The Annals of Probability
1979, Vol. 7, No. 2, 371–376

ON THE SKOROKHOD REPRESENTATION APPROACH TO MARTINGALE INVARIANCE PRINCIPLES

BY PETER HALL[1]

University of Melbourne

The Skorokhod representation method of proving martingale invariance principles has received only limited attention. This paper shows it to be a very powerful tool. We use the representation to obtain sufficient conditions for a sequence of random functions to be tight, and to give a direct proof of an invariance principle.

1. Introduction and results. The central limit theorem for martingales has attracted considerable attention since the first results were announced by Lévy in 1935. Interest has increased in recent years, and there are now several basic techniques for handling the proofs. One of these involves the Skorokhod representation for martingales, given by Strassen in 1964. Scott (1973) used the representation to extend Brown's (1971) limit theorem. However, Scott's work does not take full advantage of the representation and his method follows the traditional two-part pattern for proofs of invariance principles. Firstly he shows that the finite-dimensional distributions of his sequence of sample processes converge to those of Brownian motion, and then using this fact he establishes the tightness of the sequence by employing a result from Loynes (1970).

In this paper we show that martingale invariance principles can be obtained directly from the Skorokhod representation by exploiting the properties of Brownian motion. It is shown that tightness holds under quite general conditions, irrespective of whether or not the finite-dimensional distributions converge. We give a nontrivial example in which tightness prevails but the finite-dimensional distributions do not converge.

Let $\{(S_n, \mathcal{F}_n), n \geqslant 1\}$ be a zero-mean, square-integrable martingale, where $\mathcal{F}_n$ is the σ-field generated by $S_1, S_2, \cdots, S_n$. Define $X_n = S_n - S_{n-1}$, $U_n^2 = \Sigma_1^n X_i^2$ and $s_n^2 = E(S_n^2) = E(U_n^2)$. Let ξ_n be the random element of $C[0, 1]$ (the space of continuous functions on $[0, 1]$) defined by interpolating between the points $(0, 0)$, $(U_n^{-2}U_1^2, U_n^{-1}S_1)$, $(U_n^{-2}U_2^2, U_n^{-1}S_2)$, $\cdots$, $(1, U_n^{-1}S_n)$. Under certain conditions,

$$\xi_n \rightarrow_{\mathcal{D}} W_1 \qquad (1)$$

where W_1 is standard Brownian motion on $[0, 1]$. (See Hall (1977). Convergence in distribution will be denoted by $\rightarrow_{\mathcal{D}}$ and convergence in probability by $\rightarrow_p$. $I(E)$ will denote the indicator function of the event E.) In proving (1) it will always be

Received May 31, 1977.

[1]Now at the Australian National University.

AMS 1970 *subject classifications.* Primary 60F05; secondary 60G45.

Key words and phrases. Skorokhod representation, martingale, invariance principle.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability*. www.jstor.org

necessary to impose a negligibility condition, such as:

(2) $$U_n^{-2}\max_{j\leq n} X_j^2 \to_p 0.$$

Suppose that the sequences $\{s_n^{-2}U_n^2\}$ and $\{U_n^{-2}s_n^2\}$ are both tight; that is,

(3) $$\liminf_{n\to\infty} P(\Delta < s_n^{-2}U_n^2 < \lambda) \to 1 \quad \text{as} \quad \Delta \to 0 \quad \text{and} \quad \lambda \to \infty.$$

Then (2) is equivalent to the condition:

(4) $$s_n^{-2}\max_{j\leq n} X_j^2 \to_p 0.$$

We shall strengthen (4) to the Lindeberg condition:

(5) $$\text{for all } \varepsilon > 0,\ s_n^{-2}\Sigma_1^n E\left[X_j^2 I(|X_j| > \varepsilon s_n)\right] \to 0.$$

If $\{s_n^{-2}U_n^2\}$ is uniformly integrable then (4) and (5) are equivalent.

THEOREM 1. *If* (3) *and* (5) *hold then the sequence* $\{\xi_n\}$ *is tight.*

(All of the proofs are placed together in Section 2.)

To construct a martingale for which (3) and (5) hold but whose finite-dimensional distributions do not converge, let Y_n, $n \geq 1$ be independent $N(0, 1)$ variables; n_k be integers such that $n_1 = 1$ and $n_{k+1} = n_k + 2n_k^2$;

$$I_k = I\left(\Sigma_{n_k+1}^{n_k+n_k^2} Y_j > 0\right);\ X_1 = Y_1;$$

$$X_n = Y_n \quad \text{if} \quad n_k < n \leq n_k + n_k^2$$

$$= I_k Y_n \quad \text{if} \quad n_k + n_k^2 < n \leq n_{k+1};$$

and $S_n = \Sigma_1^n X_j$. $\{S_n\}$ is martingale, (3) and (5) hold but $\xi_n(1) = S_n/U_n$ does not converge in distribution. If S_n/c_n converges for some sequence of constants c_n then the limit is degenerate.

To see this observe that

$$n_{k+1}^{-\frac{1}{2}} S_{n_{k+1}} = n_{k+1}^{-\frac{1}{2}}\left(\Sigma_{j=1}^{n_k} Y_j + \Sigma_{j=n_k+1}^{n_k+n_k^2} Y_j + I_k \Sigma_{j=n_k+n_k^2+1}^{n_{k+1}} Y_j\right)$$

$$\sim \left(2n_k^2\right)^{-\frac{1}{2}}\left(\Sigma_{j=n_k+1}^{n_k+n_k^2} Y_j + I_k \Sigma_{j=n_k+n_k^2+1}^{n_{k+1}} Y_j\right)$$

$$=_{\mathfrak{D}} 2^{-\frac{1}{2}}(Z_1 + Z_2 I(Z_1 > 0))$$

where Z_1 and Z_2 are independent $N(0, 1)$ variables. (Here $\sim$ means "has asymptotically the same distribution as.") On the other hand,

$$\left(n_k + n_k^2\right)^{-\frac{1}{2}} S_{n_k+n_k^2} \sim n_k^{-1}\Sigma_{j=n_k+1}^{n_k+n_k^2} Y_j =_{\mathfrak{D}} Z_1.$$

Similarly it can be shown that

$$U_{n_{k+1}}^{-1} S_{n_{k+1}} \sim Z_1 I(Z_1 \leq 0) + 2^{-\frac{1}{2}}(Z_1 + Z_2)I(Z_1 > 0) \quad \text{and}$$

$$U_{n_k+n_k^2}^{-1} S_{n_k+n_k^2} \sim Z_1.$$

Hall (1977) showed that if (4) holds in L^1 and if the tightness condition (3) is replaced by the stronger condition

(6) $$s_n^{-2}U_n^2 \to_p T, \qquad \text{where } 0 < T < \infty \text{ a.s.},$$

then (1) is true. If (5) is used instead of (4) then a shorter and more elegant proof is available via the Skorokhod representation. In fact, it is very easy to prove a little more.

THEOREM 2. *Under* (5) *and* (6) *we have for any* $E \in \mathcal{F}$,

$$(\xi_n, U_n^2/s_n^2, I(E)) \rightarrow_{\mathfrak{D}} (W_1, T, I(E)),$$

where W_1 *is independent of* $(T, I(E))$. *That is, the limit theorem is mixing.*

(See Eagleson (1977).) The proof is via a limit theorem for Brownian motion, and may have application to processes other than martingales which can be embedded in the Skorokhod way:

THEOREM 3. *Let* $W(t)$, $t \geqslant 0$ *be a standard Brownian motion and* T_n, $n \geqslant 1$ *be positive random variables. Define* $\eta_n(t) = W(tT_n)/(T_n)^{\frac{1}{2}}$, $t \in [0, 1]$. *If there exist constants* c_n *such that*

(7) $$T_n/c_n \rightarrow_p T,\ c_n \rightarrow \infty \quad \textit{and} \quad 0 < T < \infty \text{ a.s.},$$

then for all events E in the probability space,

$$(\eta_n, T_n/c_n, I(E)) \rightarrow_{\mathfrak{D}} (W_1, T, I(E)),$$

where W_1 *is independent of* $(T, I(E))$.

(Theorem 3 is an extension of Billingsley's (1968) result (17.9), page 145, and can be proved using his techniques. We do not give a proof here.)

2. The proofs. Let T_n, η_n and W be the stochastic processes defined in Theorem 3. First we establish the tightness of $\{\eta_n, n \geqslant 1\}$:

LEMMA 1. *If*

(8) $$\liminf_{n\rightarrow\infty} P(\Delta < T_n/c_n < \lambda) \rightarrow 1 \quad \textit{as} \quad \Delta \rightarrow 0 \quad \textit{and} \quad \lambda \rightarrow \infty$$

for some sequence of constants c_n, *then for all* $\varepsilon > 0$,

$$\lim_{\delta\rightarrow 0} \limsup_{n\rightarrow\infty} P\big(\sup_{|u-v|<\delta} |\eta_n(u) - \eta_n(v)| > \varepsilon\big) = 0.$$

PROOF. The probability above is dominated by

$$P\big(\sup_{|u-v|<\lambda\delta} |W(u) - W(v)| > \varepsilon\Delta^{\frac{1}{2}}\big) + P(T_n/c_n < \Delta \text{ or } > \lambda).$$

Now let $n \rightarrow \infty$, $\delta \rightarrow 0$, $\Delta \rightarrow 0$ and $\lambda \rightarrow \infty$, and use the tightness of $\{W(t), t \geqslant 0\}$.

Let Z_n, $n \geqslant 1$ be random variables satisfying $0 \leqslant Z_0 \leqslant Z_1 \leqslant \cdots$. Define random elements η_n' of $C[0, 1]$ by interpolating between the points $(0, 0)$, $(Z_n^{-1}Z_1, Z_n^{-\frac{1}{2}}W(T_1)), \cdots, (1, Z_n^{-\frac{1}{2}}W(T_n))$. If the Z's are close to the T's then η_n is close to η_n':

LEMMA 2. *If* (8) *holds, if* $T_n/Z_n \rightarrow_p 1$,

$$\max_{j\leqslant n}(Z_j - Z_{j-1})/Z_n \rightarrow_p 0 \quad \textit{and} \quad \max_{j\leqslant n}|T_n^{-1}T_j - Z_n^{-1}Z_j| \rightarrow_p 0$$

then $\rho(\eta_n, \eta_n') \rightarrow_p 0$, *where* ρ *denotes the uniform metric.*

PROOF. Since the norming variables T_n and Z_n are asymptotically the same, it suffices to prove that η_n is uniformly close to the process η_n'' obtained by interpolating between the points (0, 0), $(Z_n^{-1}Z_1, T_n^{-\frac{1}{2}}W(T_1)), \cdots, (1, T_n^{-\frac{1}{2}}W(T_n))$. If $\max_{j\leqslant n}(Z_j - Z_{j-1})/Z_n \leqslant \delta$ and $\max_{j\leqslant n}|T_n^{-1}T_j - Z_n^{-1}Z_j| \leqslant \delta$ then

$$\sup_{z\in[0,1]}|\eta_n(z) - \eta_n''(z)| \leqslant \sup_{|u-v|\leqslant 2\delta}|\eta_n(u) - \eta_n(v)|.$$

In view of the tightness of $\{\eta_n\}$, $\rho(\eta_n, \eta_n'') \to_p 0$, as required.

Now we introduce the martingale theory. We approximate to the martingale $\{(S_n, \mathcal{F}_n)\}$ by a truncated martingale $\{(S_n^*, \mathcal{F}_n^*)\}$, and show that the approximation is uniformly close. Then we apply Strassen's (1964) Skorokhod representation, proving that $S_n^* = W(T_n)$ a.s., $n \geqslant 1$, for an increasing sequence of positive random variables $\{T_n\}$. If (6) holds then the T_n's satisfy $T_n/s_n^2 \to_p T$; and so by Theorem 3, $\eta_n \to_{\mathfrak{D}} W_1$, where η_n is defined by $\eta_n(t) = W(tT_n)/(T_n)^{\frac{1}{2}}$. η_n is uniformly close to ξ_n and so $\xi_n \to_{\mathfrak{D}} W_1$. Some of the techniques of our approximation are drawn from Scott (1973), and we refer the reader to this paper for details.

The Lindeberg condition (5) is equivalent to the apparently stronger condition:

$$\text{for all } \varepsilon > 0,\ s_n^{-2}\Sigma_1^n E\big[X_j^2 I(|X_j| > \varepsilon s_j)\big] \to 0. \tag{9}$$

To see this, let $0 < \delta < \varepsilon$ and $k_n = \max\{j \leqslant n | \varepsilon s_j \leqslant \delta s_n\}$. If (5) holds then the left side of (9) does not exceed

$$s_n^{-2}\Sigma_1^n E\big[X_j^2 I(|X_j| > \delta s_n)\big] + s_n^{-2}\Sigma_1^{k_n} E(X_j^2) \leqslant o(1) + \delta^2/\varepsilon^2,$$

and (9) follows on letting $n \to \infty$ and then $\delta \to 0$.

Fix $\varepsilon > 0$ and define $X_o^* = 0$ and $X_j^* = X_j I(|X_j| \leqslant \varepsilon s_j) - E[X_j I(|X_j| \leqslant \varepsilon s_j)|\mathcal{F}_{j-1}^*]$, $1 \leqslant j \leqslant n$, where $\mathcal{F}_{j-1}^*$ is the σ-field generated by $X_1^*, X_2^*, \cdots, X_{j-1}^*$. Let $S_n^* = \Sigma_1^n X_j^*$, $U_n^{*2} = \Sigma_1^n X_j^{*2}$, and ξ_n^* be the process obtained by interpolating between the points (0, 0), $(U_n^{-2}U_1^2, U_n^{-1}S_1^*), \cdots, (1, U_n^{-1}S_n^*)$.

LEMMA 3. *Under condition* (9),

$$s_n^{-1}\Sigma_1^n E\big[|X_j| I(|X_j| > \varepsilon s_j)\big] \to 0, \tag{10}$$

$$s_n^{-4}\Sigma_1^n E\big[X_j^4 I(|X_j| \leqslant \varepsilon s_j)\big] \to 0, \tag{11}$$

and

$$s_n^{-2}\max_{j\leqslant n}|U_j^2 - U_j^{*2}| \to_p 0. \tag{12}$$

Under (3) *and* (9),

$$\rho(\xi_n, \xi_n^*) \to_p 0. \tag{13}$$

PROOF. Let δ and k_n be as above. Then

$$\begin{aligned}(14)\quad & s_n^{-1}\Sigma_1^n E\big[|X_j| I(|X_j| > \varepsilon s_j)\big] \\ &\leqslant s_n^{-1}\Sigma_1^n \varepsilon^{-1}s_j^{-1}E\big[X_j^2 I(|X_j| > \varepsilon s_j)\big] \\ &\leqslant \delta\varepsilon^{-2}s_{k_n}^{-1}\Sigma_1^{k_n}s_j^{-1}E(X_j^2) + \delta^{-1}s_n^{-2}\Sigma_1^n E\big[X_j^2 I(|X_j| > \varepsilon s_j)\big].\end{aligned}$$

The last term is $o(1)$ and the first converges to $2\delta\varepsilon^{-2}$, since

$$s_n^{-1}\Sigma_1^n s_j^{-1}E(X_j^2) = s_n^{-1}\Sigma_1^n(s_j - s_{j-1})(1 + s_j^{-1}s_{j-1}) \sim 2s_n^{-1}\Sigma_1^n(s_j - s_{j-1}) = 2.$$

(10) follows on letting $n \to \infty$ and $\delta \to 0$ in (14). Condition (11) is proved as in Scott's Lemma 3, while (12) follows from (9), (10) and the inequalities:

$$E\left[s_n^{-2}\max_{j\leqslant n}|U_j^2 - U_j^{*2}|\right] \leqslant s_n^{-2}\Sigma_1^n E|X_j^2 - X_j^{*2}|$$
$$\leqslant s_n^{-2}\Sigma_1^n E\left[X_j^2 I(|X_j| > \varepsilon s_j)\right]$$
$$+ 3\varepsilon s_n^{-1}\Sigma_1^n E\left[|X_j| I(|X_j| > \varepsilon s_j)\right].$$

(13) is proved as in Scott's proof of his condition (23).

Now we introduce the Skorokhod representation (Theorem 4.3 of Strassen). Without loss of generality there exists a Brownian motion W and an increasing sequence of nonnegative random variables T_n such that $S_n^* = W(T_n)$ a.s., $n \geqslant 1$. Put $t_n = T_n - T_{n-1}$, $n \geqslant 1 (t_0 = 0)$, let $\mathcal{G}_n$ be the σ-field generated by $S_1^*, S_2^*, \cdots, S_n^*$ and $W(t)$ for $0 \leqslant t \leqslant T_n (n \geqslant 1)$ and let $\mathcal{G}_o$ and $\mathcal{F}_o^*$ denote the trivial σ-field. Strassen's Theorem 4.3 tells us that the T_n can be chosen such that t_n is $\mathcal{G}_n$-measurable, $E(t_n|\mathcal{G}_{n-1}) = E(X_n^{*2}|\mathcal{F}_{n-1}^*)$ a.s. $(n \geqslant 1)$, and for some constant $L > 0$, $E(t_n^2|\mathcal{G}_{n-1}) \leqslant LE(X_n^{*4}|\mathcal{F}_{n-1}^*)$ a.s. $(n \geqslant 1)$.

LEMMA 4. *Under condition* (9),

$$s_n^{-2}\max_{k\leqslant n}|T_k - \Sigma_1^k E(t_j|\mathcal{G}_{j-1})| \to_p 0 \tag{15}$$

and

$$s_n^{-2}\max_{k\leqslant n}|\Sigma_1^k E(t_j|\mathcal{G}_{j-1}) - U_k^{*2}| \to_p 0. \tag{16}$$

PROOF. Apply Kolmogorov's inequality to the martingale with differences $t_j - E(t_j|\mathcal{G}_{j-1})$ and σ-fields $\mathcal{G}_j$, proving that

$$P\left(s_n^{-2}\max_{k\leqslant n}|T_k - \Sigma_1^k E(t_j|\mathcal{G}_{j-1})| > \delta\right) \leqslant \delta^{-2}s_n^{-4}\Sigma_1^n E(t_j^2) \leqslant L\delta^{-2}s_n^{-4}\Sigma_1^n E(X_j^{*4}).$$

The proof of (15) is completed using (10) and (11), as in Scott's proof of his Lemma 12. (16) is proved in the same way, applying Kolmogorov's inequality to the martingale with differences $X_j^{*2} - E(X_j^{*2}|\mathcal{F}_{j-1}^*)$.

We are now in a position to prove Theorems 1 and 2. Suppose that (3) and (5) hold, and let $\eta_n(t) = W(tT_n)/(T_n)^{\frac{1}{2}}$. (Here the T_n's are the Strassen stopping times.) In view of (3), (12), (15) and (16), $U_n^{-2}\max_{j\leqslant n}|U_j^2 - T_j| \to_p 0$. For $T_n > 0$, $\max_{j\leqslant n}|T_n^{-1}T_j - U_n^{-2}U_j^2| \leqslant U_n^{-2}\max_{j\leqslant n}|U_j^2 - T_j| + |1 - U_n^{-2}T_n|$, and so $\max_{j\leqslant n}|T_n^{-1}T_j - U_n^{-2}U_j^2| \to_p 0$. Since $s_n^{-2}\max_{j\leqslant n}X_j^2 \leqslant \varepsilon^2 + s_n^{-2}\Sigma_1^n X_j^2 I(|X_j| > \varepsilon s_n)$ $\to_p \varepsilon^2$ then $\max_{j\leqslant n}X_j^2/U_n^2 \to_p 0$. It now follows from Lemma 2 that $\rho(\eta_n, \xi_n^*) \to_p 0$, and so by (13), $\rho(\eta_n, \xi_n) \to_p 0$. Theorem 1 now follows from Lemma 1. If (5) and (6) hold then $T_n/s_n^2 \to_p T$, and so (7) holds with $c_n = s_n^2$. Theorem 2 follows from Theorem 3 and the fact that $\rho(\eta_n, \xi_n) \to_p 0$.

Note added in proof. Since the preparation of this work a paper by David Aldous (*Ann. Probability* **6** 335–340) has appeared, giving an alternative method of establishing the tightness of stochastic processes related to martingales.

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

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

EAGLESON, G. K. (1977). Some simple conditions for limit theorems to be mixing. *Theor. Probability Appl.* **21** 637–643.

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

LOYNES, R. M. (1970). An invariance principle for reversed martingales. *Proc. Amer. Math. Soc.* **25** 56–64.

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. (1964). Almost sure behaviour of sums of independent random variables and martingales. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **2** 315–343. Univ. of California Press.

DEPARTMENT OF STATISTICS, S.G.S.
THE AUSTRALIAN NATIONAL UNIVERSITY
P. O. BOX 4, CANBERRA A.C.T.
AUSTRALIA 2600