SOME ONE-SIDED STOPPING RULES[1]

By David O. Siegmund[2]

Columbia University

1. Introduction and summary. Suppose that $x_1, x_2, \cdots$ are random variables with means $\mu_1, \mu_2, \cdots$ such that

$$n^{-1}(\mu_1 + \cdots + \mu_n) \to \mu \qquad (0 < \mu < \infty), \tag{1}$$

and let $s_n = x_1 + \cdots + x_n$ $(n \geqq 1)$. For any positive non-decreasing and eventually concave function f defined on the positive real numbers and $c > 0$ let

$$\begin{aligned} \tau = \tau(c) &= \text{first} \quad n \geqq 1 \quad \text{such that} \quad s_n > cf(n) \\ &= \infty \quad \text{if no such} \quad n \quad \text{exists.} \end{aligned}$$

We are interested in finding conditions on f and on the joint distribution of (x_n) which insure that if $\lambda = \lambda(c)$ is defined by

$$\mu\lambda = cf(\lambda) \tag{2}$$

(since we shall assume below that $f(n) = o(n)$, $\lambda(c)$ is unique for sufficiently large c), then

$$\lim_{c\to\infty}\lambda^{-1}E\tau = 1. \tag{3}$$

The elementary renewal theorem states that (3) holds when $x_1, x_2, \cdots$ are iid non-negative random variables and $f(n) \equiv 1$. Chow and Robbins [3] have obtained generalizations of this result to the case in which (x_n) are dependent or non-identically distributed. The case in which f is not constant has been discussed in [2], [4], and [11]. We shall assume that $x_1, x_2, \cdots$ are independent and prove the

Theorem. *Let* (x_n), (μ_n), f, τ, c, λ *be as above, and suppose that for some* $\alpha\varepsilon(0, 1)$ *and* L *slowly varying*

$$f(n) \sim n^{\alpha}L(n). \tag{4}$$

If for every $\epsilon > 0$

$$n^{-1}\sum_1^n \int_{\{x_i-\mu_i>i\epsilon\}} (x_i - \mu_i) \to 0, \qquad n \to \infty, \tag{5}$$

then

$$\limsup \lambda^{-1} E\,\tau \leqq 1. \tag{6}$$

If in addition to (5) $s_n/n \to \mu$ a.s. *or if for every* $\epsilon > 0$

$$n^{-1}\sum_1^n \int_{\{|x_i-\mu_i|>n\epsilon\}} |x_i - \mu_i| \to 0, \qquad n \to \infty, \tag{7}$$

$$\sup_n E\,|x_n - \mu_n| \leqq K < \infty, \tag{8}$$

then (3) *holds.*

Received 8 May 1967.

[1] Research supported by National Science Foundation Grant NSF GP-3694 at Columbia University.

[2] Now at Stanford University.

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

2. Proof of the theorem. The proof will be given in several lemmas, some of which may be of independent interest (see, for example, Remark (a) in Section 3). We assume throughout that (x_n), (μ_n), f, τ, c, and λ are as in the paragraph preceding the statement of the theorem. For convenience take $\mu = 1$.

LEMMA 1. *If* $f(n) \equiv 1$ *and* (5) *holds, then* (6) *holds. If in addition*

$$\sup_n n^{-1} \textstyle\sum_1^n E(x_i - \mu_i)^- < \infty, \tag{9}$$

then (3) *holds.*

PROOF. Let $0 < \delta < 1$, $t = \min(\tau, k)$, $k = 1, 2, \cdots$. By (1) and Wald's lemma

$$(1 - \delta)Et + O(1) \leqq E(\textstyle\sum_1^t \mu_i) \leqq Es_t \leqq c + E(x_t - \mu_t)^+ + E\,|\mu_t|. \tag{10}$$

Suppose $Et \to \infty$ as $k \to \infty$. From (1)

$$E\,|\mu_t| = o(Et). \tag{11}$$

Writing $y_n = (x_n - \mu_n)^+$, we have for every $\epsilon > 0$

$$\begin{aligned} Ey_t &\leqq \epsilon Et + \textstyle\int_{\{y_t > \epsilon t\}} y_t \leqq \epsilon Et + E(\sum_1^t I_{\{y_i > \epsilon i\}}\, y_i) \\ &= \epsilon Et + E(\textstyle\sum_1^t \int_{\{y_i > \epsilon i\}} y_i) \leqq 2\epsilon Et + O(1), \end{aligned}$$

where the last inequality follows from (5). Hence

$$Ey_t = o(Et). \tag{12}$$

From (10), (11), and (12)

$$(1 - \delta)Et + O(1) \leqq o(Et),$$

a contradiction, and it follows that

$$E\tau < \infty. \tag{13}$$

From (1) and (13) $E(\sum_1^\tau \mu_i) < \infty$. Since $s_\tau > c$, we also see that Es_τ and $E(s_\tau - \sum_1^\tau \mu_i)$ exist.

$$\textstyle\int_{\{\tau > n\}} (s_n - \sum_1^n \mu_i)^+ \leqq cP\{\tau > n\} \to 0,$$

and it follows (Doob [6], p. 302) that

$$E(\textstyle\sum_1^\tau \mu_i) \leqq Es_\tau\,. \tag{14}$$

Observe that $\tau \uparrow \infty$ as $c \to \infty$ and hence $E\tau \to \infty$. Thus from (1) and (14)

$$(1 - \delta)E\tau + O(1) \leqq E(\textstyle\sum_1^\tau \mu_i) \leqq Es_\tau \leqq c + E(x_\tau - \mu_\tau)^+ + E\,|\mu_\tau|$$

and as above $(1 - \delta)E\tau + O(1) \leqq c + o(E\tau)$, and (6) holds.

From (9) and (13) $E(\sum_1^\tau (x_i - \mu_i)^-) < \infty$, and again appealing to Doob [6], p. 302,

$$c \leqq Es_\tau \leqq E(\textstyle\sum_1^\tau \mu_i) \leqq O(1) + (1 + \delta)E\tau,$$

which together with (6) completes the proof.

For the applications below we state the following slight generalization of (the first part of) Lemma 1.

LEMMA 2. *Let* $0 < \alpha < 1$ *and assume that* $\alpha(\lambda) \to \alpha$, $\lambda \to \infty$. *Define*

$$\tau^* = \textit{first} \quad n \geqq 1 \quad \textit{such that} \quad s_n > \alpha(\lambda)n + \lambda(1 - \alpha(\lambda))$$

$$= \infty \quad \textit{if no such} \quad n \quad \textit{exists}.$$

If (5) *holds, then* $\limsup \lambda^{-1}E\tau^* \leqq 1$.

LEMMA 3. *If* (4) *and* (5) *hold, then* (6) *holds.*

PROOF. We may assume that f is concave and that $\lambda(c)$ is uniquely defined. Let $\alpha(\lambda) = \lambda f'(\lambda)/f(\lambda)$, where f' is the right derivative of f. Then ([8], p. 422) $\alpha(\lambda) \to \alpha$, $\lambda \to \infty$. It is easily seen that $\alpha(\lambda)x + \lambda(1 - \alpha(\lambda))$ is a line support to $cf(x)(= \lambda f(x)/f(\lambda))$ at the point (λ, λ). Hence letting τ^* be as in Lemma 2, $\tau \leqq \tau^*$, and by Lemma 2 (6) holds.

LEMMA 4. *If* (4) *holds and* $s_n/n \to 1$ a.s., *then* $\liminf \lambda^{-1}E\tau \geqq 1$.

PROOF. Obviously $\tau \to \infty$ as $c \to \infty$. From the definition of τ

$$s_{\tau-1}/\tau \leqq cf(\tau)/\tau < s_\tau/\tau.$$

Letting $c \to \infty$, $cf(\tau)/\tau \to 1$, or from (4) $c \sim \tau^{1-\alpha}/L(\tau)$ a.s. From (2) and (4) $c \sim \lambda^{1-\alpha}/L(\lambda)$. Hence by inversion ([10], p. 46) $\tau \sim \lambda$ a.s., and the result follows by Fatou's lemma.

In similar contexts (see, e.g., [11]) the technique of proof of Lemma 4 is quite useful and easy to apply. Known conditions sufficient to imply that $s_n/n \to 1$ a.s. are much stronger than (5), however, and consequently it seems interesting in the present context to attempt to establish (3) under conditions resembling (5).

LEMMA 5. *If* (7) *and*

$$\sup_n 1/n \sum_1^n E\,|x_i - \mu_i| < \infty, \tag{15}$$

then $(s_n - \sum_1^n \mu_i)/n \to 0$ *in probability.*

PROOF. The proof is an easy application of the classical method of proof of the weak law of large numbers (e.g. [8], p. 231).

LEMMA 6. *If* (4), (7), *and* (8) *hold, then* $\liminf \lambda^{-1}E\tau \geqq 1$.

PROOF. To avoid overburdening the proof, we shall assume that $\mu_n = 1$ and that $f(n) = n^\alpha$. The general case requires trivial modifications and an occasional reference to known results about slowly varying functions [10]. With $f(n) = n^\alpha$, $\lambda = c^{1/1-\alpha}$. It suffices to show that for every $\delta > 0$

$$P\{\tau > c^{1/1-\alpha}(1 - \delta)\} \to 1, \qquad c \to \infty. \tag{16}$$

But $P\{\tau \leqq n\} = P\{s_i > ci^\alpha, \text{ some } i \leqq n\}$. Letting $n \to \infty$ such that $c \sim n^{1-\alpha}(1 + \delta)$, it is easily seen that to prove (16) it suffices to show that for every $\delta > 0$

$$P\{s_i > i^\alpha n^{1-\alpha}(1 + \delta), \quad \text{some} \quad i \leqq n\} \to 0, \qquad n \to \infty. \tag{17}$$

$$(18)\quad P\{s_i > i^\alpha n^{1-\alpha}(1+\delta), \quad \text{some} \quad i \leqq n\} \leqq P\{s_i - i > i^\alpha n^{1-\alpha}\delta, \quad \text{some} \quad i \leqq n\}.$$

Let $y_n = x_n - 1$, $T_n = y_1 + \cdots + y_n$ $(n \geqq 1)$. For k fixed but arbitrary and all n define $n_i = [in/k]$, $i = 0, 1, \cdots, k$. Then by (8) and a generalization of the Hájek–Rényi inequality [1],

$$(19)\quad \begin{aligned} &P\{|T_i| > \delta i^\alpha n^{1-\alpha}, \quad \text{some} \quad i \leqq n_1\} \\ &\qquad \leqq \textstyle\sum_1^{n_1} E\,|y_i|/\delta i^\alpha n^{1-\alpha} \leqq K/\delta(1-\alpha)k^{1-\alpha}. \end{aligned}$$

Let $0 < \delta' < \delta 2^{-\alpha}$ and $\eta = \delta 2^{-\alpha} - \delta'$. Denote by u the least integer i, if any, $n_1 < i \leqq n$ such that $T_i > \delta i^\alpha n^{1-\alpha}$. Then

$$(20)\quad \begin{aligned} P\{T_i > \delta i^\alpha n^{1-\alpha} \quad \text{some} \quad i, n_1 < i \leqq n\} &\leqq P\{T_{n_i} > \delta' n_i^\alpha n^{1-\alpha} \quad \text{some} \quad i \leqq k\} \\ &+ \textstyle\sum_{i=2}^k \sum_{n_{i-1}<r<n_i} P\{u = r\} P\{T_{n_i} - T_r < n^{1-\alpha}(\delta' n_i^\alpha - \delta r^\alpha)\}. \end{aligned}$$

By Lemma 5

$$(21)\quad \begin{aligned} &P\{T_{n_i} > \delta' n_i^\alpha n^{1-\alpha}, \quad \text{some} \quad i \leqq k\} \\ &\qquad \leqq \textstyle\sum_1^k P\{T_{n_i} > \delta' n_i^\alpha n^{1-\alpha}\} \to 0, \quad n \to \infty. \end{aligned}$$

Let $\epsilon > 0$ and for each $j \leqq n$ let $y_j' = y_j I\{|y_j| \leqq \epsilon n\}$, $T_j' = y_1' + \cdots + y_j'$, $\beta_j = ET_j'$. From (7)

$$(22)\quad \begin{aligned} 1/n\,|\beta_{n_i} - \beta_r| &= 1/n\,|-\textstyle\sum_{r+1}^{n_i} \int_{\{|y_j|>\epsilon n\}} y_j| \\ &\leqq 1/n \textstyle\sum_1^n \int_{\{|y_j|>\epsilon n\}} |y_j| = o(1) \end{aligned}$$

and

$$(23)\quad \begin{aligned} &P\{y_j \neq y_j', \quad \text{some} \quad j = r+1, \cdots, n_i\} \\ &\qquad \leqq \textstyle\sum_{r+1}^{n} P\{|y_j| > \epsilon n\} \\ &\qquad \leqq 1/\epsilon n \textstyle\sum_1^n \int_{\{|y_j|>\epsilon n\}} |y_j| = o(1) \end{aligned}$$

uniformly in $i = 2, \cdots, k$, $n_{i-1} < r < n_i$. Taking n sufficiently large, we have from (8), (22), (23), and Chebyshev's inequality

$$(24)\quad \begin{aligned} P\{T_{n_i} - T_r < n^{1-\alpha}(\delta' n_i^\alpha - \delta r^\alpha)\} &\leqq P\{T_{n_i} - T_r < -\eta n k^{-\alpha}\} \\ &\leqq P\{|T'_{n_i} - T_r' - (\beta_{n_i} - \beta_r)| > (\eta n k^{-\alpha})/2\} \\ &\quad + P\{y_j \neq y_j', \quad \text{some} \quad j = r+1, \cdots, n_i\} \\ &\leqq (4k^{2\alpha}/\eta^2 n^2) \textstyle\sum_{r+1}^{n_i} Ey_j'^2 + o(1) \\ &\leqq 4Kk^{2\alpha}\epsilon/\eta^2 + o(1) \end{aligned}$$

uniformly in $i = 2, \cdots, k$, $n_{i-1} < r < n_i$. Thus from (18), (19), (20), (21), and (24)

$$\limsup_{n\to\infty} P\{s_i > i^\alpha n^{1-\alpha}(1+\delta), \quad \text{some} \quad i \leqq n\} \leqq K/\delta(1-\alpha)k^{1-\alpha} + 4Kk^{2\alpha}\epsilon/\eta^2$$

which can be made as small as desired by first taking k sufficiently large and then ϵ sufficiently small.

To complete the proof of the theorem it suffices in light of the preceding lemmas to show that (7) and (8) imply (5). Let $\epsilon > 0$. Then from (7) and (8)

$$\begin{aligned} 1/n \textstyle\sum_1^n \int_{\{|x_i-\mu_i|>\epsilon i\}} |x_i - \mu_i| \\ &\leqq 1/n \textstyle\sum_1^{[\epsilon n]} E\,|x_i - \mu_i| + 1/n \textstyle\sum_{[\epsilon n]+1}^n \int_{\{|x_i-\mu_i|>\epsilon^2 n\}} |x_i - \mu_i| \\ &\leqq \epsilon K + o(1). \end{aligned}$$

3. Remarks. (a) Lemma 1 and its proof generalize and simplify Theorem 2 of [3]. Our condition (9), for example, is implied by the Chow-Robbins conditions that either

$$\sup E(x_n - \mu_n)^- < \infty \tag{25}$$

or

$$x_n \geqq 0\ (n \geqq 1). \tag{26}$$

Obviously (25) implies (9); and if (26) holds then $x_n - \mu_n \geqq -\mu_n$, or

$$1/n \textstyle\sum_1^n E(x_i - \mu_i)^- \leqq 1/n \textstyle\sum_1^n \mu_i \rightarrow \mu < \infty.$$

(b) Although we have assumed that $x_1, x_2, \cdots$ are independent, it is easily seen that our result remains true under the assumptions of, say, Theorem 1 of [3]. This result is stronger than Chow's Theorem 1 ([2]). It should be noted, however, that although Chow assumes that $f(n) = n^\alpha$ for some $\alpha \varepsilon [0, 1)$, nevertheless, his theorem remains true under the weaker assumptions that f is non-decreasing and $f(n) = o(n)$.

(c) In the case of non-negative (x_i), we may by reflection about the line $y = \mu x$ obtain a similar result for stopping rules of the form $\tau = \inf \{n : 1 + s_n < c^{-1}f(n)\}$, where $f(n) \sim n^\alpha L(n)$, $\alpha > 1$. When $\alpha = 2$ this result is of interest in problems of sequential confidence intervals [4], [11], and it was in this context that Chernoff and Simons mentioned to me the possibility of approximating the stopping boundary by an appropriate tangent. For purposes of proving Lemma 3, this method is tantamount to use of Jensen's inequality. It does, however, have advantages (Remark (d)).

(d) In some cases it is possible to combine the method of proof of Lemma 3 with Blackwell's renewal theorem to obtain stronger results. For example, if $x_1, x_2, \cdots$ are iid with positive mean μ and finite second moment, and if $f(n) = n^\alpha$ $(0 < \alpha < 1)$, then

$$\limsup\ [E\tau - \lambda] \leqq E(x_1 - \mu\alpha)^2/2\mu^2(1 - \alpha)^2.$$

(See [8], p. 372 for the appropriate result from renewal theory.)

(e) In the presence of higher moments, one can combine the present techniques with those of Chow ([2], Theorem 2) to show that $E\tau^2 \sim \lambda^2$.

(f) The method of proof of Lemma 3 also yields the following result of interest in certain hypothesis testing problems [5], [7]: if $x_1, x_2, \cdots$ are iid with positive mean μ and variance 1, and if $\tau = \inf \{n : s_n > n^{\frac{1}{2}} L(n)\}$, where $L(n) \sim (2 \log \log n)^{\frac{1}{2}}$, then

$$E\tau \leqq 2\mu^{-2} \log \log \mu^{-1}(1 + o(1)), \mu \to 0.$$

Acknowledgment. I would like to thank Professor Richard F. Gundy, whose helpful discussions on a related problem [9] led to a substantial improvement in this note.

REFERENCES

[1] Chow, Y. S. (1960). A martingale inequality and the law of large numbers. *Proc. Amer. Math. Soc.* **11** 107–111.

[2] Chow, Y. S. (1966). On the moments of some one-sided stopping rules. *Ann. Math. Statist.* **37** 382–387.

[3] Chow, Y. S. and Robbins, H. E. (1963). A renewal theorem for random variables which are dependent or non-identically distributed. *Ann. Math. Statist.* **34** 390–396.

[4] Chow, Y. S. and Robbins, H. E. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

[5] Darling, D. A. and Robbins, H. E. (1967). Iterated logarithm inequalities. To appear in *Proc. Nat. Acad. Sci.*

[6] Doob, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[7] Farrell, R. (1964) Asymptotic behavior of expected sample size in certain one-sided tests. *Ann. Math. Statist.* **35** 36–72.

[8] Feller, W. (1966). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.

[9] Gundy, R. F. and Siegmund, D. (1967). On a stopping rule and the central limit theorem. *Ann. Math. Statist.* **38**.

[10] Karamata, J. (1930). Sur-un mode de croissance regulière. *Mathematica* (*Cluj*) **4** 38–53.

[11] Nadas, A. (1967). On the asymptotic theory of estimating the mean by sequential confidence intervals of prescribed accuracy. Unpublished dissertation. Columbia Univ.