Z. Wahrscheinlichkeitstheorie verw. Gebiete
38, 211–216 (1977)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

A Note on Convergence to Mixtures of Normal Distributions*

Holger Rootzén
University of Lund and Lund Institute of Technology, Department of Mathematical Statistics, Box 725, S-22007 Lund, Sweden

Double arrays of random variables obtained by normalizing a sequence that is asymptotically close to a martingale difference sequence are considered, and conditions ensuring that the row sums converge in distribution to a mixture of normal distributions are found. The main condition is that the sums of squares in each row converge in probability to a random variable.

1. Introduction

Let $\{X_{n,i}: 1 \leqq i \leqq k_n, n \geqq 1\}$ be an array of random variables on a probability space $(\Omega, \mathscr{B}, P)$ and let $\mathscr{B}_{n,i}$ be the sub-sigmaalgebra that is generated by $X_{n,1}, \ldots, X_{n,i}$ ($\mathscr{B}_{n,0} = \{\phi, \Omega\}$). If $E(X_{n,i} \| \mathscr{B}_{n,i-1}) = 0$, $2 \leqq i \leqq k_n$, $n \geqq 1$ then $\{X_{n,i}\}$ is a martingale difference array (m.d.a.) and if furthermore it is obtained by normalizing a single sequence of martingale differences then

$$\mathscr{B}_{n,i} \subseteq \mathscr{B}_{n+1,i} \quad \text{for } n \geqq 1, \quad 1 \leqq i \leqq \min(k_n, k_{n+1}). \tag{1}$$

It is well known that if $\sum_{i=1}^{k_n} X_{n,i}^2$ converges in probability to a constant ($=\sigma$, say) and certain other conditions are satisfied, then the distribution of $\sum_{i=1}^{k_n} X_{n,i}$ converges to a normal distribution with mean zero and variance σ (see e.g. McLeish (1974)). In this note it is shown that if $\sum_{i=1}^{k_n} X_{n,i}^2$ converges in probability to some random variable ξ, then under similar conditions the distribution of $\sum_{i=1}^{k_n} X_{n,i}$ converges to a mixture of normal distributions. Eagleson (1975) observed that if ξ is measurable $\mathscr{B}_1 = \bigcap_{n=1}^{\infty} \mathscr{B}_{n,1}$ then $\{X_{n,i}\}$ is a martingale difference array (m.d.a.) also after conditioning on ξ and thus $\sum_{i=1}^{k_n} X_{n,i} \xrightarrow{d} \Phi(\cdot/\sqrt{\xi})$[1] under $P(\circ \| \xi)$ (at least

---

* Research sponsored in part by the Office of Naval Research, Contract # N00014-75-C-0809

[1] Φ is the standard normal distribution function

along subsequences satisfying $\sum_{i=1}^{k_n} X_{n,i}^2 \xrightarrow{\text{a.s.}} \xi$). By taking expectations it follows that $\sum_{i=1}^{k_n} X_{n,i} \xrightarrow{d} \int \Phi(\cdot/\sqrt{x})\, dF(x)$, where F is the distribution of ξ. Hence the main result of this note is that the restriction that ξ is measurable $\mathscr{B}_1$ is shown to be superfluous, but also our other conditions are somewhat weaker than those of Eagleson. (In fact Eagleson's results are phrased in terms of conditional variances instead of sums of squares, but the translation is straight-forward.) In a somewhat different context interesting results on the present problem have also been obtained by Chatterji (1974). Finally, it is perhaps worth mention that the condition that $\sum_{i=1}^{k_n} X_{n,i}^2$ converges in probability cannot be weakened very much: Dvoretsky (1972) has given an example which shows that it is not enough that $\sum_{i=1}^{k_n} X_{n,i}^2$ converges in distribution.

2. Convergence to Mixtures of Normal Distributions

We are mainly concerned with sums of martingale differences, but as is argued in Rootzén (1975) the interesting case is when the truncated variables asymptotically are martingale differences. Thus we initially do not assume that $\{X_{n,i}\}$ is a m.d.a., but only that e.g.

$$\sum_{i=1}^{k_n} E\{X_{n,i} I(|X_{n,i}| \leqq 1) \| \mathscr{B}_{n,i-1}\} \xrightarrow{p} 0 \quad \text{as} \quad n \to \infty. \tag{2}$$

Furthermore define

$$\xi_{n,i} = X_{n,i} I(|X_{n,i}| \leqq 1) - E(X_{n,i} I(|X_{n,i}| \leqq 1) \| \mathscr{B}_{n,i-1}\}.$$

Then $\{\xi_{n,i}\}$ is a m.d.a. and $|\xi_{n,i}| \leqq 2$ a.s.

Theorem 1. *Assume* (1) *and* (2) *are satisfied. If furthermore* $\max_{1 \leqq i \leqq k_n} |X_{n,i}| \xrightarrow{p} 0$ *and there is a random variable* ξ *with distribution* F *such that*

$$\sum_{i=1}^{k_n} \xi_{n,i}^2 \xrightarrow{p} \xi \quad \textit{as} \quad n \to \infty, \tag{3}$$

then $\sum_{i=1}^{k_n} X_{n,i} \xrightarrow{d} \int \Phi(\circ/\sqrt{x})\, dF(x)$ *as* $n \to \infty$.

Proof. By definition

$$\sum_{i=1}^{k_n} X_{n,i} - \sum_{i=1}^{k_n} \xi_{n,i} = \sum_{i=1}^{k_n} E\{X_{n,i} I(|X_{n,i}| \leqq 1) \| \mathscr{B}_{n,i-1}\} + \sum_{i=1}^{k_n} X_{n,i} I(|X_{n,i}| > 1)$$

and thus from (2) and $\max_{1\leq i\leq k_n}|X_{n,i}|\xrightarrow{p}0$ we obtain $\sum_{i=1}^{k_n}X_{n,i}-\sum_{i=1}^{k_n}\xi_{n,i}\xrightarrow{p}0$ as $n\to\infty$. Hence it is enough to prove

$$\sum_{i=1}^{k_n}\xi_{n,i}\xrightarrow{d}\int\Phi(\cdot/\sqrt{x})\,dF(x)\quad\text{as}\quad n\to\infty. \tag{4}$$

Write $X'_{n,i}=X_{n,i}I(|X_{n,i}|\leq 1)$. Then $1\geq\max_{1\leq i\leq k_n}|X'_{n,i}|\xrightarrow{p}0$ and thus also

$$E\{(\max_{1\leq i\leq k_n}|X'_{n,i}|)\}\to 0\quad\text{as}\quad n\to\infty.$$

Since

$$\max_{1\leq i\leq k_n}|E(X'_{n,i}\|\mathscr{B}_{n,i-1})|\leq\max_{1\leq i\leq k_n}E(\max_{1\leq j\leq k_n}|X'_{n,j}|\,\|\mathscr{B}_{n,i-1})$$

and since $E(\max_{1\leq j\leq k_n}|X'_{n,j}|\,\|\mathscr{B}_{n,i-1})$, $i=1,\dots,k_n$ is a martingale it follows from Kolmogorov's inequality for martingales that $\max_{1\leq i\leq k_n}|E(X'_{n,i}\|\mathscr{B}_{n,i-1})|\xrightarrow{p}0$ and thus also $\max_{1\leq i\leq k_n}\xi_{n,i}\xrightarrow{p}0$ as $n\to\infty$.

To facilitate the remainder of the proof we will without loss of generality assume that each row in the m.d.a. $\{\xi_{n,i}\}$ is infinite and that for $i>k_n$, $\xi_{n,i}=\pm\frac{1}{n}$ with probability $\frac{1}{2}$ each and independently of $\mathscr{B}_{n,i-1}$. Hence we have $\max_{1\leq i}|\xi_{n,i}|\xrightarrow{p}0$ as $n\to\infty$ and $\sum_{i=1}^{k}\xi_{n,i}^2\xrightarrow{\text{a.s.}}\infty$ as $k\to\infty$, for each n. Let $S_n(k)=\sum_{i=1}^{k}\xi_{n,i}$ and introduce "the natural time-scale" $\tau_n(t)=\inf\left\{k:\sum_{i=1}^{k}\xi_{n,i}^2>t\right\}$ of the summation process based on $\{\xi_{n,i}\}$. Then

$$\sum_{i=1}^{k_n}\xi_{n,i}=S_n\circ\tau_n\left(\sum_{i=1}^{k_n}\xi_{n,i}^2\right)-\xi_{n,k_n+1}, \tag{5}$$

and by construction $|\xi_{n,k_n+1}|=\frac{1}{n}\to 0$ as $n\to\infty$.

The next step is to prove

$$d_n=|S_n\circ\tau_n\left(\sum_{i=1}^{k_n}\xi_{n,i}^2\right)-S_n\circ\tau_n(\xi)|\xrightarrow{p}0\quad\text{as}\quad n\to\infty. \tag{6}$$

For $\varepsilon>0$ choose K such that $P(\{\xi>K\})<\varepsilon$ and observe that for any $\delta>0$

$$P(\{d_n>\epsilon\})\leq P\left(\{d_n>\varepsilon\}\cap\left\{\left|\sum_{i=1}^{k_n}\xi_{n,i}^2-\xi\right|\leq\delta\right\}\cap\{\xi\leq K\}\right)$$
$$+P\left(\left\{\left|\sum_{i=1}^{k_n}\xi_{n,i}^2-\xi\right|>\delta\right\}\right)+\varepsilon.$$

Here

$$P\left(\{d_n>\varepsilon\}\cap\left\{\left|\sum_{i=1}^{k_n}\xi_{n,i}^2-\xi\right|\leqq\delta\right\}\cap\{\xi\leqq K\}\right)\leqq P(\{\sup_{\substack{|s-t|\leqq\delta\\0\leqq s,t\leqq K+\delta}}|S_n\circ\tau_n(t)-S_n\circ\tau_n(s)|>\varepsilon\})=P_n,\quad\text{say.}$$

Since $\max_{1\leqq i\leqq\tau_n(K+1)}|\xi_{n,i}|\leqq\max_{1\leqq i}|\xi_{n,i}|\xrightarrow{p}0$ as $n\to\infty$ it follows from Theorem 1 of Rootzén (1975) that $\{S_n\circ\tau_n(t);\ t\in[0,K+1]\}_{n=1}^{\infty}$, considered as random variables in $D(0,K+1)$ endowed with the Skorokhod topology, converge in distribution to a Brownian motion on $[0,K+1]$. Hence, by Theorem 15.2 of Billingsley (1968) we can make $\limsup_{n\to\infty}P_n<\varepsilon$ by choosing δ small enough. Furthermore, by (3), $P\left(\left\{\left|\sum_{i=1}^{k_n}\xi_{n,i}^2-\xi\right|>\delta\right\}\right)\to0$ as $n\to\infty$, so

$$\limsup_{n\to\infty}P(\{d_n>\varepsilon\})\leqq2\varepsilon,$$

which since ε is arbitrary proves (6).

We now need the following lemma, the proof of which is given on p. 215.

Lemma 2. *If for some integer k_ε the random variable $\xi_\varepsilon\geqq0$ is measurable $\mathscr{B}_{n,k_\varepsilon}$ for all large n then*

$$S_n\circ\tau_n(\xi_\varepsilon)\xrightarrow{d}\int\Phi(\cdot/\sqrt{x})\,dF_\varepsilon(x)\quad\text{as}\quad n\to\infty,\tag{7}$$

where F_ε is the distribution of ξ_ε.

Since we have assumed (1) it is for $\varepsilon>0$ possible to find an integer k_ε and a random variable $\xi_\varepsilon\geqq0$, which is measurable with respect to $\mathscr{B}_{n,k_\varepsilon}$ for all large n, such that $P(|\xi-\xi_\varepsilon|>\varepsilon)\leqq\varepsilon$ (see e.g. Doob (1953), p. 607). From Lemma 2 we obtain that (7) holds and since $F_\varepsilon\xrightarrow{d}F$ as $\varepsilon\to0$ also

$$\int\Phi(\cdot/\sqrt{x})\,dF_\varepsilon(x)\xrightarrow{d}\int\Phi(\cdot/\sqrt{x})\,dF(x)\quad\text{as}\quad\varepsilon\to0.\tag{8}$$

Moreover, by arguments precisely analogous to the proof of (6) above, we have for $\delta>0$ that

$$\lim_{\varepsilon\to0}\limsup_{n\to\infty}P(\{|S_n\circ\tau_n(\xi_\varepsilon)-S_n\circ\tau_n(\xi)|>\delta\})=0.\tag{9}$$

By Theorem 4.2 of Billingsley (1968) it follows from (7), (8), and (9) that

$$S_n\circ\tau_n(\xi)\xrightarrow{d}\int\Phi(\cdot/\sqrt{x})\,dF(x)\quad\text{as}\quad n\to\infty,$$

which by (5) and (6) proves (4) and thus the theorem. □

Proof of Lemma 2. The lemma can be reduced to the corollary on p. 561 of Eagleson (1975), but as it does not involve more work to give a direct proof we will do that instead. Put $\xi'_{n,i}=\xi_{n,i}$ if $i>k_\varepsilon$ and $\xi'_{n,i}=0$ otherwise, let $S'_n(k)=\sum_{i=1}^{k}\xi'_{n,i}$ and let $P_\varepsilon(\cdot)=P(\cdot\,\|\,\xi_\varepsilon)$ be a regular conditional probability given ξ_ε. Since $\xi_\varepsilon\in\mathscr{B}_{n,k_\varepsilon}$ for all large n, the rows of $\{\xi'_{n,i}\}$ are martingale differences for such n

also under P_ε (a.s.), cf. [4]. Moreover, for $t \geqq 0$

$$\left|\sum_{i=1}^{\tau_n(t)} \xi'^2_{n,i} - t\right| \leqq \sum_{i=1}^{k_\varepsilon} \xi^2_{n,i} + \xi^2_{n,\tau_n(t)} \leqq (k_\varepsilon + 1)(\max_{1 \leqq i} |\xi_{n,i}|)^2. \tag{10}$$

As $\max_{1 \leqq i} |\xi_{n,i}| \xrightarrow{p} 0$ we can for every subsequence of $\{1, 2, \ldots\}$ find a further (infinite) subsequence $\{n'\}$ such that $\max_{1 \leqq i} |\xi_{n',i}| \xrightarrow{\text{a.s.}} 0$. Thus also under P_ε we have a.s. that $\max_{1 \leqq i} |\xi'_{n',i}| \xrightarrow{\text{a.s.}} 0$ and $\sum_{i=1}^{\tau_n(t)} \xi'_{n',i} \xrightarrow{\text{a.s.}} t$. By Lemma 3 of [5] this proves

$$S'_{n'} \circ \tau_{n'}(t) \xrightarrow{d} \Phi(\cdot/\sqrt{t})$$

under P_ε (a.s.). Since ξ_ε is a constant under P_ε we have in particular that

$$S'_{n'} \circ \tau_{n'}(\xi_\varepsilon) \xrightarrow{d} \Phi(\cdot/\sqrt{\xi_\varepsilon})$$

(P_ε), a.s., and thus by taking expectations that

$$S'_{n'} \circ \tau_{n'}(\xi_\varepsilon) \xrightarrow{d} \int \Phi(\cdot/\sqrt{x})\, dF_\varepsilon(x) \quad \text{as} \quad n' \to \infty. \tag{11}$$

Since a subsequence $\{n'\}$ satisfying (11) can be extracted from any infinite sequence of integers it follows that $S'_n \circ \tau_n(\xi_\varepsilon) \xrightarrow{d} \int \Phi(\cdot/\sqrt{x})\, dF_\varepsilon(x)$ as $n \to \infty$, which since

$$|S_n \circ \tau_n(\xi_\varepsilon) - S'_n \circ \tau_n(\xi_\varepsilon)| \leqq k_\varepsilon \max_{1 \leqq i} |\xi_{n,i}| \xrightarrow{p} 0$$

proves (7). □

Corollary 3. *Assume $\{X_{n,i}\}$ is a m.d.a. and that it satisfies* (1). *If* $\max_{1 \leqq i \leqq k_n} |X_{n,i}| \xrightarrow{p} 0$, *if* $\sum_{i=1}^{k_n} E(X^2_{n,i} I(|X_{n,i}| > 1) \| \mathscr{B}_{n,i-1}) \xrightarrow{p} 0$, *and if furthermore* $\sum_{i=1}^{k_n} X^2_{n,i} \xrightarrow{p} \xi$ *for some random variable ξ with distribution F, then* $\sum_{i=1}^{k_n} X_{n,i} \xrightarrow{d} \int \Phi(\cdot/\sqrt{x})\, dF(x)$ *as* $n \to \infty$.

Proof. We only have to show that (2) and (3) hold. Recall the notation

$$X'_{n,i} = X_{n,i} I(|X_{n,i}| \leqq 1)$$

and put $X''_{n,i} = X_{n,i} - X'_{n,i}$. Since

$$E(X'_{n,i} \| \mathscr{B}_{n,i-1})^2 \leqq |E(X'_{n,i} \| \mathscr{B}_{n,i-1})| = |E(X''_{n,i} \| \mathscr{B}_{n,i-1})| \leqq E(X''^2_{n,i} \| \mathscr{B}_{n,i-1})$$

it follows immediately that (2) holds. Moreover, by definition

$$\begin{aligned}\sum_{i=1}^{k_n} \xi^2_{n,i} &= \sum_{i=1}^{k_n} X^2_{n,i} + \sum_{i=1}^{k_n} \{X''^2_{n,i} + E(X'_{n,i} \| \mathscr{B}_{n,i-1})^2\} \\ &\quad - 2\sum_{i=1}^{k_n} X_{n,i}\{X''_{n,i} + E(X'_{n,i} \| \mathscr{B}_{n,i-1})\} \\ &= \sum_{i=1}^{k_n} X^2_{n,i} + r_n - 2r'_n, \quad \text{say.}\end{aligned}$$

Here $r_n \xrightarrow{p} 0$ as $n \to \infty$ and using Cauchy's inequality we obtain

$$|r_n'| \leqq \left\{ 2 r_n \sum_{i=1}^{k_n} X_{n,i}^2 \right\}^{\frac{1}{2}} \xrightarrow{p} 0 \quad \text{as} \quad n \to \infty,$$

so (3) follows. □

Finally, it should perhaps be mentioned that the results of this note hold also when $\{k_n\}$ are stopping times.

References

1. Chatterji, S.D.: A principle of subsequences in probability theory: The central limit theorem. Advances in Math. **13**, 31–54 (1974)
2. Billingsley, P.: Convergence of Probability Measures. New York: Wiley 1968
3. Doob, J.L.: Stochastic Processes. New York: Wiley 1953
4. Dvoretsky, A.: Central limit theorems for dependent random variables. Proc. Sixth Berkeley Sympos. Math. Statist. Probab. Berkeley: Univ. of Calif. Press 1972.
5. Eagleson, G.K.: Martingale convergence to mixtures of infinitely divisible laws. Ann. Probability **3**, 557–562 (1975)
6. McLeish, D.L.: Dependent central limit theorems and invariance principles. Ann. Probability **2**, 620–628 (1974)
7. Rootzén, H.: On the functional central limit theorem for martingales. Z. Wahrscheinlichkeitstheorie verw. Gebiete **38**, 199–210 (1977)

Received December 23, 1975; in revised form October 25, 1976