The Annals of Mathematical Statistics
1971, Vol. 42, No. 6, 1936–1942

ON THE EXISTENCE OF THE OPTIMAL STOPPING RULE IN THE S_n/n PROBLEM WHEN THE SECOND MOMENT IS INFINITE[1]

BY M. E. THOMPSON, A. K. BASU AND W. L. OWEN

University of Waterloo, Laurentian University, and Rutgers University

Let $X_1, X_2, \cdots$ be i.i.d. random variables with mean 0, and let $S_n = \sum_{i=1}^n X_i$. The S_n/n optimal stopping problem is to maximize $E(S_\tau/\tau)$ among finite-valued stopping times τ relative to the process $(S_n, n \geqq 1)$. In this paper we prove partially Dvoretzky's (1967) conjecture that an optimal stopping time should exist when $E|X_1|^\beta < \infty$ for some $\beta > 1$, by showing that the result holds if $\lim \sup_{n\to\infty} P(S_n \geqq c\|S_n\|) > 0$ for some $c > 0$, where $\|S_n\| = (E|S_n|^\beta)^{1/\beta}$. This condition is shown to hold in some special cases, including the case where the X_i are in the domain of attraction of a stable distribution with exponent greater than one.

1. Introduction. Let $X_1, X_2, \cdots$ be independent and identically distributed random variables with mean 0, and let $S_n = \sum_{i=1}^n X_i$. If $\mathscr{M}$ is the collection of finite valued stopping times τ relative to the process $(S_n, n \geqq 1)$ for which $E(S_\tau/\tau)$ is defined (possibly infinite), then the S_n/n optimal stopping problem is to find if possible $\sigma \in \mathscr{M}$ such that

$$E(S_\sigma/\sigma) = \sup[E(S_\tau/\tau) : \tau \in \mathscr{M}]. \tag{1.1}$$

Burgess Davis (1971) has shown that if $E(X_1 \log^+ X_1) = \infty$ (where $\log^+ a = \log a$ if $a \geqq 1$, and 0 if $a < 1$), then there is a $\sigma \in \mathscr{M}$ for which $E(S_\sigma/\sigma) = \infty$. This σ is clearly optimal in the sense of (1.1). On the other hand, it is also well known (see Davis (1971)) that if $E(X_1 \log^+ X_1) < \infty$, then $E(\sup_{n\geqq 1}(S_n^+/n)) < \infty$. Therefore, if $E(X_1 \log^+ X_1) < \infty$, $\mathscr{M}$ is the class of all finite stopping times relative to $(S_n, n \geqq 1)$; and general optimal stopping theory tells us a good deal more.

The process $X = ((S_n, n), n \geqq 1)$ may be regarded as a Markov process with state space $R \times [0, \infty)$, stationary transition probabilities, and initial distribution that of X_1 on the line $\{(x, 1) : x \in R\}$. (Here R denotes the real numbers.) Suppose that $E(\sup_n (S_n^+/n)) < \infty$, and for each $(x, s) \in R \times (0, \infty)$ let

$$h(x, s) = (x/s)^+ \vee \sup[E((x+S_\tau)/(s+\tau))^+ : \tau \in \mathscr{M}]. \tag{1.2}$$

Let $D = \{(x, s) \in R \times (0, \infty) : h(x, s) = x^+/s\}$. It can be proved as in Theorem 8 of Chow and Robbins (1967) that $(h(x+S_n, s+n), n \geqq 1)$ is the minimal supermartin-

Received October 20, 1969; revised May 14, 1971.

[1] The results in this paper were obtained independently in the doctoral dissertations of the three authors. Dr. Basu's proof was slightly different from the one given here.

gale above $((x+S_n)^+/(s+n), n \geqq 1)$ for each $(x, s) \in R\times[0, \infty)$. If $\tau \in \mathscr{M}$ and $\bar{\tau} = \inf[n \geqq \tau : S_n > -x]$, then $\bar{\tau} \in \mathscr{M}$ (Feller (1966), page 380) and

$$E\left(\frac{x+S_\tau}{s+\tau}\right) \leqq E\left(\frac{x+S_{\bar{\tau}}}{s+\bar{\tau}}\right) = E\left(\left(\frac{x+S_{\bar{\tau}}}{s+\bar{\tau}}\right)^+\right). \tag{1.3}$$

Hence $D \subset (0, \infty)\times(0, \infty)$, and it follows by the Corollary to Theorem 6 of Chow and Robbins (1967) that if for $(x, s) \in R\times[0, \infty)$ the stopping time $\tau(x, s) = \inf[n \geqq 1 : (x+S_n, s+n) \in D]$ is finite, it maximizes $E(x+S_\tau/s+\tau)$ among $\tau \in \mathscr{M}$.

Dvoretzky's proof ((1967), Section 3) of the following theorem, describing the set D, is valid, as it does not use the general assumption of his paper that $EX_1{}^2 < \infty$.

THEOREM 1.1. *Let $EX_1 \log^+ X_1$ be finite. Then there is a strictly increasing positive function $f(s)$ on $[0, \infty)$ such that for $(x, s) \in R\times[0, \infty)$*

$$\begin{aligned} \frac{x}{s} &< \sup\left[E\left(\frac{x+S_\tau}{s+\tau}\right) : \tau \in \mathscr{M}\right] \quad \text{if} \quad x < f(s), \\ \frac{x}{s} &= \sup\left[E\left(\frac{x+S_\tau}{s+\tau}\right) : \tau \in \mathscr{M}\right] \quad \text{if} \quad x = f(s), \\ \frac{x}{s} &> \sup\left[E\left(\frac{x+S_\tau}{s+\tau}\right) : \tau \in \mathscr{M}\right] \quad \text{if} \quad x > f(s). \end{aligned} \tag{1.4}$$

Dvoretzky (1967) and Teicher and Wolfowitz (1966) showed that if $EX_1{}^2 < \infty$, then $\tau(0, 0)$ is finite, and therefore optimal. Dvoretzky conjectured that the same result ought to hold when $E|X_1|^\beta < \infty$ for some $\beta > 1$. In this paper we prove a partial result in this direction; we establish the existence of an optimal stopping rule when the truncated variance

$$U(x) = \int_{(-x)-}^{x} y^2\, dP(X_1 \leqq y) \tag{1.5}$$

is of "dominated variation" as $x \to \infty$, in a sense defined precisely in Section 3 below. A particular case is the situation when X_1 belongs to the domain of attraction of some stable random variable with exponent greater than one.

2. An upper bound on f. Let $\bar{\mathscr{M}}$ denote the class of all (possibly infinite) stopping times relative to $(S_n, n \geqq 1)$. If $\tau \in \bar{\mathscr{M}}$, we shall set

$$\frac{x+S_\tau}{s+\tau} = \lim_{n\to\infty} \frac{x+S_n}{s+n} = 0 \tag{2.1}$$

on the set $\{\tau = \infty\}$ for any $(x, s) \in R\times[0, \infty)$. The theorem we shall prove in this section is the following:

THEOREM 2.1. *If there is a $\beta > 1$ such that $E|X_1|^\beta < \infty$, then there is a finite constant K_0 such that for every $s \geqq 1$ and $x \geqq K_0\|S_{[s]}\|$, where $\|S_{[s]}\| = (E|S_{[s]}|^\beta)^{1/\beta}$, we have*

$$\frac{x}{s} \geqq \sup\left[E\left(\frac{x+S_\tau}{s+\tau}\right) : \tau \in \bar{\mathscr{M}}\right]. \tag{2.2}$$

The proof will be broken down into a sequence of lemmas, in which it will always be assumed that $E|X_1|^\beta < \infty$ for some $\beta > 1$. The first one is an analogue of Lemma 2 of Dvoretzky (1967).

LEMMA 2.1. *If* $s \geqq 1$ *then*

$$\sup\left[E\left(\frac{S_\tau}{\tau+s}\right): \tau \in \bar{\mathscr{M}}\right] < \frac{2\beta' \|S_{[s]}\|}{[s]} \tag{2.3}$$

where $\beta' = (1-1/\beta)^{-1}$.

PROOF. We have

$$\begin{aligned} E\left[\sup_{n\geqq 1}\left|\frac{S_n}{n+s}\right|^\beta\right] &\leqq E\left[\sup_{[s]\geqq n\geqq 1}\left|\frac{S_n}{n+s}\right|^\beta\right] + E\left[\sup_{n>[s]}\left|\frac{S_n}{n+s}\right|^\beta\right] \\ &\leqq E\left[\sup_{[s]\geqq n\geqq 1}\left|\frac{S_n}{n+s}\right|^\beta\right] + E\left[\sup_{n>[s]}\left|\frac{S_n}{n}\right|^\beta\right]. \end{aligned} \tag{2.4}$$

But the processes $(|S_1|^\beta, |S_2|^\beta, \cdots)$ and $(\cdots, |S_3/3|^\beta, |S_2/2|^\beta, |S_1|^\beta)$ are both submartingales. Therefore (see Doob (1953), page 317),

$$E\left[\sup_{n\geqq 1}\left|\frac{S_n}{n+s}\right|^\beta\right] \leqq 2(\beta')^\beta E\left[\left|\frac{S_{[s]}}{[s]}\right|^\beta\right]. \tag{2.5}$$

From Jensen's inequality

$$E\left[\sup_{n\geqq 1}\left|\frac{S_n}{n+s}\right|\right] \leqq 2\beta' \frac{\|S_{[s]}\|}{[s]}, \tag{2.6}$$

and the lemma follows.

LEMMA 2.2. *Let* $s' \geqq s > 0$, $x' \leqq x$ *and* $\tau \in \bar{\mathscr{M}}$ *satisfy*

$$E\left(\frac{x+S_\tau}{s+\tau}\right) \geqq \frac{x}{s}. \tag{2.7}$$

Then for $m = 0, 1, 2, \cdots$ *there is a stopping time* $\tau(m) \in \bar{\mathscr{M}}$ *satisfying*

$$S_{\tau(m)} > x - x' \text{ a.s. } on\ \{\tau(m) \leqq m\}, \tag{2.8}$$

and

$$E\left(\frac{x'+S_{\tau(m)}}{s'+\tau(m)}\right) \geqq \frac{x'}{s'}. \tag{2.9}$$

PROOF. This is Lemma 6 of Dvoretzky (1967).

LEMMA 2.3. *Let* (2.7) *hold for some* $x > 0$, $s > 0$, $\tau \in \bar{\mathscr{M}}$. *Then there is a* $\tau^* \in \bar{\mathscr{M}}$ *satisfying*

$$E\left(\frac{x/2+S_{\tau^*}}{s+\tau^*}\right) \geqq \frac{x}{2s} \tag{2.10}$$

and

$$E\left(\frac{1}{s+\tau^*}\right) \leqq \frac{1}{2s} + \frac{\|S_{[s]}\|}{sx}. \tag{2.11}$$

PROOF. Let τ^* be $\tau([s])$ from Lemma 2.2, with $s' = s$ and $x' = x/2$. Then (2.10) holds, and $\max(S_1, S_2, \cdots, S_{[s]}) > x/2$ a.s. on $\{\tau^* \leqq s\}$. Thus $P(\tau^* \leqq s) \leqq 2E|S_{[s]}|/x$, and hence by Jensen's inequality, $P(\tau^* \leqq s) \leqq 2\|S_{[s]}\|/x$. But

$$\begin{aligned} E\left(\frac{1}{s+\tau^*}\right) &\leqq \frac{P(\tau^* \leqq s)}{s+1} + \frac{P(\tau^* > s)}{s+[s]+1} \\ &\leqq \frac{P(\tau^* \leqq s)}{s} + \frac{P(\tau^* > s)}{2s} \\ &= \frac{1}{2s} + \frac{P(\tau^* \leqq s)}{2s}, \end{aligned} \tag{2.12}$$

and the result follows.

To prove Theorem 2.1 we observe that if (2.7) holds for some $\tau \in \bar{\mathscr{M}}$, then by Lemmas 2.3 and 2.1,

$$\begin{aligned} \frac{x}{2s} \leqq E\left(\frac{x/2+S_{\tau^*}}{s+\tau^*}\right) &= \frac{x}{2}E\left(\frac{1}{s+\tau^*}\right) + E\left(\frac{S_{\tau^*}}{s+\tau^*}\right) \\ &\leqq \frac{x}{4s} + \frac{\|S_{[s]}\|}{2s} + \frac{2\beta'\|S_{[s]}\|}{[s]}, \end{aligned} \tag{2.13}$$

or

$$\frac{x}{4s} \leqq (2\beta' + \tfrac{1}{2})\frac{\|S_{[s]}\|}{[s]} \leqq (4\beta'+1)\frac{\|S_{[s]}\|}{s}. \tag{2.14}$$

Theorem 2.1 now follows with $K_0 = 16\beta'+4$.

COROLLARY. *If* $E|X_1|^\beta < \infty$ *for some* $\beta > 1$, *then* $f(s) \leqq K_0\|S_{[s]}\|$ *for* $s \geqq 1$; *and if in addition*

$$\limsup_{n\to\infty} S_n/\|S_n\| > K_0 \text{ a.s.}, \tag{2.15}$$

then an optimal stopping time exists for the S_n/n *problem.*

It is easy to show using the Hewitt-Savage Zero-One Law (see Feller (1966), page 122) that inequality (2.15) holds if $\limsup_{n\to\infty} P(S_n > K_0\|S_n\|) > 0$. Moreover, by using an inequality of Marcinkiewicz and Zygmund (1938) (see Theorem 5) we can strengthen this result as follows.

COROLLARY. *If* $E|X_1|^\beta < \infty$ *for some* $\beta > 1$, *and if*

$$\limsup_{n\to\infty} P(S_n \geqq c\|S_n\|) > 0 \tag{2.16}$$

for some $c > 0$, *then an optimal stopping time exists for the* S_n/n *problem.*

3. Dominated variation and the relative growth of S_n and $\|S_n\|$. Let F be the common distribution function of $X_1, X_2, \cdots$ and define the function U on $[0, \infty)$ by

$$U(x) = \int_{(-x)-}^{x} y^2 \, dF(y). \tag{3.1}$$

DEFINITION. (Feller (1967), Section 8). The truncated variance U is of *dominated variation* if there exist constants $v > 0$, C and $T > 0$ such that

$$U(tx)/U(t) < Cx^{2-v} \tag{3.2}$$

for all $x > 1$ and $t > T$.

The following lemma will relate the behavior of the other truncated moments of X_1 with that of U, when U is of dominated variation.

LEMMA 3.1. (Feller (1967), Section 8, Theorem 2). *Let U be of dominated variation and let*

$$V_q(x) = \int_x^{\infty} y^{-q} \, dU(y). \tag{3.3}$$

If $q > 2-v$, then for $t > T$

$$t^q V_q(t)/U(t) \leqq L_q = -1 + \frac{Cq}{q-2+v}. \tag{3.4}$$

We observe that $V_2(x) = 1-F(x)+F[(-x)-]$, and

$$V_1(x) = \int_{(-\infty,-x)\cup(x,\infty)} |y| \, dF(y).$$

A trivial consequence of (3.2) is that $\lim_{x\to\infty} x^{-2}U(x) = 0$. Therefore, there is a sequence (a_n) of positive numbers tending to infinity and a finite constant K_1 such that

$$n a_n^{-2} U(a_n) \leqq K_1 \tag{3.5}$$

for all n. From (3.2) we then have

$$n a_n^{-2} U(a_n x) < K_1 C x^{2-v} \tag{3.6}$$

for $x > 1$ and n large. The relations

$$n V_2(a_n x) \leqq L_2 K_1 C x^{-v} \tag{3.7}$$

and (for $v > 1$)

$$n V_1(a_n x) \leqq L_1 K_1 C a_n x^{1-v}, \tag{3.8}$$

again for $x > 1$ and n large, now follow from (3.4).

LEMMA 3.2. *Let U be of dominated variation with $v > 1$, and let (a_n) be a sequence of numbers tending to infinity and satisfying* (3.5). *For n and x large, there is a constant C_0 independent of n and x such that*

$$P(|S_n| \geqq x a_n) \leqq C_0 x^{-v}. \tag{3.9}$$

PROOF. Fix $x > 1$, and define the truncated variables

$$\begin{aligned} X_{kn} &= X_k \quad \text{if } -x a_n \leqq X_k \leqq x a_n, \\ &= 0 \quad \text{otherwise.} \end{aligned} \tag{3.10}$$

Let $\mu_n = EX_{kn}$ and $S_{mn} = \sum_{k=1}^{m} X_{kn}$. Then

(3.11) $P(|S_n| \geqq xa_n) \leqq n[P(X_1 < -xa_n)+P(X_1 > xa_n)]+P(|S_{nn}| \geqq xa_n)$.

Since $EX_1 = 0$,

$$|n\mu_n| \leqq nV_1(xa_n) \leqq L_1K_1Ca_nx^{1-v} \leqq \tfrac{1}{2}a_nx \tag{3.12}$$

for x and n large. Therefore

$$\begin{aligned} P(|S_{nn}| > xa_n) &\leqq P(|S_{nn}-n\mu_n| > \tfrac{1}{2}xa_n) \\ &\leqq 4x^{-2}a_n^{-2}nU(a_nx) \\ &\leqq 4K_1Cx^{-v} \end{aligned} \tag{3.13}$$

for x and n large. Moreover,

$$n[P(X_1 < -a_nx)+P(X_1 > a_nx)] = nV_2(a_nx) \leqq L_2K_1Cx^{-v} \tag{3.14}$$

for x and n large. The conclusion of the lemma is now immediate.

COROLLARY. *Let U be of dominated variation with $v > 1$, and choose (a_n) to satisfy (3.5). Let F_n be the distribution of S_n/a_n. If $1 < \beta < v$, and some subsequence $(F_{n'})$ of (F_n) converges to G, then*

$$\lim_{n' \to \infty} \int_{-\infty}^{\infty} |y|^\beta dF_{n'}(y) = \int_{-\infty}^{\infty} |y|^\beta dG(y). \tag{3.15}$$

Feller (1967) proves that since U is of dominated variation, we can choose (a_n) so that the sequence $na_n^{-2}U(a_n)$ is bounded below by some constant $\rho > 0$, in which case the sequence (F_n) is stochastically compact in the following sense.

DEFINITION. If (G_n) is a sequence of distribution functions and every subsequence of (G_n) has a further subsequence converging to a nondegenerate distribution function, the sequence (G_n) is called *stochastically compact*.

In view of the criterion (2.16), the following theorem is immediate from these remarks and (3.15).

THEOREM 3.1. *If U is of dominated variation with $v > 1$, then an optimal stopping time exists for the S_n/n problem.*

4. Structure of the optimal stopping rule. It is an easy consequence of Theorem 2.1 and (3.15) that when U is of dominated variation with $v > 1$ and (a_n) is a sequence for which (F_n) is stochastically compact, the function f satisfies

$$f(s) \leqq K'a_{[s]} \tag{4.1}$$

for some finite positive K'. We may also prove the following in direct analogy with the second part of Dvoretzky's (1967) Theorem 2.

LEMMA 4.1. *Under the conditions stated above*

$$f(s) \geqq ka_{[s]} \tag{4.2}$$

for some positive k and all s.

PROOF. Let $c > 0$ be given. For any integer $r > 0$, let $\tau(r) = \inf[n \geqq r : S_n > -ca_r]$. The stopping time $\tau(r)$ is a.s. finite. Then

$$E\left(\frac{ca_r + S_{\tau(r)}}{r+\tau(r)}\right) \geqq \frac{1}{2r}\int_{-ca_r}^{\infty} (ca_r + u)\, dP(S_r \leqq u) \tag{4.3}$$
$$= \frac{a_r}{2r}\int_{-c}^{\infty} (c+w)\, dP(S_r/a_r \leqq w).$$

If $(F_{n'})$ is a subsequence of (F_n) converging to G,

$$\liminf_{n' \to \infty} \frac{n'}{ca_{n'}} E\left(\frac{ca_{n'} + S_{\tau(n')}}{n' + \tau(n')}\right) \geqq \frac{1}{2c}\int_{-c}^{\infty} (c+w)\, dG(w). \tag{4.4}$$

Now $2c - \int_{-c}^{\infty} (c+w)dG(w)$ is negative for $c = 0$, and continuous. Hence for some finite $c_0 > 0$, $2c_0 < \int_{-c_0}^{\infty} (c_0 + w)dG(w)$, and

$$\liminf_{n' \to \infty} \frac{n'}{c_0 a_{n'}} E\left(\frac{c_0 a_{n'} + S_{\tau(n')}}{n' + \tau(n')}\right) > 1. \tag{4.5}$$

Thus $c_0 a_{n'} < f(n')$ for all but finitely many n'. The lemma now follows by the stochastic compactness of the family (F_n).

It is well known (see Feller (1966), page 305) that if X_1 belongs to the *domain of attraction* of a nondegenerate random variable Y, with norming constants (a_n), then there exists a number $\alpha > 0$ such that $a_{rn}/a_n \to r^{1/\alpha}$ as $n \to \infty$, and Y is strictly stable with exponent α. If $\alpha > 1$, then X_1 satisfies the hypothesis of Theorem 3.1 with $1 < v < \alpha$ (see Feller (1966), page 303). Because $a_{rn}/a_n \to r^{1/\alpha}$ as $n \to \infty$, a_n is of the form $n^{1/\alpha}L(n)$ where L is a slowly varying function of its argument (Feller (1966), page 269), and equations (4.1) and (4.2) suggest that f may have the same asymptotic behavior. That this is true will be shown in a subsequent paper.

REFERENCES

CHOW, Y. S. and ROBBINS, H. (1967). On values associated with a stochastic sequence. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 427–440.

DAVIS, B. (1971). Stopping rules for S_n/n, and the class $L \log L$. To appear.

DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

DVORETZKY, A. (1967). Existence and properties of certain optimal stopping rules. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 441–452.

FELLER, W. (1966). *An Introduction to Probability Theory and its Applications* **2**. Wiley, New York.

FELLER, W. (1967). On regular variation and local limit theorems. *Proc. Fifth Berkeley Symp. Math. Stat. Prob.* **2** 373–388.

MARCINKIEWICZ, J. and ZYGMUND, A. (1938). Quelques théorèmes sur les fonctions indépendantes. *Studia Math.* **7** 104–120.

TEICHER, H. and WOLFOWITZ, J. (1966). Existence of optimal stopping rules for linear and quadratic rewards. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **5** 361–368.