The Annals of Probability
1978, Vol. 6, No. 2, 335-340

STOPPING TIMES AND TIGHTNESS[1]

By David Aldous

University of Cambridge

A sufficient condition for the tightness of a sequence of stochastic processes is given in terms of their behavior after stopping times. As an application, the conditions for McLeish's invariance principle for martingales are weakened.

Let D be the space of functions on $[0, 1]$ with discontinuities of at most the first kind, with the Skorokhod J_1 topology (see [1] for the theory of weak convergence in D and the definitions of w and w').

Let $\{X_n\}$ be a sequence of random elements of D. Let $\{\tau_n, \delta_n\}$ be such that

(1) (i) for each n, τ_n is a stopping time on the process $\{X_n(t); 0 \leqq t \leqq 1\}$, with respect to the natural σ-fields, and τ_n takes only finitely many values;

(ii) for each n, δ_n is a constant, $0 \leqq \delta_n \leqq 1$, and $\delta_n \to 0$ as $n \to \infty$.

We require τ_n to take values in $[0, 1]$, but it is technically convenient to regard each $f \in D$ as extended to $[0, 2]$ by putting $f(t) = f(1)$, $1 \leqq t \leqq 2$: this enables us to write $X_n(\tau_n + \delta_n)$ instead of $X_n(\min(1, \tau_n + \delta_n))$. For $f \in D$, let $J(f)$ denote the maximum of the jumps $|f(t) - f(t-)|$.

We are interested in the following condition on $\{X_n\}$:

$$\text{(A)} \qquad X_n(\tau_n + \delta_n) - X_n(\tau_n) \to_p 0$$

for each sequence $\{\tau_n, \delta_n\}$ satisfying (1).

Observe that if (A) is satisfied, then it remains true even when $\{\tau_n\}$ are not required to take finitely many values, by approximating from the right.

The main result in this paper is a sufficient condition for tightness, whose proof we defer.

Theorem 1. *Suppose that $\{X_n\}$ satisfies (A), and that either*

(2) $\{X_n(0)\}$ *and* $\{J(X_n)\}$ *are tight on the line; or*

(3) $\{X_n(t)\}$ *is tight on the line, for each* $t \in [0, 1]$.

Then $\{X_n\}$ is tight in D.

Hypotheses (2) and (3) are certainly necessary for tightness, but (A) is not:

Received September 1, 1976; revised April 19, 1977.

[1] This research was supported by an S.R.C. studentship.

AMS 1970 *subject classifications*. Primary 60B10; Secondary 60F05.

Key words and phrases. Stopping time, tightness, weak convergence, martingale, invariance principle.

consider

$$X_n(t) = 0 \quad 0 \leqq t < \tfrac{1}{2}$$
$$= 1 \quad \tfrac{1}{2} \leqq t \leqq 1 .$$

The essential point of Theorem 1 is that we only have to look at the behavior of the processes after stopping times. If the τ_n were allowed to be arbitrary [0, 1]-valued random variables, then the hypothesis is easily seen to be equivalent to

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P(w(X_n, \delta) > \varepsilon) = 0 , \qquad \text{for each} \quad \varepsilon > 0 .$$

And this, together with the hypothesis $\{X_n(0)\}$ is tight, is equivalent to the assertion that $\{X_n\}$ is tight and each weak limit has a.s. continuous sample paths. This establishes the following result.

COROLLARY 1. *Suppose that X_0 has* a.s. *continuous sample paths, and that the finite-dimensional distributions of $\{X_n\}$ converge to those of X_0. Then $X_n \to_{\mathscr{D}} X_0$ if and only if $\{X_n\}$ satisfies* (A).

These results and their proof are somewhat similar to those of Billingsley [2]. Indeed, let us write $M(X_n) = \sup_t |X_n(t)|$, and let $\alpha_n(\lambda, \varepsilon, \delta)$ be the smallest number such that

$$P(|X_n(s) - X_n(t_m)| > \varepsilon \,\|\, X_n(t_1), \cdots, X_n(t_m))$$
$$\leqq \alpha_n(\lambda, \varepsilon, \delta) \quad \text{a.s.} \quad \text{on the set} \quad \{\max_i |X_n(t_i)| \leqq \lambda\}$$

whenever $0 \leqq t_1 \leqq \cdots \leqq t_m < s \leqq 1$ and $s - t_m \leqq \delta$. Suppose $\{\tau_n, \delta_n\}$ satisfy (1): then a simple argument shows that

$$P(|X_n(\tau_n + \delta_n) - X_n(\tau_n)| > \varepsilon) \leqq \alpha_n(\lambda, \varepsilon, \delta_n) + P(M(X_n) > \lambda) .$$

So we can deduce from Theorem 1 the following improvement of [2], Theorem 1. (This result is also mentioned in [3].)

COROLLARY 2. *Suppose that $\{M(X_n)\}$ is tight on the line, and that for each $\varepsilon > 0$, $\lambda < \infty$*

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \alpha_n(\lambda, \varepsilon, \delta) = 0 .$$

Then $\{X_n\}$ is tight in D.

This result in [2] has been applied to sequences of Markov processes. However, that would seem to be the only circumstance in which its hypothesis about conditional distributions is satisfied. It seems plausible that hypothesis (A) is satisfied in more general situations where we have control over the behavior of the processes after stopping times. As an example, let us show how the conditions of McLeish [4] for the invariance principle for martingale difference arrays can be weakened slightly. We use the notation of [4]: $\{X_{n,i}\}$ is a triangular array of random variables adapted to σ-fields $\{\mathscr{F}_{n,i}\}$, and $\{k_n(t)\}$ is a sequence of integer-valued, nondecreasing, right-continuous functions such that $k_n(0) = 0$. Let

$W_n(t) = \sum_{i=1}^{k_n(t)} X_{n,i}$, and let W denote standard Brownian motion on D. Theorem 2 is a restatement of [4], 2.3, trivially modified. Theorem 3 is a slight improvement of [4], 3.2.

THEOREM 2. *Let $s^2 < \infty$. Let $\{Y_{n,i}\}$ be a martingale difference array such that*

(4) $$\max_i |Y_{n,i}| \to_p 0$$

(5) $$\sup_n E(\max_i |Y_{n,i}|^2) < \infty$$

(6) $$\sum_i Y_{n,i}^2 \to_p s^2 .$$

Then $\sum_i Y_{n,i} \to_{\mathscr{D}} N(0, s^2)$.

THEOREM 3. *Let $\{X_{n,i}\}$ be a martingale difference array satisfying* (5) *and*

(7) $$\sum_{i=1}^{k_n(t)} X_{n,i}^2 \to_p t , \qquad \text{for each} \quad t \in [0, 1] .$$

Then $W_n \to_{\mathscr{D}} W$.

PROOF. Let us first show that

(8) $$J(W_n) = \max_i |X_{n,i}| \to_p 0 .$$

Let $B_{n,q} = \{\omega : |\sum_{i=1}^{k_n(j/q)} X_{n,i}^2(\omega) - j/q| \leqq 1/q$ for each $j = 1, \cdots, q\}$. Then

(9) $$\lim_{n\to\infty} P(B_{n,q}) = 1 \qquad \text{for each} \quad q , \quad \text{by (7).}$$

And for $\omega \in B_{n,q}$,

(10) $$\sum_{k_n(j-1/q)}^{k_n(j/q)} X_{n,i}^2(\omega) \leqq 3/q , \qquad \text{for each} \quad j = 1, \cdots, q .$$

Hence $\max_i |X_{n,i}|^2 \leqq 3/q$ on $B_{n,q}$, and (8) follows.

We will now use Theorem 1 to establish the tightness of $\{W_n\}$. Let $\{\tau_n, \delta_n\}$ satisfy (1). Then

(11) $$\begin{aligned} W_n(\tau_n + \delta_n) - W_n(\tau_n) &= \textstyle\sum_i X_{n,i} I(A_{n,i}) \\ &= \textstyle\sum_i Y_{n,i} \quad \text{say,} \end{aligned}$$

where $A_{n,i} = \{k_n(\tau_n) < i \leqq k_n(\tau_n + \delta_n)\}$. We assert

(12) $$\{k_n(\tau_n) < i\} \in \mathscr{F}_{n,i-1} .$$

For if $r \leqq s$ then

$$\{k_n(s) < i\} \cap \{W_n(r) < a\} = \textstyle\bigcup_{j=0}^{i-1} \{k_n(s) = j\} \cap \{\sum_1^{k_n(r)} X_{n,i} < a\} \in \mathscr{F}_{n,i-1}$$

and so

$$\{k_n(s) < i\} \cap \mathscr{F}(W_n(r) : 0 \leqq r \leqq s) \subset \mathscr{F}_{n,i-1}$$

and so

$$\{k_n(s) < i\} \cap \{\tau_n = s\} \in \mathscr{F}_{n,i-1}$$

from which (12) follows. Using a similar argument,

$$\{k_n(\tau_n + \delta_n) < i\} \in \mathscr{F}_{n,i-1}$$

and hence $A_{n,i} \in \mathscr{F}_{n,i-1}$.

So $\{Y_{n,i}\}$ is a martingale difference array, which certainly satisfies (4) and (5). Now if $\omega \in B_{n,q}$ and $\delta_n < 1/q$, then (10) shows that $\sum_i Y^2_{n,i}(\omega) \leqq 6/q$, and so it follows from (9) that

$$\sum_i Y^2_{n,i} \to_p 0 .$$

Applying Theorem 2, we see that $\sum_i Y_{n,i} \to_p 0$. Now (11) shows that (A) is satisfied, and then (8) and Theorem 1 establish the tightness of $\{W_n\}$. The convergence of the finite-dimensional distributions of W_n to those of W follows from Theorem 2 and the usual Cramér–Wold argument.

REMARK. The reader will observe that Theorem 1 leads to a tightness argument quite different from the usual ones involving maximal inequalities and moment conditions.

McLeish [4] remarks that his results extend to the situation where $k_n(t)$ is, for each fixed t, a stopping time on $\{\mathscr{F}_{n,i}\}$. The proof of Theorem 3 goes over unchanged in this situation.

PROOF OF THEOREM 1. Hypothesis (A) is equivalent to the assertion that, given $\varepsilon > 0$, there exists $\delta > 0$ and n_0 such that, for $n > n_0$,

$$P(|X_n(\tau + \delta') - X_n(\tau)| \geqq \varepsilon) < \varepsilon \tag{13}$$

for each $0 \leqq \delta' \leqq 2\delta$, and each stopping time τ on X_n. So given $\varepsilon > 0$, choose δ and n_0 as above and consider these quantities fixed. Let $q > 2/\delta$ be a fixed integer. Then similarly there exist $\sigma > 0$, $n_1 \geqq n_0$ such that, for $n \geqq n_1$,

$$P(|X_n(\tau + \delta') - X_n(\tau)| \geqq \varepsilon) < \varepsilon/q \tag{14}$$

for each $0 \leqq \delta' \leqq 2\sigma$, and each stopping time τ on X_n.

As in [2], define stopping times $\{T_{n,i}\}$, $i = 0, \cdots, q$, by $T_{n,0} = 0$

$$\begin{aligned} T_{n,i+1} &= \min\{t\colon T_{n,i} < t < 1, |X_n(t) - X_n(T_{n,i})| \geqq 2\varepsilon\} \\ &= 2 \qquad \text{if no such} \quad t \quad \text{exists.} \end{aligned}$$

We shall prove later that, for $n \geqq n_1$,

$$P(T_{n,q} < 1) \leqq 16\varepsilon \tag{15}$$

$$P(T_{n,i} < \min(1, T_{n,i-1} + \sigma), \text{ for some } i = 1, \cdots, q) \leqq 8\varepsilon . \tag{16}$$

Assume (15) and (16) for the moment. Think of $\{X_n\}$ as being extended to $D[0, 2]$ as described earlier. Write $w'(x, \sigma)$ for the modulus $w_x'(\sigma)$ of [1], page 110, modified for $D[0, 2]$ by taking the infimum over $\{t_i\}$ satisfying $0 = t_0 < t_1 < \cdots < t_r = 2$, $t_i \geqq t_{i-1} + \sigma$. Then it follows from (15), (16) and the definition of $\{T_{n,i}\}$ that

$$P(w'(X_n, \sigma) > 4\varepsilon) \leqq 24\varepsilon , \quad n \geqq n_1 . \tag{17}$$

Now suppose that hypothesis (2) is in force. Then we can choose λ such that, for each n,

$$P(|X_n(0)| \geqq \lambda) < \varepsilon$$
$$P(J(X_n) \geqq \lambda) < \varepsilon .$$

Then from (15) and the definition of $\{T_{n,i}\}$ it follows that

(18) $$P(\sup_t |X_n(t)| \geqq (q+1)\lambda + 2q\varepsilon) \leqq 18\varepsilon\,, \quad n > n_1\,.$$

Alternatively, suppose that hypothesis (3) holds. Let S_λ denote the set of t such that

(19) $$\sup_n P(|X_n(t)| \geqq \lambda) \leqq \varepsilon\delta\,.$$

By (3), $\bigcup_\lambda S_\lambda = [0, 2]$. Choose λ so large that $[0, 2]\backslash S_\lambda$ has Lebesgue measure at most $\varepsilon\delta$. We shall prove later that, for $n \geqq n_1$,

(20) $$P(\sup_t |X_n(t)| \geqq \lambda + \varepsilon) \leqq 3\varepsilon\,.$$

Now we can apply [1] Theorem 15.2 to (17), together with (18) or (20), and deduce that $\{X_n\}$ is tight in $D[0, 2]$. To show that $\{X_n\}$ is also tight when considered as elements of $D = D[0, 1]$ is straightforward. Suppose $X_{j_n} \to_{\mathscr{D}} X_\infty$ in $D[0, 2]$: to prove convergence in D it is sufficient to show that X_∞ is continuous in probability at 1, and this is true because, from (13),

$$\limsup_{n\to\infty} P(|X_n(s) - X_n(t)| \geqq \varepsilon) \leqq \varepsilon \qquad \text{if} \quad |s - t| < 2\delta\,.$$

Thus the proof of Theorem 1 has been reduced to the proof of (15), (16), and (20).

Fix $n \geqq n_1$, and for typographical convenience drop subscripts n from $T_{n,i}$ and X_n. Let $\mathscr{F}$ be the σ-field generated by X. Let θ be a random variable distributed uniformly on $[0, 2\delta]$ independent of $\mathscr{F}$.

Temporarily fix $f \in D$ and $0 \leqq t_1 \leqq t_2 \leqq 1$. Suppose

(21) (i) $$t_2 - t_1 < \delta$$

(ii) $$P(|f(t_j + \theta) - f(t_j)| < \varepsilon) > \tfrac{3}{4}\,, \qquad j = 1, 2\,.$$

Then there exists $\theta_0 \in [t_1 + \delta, t_1 + 2\delta]$ such that

$$|f(\theta_0) - f(t_j)| < \varepsilon\,, \qquad j = 1, 2\,,$$

and so $|f(t_2) - f(t_1)| < 2\varepsilon$. In other words, the set

$$\{(f, t_1, t_2)\colon |f(t_2) - f(t_1)| \geqq 2\varepsilon \text{ and } t_2 < t_1 + \delta\}$$

is contained in the set

$$\{(f, t_1, t_2)\colon P(|f(t_j + \theta) - f(t_j)| \geqq \varepsilon) \geqq \tfrac{1}{4} \text{ for } j = 1 \text{ or } 2\}\,.$$

Hence for $i = 1, \cdots, q$

(22) $$P(|X(T_i) - X(T_{i-1})| \geqq 2\varepsilon, T_i < T_{i-1} + \delta) \leqq P(A_i) + P(A_{i-1})$$

where $A_i = \{\omega\colon P(|X(T_i + \theta) - X(T_i)| \geqq \varepsilon \,\|\, \mathscr{F}) \geqq \tfrac{1}{4}\}$.

Now $P(A_i) \leqq 4P(|X(T_i + \theta) - X(T_i)| \geqq \varepsilon)$. And from (13),

(23) $$P(|X(\tau + \theta) - X(\tau)| \geqq \varepsilon) < \varepsilon$$

for each stopping time τ on X. Also, $|X(T_i) - X(T_{i-1})| \geqq 2\varepsilon$ on $\{T_i < 1\}$ by definition. So (22) implies that

$$P(T_i < 1,\, T_i < T_{i-1} + \delta) \leqq 8\varepsilon\,, \qquad i = 1, \cdots, q\,. \tag{24}$$

A similar argument using (14) shows that

$$P(T_i < 1,\, T_i < T_{i-1} + \sigma) \leqq 8\varepsilon/q\,, \qquad i = 1, \cdots, q\,. \tag{25}$$

Now (16) follows immediately from (25). And

$$\begin{aligned} E(T_i - T_{i-1} \mid T_q < 1) &\geqq \delta P(T_i - T_{i-1} \geqq \delta \mid T_q < 1) \\ &\geqq \delta\{1 - P(T_i - T_{i-1} < \delta,\, T_q < 1)/P(T_q < 1)\} \\ &\geqq \delta\{1 - 8\varepsilon/P(T_q < 1)\} \end{aligned}$$

by (24). So $1 \geqq E(T_q \mid T_q < 1) = \sum E(T_i - T_{i-1} \mid T_q < 1) \geqq q\delta\{1 - 8\varepsilon/P(T_q < 1)\}$. Since we chose $q > 2/\delta$, (15) follows.

It remains only to prove (20). Let ϕ be a random variable distributed uniformly on [0, 1] independent of $\mathscr{F}$. Suppose that $0 \leqq t \leqq 1$, $f \in D$, and A is a measurable subset of the line. Then by considering densities,

$$P(f(t + \theta) \in A) \leqq P(f(2\phi) \in A)/\delta\,. \tag{26}$$

So if τ is a stopping time on X, then

$$\begin{aligned} P(|X(\tau)| \geqq \lambda + \varepsilon) &\leqq \varepsilon + P(|X(\tau + \theta)| \geqq \lambda) \quad \text{by (23);} \\ &\leqq \varepsilon + P(|X(2\phi)| \geqq \lambda)/\delta \quad \text{by (26);} \\ &\leqq \varepsilon + \{\varepsilon\delta + P(\phi \notin S_\lambda)\}/\delta \quad \text{by (19);} \\ &\leqq 3\varepsilon \quad \text{by choice of } \lambda\,. \end{aligned}$$

And (20) follows by considering

$$\begin{aligned} \tau &= \min\{t : |X(t)| \geqq \lambda + \varepsilon\} \\ &= 1 \quad \text{if no such } t \text{ exists.} \end{aligned}$$

REFERENCES

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

[2] BILLINGSLEY, P. (1974). Conditional distributions and tightness. *Ann. Probability* **2** 480–485.

[3] MACKEVICIUS, V. (1974). On the weak convergence of stochastic processes on the spaces $D_{[0,\infty)}X$. *Litovsk. Mat. Sb.* **14** 117–121.

[4] MCLEISH, D. L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probability* **2** 620–628.

THE STATISTICAL LABORATORY
UNIVERSITY OF CAMBRIDGE
16 MILL LANE
CAMBRIDGE, CB2 1SB
ENGLAND