*The Annals of Mathematical Statistics*
1972, Vol. 43, No. 5, 1559-1565

# SHARP BOUNDS FOR THE TOTAL VARIANCE OF UNIFORMLY BOUNDED SEMIMARTINGALES[1]

BY LESTER E. DUBINS[2]

*University of California, Berkeley*

Let $S_n = f + X_1 + \cdots + X_n$ be an expectation-decreasing semimartingale with values in the unit interval, and let $V_n$ be the conditional variance of $X_n$ given the past. Then $E(\sum V_n)$ is less than $f(2-f)$, and this bound is sharp. Sharper bounds are available if the process $S_0, S_1, \cdots$ satisfies suitable additional constraints.

**1. Introduction.** For each stochastic process $S = \{S_0, S_1, \cdots\}$ for which $V_i$, the conditional variance of the increment $S_i - S_{i-1}$, given the past is meaningful for each $i$, let $\bar{V} = \bar{V}(S)$, the *total variance* of $S$, be defined as the sum $V_0 + V_1, \cdots$, where $V_0$, the variance of $S_0$, will, in this note, be zero.

The size of $\bar{V}$ reflects the size of $S$, and conversely. For example, as is well known when the increments have mean 0 and are independent, $\bar{V} < \infty$ if and only if $\lim S_n$ exists and is finite. Also, for martingale increments, the size of $\bar{V}$ and the growth of $S$ are known to be related. (See e.g., [2] Theorem 4.1 (5), [3] and [5].)

One obvious measure of the size of the total variance $\bar{V}$ is its expected value $E(\bar{V})$. It is the purpose of this note to give sharp upper bounds for $E(\bar{V})$ when $S$ ranges over a sufficiently simple class of stochastic processes. This note is closely related to [1]. Some notation is helpful.

Let $I$ be the closed unit interval $[0, 1]$. For $f \in I$, let $\bar{S}(f)$ be the class of all expectation-decreasing semimartingales $S$ for which $S_0 = f$ and $S_j \in I$ for all $j$, and let $\bar{M}(f)$ be the set of all martingales $S \in \bar{S}(f)$.

Here is a simple preliminary observation.

PROPOSITION 1. *For $S \in \bar{M}(f)$,*

$$E(\bar{V}) \leqq f(1-f)\,. \tag{1}$$

*For every $f$ the bound is attained, and is attained by $S \in \bar{M}(f)$ if and only if* $\lim S_n$ *is, with probability one, an endpoint of $I$.*

Proposition 1 is an easy consequence of this easily established lemma.

LEMMA 1. *For $S \in \bar{M}(f)$, $E(\bar{V})$ is the variance of the limit of the $S_n$.*

A related inequality, typical of the contents of this note, is this.

Received July 19, 1971; revised January 27, 1972.

[1] Research sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, USAF, under Grant No. AFOSR-1312-67. The United States Government is authorized to reproduce and distribute reprints for Govermental purposes notwithstanding any copyright notation hereon.

[2] The hospitality of the University of Rome, the University of Florence, and the Consiglio Nazionale delle Ricerche during the author's sabbatical year facilitated this research.



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

PROPOSITION 2. *For* $S \in \bar{S}(f)$,

$$E(\bar{V}) < f(2-f)\,, \tag{2}$$

*and this bound is sharp.*

To make inequality (2) plausible and to see how it was conjectured, consider a standard Brownian motion with $f$ as its initial state and with a reflecting barrier at 1 and an absorbing barrier at 0. An easy computation would show that the expected value of the time until absorption for this expectation-decreasing semimartingale in continuous time is indeed the right-hand side of (2). This suggests that for any discrete-time $S \in \bar{S}(f)$ which approximates this continuous-time semimartingale, $E(\bar{V})$ is close to $f(2-f)$.

For perhaps the simplest $S \in \bar{S}(f)$ for which $E(\bar{V})$ is close to $f(2-f)$, proceed thus. Fix $\varepsilon > 0$. For each $g$, $0 \leqq g < 1$, let $\gamma = \gamma_g$ be that probability measure on $I$ of mean $g$ that "lives" on the two-point set $\{0, 1\}$. Thus $\gamma_g$ assigns to $\{1\}$ the probability $g$ and assigns probability $1 - g$ to $\{0\}$. For $g = 1$, let $\gamma_g = \gamma_1$ be the one-point, or Dirac-delta, measure that lives on the one-point set $\{1 - \varepsilon\}$. Consider the family $\gamma_g$ as the system of transition probabilities for a Markov process. Plainly, for each initial state $f$, this process $S$ is an expectation-decreasing semimartingale, and, therefore, an element of $\bar{S}(f)$. Plainly, for each $g$, the variance of $\gamma_g$ is $g(1-g)$. A trite calculation shows that if the initial state is 1, then the expectation of $N$, the number of visits to $1 - \varepsilon$, is $\varepsilon^{-1}$. Since the total variance $\bar{V}_1$, when the initial state is 1, is plainly $\varepsilon(1-\varepsilon)N$,

$$E(\bar{V}_1) = \varepsilon(1-\varepsilon)E(N) = 1 - \varepsilon\,. \tag{3}$$

Plainly, for an arbitrary initial state $f \in I$, $\bar{V}$ is the variance $\gamma_f$ of the first gamble plus the variance of the later gambles, so

$$\begin{aligned} E(\bar{V}) &= \sigma^2(\gamma_f) + \gamma_f\{1\}E(\bar{V}_1) \\ &= \sigma^2(\gamma_f) + f(1-\varepsilon) \\ &= f(1-f) + f(1-\varepsilon)\,. \end{aligned} \tag{4}$$

As is now evident, the bound $q(f) = f(2-f)$ in Proposition 2 cannot be lowered. That $q$ is indeed a bound, is obviously a consequence of these two lemmas.

LEMMA 2. *Let* $q(x) = x(2-x)$ *for* $x$ *in the unit interval* $I$, *let* $f \in I$, *let* $\gamma$ *be a probability measure on the unit interval, of mean at most* $f$, *and let* $\sigma^2(\gamma)$ *be the variance of* $\gamma$. *Then*

$$\int q(x)\, d\gamma(x) \leqq q(f) - \sigma^2(\gamma)\,. \tag{5}$$

The computation needed to verify (5) is straightforward, and omitted here, but is included in the slightly less simple computation given in Section 3.

LEMMA 3. *Let* $q$ *be any nonnegative, real-valued, (measurable) function defined on* $I$ *which satisfies* (5) *for all* $f \in I$ *and all* $\gamma$ *on* $I$ *of mean at most* $f$. *Then, for every* $S \in \bar{S}(f)$,

$$E(\bar{V}) \leqq q(f)\,. \tag{6}$$

Though Lemma 3 is a special case of the basic Theorem 2.12.1 in [5], as will be seen in Section 2 below, a direct proof is given here, in part to make this introduction self-contained.

PROOF OF LEMMA 3. For $S \in \bar{S}(f)$, the hypotheses imply:

$$E(q(S_1)) \leqq q(f) - V_1 ; \tag{7}$$

and, for every positive integer $n$,

$$E(q(S_{n+1}) \mid S_0, \cdots, S_n) \leqq q(S_n) - V_{n+1} . \tag{8}$$

Assume, by induction, that

$$E(q(S_n)) \leqq q(f) - E(V_1 + \cdots + V_n) , \tag{9}$$

which certainly holds for $n = 1$ by (7), since $V_1$, being constant, equals $E(V_1)$.

Now, take expectations in (8), and use (9) to see that (9) holds again when $n$ is replaced by $n + 1$. Since $q \geqq 0$, the left-hand, and hence the right-hand, side of (9) is nonnegative. This completes the proof of Lemma 3, and hence of Proposition 2.

Proposition 2 is subsumed under Theorem 2 below, and Theorem 2 itself follows from the (gambling) theorem presented in the next section.

**2. Specialization of the basic gambling theorem.** Each gambling house $\Gamma$ defined on an abstract set $F$ of fortunes determines, and is determined by, the set of all $(\gamma, f)$ such that $\gamma \in \Gamma(f)$. There is no great ambiguity, and some economy of notation, if the symbol "$\Gamma$" is used to designate this set of ordered couples too.

Let $w$ be a nonnegative, real-valued function defined on $\Gamma$, that is, $w(\gamma, f) \geqq 0$ for each $(\gamma, f)$ such that $\gamma \in \Gamma(f)$. In this note, principal interest focuses on a $w$ which is a function of $\gamma$ only, indeed where $w(\gamma)$ is simply the variance of $\gamma$. But having in mind applications where $w(\gamma, f)$ could be, for example when $F$ is a subset of the reals, the second moment of the lottery $[\gamma - f]$, that is, of the distribution of the displacement about $f$, the more general case will be treated here, which is hardly less simple.

The immediate program is to describe a nonnegative, real-valued function defined on $F$, and determined by $\Gamma$ and $w$, here to be designated by $\Gamma w$.

As a preliminary, for each initial fortune $f$, strategy $\sigma = (\sigma_0, \sigma_1, \cdots)$, and history $h = (f_1, f_2, \cdots)$, let

$$w(\sigma, f, h, n) = w(\sigma_0, f) + \cdots + w(\sigma_n(f_1, \cdots, f_n), f_n) , \tag{1}$$

where $n$ can be any positive integer, or more generally, any stop rule. For fixed $(\sigma, f, n)$, this is a nonnegative, finitary function of $h$, and hence has an expectation under $\sigma$; call this expectation $w(\sigma, f, n)$. Now designate the supremum over all $n$ of $w(\sigma, f, n)$ by $\hat{w}(\sigma, f)$ and define $\Gamma w$ thus.

$$(\Gamma w)(f) = \sup \hat{w}(\sigma, f) \tag{2}$$

where the sup is taken over all strategies $\sigma$ available in $\Gamma$ at $f$.

THEOREM 1. *If $q$ is a nonnegative, real-valued function defined on $F$ which satisfies*

$$w(\gamma, f) + \gamma q \leqq q(f) \qquad \text{for all } f \text{ and all } \quad \gamma \in \Gamma(f)\,, \tag{3}$$

*then* $q \geqq \Gamma w$.

The special case of constant $w$, $w(\gamma, f) = 1$ for all $(\gamma, f)$ includes "a general theorem" in [1].

PROOF OF THEOREM 1. The proof is a simple application of Theorem 2.12.1 in [5] to a gambling problem $(\Gamma_w, u)$ now to be described.

For each point $y$, let $F_y$ be the set of $(f, y)$ for $f \in F$, let $\delta(y)$ be the one-point measure that lives at $y$, and for each $\gamma$ on $F$, let $\gamma(y)$, defined for subsets of $F_y$, be the product measure $\gamma \times \delta(y)$. Of course, $\gamma(y)$ can be extended to the subsets of any set that contains $F_y$ (and, in particular, to $F \times R$ if $y \in R$) by assigning measure zero to any subset of the complement of $F_y$.

Let the fortune space of $\Gamma_w$ be $F \times R$ where $R$ is the set of nonnegative real numbers, and let $\gamma' \in \Gamma_w(f, x)$ if, and only if, for some $\gamma \in \Gamma(f)$, $\gamma'$ is $\gamma$ transferred to the section of $F \times R$ whose second coordinate is $x + w(\gamma, f)$, or more formally,

$$\gamma' = \gamma(x + w(\gamma, f))\,. \tag{4}$$

Now that $\Gamma_w$ has been defined, define $u$ as the projection of $F \times R$ onto $R$, that is $u(f, x) = x$, and define $Q(f, x)$ as $q(f) + x$. Since $q \geqq 0$, $Q \geqq u$, and since $q$ satisfies (3), $Q$ is excessive for $\Gamma_w$. Though $u$ and $Q$ are not bounded, they are bounded from below, and this is adequate to conclude from the basic Theorem 2.12.1 in [5] that $Q$ majorizes $\Gamma_w u$. Consequently,

$$\begin{aligned} q(f) &= Q(f, 0) \\ &\geqq (\Gamma_w u)(f, 0) \\ &= (\Gamma_w)(f)\,, \end{aligned} \tag{5}$$

and the theorem is proven.

**3. Return to the unit interval when the variances are bounded from below.** In this section $F$ is specialized to be the closed unit interval, and for every $\gamma$, $\gamma(F) = 1$. Let $s$ be a nonnegtive number, and let $\Gamma_s(f)$ consist of all $\gamma$ such that the mean of $\gamma$ is at most $f$ and the variance of $\gamma$ is at least $s$. There is no loss in supposing that $s$ does not exceed one-fourth, for otherwise, there is no $\gamma$ available in $\Gamma_s$. Also, as is easily seen, for any $\gamma$ on $F$ of mean $m$, the variance of $\gamma$ is at most $m(1 - m)$. Hence, for $\gamma$ to be available in $\Gamma_s$, $m(1 - m) \geqq s$, or, equivalently,

$$1 - \beta(s) \leqq m(\gamma) \leqq \beta(s)\,, \tag{1}$$

where $\beta(s)$ is the maximum of the two solutions to $x(1 - x) = s$, and $m(\gamma)$ is the mean of $\gamma$.

In particular, for $0 \leqq f < 1 - \beta(s)$, there is no $\gamma$ available in $\Gamma_s$. There is, therefore, some technical advantage in modifying the definition of $\Gamma_s$ by also permitting $\delta(f) \in \Gamma_s(f)$ for each $f$.

Let $w(\gamma)$ be the variance of $\gamma$, and the problem is to determine $\Gamma_s w$. To conjecture $\Gamma_s w$, one finds, for each $g$, a gamble $\gamma(g) \in \Gamma_s(g)$ such that, for each initial state $f$, the Markov process with stationary transition probabilities $\gamma(g)$ has a large total variance. Of course, for $0 \leqq g < 1 - \beta(s)$, $\gamma(g)$ must be $\delta(g)$. For $1 - \beta(s) \leqq g \leqq \beta(s)$, there are nontrivial fair gambles available in $\Gamma_s(g)$, and among them the simplest is $\gamma(g)$ which lives on the two endpoints $\{0, 1\}$, and assigns probability $g$ to $\{1\}$ and probability $1 - g$ to $\{0\}$. For $\beta(s) < g \leqq 1$, there are no fair gambles available; in fact, the mean of every available $\gamma$ is at most $\beta(s)$, as (1) implies. Moreover, for each $g$ in this interval, there is precisely one $\gamma$ available whose mean is $\beta(s)$, namely, the $\gamma$ of mean $\beta(s)$ that "lives" on the two-points $\{0, 1\}$. Therefore, for $\beta(s) < g \leqq 1$, let $\gamma(g) = \gamma(\beta(s))$. Let $\sigma(f)$ be the strategy that corresponds to the Markov process with initial state $f$ and transition probabilities $\gamma(f)$. For $s > 0$, let $q_s(f)$ be the total variance of the strategy $\sigma(f)$, that is, $q_s(f) = \hat{w}(\sigma(f), f)$, as in Section 2. A simple computation, like the one given in Section 1, determines the analytic form of $q_s$. Namely,

$$
\begin{aligned}
q_s(f) &= 0 & 0 \leqq f < 1 - \beta(s)\,, \\
&= f(1 + \beta(s) - f) & 1 - \beta(s) \leqq f \leqq \beta(s)\,, \\
&= \beta(s) & \beta(s) \leqq f \leqq 1\,.
\end{aligned}
\tag{2}
$$

Plainly, for $s = 0$, $q_s(f) = f(2 - f)$, which is the bound in Proposition 2.

Of course, $q_s \leqq \Gamma w$. For the reverse inequality, this extension of Lemma 2 is needed.

LEMMA 4. *For all $f$ and all $\gamma \in \Gamma_s(f)$,*

$$\sigma^2(\gamma) + \gamma q_s \leqq q_s(f)\,. \tag{3}$$

PROOF OF LEMMA 4. For $0 \leqq f < 1 - \beta(s)$, only $\delta(f) \in \Gamma_s(f)$, so the inequality is trivial. For $1 - \beta(s) \leqq f \leqq \beta(s)$, and any $\gamma \in \Gamma_s(f)$, indeed even for any $\gamma \in \Gamma_0(f)$, let $m$ be the mean of $\gamma$ and $\sigma^2$ its variance, and verify that

$$\sigma^2 + \gamma Q_s = Q_s(m) \leqq Q_s(f)\,, \tag{4}$$

where

$$Q_s(g) = g(1 + \beta(s) - g) \qquad 0 \leqq g \leqq 1\,. \tag{5}$$

Therefore,

$$
\begin{aligned}
\sigma^2 + \gamma q_s &\leqq \sigma^2 + \gamma Q_s \\
&\leqq Q_s(f) \\
&= q_s(f)\,.
\end{aligned}
\tag{6}
$$

Finally, to verify (3) for $f$ in the interval $(\beta(s), 1]$, first observe that for any $\gamma$ available in $\Gamma_s$, $\gamma$ is available in $\Gamma_s(m)$, where $m$ is the mean of $\gamma$. Moreover, unless $\gamma$ is trivial, $m \leqq \beta(s)$, as (1) implies. Therefore, for nontrivial $\gamma \in \Gamma_s(f)$,

(7) $$\sigma^2\gamma + \gamma q_s \leqq q_s(m) \leqq q_s(f)\,,$$

where the first inequality holds because (3) has already been established for $f \leqq \beta(s)$, and the second holds because $q_s$ is nondecreasing.

Now that Lemma 4 has been established, Theorem 1 applies to prove this extension of Proposition 2.

THEOREM 2. *For $w(\gamma)$ equal to the variance of $\gamma$, $\Gamma_s w = q_s$ for $0 \leqq s \leqq \frac{1}{4}$, and, for $s > 0$, $q_s$ is attained.*

**4. Applications of Theorem** 2. With only slight loss, Theorem 2 can be recast into the usual language of countably additive stochastic processes. Suppose, for example, that $S = \{S_0, S_1, \cdots\}$ is an expectation-decreasing semimartingale for which $S_0 = f$, $0 < f < 1$. Suppose, too, that $\tau$ is a stopping time for $S$ such that $S_i$ is in the unit interval [0, 1] for all $i \leqq \tau$, and, for all $i < \tau$, $V_i \geqq s$, where $V_i$ is the conditional variance of $S_i - S_{i-1}$ given the past. Then $E(\bar{V}) \leqq q_s(f)$, where $\bar{V}$ is the total variance of $S$ before time $\tau$. Moreover, since $s\tau \leqq \bar{V}$,

(1) $$E(\tau) \leqq s^{-1}q_s(f)\,,$$

and, for each $s > 0$, this bound is attained.

For $0 < c < d < 1$, consider the interesting example in which $\tau$ is the least $i$ such that $S_i$ is outside $[c, d]$. Then the first hypothesis on $\tau$ is satisfied if, and only if, $0 \leqq S_\tau \leqq 1$, and this plainly obtains if

(2) $$-c \leqq S_n - S_{n-1} \leqq 1 - d \qquad \text{for all} \quad n\,.$$

For such $\tau$, a bound on $E(\tau)$ closely related to the right-hand side of (1), but for processes with uniformly bounded increments and with a constraint on the second moment rather than on the variance, was obtained by Blackwell ([1], Inequality 4).

In a forthcoming joint paper with Isaac Meilijson, I expect even the simple Proposition 2 to find application to the proof that if a subfair casino with a fixed goal is perturbed a little, then the optimal probability of reaching that goal undergoes a correspondingly small alteration.

**5. Change of scale.** Of course, if the fortune space $F = [0, 1]$ is replaced by $F^* = [a, b]$, and $\Gamma_s$ is correspondingly replaced by $\Gamma_s{}^*$, then the upper bound $q_s$ must be replaced by $q_s{}^*$, where

(1) $$q_s{}^*(f) = q_{s^*}(f^*)\,,$$

and where

(2) $$s^* = s/(b - a)^2 \qquad \text{and} \qquad f^* = (f - a)/(b - a)\,.$$

REFERENCES

[1] BLACKWELL, DAVID (1964). Probability bounds via dynamics programming. *Proc. Symp. Appl. Math.*, American Mathematics Society, Providence. **16** 277-280.

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

[3] DUBINS, L. E. and FREEDMAN, DAVID A. (1965). A sharper form of the Borel-Cantelli Lemma and the strong law. *Ann. Math. Statist.* **36** 800-807.

[4] DUBINS, L. E. and SAVAGE, LEONARD J. (1965a). A Tchebycheff-like inequality for stochastic processes. *Proc. Nat. Acad. Sci.* **51** 274-275.

[5] DUBINS, L. E. and SAVAGE, LEONARD J. (1965b). *How to Gamble if You Must.* McGraw-Hill, New York.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720