B}\left(t+\sum_{i=1}^{k}\tau_i\right)-\boldsymbol{B}\left(\sum_{i=1}^{k}\tau_i\right)\right|\leqq K_1 n^{-1/2} \text{ a.s.}, \tag{2.6}$$

and hence

$$F_n(\lambda)\leqq P\Big(\max_{0\leqq t\leqq \sum_{i=1}^{n}\tau_i}|\boldsymbol{B}(t)|\leqq\lambda+K_1n^{-1/2}\Big)\,, \tag{2.7}$$

$$F_n(\lambda)\geqq P\Big(\max_{0\leqq t\leqq \sum_{i=1}^{n}\tau_i}|\boldsymbol{B}(t)|\leqq\lambda-K_1n^{-1/2}\Big)\,. \tag{2.8}$$

Write $\xi_i=n^{1/2}(\tau_i-E(\tau_i|\mathscr{B}_{n,i-1}))$ for each i. From the inequality (2.7), we have

$$\begin{aligned} F_n(\lambda)\leqq & P\Big(\max_{0\leqq t\leqq \sum_{i=1}^{n}\tau_i}|\boldsymbol{B}(t)|\leqq\lambda+K_1n^{-1/2},\ \Big|\sum_{i=1}^{n}\tau_i-1\Big|<2g(n)\Big) \\ & +P\Big(\Big|\sum_{i=1}^{n}\tau_i-1\Big|\geqq 2g(n)\Big) \\ \leqq & P\Big(\max_{0\leqq t\leqq 1-2g(n)}|\boldsymbol{B}(t)|\leqq\lambda+K_1n^{-1/2}\Big)+P\Big(\Big|\sum_{i=1}^{n}n^{-1/2}\xi_i\Big|\geqq g(n)\Big) \\ & +P\Big(\Big|\sum_{i=1}^{n}E(\tau_i|\mathscr{B}_{n,i-1})-1\Big|\geqq g(n)\Big)\,. \end{aligned} \tag{2.9}$$

We first estimate the second part of the extreme right hand side of (2.9). From the definition of ξ_i and the inequalities (1.4), (2.4) and (2.5), we have

$$\begin{aligned} & E(\xi_i|\mathscr{B}_{n,i-1})=0 \text{ a.s.}\,, \\ & E(\xi_i^2|\mathscr{B}_{n,i-1})\leqq nE(\tau_i^2|\mathscr{B}_{n,i-1})\leqq K_8nE(X_{n,i-1}^4|\mathscr{B}_{n,i-1})\leqq K_9n^{-1} \quad \text{a.s.}\,, \\ & E(\xi_i^4)\leqq n^2E(\tau_i^4)\leqq K_{10}n^2E(Y_{n,i-1}^8)\leqq K_{11}n^{-2}\,, \end{aligned}$$

for each i. Thus we can use Lemma 1 to obtain, from the assumption (1.7), that

$$P\Big(\Big|\sum_{i=1}^{n}n^{-1/2}\xi_i\Big|\geqq g(n)\Big)\leqq P\Big(\Big|\sum_{i=1}^{n}\xi_i\Big|\geqq \log n\Big)\leqq n^{-1}\{2\exp(K_9(1+e^2/2))+K_{11}\}\,. \tag{2.10}$$

As to the last term of the right hand side of (2.9), we have from (2.4) and (1.8) that

$$P\Big(\Big|\sum_{i=1}^{n}E(\tau_i|\mathscr{B}_{n,i-1})-1\Big|\geqq g(n)\Big)=P\Big(\Big|\sum_{i=1}^{n}E(Y_{n,i}^2|\mathscr{G}_{n,i-1})-1\Big|\geqq g(n)\Big)=O(g(n))\,, \tag{2.11}$$

and as to the first term we have, from the scaling property of the Wiener process, that

$$\begin{aligned} & P\Big(\max_{0\leqq t\leqq 1-2g(n)}|\boldsymbol{B}(t)|\leqq\lambda+K_1n^{-1/2}\Big) \\ & =P\Big(\max_{0\leqq t\leqq 1}|B((1-2g(n))t)|\leqq\lambda+K_1n^{-1/2}\Big) \end{aligned} \tag{2.12}$$

$$=P\left(\max_{0\le t\le 1}|\boldsymbol{B}(t)|\leqq(\lambda+K_1n^{-1/2})(1-2g(n))^{-1/2}\right)$$
$$=T((\lambda+K_1n^{-1/2})(1-2g(n))^{-1/2}) .$$

Thus we have from (2.9)-(2.12) that

$$F_n(\lambda)\leqq T((\lambda+K_1n^{-1/2})(1-2g(n))^{-1/2})+O(g(n)) . \tag{2.13}$$

We can also obtain, by a similar argument, that

$$\begin{aligned}F_n(\lambda)&\geqq P\left(\sup_{0\le t\le 1+2g(n)}|\boldsymbol{B}(t)|\leqq\lambda-K_1n^{-1/2}\right)-P\left(\left|\sum_{i=1}^n\tau_i-1\right|\geqq 2g(n)\right)\\&\geqq T((\lambda-K_1n^{-1/2})(1+2g(n))^{-1/2})-O(g(n)) .\end{aligned} \tag{2.14}$$

On the other hand, we have, from the inequality (3.1) in Sawyer [4], that

$$\begin{aligned}&\sup_{\lambda}|T((\lambda+K_1n^{-1/2})(1-2g(n))^{-1/2})-T((\lambda-K_1n^{-1/2})(1+2g(n))^{-1/2})|\\&\quad\leqq 8\{3/(\pi e)\}^{1/2}g(n)+8(1+g(n))K_1n^{-1/2}=O(g(n)) .\end{aligned} \tag{2.15}$$

The combination of this with (2.13) and (2.14) yields the theorem.

3. Example. With a slight modification of the Hall and Heyde example (Example 5 in [1], p. 82), we can give an example of a triangular array of independent random variables which satisfies all assumptions of Theorem 1 and the lower bound (1.11).

Example. *Let* $\{Z_i, i\geqq 1\}$ *be i.i.d. random variables with* $P(Z_1=1)=P(Z_1=-1)=1/2$. *Given a function* $g(\cdot)$ *satisfying* (1.7), *we set* $m_n=[n(1-g(n)/2)]$ *and*

$$X_{n,i}=\begin{cases}n^{-1/2}Z_i , & 1\leqq i\leqq m_n ,\\ 0 , & m_n<i\leqq n ,\end{cases}$$

here [**a**] *denotes the integral part of* **a**.

Then for each n, $\boldsymbol{X}_n=\{X_{n,i}, 1\leqq i\leqq n\}$ is a martingale difference satisfying (1.4). It also satisfies (1.8), because

$$V_n{}^2=\sum_{i=1}^n E(X_{n,i}^2|\mathscr{F}_{n,i-1})=m_n/n=1-g(n)/2+O(n^{-1}) \text{ a.s.}$$

We now prove that $\boldsymbol{X}_n$ satisfies (1.11). Since the derivative $T'(\lambda)$ is continuous and strictly positive for $\lambda>0$, we have $T(1+\varepsilon)-T(1)\sim K_{12}\varepsilon$, $\varepsilon\downarrow 0$, where $K_{12}=T'(1)>0$. By the scaling property of the Wiener process we have

$$\begin{aligned}\max_{0\le t\le m_n/n}|\boldsymbol{B}(t)|&\overset{\mathrm{d}}{=}(m_n/n)^{1/2}\max_{0\le t\le 1}|\boldsymbol{B}(t)|\\&=\{1-g(n)/2+O(n^{-1})\}^{1/2}\max_{0\le t\le 1}|\boldsymbol{B}(t)|\\&=\{1+g(n)/4+O(n^{-1})\}^{-1}\max_{0\le t\le 1}|\boldsymbol{B}(t)| ,\end{aligned}$$

and hence

$$\Delta_n{}^* \equiv \sup_{\lambda} \left| P\left(\max_{0 \leq t \leq m_n/n} |\boldsymbol{B}(t)| \leqq \lambda \right) - T(\lambda) \right| \tag{3.1}$$
$$\geqq \left| P\left(\max_{0 \leq t \leq m_n/n} |\boldsymbol{B}(t)| \leqq 1 \right) - T(1) \right|$$
$$= T(1+g(n)/4+O(n^{-1})) - T(1) \sim K_{12}g(n) .$$

But from Theorem 1 in Sawyer [3] we know that

$$\tilde{\Delta}_n \equiv \sup_{\lambda} \left| P\left(\max_{1 \leq k \leq m_n} |S_{n,k}| \leqq \lambda \right) - P\left(\max_{0 \leq t \leq m_n/n} |\boldsymbol{B}(t)| \leqq \lambda \right| \right) \tag{3.2}$$
$$= O(n^{-1/2}(\log n)^{1/2}) = o(g(n)) ,$$

and hence from (3.1) and (3.2) we have

$$\Delta_n \geqq \Delta_n{}^* - \tilde{\Delta}_n \geqq K_{13}g(n) - o(g(n)) ,$$

proving (1.11).

Acknowledgement

The author would like to express his thanks to Professor T. Kawata of Keio University for his attention to the work.

References

[1] P. Hall and C. C. Heyde, *Martingale Limit Theory and its Applications*, Academic, New York (1980).

[2] Y. Kato, *Convergence rates in cetral limit theorems for martingale differences*, Bulletin of Math. Statist., **18** (1979), 1-9.

[3] S. Sawyer, *A uniform rate of convergence for the maximum absolute value of partial sums in probability*, Comm. Pure and Appl. Math., **20** (1967), 647-659.

[4] A. V. Skorokhod, *Studies in the Theory of Random Processes*, Addison-Wesley, Massachusetts (1965).

[5] V. Strassen, *Almost sure behavior of sums of independent random variables and martingales*, Proc. of the 5th Berkeley Symp. of Math. Statist. and Probability, (1965), 315-343.

Department of Mathematics
Faculty of Science and Technology
Keio University
3-14-1, Hiyoshi, Kohoku-ku
Yokohama 223, Jadan