AN INVARIANCE PRINCIPLE FOR REVERSED MARTINGALES[1]

R. M. LOYNES

ABSTRACT. Let X_n, $n=1, 2, \cdots$, be a reversed martingale with zero mean and for each n construct a random function $W_n(t)$, $0 \leqq t \leqq 1$, by a suitable method of interpolation between the values $X_k/(EX_n^2)^{1/2}$ at times EX_k^2/EX_n^2; these are the natural times to use. Then it is shown that the distribution of W_n (in function space C or D) converges weakly to that of the Wiener process, if the finite-dimensional distributions converge appropriately. It is also shown that the sufficient conditions recently given by the author for the central limit theorem for such martingales also imply convergence of finite-dimensional distributions. Illustrations of the use of these results are given in applications to U-statistics and sums of independent random variables.

A result for forward martingales exactly analogous to the first result above is also given, but is given no emphasis.

1. Introduction and summary. It was shown in Loynes [4] that, under suitable conditions, if $\{X_n : n \geqq 1\}$ is a reverse martingale there exists a function $\phi(n)$ such that the distribution of $\phi(n)^{1/2}X_n$ converges to $N(0, 1)$ weakly as $n \to \infty$. (For convenience, we take the limit of X_n as 0 in the present paper.)

Under the same conditions, define for each n a random function $W_n(t)$ for $0 \leqq t \leqq 1$ as follows:

(1) $W_n(0)=0$;

(2) $W_n(t_k^n)=\phi(n)^{1/2}X_k$, where $t_k^n=\phi(n)\{\phi(k)\}^{-1}$, for $k \geqq n$;

(3) $W_n(t)=W_n(t_k^n)$ if $t_{k+1}^n < t < t_k^n$.

(Condition (3) ensures that $W_n(t)$ is a step-function which is continuous on the left: as in other situations the theorems below would remain true under slight variations in the definition of W_n.) Then it will be shown, in Theorem 1, that the finite-dimensional distributions of W_n converge to those of the Wiener process W.

In general it is not sufficient, in order to prove weak convergence of the distribution of random functions such as W_n to that of W, to prove convergence of the finite-dimensional distributions. It will be

Received by the editors July 21, 1969.

AMS Subject Classifications. Primary 6030; Secondary 6066, 6215.

Key Words and Phrases. Invariance theorem, martingale, reversed martingale, U-statistic, weak convergence, sums of independent identically distributed random variables, tail sums of independent random variables.

[1] Research supported by Office of Naval Research, Contract Nonr 855(09).

shown in Theorem 2 that in the present case, in which reverse martingales appear, this condition is sufficient provided that $\phi(n) = [E(X_n^2)]^{-1}$. This rather striking result has an obvious analogue for forward martingales, which is stated as Theorem 3. There is an analogue of Theorem 1 also, but it will not be presented here.

The theorems are stated in §2, after the definitions and notation, and their proofs follow in §§3 and 4. §5 contains some applications.

2. **Definitions, notation, and statement of results.** The notation will be as in [4]. Thus on a fixed probability space $(\Omega, \mathcal{A}, P)$ we have a sequence of random variables X_n and a sequence of σ-fields $\mathcal{G}_n$; in both cases, $n \geqq 1$. Then for all n: $\mathcal{G}_n \supset \mathcal{G}_{n+1}$; X_n is measurable with respect to $\mathcal{G}_n$; $E|X_n| < \infty$; $E(X_n | \mathcal{G}_{n+1}) = X_{n+1}$ with probability 1. We suppose $X_n \to 0$ with probability 1, and write

(4) $Y_n = X_n - X_{n+1}$,

(5) $\sigma_n^2 = E(Y_n^2 | \mathcal{G}_{n+1})$,

and

(6) $s_n^2 = \sum_{r=n}^{\infty} \sigma_r^2$.

Then it follows that

(7) $E(Y_n | \mathcal{G}_{n+1}) = 0$,

(8) $X_n = \sum_{r=n}^{\infty} Y_r$,

and

(9) $EX_n^2 = \sum_{r=n}^{\infty} EY_r^2 = Es_n^2$.

Now we may state the results: the notation $W_n \Rightarrow W$ means that the distribution of the random function W_n converges weakly to that of the Wiener process W (cf. Billingsley [1, p. 7]).

THEOREM 1. *If* (a) $\phi(n)s_n^2 \to 1$, *for some function* ϕ *satisfying*

$$\phi(n+1) \geqq \phi(n) \quad \textit{and} \quad \phi(n+1)/\phi(n) \to 1,$$

and (b) *either*

$$\frac{1}{s_n^2} \sum_{r=n}^{\infty} Y_r^2 \to 1$$

or

$$\phi(n) E(s_n^2) \to 1$$

and

$$\frac{1}{s_n^2} \sum_{r=n}^{\infty} E[Y_r^2 I(|Y_r| > \epsilon s_n) | \mathcal{G}_{r+1}] \to 0 \quad \textit{for all } \epsilon > 0,$$

all convergence being with probability 1, *and if* $W_n(t)$ *is defined by* (1),

(2) *and* (3), *then the finite-dimensional distributions of* $W_n(t)$ *converge to those of* $W(t)$.

REMARK. In [4], some stronger conditions are given which imply the conditions required here and which are probably simpler to verify in practice.

THEOREM 2. *If* X_n *is a reversed martingale, and if the finite-dimensional distributions of* W_n, *as defined by* (1), (2), *and* (3) *with* $\phi(n) = [E(X_n^2)]^{-1}$, *converge to those of* W, *then* $W_n \Rightarrow W$.

For the next theorem, suppose $\{X_n : n \geqq 1\}$ is a forward martingale, and for each n define a random function $W_n(t)$ for $0 \leqq t \leqq 1$ by

(10 $W_n(0) = 0$;

(11) $W_n(t_k^n) = \phi(n)^{1/2}\ X_k$, where $t_k^n = \phi(n)\{\phi(k)\}^{-1}$, for $k \leqq n$;

(12) $W_n(t) = W_n(t_k^n)$, if $t_k^n < t < t_{k+1}^n$.

THEOREM 3. *If* X_n *is a (forward) martingale, and if the finite-dimensional distributions of* W_n, *as defined by* (10), (11) *and* (12) *with* $\phi(n) = [E(X_n^2)]^{-1}$, *converge to those of* W, *then* $W_n \Rightarrow W$.

Although Theorem 3 would appear to be useful, there seems less need for an analogue to Theorem 1; in any case, none will be given here. The possibility of embedding a forward martingale in a Wiener process, noted by Strassen [6], allows an approach to convergence problems (see Theorem 4.4 of [6]) which is capable of giving much more delicate results. Now one may speculate that a reversed martingale might similarly be embedded: presumably a sequence of stopping-times $0 \leqq \cdots \leqq T_n \leqq \cdots \leqq T_1$ would be constructed in such a way that $\{W(T_n)\}$ has the same distribution as $\{X_n\}$. However, although it is very plausible that such an embedding is possible, and that if it were it would allow an elegant treatment of the present results, it appears that a proof would be difficult, and none has so far been given.

We are therefore thrown back on to more direct methods, of the type described in the recent monograph by Billingsley [1]. A weak convergence result is given there (in §23) for (forward) martingales, but the proof leans heavily on the assumed stationarity, and our proofs will return to first principles.

3. **Proof of Theorem 1.** To prove convergence of the finite-dimensional distributions is rather straightforward, most of the work having already been done in [4]; in view of this, we shall assume familiarity with its contents. We shall consider the joint distribution of $W_n(a)$

and $W_n(b)$ only, where $0<a<b\leqq 1$, but the joint distribution for three or more times can be treated quite similarly. In [4], equation (17) onwards, various quantities are defined as functions of a single variable t; we shall need to deal simultaneously with two values at and bt, since our first aim is to prove that the distribution of $(t^{-1/2}Z_{at}, t^{-1/2}Z_{bt})$ converges weakly to that of $(W(a), W(b))$. We shall, therefore, have two sets of quantities which will be denoted by m_{at}, m_{bt}, $\tilde{Y}_{a,k}$, $\tilde{Y}_{b,k}$ and so on: notice that $m_{at}\geqq m_{bt}$, so that $\tilde{Y}_{ak}=\tilde{Y}_{bk}$ if $m_{at}<k$, and that we have $\sum\tilde{\sigma}_{ak}^2=at$, $\sum\tilde{\sigma}_{bk}^2=bt$. In the definition of η_n, a sequence of independent Normal random variables is needed: we choose the *same* sequence for both η_{an} and η_{bn}. It will be shown that as $t\to 0$

$$(13)\quad E[\exp(iua^{1/2}\eta_{a1}+ivb^{1/2}\eta_{b1})]-E[\exp(iua^{1/2}\eta_{a\infty}+ivb^{1/2}\eta_{b\infty})]\to 0:$$

the first term of (13) is the characteristic function of $(t^{-1/2}Z_{at}, t^{-1/2}Z_{bt})$ and the second term tends, as $t\to 0$, to that of $(W(a), W(b))$. (The means and variances of the second term are correct for all t, and the distribution is Normal, but the covariance will not in general be exactly equal to a for finite t; however, convergence to the correct value as $t\to 0$ is a consequence of the fact that $\sigma_n^2/s_n^2\to 0$.) Now following the same argument as in [4], we find

$$(14)\quad \begin{aligned} &\left| E(\exp(iua^{1/2}\eta_{an}+ivb^{1/2}\eta_{bn})-\exp(iua^{1/2}\eta_{an+1}+ivb^{1/2}\eta_{bn+1}))\right| \\ &\quad\leqq E\left| E\{(\exp(it^{-1/2}(u\tilde{Y}_{an}+v\tilde{Y}_{bn})) \right. \\ &\qquad\qquad \left. -\exp(it^{-1/2}\xi_n(u\tilde{\sigma}_{an}+v\tilde{\sigma}_{bn})))\mid \mathcal{J}_{n+1}\}\right| \\ &\quad\leqq E\left\{\frac{(|u|+|v|)^2}{t}\tilde{Y}_{bn}^2 g((|u|+|v|)t^{-1/2}|\tilde{Y}_{bn}|)\right. \\ &\qquad\qquad \left.+\frac{(|u|+|v|)^2}{t}\tilde{\sigma}_{bn}^2 h((|u|+|v|)t^{-1/2}\tilde{\sigma}_{bn})\right\} \end{aligned}$$

on using the fact that $|\tilde{Y}_{an}|\leqq|\tilde{Y}_{bn}|$, $\tilde{\sigma}_{an}^2\leqq\tilde{\sigma}_{bn}^2$. (Note that (29) of [4] contains two misprints: an exponent is obviously misplaced, and the interior expectation on the right-hand side should be conditional on $\mathcal{J}_{n+1}$.) As the bound obtained in (14) is essentially the same as that entering into (30) of [4], this first part of the proof is complete.

Next we shall show that for fixed a $(0<a\leqq 1)$ $W_n(a)-t_n^{-1/2}Z_{at_n}\to 0$ in probability, where $t_n=\phi(n)^{-1}$; from this and the previous results will follow convergence of the finite-dimensional distributions. Observe that $W_n(a)=\phi(n)^{1/2}X_{k(n)}$, where $k(n)$ is an integer such that

$$(15)\quad \phi(n)/\phi(k(n)+1)<a\leqq\phi(n)/\phi(k(n)).$$

If we write $m_{at_n} = m_{an}$ for convenience, it follows as in (45) of [4] that

$$\phi(m_{an})/\phi(k(n)) \to 1 \quad \text{with probability 1,}$$

and then the argument used to prove (47) of [4] (in which the suffix of Z should be t_n) will complete the proof.

4. **Proof of Theorems 2 and 3.** According to Theorem 15.1 of Billingsley [1], we have only to show that the sequence W_n is tight, and we can conveniently separate some parts of the proof into preliminary lemmas.

LEMMA 1. *If the finite-dimensional distributions of W_n converge to those of W, then*

$$\sup_{k \geq n} \left| \frac{\phi(n)}{\phi(k)} - \frac{\phi(n)}{\phi(k+1)} \right| \to 0 \quad \textit{as } n \to \infty.$$

PROOF. It is clearly sufficient to show that $\phi(n+1)/\phi(n) \to 1$. Suppose the contrary: then along some subsequence we have $\phi(n)/\phi(n+1) < 1-\epsilon$ for some $\epsilon > 0$, so that $W_n(1-\epsilon) = W_n(1)$. Thus along this subsequence $P[W_n(1-\epsilon) = W_n(1)] = 1$, which contradicts the hypothesis of the lemma.

LEMMA 2. *If for some random variable X, $E(X^2) < \infty$, and h is a function with values in $[0, 1]$ then*

$$E[X^2 h(X)] \leqq E[X^2 \lambda(X)]$$

where

$$\begin{aligned} 0 \leqq \lambda(x) &\leqq 1, \quad \textit{for all } x, \\ \lambda(x) &= 0, \quad |x| < \alpha(h), \\ &= 1, \quad |x| > \alpha(h), \end{aligned}$$

and $E[\lambda(X)] = E[h(X)]$.

PROOF. This is an immediate consequence of the Neyman Pearson Lemma. The constant $\alpha(h)$ does of course depend on the distribution of X.

We may now begin the main part of the proof of Theorem 2. We shall apply Theorem 15.5 of [1], proving the second condition by applying the Corollary to Theorem 8.3; the first condition is already satisfied. It is sufficient to show that given ϵ, η, r may be chosen to make

$$(16) \qquad p = \sum_1^r P\left[\sup_{(i-1)/r \leqq s \leqq i/r} \left| W_n(s) - W_n\left(\frac{i-1}{r}\right) \right| \geqq \epsilon \right] < \eta.$$

Let us for convenience write

$$A_r(n, i) = \left\{ \sup_{(i-1)/r \leqq s \leqq i/r} \left| W_n(s) - W_n\left(\frac{i-1}{r}\right) \right| \geqq \epsilon \right\} \tag{17}$$

and

$$Z_r(i, n) = r^{1/2}\{W_n(i/r) - W_n((i-1)/r)\}. \tag{18}$$

Then for fixed i and r, $Z_r(i, n) \Rightarrow N(0, 1)$ by hypothesis, while from Lemma 1 it follows that $E[Z_r(i, n)^2] \to 1$, so that if $n \geqq n_r$ (to be chosen suitably) $E[Z_r(i, n)^2] \leqq 2$.

If we apply Theorem 3.2 in Chapter VII of Doob [**2**], recalling the definition of W_n, we find

$$\begin{aligned} P(A_r(n, i)) &\leqq \frac{1}{\epsilon^2 r} \int_{A_r(n,i)} Z_r(i, n)^2 dP \\ &\leqq \frac{1}{\epsilon^2 r} E[Z_r(i, n)^2] \\ &\leqq \frac{2}{\epsilon^2 r} \quad \text{if } n \geqq n_r. \end{aligned} \tag{19}$$

According to Lemma 2, we have

$$\begin{aligned} \int_{A_r(n,i)} Z_r(i, n)^2 dP &= E[Z_r(i, n)^2 P\{A_r(n, i) \mid Z_r(i, n)\}] \\ &\leqq E[Z_r(i, n)^2 \lambda(Z_r(i, n))] \end{aligned} \tag{20}$$

with $\lambda(x) = 0$ if $|x| < \alpha_r(i, n)$, 1 if $|x| > \alpha_r(i, n)$, and

$$E[\lambda(Z_r(i, n))] = P\{A_r(n, i)\}. \tag{21}$$

Next choose θ so that

$$\int_{|X| \geqq \theta} X^2 dP = \eta\epsilon^2/2 \tag{22}$$

when X is $N(0, 1)$, and r so that

$$2/\epsilon^2 r < \tfrac{1}{2} P[\,|X| \geqq \theta]. \tag{23}$$

It follows from (19), (21) and (23), since $Z_r(i, n)$ converges to $N(0, 1)$, that if n is sufficiently large $\theta \leqq \alpha_r(i, n)$, for $1 \leqq i \leqq r$. Then from (20) we find

$$\int_{A_r(n,i)} Z_r(i, n)^2 dP \leqq \int_{\{|Z_r| \geqq \theta\}} Z_r(i, n)^2 dP, \tag{24}$$

and because of (22), the convergence of $Z_r(i, n)$ to $N(0, 1)$, and the convergence of $E[Z_r(i, n)^2]$ to 1, the right-hand side of (24) is less than $\eta\epsilon^2$ for sufficiently large n. (Compare the argument in [**1**, p. 33].) But now (16) follows from (19).

The proof of Theorem 3 is almost identical, and will therefore not be given.

5. **Applications.** Apart from observing that Theorem 3 is especially easy to apply when X_n is the partial sum of independent random variables, and that in particular Donsker's original theorem follows from it, by using the Central Limit Theorem, we shall make no reference to forward martingales.

Three types of reversed martingale were considered in [**4**]:

(a) S_n/n, where S_n is the nth partial sum of a sequence of independent identically distributed random variables with mean 0 and variance $\sigma^2 < \infty$;

(b) U_n, the nth U-statistic formed from a given kernel with mean 0 and finite variance;

(c) $R_n = \sum_{r=n}^{\infty} c_r P_r$, where P_r are independent and identically distributed with mean 0 and variance σ^2, and $c_n^2 = o(\sum_{r=n}^{\infty} c_r^2)$.

(The notation is slightly different from [**4**].) Both S_n/n and R_n were shown in [**4**] to satisfy the conditions of the present Theorem 1, and weak convergence for the random functions W_n constructed as in (1), (2) and (3) then follows from Theorem 2. Weak convergence in the case of U_n also follows from Theorem 2, although it has not been shown that the conditions of Theorem 1 are satisfied, for convergence of finite-dimensional distributions follows easily from Hoeffding [**3**].

If we now apply Theorem 5.1 of [**1**] with some admissible function h, we obtain limit theorems about S_n/n, U_n, and R_n. The following are examples. (The variances of U_n and R_n are denoted by $v_U(n)$ and $v_R(n)$ respectively for brevity.)

(i) $h(x) = \sup\{x(t) : 0 \leqq t \leqq 1\}$.

THEOREM 4. *Each of*

$$P\left[\sup_{k \geqq n} S_k/k \geqq n^{-1/2} x\sigma\right], \qquad P\left[\sup_{k \geqq n} U_k \geqq v_U^{1/2}(n) x\right],$$

$$P\left[\sup_{k \geqq n} R_k \geqq v_R^{1/2}(n) x\right]$$

tends to

$$P\left[\sup_{0 \leqq t \leqq 1} W(t) \geqq x\right] = 2(1 - \Phi(x)) \quad \textit{as } n \to \infty,$$

for all $x>0$, *where* $\Phi(x)$ *is the distribution function of* $N(0, 1)$.

From this, Theorem 1 of [5] follows immediately and Theorem 3 is also easy to prove. Similar results may be obtained from $h(x) = \sup\{|x(t)| : 0 \leqq t \leqq 1\}$.

The functionals h used in the next two results are not almost everywhere continuous on D, and thus Theorem 2 as stated can not be applied.[2] They are, however, admissible if attention is restricted to C, and since as already observed Theorem 2 remains true if the definition of W_n is changed slightly, for example by requiring linear interpolation between successive points t_k^n for fixed n, the results are valid.

(ii) $h(x) = \sup\{t : t \in [0, 1] \text{ and } x(t) = 0\}$.

For a reversed martingale X_n, let us say that a 0-crossing takes place at i if the event

$$E_i = \{X_i = 0\} \cup \{X_{i-1} > 0 > X_i\} \cup \{X_{i-1} < 0 < X_i\}$$

occurs (cf. Billingsley [1] equation (11.28)). Define K_n with appropriate superfixes as the minimum $i \geqq n$ for which a 0-crossing takes place at i.

THEOREM 5. *Each of*

$$P[\overset{S}{K_n} \geqq \lambda n], \qquad P[v_U(\overset{U}{K_n})^{-1} \geqq \lambda v_U(n)^{-1}], \qquad P[v_R(\overset{R}{K_n})^{-1} \geqq \lambda v_R(n)^{-1}]$$

tends to $(2/\pi) \arcsin (1/\lambda)^{1/2}$ *as* $n \to \infty$.

(iii) $h(x) = \inf\{t : x(t) = \sup_{0 \leqq s \leqq 1} x(s)\}$.

The results will not be given in detail, but they determine the asymptotic distribution of the largest k for which $X_k = \sup_{j \geqq n} X_j$.

One may of course investigate the joint distributions of such quantities, and other choices of h may be made. Because of the nonuniform spacing of the t_k^n for a given n the results are sometimes intuitively less appealing than the corresponding ones for the partial sums derived from the Donsker principle: as an example the choice of h to be the Lebesgue measure of the set of times that $x(t)$ is positive may be cited.

REFERENCES

1. Patrick Billingsley, *Convergence of probability measures*, Wiley, New York, 1968. MR 38 #1718.

2. J. L. Doob, *Stochastic processes*, Wiley, New York, 1953. MR 15, 445.

[2] A referee has pointed out that Theorem 5 at least can be obtained from the stated version of Theorem 2 if h is redefined by

$$h(x) = \sup\{t : t \in [0, 1] \quad \text{and} \quad x(t)x(t+) \leqq 0\}.$$

3. W. Hoeffding, *A class of statistics with asymptotically normal distribution*, Ann. Math. Statist. **19** (1948), 293–325. MR **10,** 134.

4. R. M. Loynes, *The central limit theorem for backwards martingales*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **13** (1969), 1–8.

5. H. Robbins, D. Siegmund and J. Wendel, *The limiting distribution of the last time* $s_n \geqq n\epsilon$, Proc. Nat. Acad. Sci. U.S.A. **61** (1968), 1228–1230.

6. Volker Strassen, *Almost sure behavior of sums of independent random variables and martingales*, Proc. Fifth Berkeley Sympos. Math. Statist. Prob. (Berkeley, Calif., 1965/66) vol. II: Contributions to Probability Theory, part 1, Univ. of California Press, Berkeley, Calif., 1967, pp. 315–343. MR **35** #4969.

University of North Carolina, Chapel Hill, North Carolina 27514 and University of Cambridge, Cambridge, England