The Annals of Probability
1981, Vol. 9, No. 3, 395–404

RATES OF CONVERGENCE IN THE MARTINGALE CENTRAL LIMIT THEOREM

BY PETER HALL AND C. C. HEYDE

Australian National University and CSIRO Division of Mathematics and Statistics, Canberra

We obtain a nonuniform estimate of the rate of convergence in the martingale central limit theorem for convergence to mixtures of normal distributions. The uniform rates of convergence obtained by several other authors are special cases of our nonuniform estimate. We also obtain a rate of convergence in B. M. Brown's central limit theorem, assuming only Brown's elementary conditions. This result is a martingale analogue of Feller's generalization of the Berry-Esseen theorem.

1. Introduction and summary. Martingale limit theory has grown steadily in importance since it was first brought to prominence by the work of Doob. It has merit in its own right as an abstract mathematical theory, but its principal relevance to probability and statistics is its ease of application to a diverse range of problems. From the point of view of applications a martingale limit theorem is of little use without some knowledge of the rate of convergence. In this paper we obtain improved rates of convergence in the martingale central limit theorem.

Heyde and Brown (1970) provided a uniform bound on the rate of convergence in the martingale central limit theorem. They used a martingale version of the Skorokhod representation (see, for example, Strassen (1967, Theorem 4.3) and Dubins (1968)). More recently Erickson, Quine and Weber (1978) obtained bounds of the same order of magnitude using charcteristic function techniques. The virtue of results of this type is that they provide rates of convergence under basic conditions of the martingale central limit theorem, asking only that these conditions hold in an L^p space for some $p > 2$. Faster rates of convergence may be obtained under more stringent conditions; see the work of Ibragimov (1963), Grams (1972), Nakata (1976) and Kato (1979). However, in the field of applications it is often difficult to check even the basic sufficient conditions of the central limit theorem. For this reason we shall confine ourselves to rates under these conditions alone; faster rates under less general conditions are presented in Hall and Heyde (1980, Section 3.6).

In Section 2 we present nonuniform rates of convergence in the central limit theorem for convergence to mixtures of normal laws. (Sufficient conditions for the limit theorem were obtained by Chatterji (1974), Eagleson (1975), Hall (1977, 1979), Rootzén (1977) and Aldous and Eagleson (1978).) We believe that ours are the only nonuniform rates available for the martingale central limit theorem, and they contain the uniform rates established by Heyde and Brown (1970) and Erickson, Quine and Weber (1978).

Feller (1968) generalized the Berry-Esseen theorem by obtaining a uniform rate of convergence in the central limit theorem for sums of independent variables with only second moments assumed finite. His result contains the principal rates of convergence in the central limit theorem for sums of independent variables, and in Section 2 we present an analogue of it for martingales which contains the principal rates in the martingale central limit theorem. Our result provides rates under the very basic conditions assumed by Brown (1971).

Received November 16, 1979.

AMS 1970 *subject classification.* Primary 60G42, G0F05.

Key words and phrases. Martingale central limit theorem, mixtures of normal distributions, nonuniform bound, rate of convergence, uniform bound.

A version of Theorem 1, with $\eta^2 = 1$ and $m = 0$, is given in Hall and Heyde (1980, Theorem 3.9). Rates of convergence in the martingale invariance principle are presented in Section 4.3 of the same monograph, and it should be noted that these can be generalized using the techniques of proof of Theorem 1 in this paper.

2. The results. Let $\{S_{ni} = \sum_{j=1}^{i} X_{nj},\ \mathscr{F}_{ni},\ 1 \le i \le k_n\}$ be a zero mean martingale for each $n \ge 1$. Suppose that the martingale differences satisfy the conditions

$$\text{for all} \quad \varepsilon > 0, \qquad \textstyle\sum_{i=1}^{k_n} E[X_{ni}^2 I(|X_{ni}| > \varepsilon)] \to 0$$

and

$$U_n^2 = \textstyle\sum_{i=1}^{k_n} X_{ni}^2 \to_p \eta^2,$$

where the random variable $\eta > 0$ a.s. If the σ-fields are nested—that is, if

$$\mathscr{F}_{ni} \subseteq \mathscr{F}_{n+1,i} \quad \text{for all} \quad i \le k_n, \tag{1}$$

then $P(S_{nk_n} \le x) \to E[\Phi(\eta^{-1}x)]$ for all real x, where Φ is the standard normal distribution function (Hall (1977), Rootzén (1977)). The result is also true if instead of (1), η^2 is measurable in $\bigcap_1^\infty \mathscr{F}_{n1}$ (Eagleson (1975)). Our first result provides a rate of convergence in this type of limit theorem. (In this and some later work we suppress the dependence of the martingale on n.)

THEOREM 1. *Let $\{S_i = \sum_{j=1}^{i} X_j,\ \mathscr{F}_i,\ 1 \le i \le n\}$ be a zero mean, square integrable martingale. Set*

$$U_i^2 = \textstyle\sum_{j=1}^{i} X_j^2 \quad \text{and} \quad V_i^2 = \sum_{j=1}^{i} E(X_j^2 \mid \mathscr{F}_{j-1}), \qquad 1 \le i \le n; U_0 = 0;$$

and suppose that $0 < \delta \le 1$. For $0 \le m \le n$ define

$$L_{nm} = E(U_m^{2+2\delta}) + \textstyle\sum_{m+1}^{n} E \mid X_i \mid^{2+2\delta} + E \mid U_n^2 - \eta^2 \mid^{1+\delta} + E \mid \eta^2 - E(\eta^2 \mid \mathscr{F}_m) \mid^{1+\delta},$$

where η is a random variable and $\mathscr{F}_0$ is taken to be the trivial σ-field. Assume that for positive constants c and c', $\eta \ge$ a.s. and $E(\eta^{2+2\delta}) < c'$. There exists a constant A depending only on c, c' and δ such that for all x and m, and whenever $L_{nm} \le 1$,

$$\mid P(S_n \le x) - E[\Phi(\eta^{-1})] \mid \le A L_{nm}^{1/(3+2\delta)} [1 + \mid x \mid^{4(1+\delta)^2/(3+2\delta)}]^{-1}. \tag{2}$$

The result is also true if the term $E \mid U_n^2 - \eta^2 \mid^{1+\delta}$ in the definition of L_{nm} is replaced by $E \mid V_n^2 - \eta^2 \mid^{1+\delta}$.

The next result is immediate.

COROLLARY. *In the notation of Theorem 1, assume $\eta \ge c$ a.s., $E(\eta^{2+2\delta}) < c'$ and $0 < \delta \le 1$. There exists a constant A depending only on c, c' and δ such that for all $0 \le m \le n$,*

$$\sup_x \mid P(S_n \le x) - E[\Phi(\eta^{-1}x)] \mid \le A L_{nm}^{1/(3+2\delta)}.$$

The result continues to hold if $E \mid U_n^2 - \eta^2 \mid^{1+\delta}$ is replaced by $E \mid V_n^2 - 1 \mid^{1+\delta}$.

If we set $m = 0$ and $\eta = 1$, we obtain the theorem of Heyde and Brown (1970). In the case of a general triangluar array with

$$\textstyle\sum_1^{k_n} E \mid X_{ni} \mid^{2+2\delta} \to 0 \qquad \text{and} \qquad E \mid U_n^2 - \eta^2 \mid^{1+\delta} \to 0,$$

and for which condition (1) holds, a judicious choice of $m = m(n) \to \infty$ permits both

$$E \mid \textstyle\sum_1^m X_{ni}^2 \mid^{1+\delta} \to 0 \qquad \text{and } E \mid \eta^2 - E(\eta^2 \mid \mathscr{F}_{nm}) \mid^{1+\delta} \to 0,$$

and then (2) provides a proper rate of convergence.

The corollary may also be deduced from a rate of convergence in the central limit theorem with only second order moments assumed finite. The form of the error bound for this result is quite complicated in the general setting of Theorem 1, and so we shall assume that $\eta = 1$ and obtain only a uniform rate of convergence.

THEOREM 2. *In the notation of Theorem* 1, *there exists an absolute constant* $A > 0$ *such that whenever* $0 < \delta \le 1$ *and* $\varepsilon > 0$,

$$\sup_x | P(S_n \le x) - \Phi(x) | \le A\{[\textstyle\sum_1^n E(X_i^2 I(|X_i| > 1))]^{1/3} + \delta^{-(1/2)(1+\delta)}[\sum_1^n E(|x_i|^{2+2\delta} I(|X_i| \le 1))]^{1/(3+2\delta)} + \varepsilon^{1/2} + P(|U_n^2 - 1| > \varepsilon)\}.$$

The result is also true if $P(|U_n^2 - 1| > \varepsilon)$ *is replaced by* $P(|V_n^2 - 1| > \varepsilon)$.

To obtain the corollary from Theorem 2 in the case $\eta = 1$ and $m = 0$, note that

$$[\textstyle\sum_1^n E(X_i^2 I(|X_i| > 1))]^{1/3} \le (\sum_1^n E |X_i|^{2+2\delta})^{1/3} \le (\sum_1^n E |X_i|^{2+2\delta})^{1/(3+2\delta)}$$

if $\sum_1^n E | X_i|^{2+2\delta} \le 1$, and

$$\varepsilon^{1/2} + P(|U_n^2 - 1| > \varepsilon) \le \varepsilon^{1/2} + \varepsilon^{-1-\delta} E | U_n^2 - 1|^{1+\delta} = 2(E | U_n^2 - 1|^{1+\delta})^{1/(3+2\delta)}$$

if $\varepsilon = (E | U_n^2 - 1|)^{2/(3+2\delta)})$.

Let $\{S_n = \sum_1^n X_i, \mathscr{F}_n, n \ge 1\}$ be a zero mean, square integrable martingale and set $s_n^2 = E(S_n^2)$ and $V_n^2 = \sum_1^n E(X_i^2 | \mathscr{F}_{i-1})$. The conditions

$$\text{(3)} \qquad \text{for all} \quad \varepsilon > 0, \qquad s_n^{-2} \textstyle\sum_1^n E[X_i^2 I(|X_i| > \varepsilon s_n) | \mathscr{F}_{i-1}] \to_p 0$$

and

$$\text{(4)} \qquad s_n^{-2} V_n^2 \to_p 1$$

are sufficient to imply the asymptotic normality of $s_n^{-1} S_n$ (Brown (1971)), and they are necessary if the differences $X_1, X_2, \cdots$ are independent and asymptotically negligible (Feller (1971), page 520). Theorem 2 provides a rate of convergence under only these basic conditions. To see this, let $\delta = 1$ and note that under (3) and (4),

$$s_n^{-2} \textstyle\sum_1^n E[X_i^2 I(|X_i| > 1)] \to 0, \quad s_n^{-4} \sum_1^n E[X_i^4 I(|X_i| \le s_n)] \to 0 \quad \text{and} \quad E | s_n^{-2} V_n^2 - 1| \to 0$$

(see Brown (1971) and Scott (1973)). Set $\delta = 1$ and $\varepsilon = (E | s_n^{-2} V_n^2 - 1|)^{2/3}$, and observe that

$$\varepsilon^{1/2} + P(|s_n^{-2} V_n^2 - 1| > \varepsilon) \le 2(E | s_n^{-2} V_n^2 - 1|)^{1/3}.$$

From Theorem 2 we obtain the bound

$$\sup_x | P(s_n^{-1} S_n \le x) - \Phi(x) | \le A\{[s_n^{-2} \textstyle\sum_1^n E(X_i^2 I(|X_i| > s_n)) + E | s_n^{-2} V_n^2 - 1|]^{1/3} + [s_n^{-4} \sum_1^n E(X_i^4 I(|X_i| \le s_n))]^{1/5}\},$$

where A is an absolute constant.

The above results have application far beyond the context of the martingale central limit theorem. They are potentially useful in most contexts of sums of weakly dependent random variables, for most such processes have naturally approximating martingales. A wide variety of such examples are given in Philipp and Stout (1975) and Hall and Heyde (1980, Chapter 5). The errors of approximation can be straightforwardly bounded using the inequality

$$\text{(5)} \qquad | P(X \le x) - P(Y \le x) | \le P(|X - x| \le \delta) + P(|X - Y| > \delta),$$

which holds for arbitrary random variables X, Y and $\delta > 0$.

To illustrate we shall consider the stationary linear process

$$X_n = \sum_{j=0}^{\infty} \beta(j)\varepsilon(n-j), \qquad \sum_{j=0}^{\infty} \beta^2(j) < \infty,$$

where the $\{\varepsilon(n), \mathscr{F}_n, -\infty < n < \infty\}$ are stationary and ergodic martingale differences, $\mathscr{F}_n$ being the σ-field generated by $\varepsilon(m)$, $m \le n$, and $E(\varepsilon^2(n) \mid \mathscr{F}_{n-1}) = \sigma^2 < \infty$ for each n. This model is important in time series analysis, the martingale condition corresponding to the condition that the best linear predictor is the best predictor (both in the least squares sense; see Hannan and Heyde (1972)).

Suppose that the spectral density

$$f(\lambda) = \sigma^2(2\pi)^{-1} \mid \sum_{j=0}^{\infty} \beta(j)e^{i\lambda j} \mid^2$$

is uniformly bounded and continuous at $\lambda = 0$ with $f(0) > 0$. Then

$$n^{-1/2} \sum_{j=1}^{n} X_j \to_d N(0, 2\pi f(0))$$

(Heyde (1974)). If we suppose in addition that

$$\sum_{n=1}^{\infty} (\sum_{j=n}^{\infty} \beta(j))^2 < \infty$$

and $E\varepsilon^4(n) < \infty$ for each n, we can apply the corollary of the present paper.

Put $C = \sum_{j=0}^{\infty} \beta(j)$, $Y_j = C\varepsilon(j)$, $0 \le j < \infty$, $S_n = \sum_{j=1}^{n} X_j$ and $T_n = \sum_{j=1}^{n} Y_j$. The martingale $\{T_n, n \ge 1\}$ closely approximates the process $\{S_n, n \ge 1\}$. Indeed, we may write

$$X_k = Y_k + Z_k - Z_{k+1}, \qquad -\infty < k < \infty,$$

where $\{Z_k, -\infty < k < \infty\}$ is a stationary sequence with

$$EZ_0^2 = \sum_{n=1}^{\infty} (\sum_{j=n}^{\infty} \beta(j))^2 < \infty$$

(Hall and Heyde (1980), Chapter 5, Section 5.4). Specifically,

$$Z_n = \sum_{j=-\infty}^{-1} \varepsilon(n+j) \sum_{k=-j}^{\infty} \beta(k).$$

Then, using (5),

$$\begin{aligned} &| P((2\pi f(0)n)^{-1/2}S_n \le x) - P((2\pi f(0)n)^{-1/2}T_n \le x) | \\ (6) \qquad &\le P(| (2\pi f(0)n)^{-1/2}T_n - x | \le \delta) + P((2\pi f(0)n)^{-1/2} | Z_1 - Z_{n+1} | > \delta) \\ &\le P(x - \delta \le (2\pi f(0)n)^{-1/2}T_n \le x + \delta) + O(n^{-1}\delta^2) \end{aligned}$$

since $EZ_0^2 < \infty$. Furthermore, the corollary gives (setting $m = 0$, $\delta = 1$, and noting that $\eta = 1$),

$$(7) \qquad \sup_x | P((2\pi f(0)n)^{-1/2}T_n \le x) - \Phi(x) | = O(n^{-1/5}).$$

Taking $\delta = \delta_n = n^{-1/3}$ in (6) and using (7) we finally obtain

$$\sup_x | P((2\pi f(0)n)^{-1/2}S_n \le x) - \Phi(x) | = O(n^{-1/5}).$$

3. The proofs. We first prove the following lemma.

LEMMA. *Let $W(t)$, $t \ge 0$, be standardized Brownian motion and let T be a nonegative random variable. Then for all real x and all $0 < \varepsilon \le 1/2$,*

$$| P(W(T) \le x) - \Phi(x) | \le (2.65)\varepsilon^{1/2} \exp(-x^2/4) + P(| T - 1 | > \varepsilon).$$

PROOF. Using the techniques of Heyde and Brown (1970) we obtain the bound

$$(8) \qquad | P(W(T) \le x) - \Phi(x) | \le \pi^{-1/2} \int_0^{\infty} \psi(x, y)e^{-y^2/4}\, dy + P(| T - 1 | > \varepsilon)$$

where $\psi(x, y) = \Phi((1-\varepsilon)^{-1/2}(x+\varepsilon^{1/2}y)) - \Phi((1-\varepsilon)^{-1/2}(x-\varepsilon^{1/2}y))$. For $x \geq 0$ and $0 < \varepsilon \leq \frac{1}{2}$ we have the bounds

$$\int_0^1 \psi(x, y)e^{-y^2/4}\,dy \leq \sup_{0\leq y\leq 1} \psi(x, y) \leq [2\varepsilon/\pi(1-\varepsilon)]^{1/2} \exp(-u^2/2(1-\varepsilon)), \tag{9}$$

where

$$u^2 = \inf_{|y|\leq 1} |x+\varepsilon^{1/2}y|^2 > \tfrac{1}{2}(1-\varepsilon)x^2 - (1-\varepsilon),$$

and

$$\int_1^\infty \psi(x, y)e^{-y^2/4}\,dy \leq \int_0^\infty \psi(x, y)\,ye^{-y^2/4}\,dy$$

$$= \int_{-\infty}^{(1-\varepsilon)^{-1/2}x} (2\pi)^{-1/2}e^{-t^2/2}\,dt \int_{-\varepsilon^{1/2}[(1-\varepsilon)^{1/2}t-x]}^\infty ye^{-y^2/4}\,dy$$

$$+ \int_{(1-\varepsilon)^{-1/2}x}^\infty (2\pi)^{-1/2}e^{-t^2/2}\,dt \int_{\varepsilon^{-1/2}[(1-\varepsilon)^{1/2}t-x]}^\infty ye^{-y^2/4}\,dy.$$

The first term on the right-hand side equals

$$(2/\pi)^{1/2} \int_{-\infty}^{(1-\varepsilon)^{-1/2}x} \exp\{-\tfrac{1}{2}t^2 - [(1-\varepsilon)^{1/2}t - x]^2/4\varepsilon\}\,dt$$

$$\leq [2/\pi(1-\varepsilon)]^{1/2} \int_{-\infty}^{x} \exp\{-\tfrac{1}{2}u^2 - (u-x)^2/4\,\varepsilon\}\,du$$

$$= [2/\pi(1-\varepsilon)]^{1/2} \int_{-\infty}^{x} \exp\{-(1+2\varepsilon)[u - x/(1+2\varepsilon)]^2/4\varepsilon - x^2/2(1+2\varepsilon)\}\,du.$$

A similar bound applies to the second term and so

$$\int_1^\infty \psi(x, y)e^{-y^2/4}\,dy$$

$$\leq [2/\pi(1-\varepsilon)]^{1/2} \int_{-\infty}^{\infty} \exp\{-(1+2\varepsilon)[u - x/(1+2\varepsilon)]^2/4\varepsilon - x^2/2(1+2\varepsilon)\}\,du$$

$$= 2^{3/2}[\varepsilon/(1+2\varepsilon)(1-\varepsilon)]^{1/2} \exp[-x^2/2(1+2\varepsilon)].$$

Combining this estimate with (8) and (9) we see that, for $0 < \varepsilon \leq \frac{1}{2}$,

$$|P(W(T) \leq x) - \Phi(x)| \leq 2\pi^{-1}\varepsilon^{1/2}\exp(-x^2/4 + \tfrac{1}{2}) + 2^{3/2}\pi^{-1/2}\varepsilon^{1/2}e^{-x^2/4} + P(|T-1| > \varepsilon),$$

and a simple computation completes the proof of the lemma.

Proof of Theorem 1. (The symbol C denotes a generic positive constant, depending only on c, c' and δ.) We observe that

$$P(S_n \leq x) - E[\Phi(\eta^{-1}x)] \leq P(S_n - S_m \leq x + \varepsilon) - E[\Phi(\eta^{-1}(x+\varepsilon))]$$
$$+ P(|S_m| > \varepsilon) + E|\Phi(\eta^{-1}x) - \Phi(\eta^{-1}(x+\varepsilon))|.$$

From this and a similar inequality we see that

$$\begin{aligned} |P(S_n \le x) - E[\Phi(\eta^{-1}x)]| \le \max\nolimits_{\beta=\pm 1} \{ |P(S_n - S_m \le x + \beta\varepsilon) \\ - E[\Phi(\eta^{-1}(x + \beta\varepsilon))]| \} \\ + P(|S_m| > \varepsilon) + \max\nolimits_{\beta=\pm 1} \{ |E[\Phi(\eta^{-1}x)] - E[\Phi(\eta^{-1}(x + \beta\varepsilon))]| \}. \end{aligned} \tag{10}$$

Let

$$\varepsilon = \varepsilon(n, m, x) = L_{nm}^{1/(3+2\delta)} (1 + |x|^{\alpha})^{1/2(1+\delta)}, \tag{11}$$

where α is any fixed number subject to the constraint

$$0 < \alpha < 2(1 + \delta). \tag{12}$$

Using Markov's inequality followed by Burkholder's inequality (see Burkholder (1973)),

$$\begin{aligned} P(|S_m| > \varepsilon) &\le \varepsilon^{-2-2\delta} E|S_m|^{2+2\delta} \le C\varepsilon^{-2-2\delta} E(U_m^{2+2\delta}) \\ &\le C\varepsilon^{-2-2\delta} L_{nm} = CL_{nm}^{1/(3+2\delta)} (1 + |x|^{\alpha})^{-1}. \end{aligned} \tag{13}$$

With $m(\varepsilon, x) = \min\{|z| : x - \varepsilon \le z \le x + \varepsilon\}$ we have

$$|\Phi(\eta^{-1}x) - \Phi(\eta^{-1}(x \pm \varepsilon))| \le (2\pi)^{-1/2}\eta^{-1}\varepsilon \exp[-\tfrac{1}{2}\eta^{-2}m^2(\varepsilon, x)].$$

From (12) we see that there exist constants C_1 and C_2 such that whenever $L_{nm} \le 1$ and $|x| > C_1$, $m(\varepsilon, x) > C_2|x|$. Therefore

$$E|\Phi(\eta^{-1}x) - \Phi(\eta^{-1}(x \pm \varepsilon))| \le CL_{nm}^{1/(3+2\delta)} (1 + |x|^{\alpha})^{-1}. \tag{14}$$

In view of the formula

$$\sup\nolimits_{z>0} z^{\alpha+1} e^{-\beta z^2} = [(\alpha + 1)/2\beta]^{(\alpha+1)/2} e^{-(\alpha+1)/2}$$

and condition (12) we may write

$$(1 + |x|^{\alpha})\, |x \pm \varepsilon| \exp(-\tfrac{1}{2}\eta^2(x \pm \varepsilon)^2) \le C\eta^{\alpha+1}.$$

The same inequality holds if η is replaced by $\eta_m = [E(\eta^2 | \mathscr{F}_m)]^{1/2}$, and so

$$\begin{aligned} &|\Phi(\eta^{-1}(x \pm \varepsilon)) - \Phi(\eta_m^{-1}(x \pm \varepsilon))| \\ &\le |(x \pm \varepsilon)(\eta^{-1} - \eta_m^{-1})| [\exp(-\tfrac{1}{2}\eta^{-2}(x \pm \varepsilon)^2) + \exp(-\tfrac{1}{2}\eta_m^{-2}(x \pm \varepsilon)^2)] \\ &\le C(1 + |x|^{\alpha})^{-1} |\eta - \eta_m| (\eta\eta_m)^{-1}(\eta^{\alpha+1} + \eta_m^{\alpha+1}) \\ &\le Cc^{-1}(1 + |x|^{\alpha})^{-1} |\eta^2 - \eta_m^2| (\eta^{\alpha-1} + \eta_m^{\alpha-1}). \end{aligned}$$

Therefore by Hölder's and Jensen's inequalities

$$\begin{aligned} \Delta(x) &= (1 + |x|^{\alpha}) E[|\Phi(\eta^{-1}(x \pm \varepsilon)) - \Phi(\eta_m^{-1}(x \pm \varepsilon))| I(|\eta^2 - \eta_m^2| \le 1)] \\ &\le C_1 E[|\eta^2 - \eta_m^2| (\eta^{\alpha-1} + \eta_m^{\alpha-1}) I(|\eta^2 - \eta_m^2| \le 1)] \\ &\le C_2 \{E[|\eta^2 - \eta_m^2|^{3+2\delta} I(|\eta^2 - \eta_m^2| \le 1)]\}^{1/(3+2\delta)} \{E(\eta^{p(\alpha-1)})\}^{1/p} \end{aligned}$$

where $p^{-1} = 1 - 1/(3 + 2\delta) = 2(1 + \delta)/(3 + 2\delta)$. Let us choose

$$\alpha = 4(1 + \delta)^2/(3 + 2\delta);$$

then $p(\alpha - 1) = 2(1 + \delta) - (3 + 2\delta)/2(1 + \delta) < 2(1 + \delta)$, and so

$$\Delta(x) \le C\{E|\eta^2 - \eta_m^2|^{1+\delta}\}^{1/(3+2\delta)} \le CL_{nm}^{1/(3+2\delta)}.$$

From a similar argument we deduce the inequalities

$$(1 + |x|^{\alpha})[1 - \Phi(\eta^{-1}|x \pm \varepsilon|)] \le C\eta^{\alpha},$$

$$|\Phi(\eta^{-1}(x \pm \varepsilon)) - \Phi(\eta_m^{-1}(x \pm \varepsilon))|$$
$$\le [1 - \Phi(\eta^{-1}|x \pm \varepsilon|)] + [1 - \Phi(\eta_m^{-1}|x \pm \varepsilon|)] \le C(1 + |x|^{\alpha})^{-1}(\eta^{\alpha} + \eta_m^{\alpha})$$

and

$$\begin{aligned}
&(1 + |x|^{\alpha})E[|\Phi(\eta^{-1}(x \pm \varepsilon)) - \Phi(\eta_m^{-1}(x \pm \varepsilon))| I(|\eta^2 - \eta_m^2| > 1)] \\
&\le C_1 E[(\eta^{\alpha} + \eta_m^{\alpha}) I(|\eta^2 - \eta_m^2| > 1)] \\
&\le C_2\{P(|\eta^2 - \eta_m^2| > 1)\}^{1-\alpha/(2+2\delta)}\{E(\eta^{2+2\delta})\}^{\alpha/(2+2\delta)} \\
&\le C_3\{E|\eta^2 - \eta_m^2|^{1+\delta}\}^{1-\alpha/(2+2\delta)} \\
&= C_3\{E|\eta^2 - \eta_m^2|^{1+\delta}\}^{1/(3+2\delta)} \le C_3 L_{nm}^{1/(3+2\delta)}.
\end{aligned}$$

Combining the estimates above we obtain the bound

$$E|\Phi(\eta^{-1}(x \pm \varepsilon)) - \Phi(\eta_m^{-1}(x \pm \varepsilon))| \le C(1 + |x|^{\alpha})^{-1} L_{nm}^{1/(3+2\delta)}, \tag{15}$$

and from (10) and (13)–(15) we deduce that

$$\begin{aligned}
|P(S_n \le x) - E[\Phi(\eta^{-1}x)]| \le \max\nolimits_{\beta=\pm1} \{|P(S_n - S_m \le x + \beta\varepsilon) \\
- E[\Phi(\eta_m^{-1}(x + \beta\varepsilon))]|\} + CL_{nm}^{1/(3+2\delta)}(1 + |x|^{\alpha})^{-1}.
\end{aligned} \tag{16}$$

Let $Y_{m+1}, \cdots, Y_n$ have the distribution of $X_{m+1}, \cdots, X_n$ conditional on $\mathscr{F}_m$, and let $\mathscr{G}_\iota$ be the σ-field generated by $Y_{m+1}, \cdots, Y_\iota$. Then $\{R_i = \sum_{m+1}^{\iota} Y_j, \mathscr{G}_i, m < i \le n\}$ is a martingale, and the Skorokhod embedding theorem (Strassen (1967, Theorem 4.3)) permits us to assume without loss of generality that $R_i = W(T_i)$ a.s., $m < i \le n$, where conditional on $\mathscr{F}_m$, W is standardized Brownian motion and $\{T_i\}$ is a nondecreasing sequence of nonnegative variables. Let P_m and E_m denote operators conditional on $\mathscr{F}_m$, and define

$$\begin{aligned}
\Delta_{nm}(x) &= |P_m(S_n - S_m \le x) - \Phi(\eta_m^{-1}x)| \\
&= |P_m(W(T_n) \le x) - \Phi(\eta_m^{-1}x)| \\
&= |P_m(V(T_n/\eta_m^2) \le \eta_m^{-1}x) - \Phi(\eta_m^{-1}x)|,
\end{aligned}$$

where conditional on $\mathscr{F}_m$, $V(t) = W(\eta_m^2 t)/\eta_m$ is standardized Brownian motion. From the lemma we see that for any $\mathscr{F}_m$-measurable variable Z satisfying $0 < Z \le 1/2$,

$$\Delta_{nm}(x) \le 3Z^{1/2}\exp(-x^2/4\eta_m^2) + P_m(|T_n/\eta_m^2 - 1| > Z).$$

Let $Z = \varepsilon^2/\eta_m^2$ where ε is defined by (11). Assume for the time being that $\varepsilon^2 \le 1/2 c^2$, so that $Z \le 1/2$. Then

$$\Delta_{nm}(x) \le CL_{nm}^{1/(3+2\delta)}(1 + |x|^{\alpha})^{-1} + \varepsilon^{-2-2\delta}E_m|T_n - \eta_m^2|^{1+\delta}$$

and so

$$\begin{aligned}
|P(S_n - S_m \le x) - E[\Phi(\eta_m^{-1}x)]| &\le E[\Delta_{nm}(x)] \\
&\le CL_{nm}^{1/(3+2\delta)}(1 + |x|^{\alpha})^{-1} + \varepsilon^{-2-2\delta}E|T_n - \eta_m^2|^{1+\delta}.
\end{aligned} \tag{17}$$

Let $A_n^2 = \sum_{m+1}^n E(Y_i^2 | \mathscr{G}_{i-1})$ and $B_n^2 = \sum_{m+1}^n Y_i^2$. As on page 2164 of Heyde and Brown (1970) we obtain the estimates

$$E_m|T_n - A_n^2|^{1+\delta} + E_m|A_n^2 - B_n^2|^{1+\delta} \le C\sum\nolimits_{m+1}^n E_m|Y_i|^{2+2\delta} \tag{18}$$

and

$$E|U_n^2 - V_n^2|^{1+\delta} \le C\sum\nolimits_1^n E|X_i|^{2+2\delta}. \tag{19}$$

From (18) we see that $E|T_n - \eta_m^2|^{1+\delta} \le CL_{nm}$, and in view of (11) and (12) we deduce from (16) and (17) that (for $\varepsilon^2 \le \frac{1}{2}c^2$)

$$|P(S_n \le x) - E[\Phi(\eta^{-1}x)]| \le CL_{nm}^{1/(3+2\delta)}(1+|x|^\alpha)^{-1}.$$

Next consider the case $\varepsilon^2 > \frac{1}{2}c^2$ and $|x| > 1$. Then

$$|x|^{-2-2\delta} < [\tfrac{1}{2}(1+|x|^\alpha)]^{-2(1+\delta)/\alpha},$$

$$1 < 2^{1/2}c^{-1}L_{nm}^{1/(3+2\delta)}(1+|x|^\alpha)^{(1+\delta)/2},$$

$$\begin{aligned}|P(S_n - S_m \le x) - E[\Phi(\eta_m^{-1}x)]| &\le P(|S_n - S_m| > |x|) + E[1 - \Phi(\eta_m^{-1}|x|)] \\ &\le |x|^{-2-2\delta}[E|S_n - S_m|^{2+2\delta} + E(\eta_m^{2+2\delta})E|N|^{2+2\delta}]\end{aligned}$$

where N is a standard normal variable, and by Burkholder's inequality,

$$\begin{aligned}E|S_n - S_m|^{2+2\delta} \le C_1(\textstyle\sum_{m+1}^n X_i^2)^{1+\delta} &= C_1E|U_n^2 - U_m^2|^{1+\delta} \\ &\le C_2(L_{nm} + E(\eta_m^{2+2\delta})).\end{aligned}$$

Therefore if $L_{nm} \le 1$,

$$|P(S_n - S_m \le x) - E[\Phi(\eta_m^{-1}x)]| \le CL_{nm}^{1/(3+2\delta)}(1+|x|^\alpha)^{(1+\delta)/2-2(1+\delta)/\alpha}. \tag{20}$$

Comparing this with (17) we see that the best overall rate of convergence is obtained when $(1+\delta)/2 - 2(1+\delta)/\alpha = -1$; that is, $\alpha = 4(1+\delta)^2/(3+2\delta)$.

Finally, suppose that $\varepsilon^2 > \frac{1}{2}c^2$ and $|x| \le 1$. Then

$$4c^{-1}L_{nm}^{1/(3+2\delta)}(1+|x|^\alpha)^{-1} > 1 > |P(S_n - S_m \le x) - E[\Phi(\eta_m^{-1}x)]|,$$

and from this, (16) and (20) we deduce that (2) holds for $\varepsilon^2 > \frac{1}{2}c^2$ and all x. This completes the proof of (2). The result when U_n^2 is replaced by V_n^2 follows from (19).

PROOF OF THEOREM 2. (The symbol C denotes a generic *absolute* positive constant.) Set $Z_i = X_iI(|X_i| \le 1)$, $Y_i = Z_i - E(Z_i | \mathcal{F}_{i-1})$, $R_i = \sum_{j=1}^i Y_j$ and let $\mathcal{G}_i$ be the σ-field generated by $Y_1, \ldots, Y_i$, $1 \le i \le n$. Then $\{R_i, \mathcal{G}_i, 1 \le i \le n\}$ is a martingale. Instead of (10) we use the bound

$$\sup_x|P(S_n \le x) - \Phi(x)| \le \sup_x|P(R_n \le x) - \Phi(x)| + P(|S_n - R_n| > \Delta) + (2\pi)^{-1/2}\Delta$$

where $\Delta > 0$ (replace S_m by $S_n - R_n$ in (10)). Since

$$E(S_n - R_n)^2 = \textstyle\sum_1^n E(X_i - Y_i)^2 \le \sum_1^n E[X_i^2I(|X_i| > 1)]$$

then

$$\begin{aligned}&\sup_x|P(S_n \le x) - \Phi(x)| \\ &\quad\le \sup_x|P(R_n \le x) - \Phi(x)| + \Delta^{-2}\textstyle\sum_1^n E[X_i^2I(|X_i| > 1)] + (2\pi)^{-1/2}\Delta.\end{aligned} \tag{21}$$

Embed the martingale $\{R_i, \mathcal{G}_i, 1 \le i \le n\}$ in Brownian motion as in the proof of Theorem 1, so that $R_i = W(T_i)$ a.s., $1 \le i \le n$. Set $\tau_i = T_i - T_{i-1}$ ($T_0 \equiv 0$) and let $\mathcal{G}_i^*$ be the σ-field generated by $Y_1, \ldots, Y_i$ and $W(t)$ for $t \le T_i$. The sequence

$$\{\textstyle\sum_1^i [\tau_j - E(\tau_j | \mathcal{G}_{j-1}^*)], \mathcal{G}_i^*, 1 \le i \le n\}$$

is a martingale, $E(Y_i^2 | \mathcal{G}_{i-1}) = E(\tau_i | \mathcal{G}_{i-1}^*)$ a.s., and

$$E(\tau_i^{1+\delta} | \mathcal{G}_{i-1}^*) \le C_\delta E(|Y_i|^{2+2\delta} | \mathcal{G}_{i-1}) \text{ a.s.,}$$

where $\sup_{0<\delta<1} C_\delta < \infty$. (See Strassen (1967), Theorem 4.3. An explicit form for the constants C_δ is given in Theorem A.1, Appendix I of Hall and Heyde (1980).) Letting C_1, $C_2 \cdots$ denote absolute constants and using an explicit form for the constants in Burkholder's inequality (see Burkholder ((1973), pages 22–23) we deduce that

$$
\begin{aligned}
&E\,|\textstyle\sum_1^n [\tau_i - E(\tau_i\,|\,\mathscr{G}_{i-1}^*)]\,|^{1+\delta} \le C_1\delta^{-(1+\delta)/2}E\{\sum_1^n [\tau_i - E(\tau_i\,|\,\mathscr{G}_{i-1}^*)]^2\}^{(1+\delta)/2} \\
(22)\qquad &\le C_1\delta^{-(1+\delta)/2}\, E\textstyle\sum_1^n |\,\tau_i - E(\tau_i\,|\,\mathscr{G}_{i-1}^*)\,|^{1+\delta} \le C_2\delta^{-(1+\delta)/2}\, E(\sum_1^n \tau_i^{1+\delta}) \\
&\le C_3\delta^{-(1+\delta)/2}\textstyle\sum_1^n E\,|\,Y_i\,|^{2+2\delta} \le C_4\delta^{-(1+\delta)/2}\sum_1^n E[\,|X_i|^{2+2\delta}I(\,|X_i|\le 1)].
\end{aligned}
$$

(Note that $0 < \delta \le 1$.) Similarly

$$
(23)\qquad E\,|\textstyle\sum_1^n [Y_i^2 - E(Y_i^2\,|\,\mathscr{G}_{i-1})]\,|^{1+\delta} \le C\,\delta^{-(1+\delta)/2}\sum_1^n E[\,|X_i|^{2+2\delta}I(\,|X_i|\le 1)],
$$

$$
\begin{aligned}
&E\,|\textstyle\sum_1^n \{Y_i^2 - Z_i^2 + [E(Z_i\,|\,\mathscr{F}_{i-1})]^2\}\,|^{1+\delta} \\
(24)\qquad &= 2^{1+\delta}\, E\,|\textstyle\sum_1^n [Z_i - E(Z_i\,|\,\mathscr{F}_{i-1})]E(Z_i\,|\,\mathscr{F}_{i-1})\,|^{1+\delta} \\
&\le C_1\delta^{-(1+\delta)/2}\, E\,|\textstyle\sum_1^n [Z_i - E(Z_i\,|\,\mathscr{F}_{i-1})]^2[E(Z_i\,|\,\mathscr{F}_{i-1})]^2\,|^{(1+\delta)/2} \\
&\le C_2\delta^{-(1+\delta)/2}\textstyle\sum_1^n E[\,|X_i|^{2+2\delta}I(\,|X_i|\le 1)]
\end{aligned}
$$

and

$$
(25)\qquad E\,|\textstyle\sum_1^n \{X_i^2 - Z_i^2 + [E(Z_i\,|\,\mathscr{F}_{i-1})]^2\}\,| \le 2\sum_1^n E[X_i^2\, I(\,|X_i| > 1)].
$$

From Lemma 1 we see that for any $\Delta > 0$,

$$
\sup_x |\,P(W(T_n) \le x) - \Phi(x)\,| \le C\,\Delta + P(\,|\,T_n - 1\,| > 5\,\Delta^2),
$$

and so

$$
\begin{aligned}
\sup_x |\,P(R_n \le x) - \Phi(x)\,| &\le C[\Delta + P(\,|\,T_n - \textstyle\sum_1^n E(\tau_i\,|\,\mathscr{G}_{i-1}^*)\,| > \Delta^2) \\
&+ P(\,|\textstyle\sum_1^n [E(Y_i^2\,|\,\mathscr{G}_{i-1}) - Y_i^2]\,| > \Delta^2) + P(\,|\sum_1^n \{Y_i^2 - Z_i^2 + [E(Z_i\,|\,\mathscr{F}_{i-1})]^2\}\,| > \Delta^2) \\
&+ P(\,|\textstyle\sum_1^n \{X_i^2 - Z_i^2 + [E(Z_i\,|\,\mathscr{F}_{i-1})]^2\}\,| > \Delta^2) + P(\,|\,U_n^2 - 1\,| > \Delta^2)].
\end{aligned}
$$

Combining this with (21)–(25) we find that for any $\Delta > 0$,

$$
\begin{aligned}
&\sup_x |\,P(S_n \le x) - \Phi(x)\,| \le C\{\Delta + \Delta^{-2-2\delta}\delta^{-(1+\delta)/2}\textstyle\sum_1^n E[\,|X_i|^{2+2\delta}I(\,|X_i|\le 1)] \\
(26)\qquad &\qquad + \Delta^{-2}\textstyle\sum_1^n E[X_i^2\, I(\,|X_i| > 1)] + P(\,|\,U_n^2 - 1\,| > \Delta^2)\}.
\end{aligned}
$$

Given $\varepsilon > 0$, let

$$
\Delta = \max\{[\textstyle\sum_1^n E(X_i^2\, I(\,|X_i| > 1))]^{1/3}, [\sum_1^n E(\,|X_i|^{2+2\delta}I(\,|X_i|\le 1))]^{1/(3+2\delta)}, \varepsilon^{1/2}\},
$$

to prove the first part of Theorem 2.

To see that U_n^2 may be replaced by V_n^2, note that

$$
\begin{aligned}
P(\,|\,U_n^2 - 1\,| > 4\,\Delta^2) &\le P(\textstyle\sum_1^n X_i^2\, I(\,|X_i| > 1) > \Delta^2) \\
&+ P(\,|\textstyle\sum_1^n [X_i^2\, I(\,|X_i|\le 1) - E(X_i^2\, I(\,|X_i|\le 1)\,|\,\mathscr{F}_{i-1})]\,| > \Delta^2) \\
&+ P(\textstyle\sum_1^n E[X_i^2\, I(\,|X_i| > 1)\,|\,\mathscr{F}_{i-1}] > \Delta^2) + P(\,|\,V_n^2 - 1\,| > \Delta^2).
\end{aligned}
$$

The first and third terms on the right-hand side are dominated by $\Delta^{-2}\sum_1^n E[X_i^2\, I(\,|X_i| > 1)]$, and using the argument leading to (22) we see that the second is dominated by

$$
C\,\delta^{-(1+\delta)/2}\,\Delta^{-2-2\delta}\textstyle\sum_1^n E[\,|X_i|^{2+2\delta}I(\,|X_i|\le 1)].
$$

From these estimates and (26) we deduce the desired result.

REFERENCES

ALDOUS, D. J. and EAGLESON, G. K. (1978). On mixing and stability of limit theorems. *Ann. Probability* **6** 325–331.

BROWN, B. M. (1971). Martingale central limit theorems. *Ann. Math. Statist.* **42** 59–66.

BURKHOLDER, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probability* **1** 19–42.

CHATTERJI, S. D. (1974). A principle of subsequences in probability theory: the central limit theorem. *Advances in Math.* **13** 31–54.

DUBINS, L. E. (1968). On a theorem of Skorokhod. *Ann. Math. Statist.* **39** 2094–2097.

EAGLESON, G. K. (1975). Martingale convergence to mixtures of infinitely divisible laws. *Ann. Probability* **3** 557–562.

ERICKSON, R. V., QUINE, M. P. and WEBER, N. C. (1979). Explicit bounds for the departure from normality of sums of dependent random variables. *Acta Math. Acad. Sci. Hungar.* **34** 27–32.

FELLER, W. (1968). On the Berry-Esseen theorem. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **10** 261–268.

FELLER, W. (1971). *An Introduction to Probability Theory and its Applications. Vol. II, 2nd ed.* Wiley, New York.

GRAMS, W. F. (1972). Rates of convergence in the central limit theorem for dependent variables. Ph.D. Dissertation, Florida State Univ.

HALL, P. (1977). Martingale invariance princples. *Ann. Probability* **5** 875–887.

HALL, P. (1979). On the Skorokhod representation approach to martingale invariance principles. *Ann. Probability* **7** 371–376.

HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and its Application.* Academic, New York.

HANNAN, E. J. and HEYDE, C. C. (1972). On limit theorems for quadratic functions of discrete time series. *Ann. Math. Statist.* **43** 2058–2066.

HEYDE, C. C. (1974). On the central limit theorem for stationary processes. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **38** 315–320.

HEYDE, C. C. and BROWN, B. M. (1970). On the departure from normality of a certain class of martingales. *Ann. Math. Statist.* **41** 2161–2165.

IBRAGIMOV, I. A. (1963). A central limit theorem for a class of dependent random variables. *Theor. Probability Appl.* **8** 83–89.

KATO, Y. (1979). Rates of convergence in central limit theorem for martingale differences. *Bull. Math. Statist.* **18** 1–8.

NAKATA, T. (1976). On the rate of convergence in mean central limit theorems for martingale differences. *Rep. Statist. Appl. Res. Un. Japan Sci. Engrs.* **23** 10–15.

PHILIPP, W. and STOUT, W. (1975). Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.* **161**.

ROOTZÉN, H. (1977). A note on convergence to mixtures of normal distributions. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **38** 211–216.

SCOTT, D. J. (1973). Central limit theorems for martingales and for processes with stationary increments using a Skorokhod representation approach. *Adv. Appl. Probability* **5** 119–137.

STRASSEN, V. (1967). Almost sure behaviour of sums of independent random variables and martingales. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **2** 315–343.

DEPARTMENT OF STATISTICS, S.G.S.,
AUSTRALIAN NATIONAL UNIVERSITY,
P.O. BOX 4
CANBERRA, ACT 2600
AUSTRALIA

CSIRO DIVISION OF MATHEMATICS
AND STATISTICS,
P.O. BOX 1965
CANBERRA, ACT 2601
AUSTRALIA