

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

AN INVARIANCE PRINCIPLE FOR REVERSED MARTINGALES

by

R. M. LOYNES¹

Department of Statistics

University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 631

JUNE 1969

¹ Work done while visiting from the University of Cambridge, England.

AN INVARIANCE PRINCIPLE FOR REVERSED MARTINGALES

by

R. M. LOYNES

University of North Carolina at Chapel Hill¹ and University of Cambridge

1. Introduction and Summary. It was shown in Loynes [4] that, under suitable conditions, if $\{X_n: n \geq 1\}$ is a reversed martingale there exists a function $\phi(n)$ such that the distribution of $\phi(n)^{\frac{1}{2}}X_n$ converges to $N(0,1)$ weakly as $n \rightarrow \infty$. (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 $0 \leq t \leq 1$ as follows:

- (1) $W_n(0) = 0$;
- (2) $W_n(t_k^n) = \phi(n)^{\frac{1}{2}}X_k$, where $t_k^n = \phi(n)\{\phi(k)\}^{-1}$, for $k \geq n$;
- (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 right: 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 shown in Theorem 2 that in the present case, in which reverse martingales appear, this condition is sufficient provided that $\phi(n) = E(X_n^2)$. This rather striking result has an obvious analogue for forward martin-

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

gales, 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 Section 2, after the definitions and notation, and their proofs follow in Sections 3 and 4. Section 5 contains some applications.

2. Definitions, Notation, and Statement of Results. The notation will be as in [4]. Thus on a fixed probability space (Ω, A, P) we have a sequence of random variables X_n and a sequence of σ -fields I_n ; in both cases, $n \geq 1$. Then for all n : $I_n \supset I_{n+1}$; X_n is measurable with respect to I_n ; $E|X_n| < \infty$; $E(X_n | I_{n+1}) = X_{n+1}$ with probability 1. We suppose $X_n \rightarrow 0$ with probability 1, and write

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

$$(5) \quad \sigma_n^2 = E(Y_n^2 | I_{n+1}),$$

and

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

Then it follows that

$$(7) \quad E(Y_n | I_{n+1}) = 0,$$

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

and

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

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) $\phi(n)s_n^2 \rightarrow 1$, for some function ϕ satisfying

$$\phi(n+1) \geq \phi(n) \quad \text{and} \quad \phi(n+1)/\phi(n) \rightarrow 1,$$

and (b) either

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

or

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

and

$$\frac{1}{s_n^2} \sum_{r=n}^{\infty} E[Y_r^2 I(|Y_r| > \varepsilon s_n) | I_{r+1}] \rightarrow 0$$

for all $\varepsilon > 0$,

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 reverse 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 $\{X_n; n \geq 1\}$ is a forward martingale, and for each n define a random function $W_n(t)$ for $0 \leq t \leq 1$ by

$$(10) \quad W_n(0) = 0;$$

$$(11) \quad W_n(t_k^n) = \phi(n)^{\frac{1}{2}} X_k, \quad \text{where } t_k^n = \phi(n)\{\phi(k)\}^{-1}, \quad \text{for } k \leq n;$$

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

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 reverse martingale might similarly be embedded: presumably a sequence of stopping-times $0 \leq \dots \leq T_n \leq \dots \leq T_1$ 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 $0 < a < b \leq 1$, 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 $(t^{-\frac{1}{2}}Z_{at}, t^{-\frac{1}{2}}Z_{bt})$ 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 $m_{at} \geq m_{bt}$, so that $\tilde{Y}_{ak} = \tilde{Y}_{bk}$ if $m_{at} < k$, and that we have $\Sigma \tilde{\sigma}_{ak}^2 = at$, $\Sigma \tilde{\sigma}_{bk}^2 = bt$. 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 $t \rightarrow 0$

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

the first term of (13) is the characteristic function of $(t^{-\frac{1}{2}}Z_{at}, t^{-\frac{1}{2}}Z_{bt})$ and the second term tends, as $t \rightarrow 0$, 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 $t \rightarrow 0$ is a consequence of the fact that $\sigma_n^2/s_n^2 \rightarrow 0$.) Now following the same argument as in [4], we find

$$\begin{aligned}
(14) \quad & |E(e^{iua^{\frac{1}{2}}n_{an}+ivb^{\frac{1}{2}}n_{bn}} - e^{iua^{\frac{1}{2}}n_{an+1}+ivb^{\frac{1}{2}}n_{bn+1}})| \\
& \leq E|E\{(e^{it^{-\frac{1}{2}}(u\tilde{Y}_{an}+v\tilde{Y}_{bn})} - e^{it^{-\frac{1}{2}}\xi_n(u\tilde{\sigma}_{an}+v\tilde{\sigma}_{bn})})|I_{n+1}\}| \\
& \leq E\left\{\frac{(|u|+|v|)^2}{t} \tilde{Y}_{bn}^2 g((|u|+|v|)t^{-\frac{1}{2}}|\tilde{Y}_{bn}|) \right. \\
& \quad \left. + \frac{(|u|+|v|)^2}{t} \tilde{\sigma}_{bn}^2 h((|u|+|v|)t^{-\frac{1}{2}}\tilde{\sigma}_{bn})\right\}
\end{aligned}$$

on using the fact that $|\tilde{Y}_{an}| \leq |\tilde{Y}_{bn}|$, $\tilde{\sigma}_{an}^2 \leq \tilde{\sigma}_{bn}^2$. (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 I_{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 ($0 < a \leq 1$) $W_n(a) - t^{-\frac{1}{2}}Z_{at_n} \rightarrow 0$ in probability, where $t_n = \phi(n)^{-1}$; from this and the previous results will follow convergence of the finite-dimensional distributions. Observe that $W_n(a) = \phi(n)^{\frac{1}{2}}X_{k(n)}$, where $k(n)$ is an integer such that

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

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

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

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

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

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

where

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

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

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

Let us for convenience write

$$(17) \quad A_r(n,i) = \left\{ \sup_{\frac{i-1}{r} \leq s \leq \frac{i}{r}} |W_n(s) - W_n\left(\frac{i-1}{r}\right)| \geq \epsilon \right\}$$

and

$$(18) \quad Z_r(i,n) = r^{\frac{1}{2}} \{W_n\left(\frac{i}{r}\right) - W_n\left(\frac{i-1}{r}\right)\}.$$

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

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

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

According to Lemma 2, we have

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

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

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

Next choose θ so that

$$(22) \quad \int_{|X| \geq \theta} X^2 dP = \frac{\eta \epsilon^2}{2}$$

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

$$(23) \quad \frac{2}{\epsilon^2 r} < \frac{1}{2} P[|X| \geq \theta].$$

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

$$(24) \quad \int_{A_r(n,i)} Z_r(i,n)^2 dP \leq \int_{\{|Z_r| \geq \theta\}} Z_r(i,n)^2 dP,$$

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 reverse martingale were considered in [4]:

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

$$(i) \quad h(x) = \sup\{x(t) : 0 \leq t \leq 1\}$$

THEOREM 4. Each of $P[\sup_{k \geq n} S_k/k \geq n^{-\frac{1}{2}} x \sigma]$,

$$P[\sup_{k \geq n} U_k \geq v_U^{\frac{1}{2}}(n)x], \quad P[\sup_{k \geq n} R_k \geq v_R^{\frac{1}{2}}(n)x] \quad \text{tends to}$$

$$P[\sup_{0 \leq t \leq 1} W(t) \geq x] = 2(1 - \Phi(x)) \quad \text{as } n \rightarrow \infty, \quad \text{for all } x > 0, \quad \text{where}$$

$\Phi(x)$ 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 $h(x) = \sup\{|x(t)|: 0 \leq t \leq 1\}$.

The next two results do not follow from Theorem 2 as stated, since the functionals h are not almost everywhere continuous on D . 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 reverse 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 $i \geq n$ for which a 0-crossing takes place at i .

THEOREM 5. Each of $P[K_n^S \geq \lambda n]$, $P[v_U(K_n^U)^{-1} \geq \lambda v_U(n)^{-1}]$, $P[v_R(K_n^R)^{-1} \geq v_R(n)^{-1}]$ tends to $(2/\pi) \arcsin \sqrt{1/\lambda}$ as $n \rightarrow \infty$.

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

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

One may of course investigate the joint distributions of such quantities, and other choices of h may be made. Because of the non -

uniform 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] Billingsley, Patrick. Convergence of Probability Measures.
New York: Wiley 1968.
- [2] Doob, J.L. Stochastic Processes.
New York: Wiley 1953.
- [3] Hoeffding, W. A class of statistics with asymptotically Normal distribution. Ann. Math. Statist. 19 (1948), 293-325.
- [4] Loynes, R.M. The Central Limit Theorem for backwards martingales.
Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 13 (1969), 1-8.
- [5] Robbins, H., Siegmund, D. and Wendel, J. The limiting distribution of the last time $s_n \geq n\epsilon$. Proc. Nat. Acad. Sci. 61 (1968), 1228-1230.
- [6] Strassen, Volker. Almost sure behavior of sums of independent random variables and martingales. Proceedings of Fifth Berkeley Symposium on Mathematical Statistics and Probability 2 (1), 315-343.

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)		2a. REPORT SECURITY CLASSIFICATION	
Department of Statistics The University of North Carolina at Chapel Hill Chapel Hill, North Carolina 27514		UNCLASSIFIED	
3. REPORT TITLE		2b. GROUP	
AN INVARIANCE PRINCIPLE FOR REVERSED MARTINGALES			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)			
TECHNICAL REPORT			
5. AUTHOR(S) (First name, middle initial, last name)			
Robert M. Loynes			
6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS	
June 1969	12	6	
8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)		
NONR-855-(09)	INSTITUTE OF STATISTICS MIMEO SERIES		
b. PROJECT NO.	NUMBER 631		
RR-003-05-01	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)		
c.			
d.			
10. DISTRIBUTION STATEMENT			
The Distribution of this Report is Unlimited			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY	
		Logistics and Mathematical Statistics Branch Office of Naval Research Washington, D.C. 20360	
13. ABSTRACT			
<p>Let $\{X_n : n \geq 1\}$ be a martingale, and for each n construct a random function W_n by plotting X_k at $t = EX_k^2 / EX_n^2$, ($1 \leq k \leq n$), and scaling.</p> <p>If the finite-dimensional distributions of W_n converge to those of the Wiener process W, then $W_n \Rightarrow W$. Analogously, if $\{X_n : n \geq 1\}$ is a reverse martingale, construct W_n by plotting X_k ($k \geq n$) at appropriate points; the same result holds. Sufficient conditions for the required convergence, and applications, are given for the reversed martingale case.</p>			