

A NOTE ON CONVERGENCE TO MIXTURES
OF NORMAL DISTRIBUTIONS*

BIOMATHEMATICS TRAINING PROGRAM

by

Holger Rootzén

Department of Statistics

University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series #1045

December, 1975

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

A NOTE ON CONVERGENCE TO MIXTURES
OF NORMAL DISTRIBUTIONS*

by

Holger Rootzén

ABSTRACT: 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.

AMS 1970 Subject classifications: Primary 60F05; Secondary 60G45.

Key words and phrases: Central limit theorem, martingales, mixtures of normal distributions.

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

A NOTE ON CONVERGENCE TO MIXTURES
OF NORMAL DISTRIBUTIONS

1. INTRODUCTION.

Let $\{X_{n,i} : 1 \leq i \leq k_n, n \geq 1\}$ be an array of random variables on a probability space (Ω, \mathcal{B}, P) and let $\mathcal{B}_{n,i}$ be the sub-sigmaalgebra that is generated by $X_{n,1}, \dots, X_{n,i}$ ($\mathcal{B}_{n,0} = \{\phi, \Omega\}$). If $E(X_{n,i} | \mathcal{B}_{n,i-1}) = 0$, $2 \leq i \leq k_n, n \geq 1$ 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

$$(1) \quad \mathcal{B}_{n,i} \subseteq \mathcal{B}_{n+1,i} \quad \text{for } n \geq 1, \quad 1 \leq i \leq \min(k_n, k_{n+1}).$$

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 $\mathcal{B}_1 = \bigcap_{n=1}^{\infty} \mathcal{B}_{n,1}$ then $\{X_{n,i}\}$ is a martingale difference array (m.d.a.) also after conditioning on ξ and thus $\sum_{i=1}^{k_n} X_{n,i} \xrightarrow{d} \phi(\cdot/\sqrt{\xi})^*$ under $P(\cdot | \xi)$ (at least 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 \mathcal{B}_1 is shown to be superfluous, but also our other conditions are somewhat weaker

* ϕ is the standard normal distribution function.

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.) 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: Dvoretzky (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 don't assume that $\{X_{n,i}\}$ is a m.d.a., but only that e.g.

$$(2) \quad \sum_{i=1}^{k_n} E\{X_{n,i} I(|X_{n,i}| \leq 1) \mid \mathcal{B}_{n,i-1}\} \xrightarrow{P} 0 \text{ as } n \rightarrow \infty.$$

Furthermore define

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

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

THEOREM 1. Assume (1) and (2) are satisfied. If furthermore

$\max_{1 \leq i \leq k_n} |X_{n,i}| \xrightarrow{P} 0$ and there is a random variable ξ with distribution F such that

$$(3) \quad \sum_{i=1}^{k_n} \xi_{n,i}^2 \xrightarrow{P} \xi \text{ as } n \rightarrow \infty,$$

then $\sum_{i=1}^{k_n} X_{n,i} \xrightarrow{d} \int \phi(\cdot/\sqrt{x}) dF(x)$ as $n \rightarrow \infty$.

PROOF. By definition $\sum_{i=1}^{k_n} X_{n,i} - \sum_{i=1}^{k_n} \xi_{n,i} = \sum_{i=1}^{k_n} E\{X_{n,i} I(|X_{n,i}| \leq 1) \mid \mathcal{B}_{n,i-1}\} + \sum_{i=1}^{k_n} X_{n,i} I(|X_{n,i}| > 1)$ 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 $n \rightarrow \infty$. Hence it is enough to prove

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

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

$E\{(\max_{1 \leq i \leq k_n} |X'_{n,i}|)^2\} \rightarrow 0$ as $n \rightarrow \infty$. Since $\max_{1 \leq i \leq k_n} |E(X'_{n,i} \mid \mathcal{B}_{n,i-1})| \leq$

$\leq \max_{1 \leq i \leq k_n} E(\max_{1 \leq i \leq k_n} |X'_{n,i}| \mid \mathcal{B}_{n,i-1})$ and since $E(\max_{1 \leq i \leq k_n} |X'_{n,i}| \mid \mathcal{B}_{n,i-1})$,

$i = 1, \dots, k_n$ is a martingale it follows from Kolmogorov's inequality for

martingales that $\max_{1 \leq i \leq k_n} |E(X'_{n,i} \mid \mathcal{B}_{n,i-1})| \xrightarrow{P} 0$ and thus also $\max_{1 \leq i \leq k_n} \xi_{n,i} \xrightarrow{P} 0$ as $n \rightarrow \infty$.

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 $\mathcal{B}_{n,i-1}$. Hence we

have $\max_{1 \leq i} |\xi_{n,i}| \xrightarrow{P} 0$ as $n \rightarrow \infty$ and $\sum_{i=1}^{k_n} \xi_{n,i}^2 \xrightarrow{a.s.} \infty$ as $k_n \rightarrow \infty$, for each n .

Let $S_n(k) = \sum_{i=1}^k \xi_{n,i}$ and introduce "the natural time-scale"

$\tau_n(t) = \inf\{k: \sum_{i=1}^k \xi_{n,i}^2 > t\}$ of the summation process based on $\{\xi_{n,i}\}$. Then

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

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

The next step is to prove

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

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

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

Here

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

Since $\max_{1 \leq i \leq \tau_n(K+1)} |\xi_{n,i}| \leq \max_{1 \leq i} |\xi_{n,i}| \xrightarrow{P} 0$ as $n \rightarrow \infty$ 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 $\limsup P_n < \epsilon$ by choosing δ small enough.

Furthermore, by (3), $P(\{|\sum_{i=1}^{k_n} \xi_{n,i}^2 - \xi| > \delta\}) \rightarrow 0$ as $n \rightarrow \infty$, so

$$\limsup_{n \rightarrow \infty} P(\{d_n > \epsilon\}) \leq 2\epsilon,$$

which since ϵ is arbitrary proves (6).

We now need the following lemma, the proof of which is given on p. 5.

LEMMA 2. If for some integer k_ϵ the random variable $\xi_\epsilon \geq 0$ is measurable

$\mathcal{B}_{n, k_\epsilon}$ for all large n then

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

where F_ϵ is the distribution of ξ_ϵ .

Since we have assumed (1) it is for $\epsilon > 0$ possible to find an integer k_ϵ and a random variable $\xi_\epsilon \geq 0$, which is measurable with respect to $\mathcal{B}_{n, k_\epsilon}$

for all large n , such that $P(|\xi - \xi_\epsilon| > \epsilon) \leq \epsilon$ (see e.g. Doob (1953), p. 607). From Lemma 2 we obtain that (7) holds and since $F_\epsilon \xrightarrow{d} F$ as $\epsilon \rightarrow 0$ also

$$(8) \quad \int \phi(\cdot/\sqrt{x}) dF_\epsilon(x) \xrightarrow{d} \int \phi(\cdot/\sqrt{x}) dF(x) \quad \text{as } \epsilon \rightarrow 0.$$

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

$$(9) \quad \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} P(\{|S_{n \circ \tau_n}(\xi_\epsilon) - S_{n \circ \tau_n}(\xi)| > \delta\}) = 0.$$

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

$$S_{n \circ \tau_n}(\xi) \xrightarrow{d} \int \phi(\cdot/\sqrt{x}) dF(x) \quad \text{as } n \rightarrow \infty,$$

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

PROOF OF LEMMA 2. The lemma can be reduced to the corollary on p. 561 of Eagleson (1975), but as it doesn't involve more work to give a direct proof we will do that instead. Put $\xi'_{n,i} = \xi_{n,i}$ if $i > k_\epsilon$ and $\xi'_{n,i} = 0$ otherwise, let $S'_n(k) = \sum_{i=1}^k \xi'_{n,i}$ and let $P_\epsilon(\cdot) = P(\cdot | \xi_\epsilon)$ be a regular conditional probability given ξ_ϵ . Since $\xi_\epsilon \in B_{n, k_\epsilon}$ for all large n , the rows of $\{\xi'_{n,i}\}$ are martingale differences for such n also under P_ϵ (a.s.), cf. ref [4]. Moreover, for $t \geq 0$

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

As $\max_{1 \leq i} |\xi_{n,i}| \xrightarrow{P} 0$ we can for every subsequence of $\{1, 2, \dots\}$ find a further (infinite) subsequence $\{n'\}$ such that $\max_{1 \leq i} |\xi_{n',i}| \xrightarrow{\text{a.s.}} 0$. Thus also under P_ϵ we have a.s. that $\max_{1 \leq i} |\xi'_{n',i}| \xrightarrow{\text{a.s.}} 0$ and $\sum_{i=1}^{\tau_n(t)} \xi'^2_{n',i} \xrightarrow{\text{a.s.}} t$. By

Lemma 3 of [5] this proves $S'_{n', \tau_{n'}}(t) \xrightarrow{d} \phi(\cdot/\sqrt{t})$ under P_ϵ (a.s.). Since ξ_ϵ is a constant under P_ϵ we have in particular that $S'_{n', \tau_{n'}}(\xi_\epsilon) \xrightarrow{d} \phi(\cdot/\sqrt{\xi_\epsilon})$ (P_ϵ), a.s., and thus by taking expectations that

$$(11) \quad S'_{n', \tau_{n'}}(\xi_\epsilon) \xrightarrow{d} \int \phi(\cdot/\sqrt{x}) dF_\epsilon(x) \quad \text{as } n' \rightarrow \infty.$$

Since a subsequence $\{n'\}$ satisfying (11) can be extracted from any infinite sequence of integers it follows that $S'_{n, \tau_n}(\xi_\epsilon) \xrightarrow{d} \int \phi(\cdot/\sqrt{x}) dF(x)$ as $n \rightarrow \infty$, which since $|S_{n, \tau_n}(\xi_\epsilon) - S'_{n, \tau_n}(\xi_\epsilon)| \leq k_\epsilon \max_{1 \leq i} |\xi_{n,i}| \xrightarrow{P} 0$ proves (7). \square

COROLLARY 3. Assume $\{X_{n,i}\}$ is a m.d.a. and that it satisfies (1). If $\max_{1 \leq i \leq k_n} |X_{n,i}| \xrightarrow{P} 0$, if $\sum_{i=1}^{k_n} E(X_{n,i}^2 I(|X_{n,i}| > 1) | \mathcal{B}_{n,i-1}) \xrightarrow{P} 0$, and if furthermore $\sum_{i=1}^{k_n} X_{n,i}^2 \xrightarrow{P} \xi$ 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 $n \rightarrow \infty$.

PROOF. We only have to show that (2) and (3) hold. Recall $X'_{n,i} = X_{n,i} I(|X_{n,i}| \leq 1)$ and put $X''_{n,i} = X_{n,i} - X'_{n,i}$. Since $E(X'_{n,i} | \mathcal{B}_{n,i-1})^2 \leq |E(X'_{n,i} | \mathcal{B}_{n,i-1})| = |E(X''_{n,i} | \mathcal{B}_{n,i-1})| \leq E(X''_{n,i}{}^2 | \mathcal{B}_{n,i-1})$ it follows immediately that (2) holds. Moreover, by definition

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

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

$$|r'_n| \leq \{2r_n \sum_{i=1}^{k_n} X_{n,i}^2\}^{1/2} \xrightarrow{P} 0 \text{ as } n \rightarrow \infty, \text{ so (3) follows} \quad \square$$

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

REFERENCES

- [1] Billingsley, P.: *Convergence of Probability Measures*. New York: Wiley 1968.
- [2] Doob, J.L.: *Stochastic Processes*. New York: Wiley 1953.
- [3] Dvoretzky, A.: Central limit theorems for dependent random variables. Proc. ~~Sixth~~ Berkeley Symp. Math. Statist. Probability. Berkeley: Univ. of Calif. Press 1972.
- [4] Eagleson, G.K.: Martingale convergence to mixtures of infinitely divisible laws. *Ann. Probability* 3 557-562 (1975).
- [5] Rootzén, H.: On the functional central limit theorem for martingales. Institute of Statistics Mimeo Series #1043, University of North Carolina at Chapel Hill, 1975.