TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 278, Number 1, July 1983

STOP RULE INEQUALITIES FOR UNIFORMLY BOUNDED SEQUENCES OF RANDOM VARIABLES

BY

THEODORE P. HILL AND ROBERT P. KERTZ

ABSTRACT. If $X_0, X_1, \ldots$ is an arbitrarily-dependent sequence of random variables taking values in $[0, 1]$ and if $V(X_0, X_1, \ldots)$ is the supremum, over stop rules t, of EX_t, then the set of ordered pairs $\{(x, y)\colon x = V(X_0, X_1, \ldots, X_n)$ and $y = E(\max_{j \leq n} X_j)$ for some $X_0, \ldots, X_n\}$ is precisely the set

$$C_n = \left\{(x, y)\colon x \leq y \leq x\left(1 + n(1 - x^{1/n})\right); 0 \leq x \leq 1\right\};$$

and the set of ordered pairs $\{(x, y)\colon x = V(X_0, X_1, \ldots)$ and $y = E(\sup_n X_n)$ for some $X_0, X_1, \ldots\}$ is precisely the set

$$C = \bigcup_{n=1}^{\infty} C_n.$$

As a special case, if $X_0, X_1, \ldots$ is a martingale with $EX_0 = x$, then $E(\max_{j \leq n} X) \leq x + nx(1 - x^{1/n})$ and $E(\sup_n X_n) \leq x - x \ln x$, and both inequalities are sharp.

1. Introduction. The subject of this paper is comparisons between the expected supremum of a uniformly bounded process and the optimal expected return (using stop rules) of the process.

Let $X_0, X_1, \ldots$ be random variables (on some common probability space $(\Omega, \mathcal{A}, P)$) taking values in $[0, 1]$ and let $V(X_0, X_1, \ldots)$ denote the value (supremum, over stop rules t, of EX_t) of the process $X_0, X_1, \ldots$ (for a formal definition, see §2).

The first main result of this paper (Theorem 3.2) gives a complete description of the possible values of the ordered pairs $(V(X_0, \ldots, X_n), E(\max_{j \leq n} X_j))$ for *all* processes uniformly bounded in $[0, 1]$, and the second main result (Theorem 4.2) gives the corresponding result for infinite sequences.

Comparisons of the value $V(X_0, X_1, \ldots)$ and $E(\sup_n X_n)$ have been called "prophet" problems because of the natural identification of $E(\sup_n X_n)$ with the optimal expected return of a prophet or player endowed with complete foresight. Such comparisons for sequences of independent random variables have been given by Krengel and Sucheston [**12, 13**], Garling, and Dvoretzky (both in [**13**]), and Hill and Kertz [**8–10**]. Extending these results, Hill [**7**] has shown that

(1) the set of ordered pairs $\{(x, y)\colon x = V(X_0, X_1, \ldots)$ and $y = E(\sup_n X_n)$ for some sequence of independent random variables $X_0, X_1, \ldots$ taking values in $[0, 1]\}$ is precisely the set $\{(x, y)\colon x \leq y \leq 2x - x^2; 0 \leq x \leq 1\}$.

Received by the editors March 9, 1982.

1980 *Mathematics Subject Classification*. Primary 60G40, 60E15; Secondary 60G42, 62L15, 90C99.

Key words and phrases. Optimal stopping, extremal distributions, inequalities for stochastic processes, prophet inequalities, martingales, Markov processes.

©1983 American Mathematical Society
0002-9947/82/0000-0937/$03.25

For other types of processes, few results on such comparisons have been available in the literature. A notable exception is the collection of averages of nonnegative independent random variables and generalizations of this collection investigated by Krengel and Sucheston [**13**] and Brunel and Krengel [**3**]. The main results of the present paper give a complete solution of the comparison of value and expected supremum for the collections of uniformly bounded processes, of uniformly bounded Markov processes, and of uniformly bounded martingales.

These results generalize inequalities (e.g. [**6**, §4]) for uniformly bounded martingales: if $X_0, X_1, \ldots$ is any martingale taking values in $[0, 1]$, then $E(\max_{j \leq n} X_j) \leq x + nx(1 - x^{1/n})$ and $E(\sup_n X_n) \leq x - x \ln x$, where $x = EX_0$ (Corollaries 3.4 and 4.3).

The paper is organized as follows. §2 recalls definitions and background results on the value of a process from classical optimal stopping theory, and presents two propositions which allow reduction from general uniformly bounded sequences to special types of martingales. §§3 and 4 develop the main results on comparison of value and expected supremum for uniformly bounded sequences of finite and of infinite lengths, respectively. In §5 an application of these results is made to determine the advantage of order selection for a gambler in optimal stopping problems.

2. Preliminaries. Throughout this paper the assumption is made that *all random variables take values in* $[0, 1]$. For the extension of the main results to general uniformly bounded sequences, see the Remarks at the end of §§3 and 4.

For subsets A of Ω, I_A denotes the indicator function of A. For random variables X and Y, $X \vee Y$ and $X \wedge Y$ denote the maximum and minimum, respectively, of X and Y. EX is the expectation of X; $E(X \mid Y)$ is the conditional expectation of X given Y; and $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$ is the sigma-field generated by $X_0, \ldots, X_n$. The essential supremum of a collection of random variables $\{X_s\colon s \in S\}$ is a random variable Y (written $Y = \operatorname{ess\,sup}\{X_s\colon s \in S\}$) satisfying (i) $P(Y \geq X_s) = 1$ for every $s \in S$, and (ii) if Y' is any random variable such that $P(Y' \geq X_s) = 1$ for every $s \in S$, then $P(Y' \geq Y) = 1$. From Theorem 1.5 of [**4**], $\operatorname{ess\,sup}\{X_s\colon s \in S\}$ always exists, and if X_s is measurable with respect to the same σ-algebra $\mathcal{F}$ for each $s \in S$, then $\operatorname{ess\,sup}\{X_s\colon x \in S\}$ may also be taken to be $\mathcal{F}$-measurable.

We now recall several definitions and results from classical optimal stopping theory; as a reference see Chapters 3 and 4 of [**4**].

Given the ordered collection of random variables $X_0, X_1, \ldots$, let T denote the set of (a.e.) finite stop rules for $X_0, X_1, \ldots$, and let $T_m = \{t \in T\colon t \geq m\}$.

DEFINITION 2.1. The *value* $V(X_0, X_1, \ldots)$ of $X_0, X_1, \ldots$ is $V(X_0, X_1, \ldots) = \sup\{EX_t\colon t \in T\}$, and $V(X_0, \ldots, X_n) = \sup\{EX_t\colon t \in T,\ t \leq n\}$. The *conditional value of* $X_m, X_{m+1}, \ldots$ *given* $X_0, \ldots, X_j$, $V(X_m, X_{m+1}, \ldots \mid \mathcal{F}_j)$, is given by $V(X_m, X_{m+1}, \ldots \mid \mathcal{F}_j) = \operatorname{ess\,sup}\{E(X_t \mid \mathcal{F}_j)\colon t \in T_m\}$, and $V(X_m, X_{m+1}, \ldots, X_n \mid \mathcal{F}_j) = \operatorname{ess\,sup}\{E(X_t \mid \mathcal{F}_j)\colon t \in T_m,\ t \leq n\}$.

The following two standard results relating these concepts will be useful:

$$V(X_{m+1}, X_{m+2}, \ldots \mid \mathcal{F}_m) = E\big(V(X_{m+1}, X_{m+2}, \ldots \mid \mathcal{F}_{m+1}) \mid \mathcal{F}_m\big) \quad \text{a.e. for all } m \geq 0 \tag{2}$$

and

(3) $V(X_m, X_{m+1}, \ldots \mid \mathcal{F}_m) = X_m \vee V(X_{m+1}, X_{m+2}, \ldots \mid \mathcal{F}_m)$ a.e. for all $m \geqslant 0$.

The first proposition allows a reduction to martingales from arbitrarily-dependent sequences of random variables for the purpose of determining how much larger than the value the expected supremum of a process can be.

PROPOSITION 2.2. *Given random variables $X_0, X_1, \ldots$, there exists a martingale $\hat{X}_0, \hat{X}_1, \ldots$ for which $V(X_0, X_1, \ldots) = V(\hat{X}_0, \hat{X}_1, \ldots)$ and $E(\sup_n X_n) \leqslant E(\sup_n \hat{X}_n)$.*

The proof of Proposition 2.2 is based on three lemmas. The assumption that all random variables take values in $[0, 1]$ is used for the first time in Lemma 2.4 for purposes other than as a guarantee of integrability of random variables.

LEMMA 2.3. *Given $X_0, X_1, \ldots$ and $m \geqslant 0$, define*

$$X'_m = X_m \vee V(X_{m+1}, X_{m+2}, \ldots \mid \mathcal{F}_m).$$

Then $X_0, \ldots, X_{m-1}, X'_m, X_{m+1}, \ldots$ satisfies

(i) $X'_m \geqslant V(X_{m+1}, X_{m+2}, \ldots \mid X_0, \ldots, X_{m-1}, X'_m)$ *a.e.*;

(ii) $V(X_0, X_1, \ldots) = V(X_0, \ldots, X_{m-1}, X'_m, X_{m+1}, \ldots)$; *and*

(iii) $E(\sup_n X_n) \leqslant E(X_0 \vee \cdots \vee X_{m-1} \vee X'_m \vee X_{m+1} \vee \cdots)$.

PROOF. For (iii), notice that $X'_m \geqslant X_m$; (i) and (ii) follow routinely from standard arguments using (2) and (3). □

LEMMA 2.4. *Given $X_0, X_1, \ldots$ and $m \geqslant 0$, define $\beta = \beta_m(X_0, X_1, \ldots, X_m)$ by*

$$\beta = \left[\left(X_m - V(X_{m+1}, \ldots \mid \mathcal{F}_m)\right) / \left(1 - V(X_{m+1}, \ldots \mid \mathcal{F}_m)\right)\right] \cdot I_{(X_m > V(X_{m+1}, \ldots \mid \mathcal{F}_m))}$$

if $V(X_{m+1}, \ldots \mid \mathcal{F}_m) < 1$, and $= 0$ otherwise; and for $k \geqslant m + 1$, define $\tilde{X}_k = \beta + (1 - \beta)X_k$. Then $X_0, \ldots, X_m, \tilde{X}_{m+1}, \tilde{X}_{m+2}, \ldots$ satisfies

(i) $V(\tilde{X}_{m+1}, \tilde{X}_{m+2}, \ldots \mid \mathcal{F}_m) = X_m$ *a.e. on* $\{X_m \geqslant V(\tilde{X}_{m+1}, \tilde{X}_{m+2}, \ldots \mid \mathcal{F}_m)\}$;

(ii) $V(X_0, X_1, \ldots) = V(X_0, \ldots, X_m, \tilde{X}_{m+1}, \tilde{X}_{m+2}, \ldots)$; *and*

(iii) $E(\sup_n X_n) \leqslant E(X_0 \vee \cdots \vee X_m \vee \tilde{X}_{m+1} \vee \tilde{X}_{m+2} \vee \cdots)$.

PROOF. Conclusion (iii) follows since $0 \leqslant \beta \leqslant 1$ implies $\tilde{X}_k \geqslant X_k$ for all $k > m$; (i) and (ii) follow routinely using (2) and (3), and noting that

$$\beta + (1 - \beta)V(X_{m+1}, \ldots \mid \mathcal{F}_m) = X_m$$

a.e. on $\{X_m \geqslant V(X_{m+1}, \ldots \mid \mathcal{F}_m)\} = \{X_m \geqslant V(\tilde{X}_{m+1}, \ldots \mid \mathcal{F}_m)\}$. □

LEMMA 2.5. *Given random variables $X_0, X_1, \ldots$, the following are equivalent*:

(i) $X_0, X_1, \ldots$ *is a martingale*;

(ii) $EX_t = EX_0 = V(X_0, X_1, \ldots)$ *for all $t \in T$; and*

(iii) $X_m = V(X_{m+1}, X_{m+2}, \ldots \mid \mathcal{F}_m)$ *a.e. for all $m \geqslant 0$.*

PROOF. The equivalence of (i) and (ii) is well known (see, for example, [**13**, p. 200]). The equivalence of (ii) and (iii) follows routinely from (2) and (3) (using regular conditional distributions, if necessary, as in §4.3 of [**2**]). □

PROOF OF PROPOSITION 2.2. Given random variables $X_0, X_1, \ldots$, apply Lemmas 2.3 and 2.4 for each $m = 0, 1, 2, \ldots$ to obtain random variables $\hat{X}_0, \hat{X}_1, \ldots$ satisfying: $\hat{X}_k = V(\hat{X}_{k+1}, \hat{X}_{k+2}, \ldots \mid \hat{X}_0, \ldots, \hat{X}_k)$ a.e. for all $k \geqslant 0$; $V(X_0, X_1, \ldots) = V(\hat{X}_0, \hat{X}_1, \ldots)$; and $E(\sup_n X_n) \leqslant E(\sup_n \hat{X}_n)$. By Lemma 2.5, $\hat{X}_0, \hat{X}_1, \ldots$ is a martingale. □

REMARK. It will be seen from Theorem 4.2 and Proposition 4.5 that the inequality in Proposition 2.2 can even be taken to be strict.

The next proposition allows a reduction to martingales of a particularly simple form for the purpose of determining how much larger than EX_0 the expected supremum of a martingale can be.

PROPOSITION 2.6. *Given any martingale $X_0, X_1, \ldots, X_n$ there is a martingale $\hat{X}_0 = X_0, \hat{X}_1, \ldots, \hat{X}_n$ with $P(\hat{X}_{m+1} \geqslant \hat{X}_m) + P(\hat{X}_{m+1} = 0) = 1$ for all $m = 0, 1, \ldots, n-1$, and satisfying $E(\max_{j \leqslant n} \hat{X}_j) \geqslant E(\max_{j \leqslant n} X_j)$.*

PROOF. First replace $X_0, X_1, \ldots, X_n$ by $X_0', \ldots, X_n'$, where $X_j' = X_j$ for $j < n$, and X_n' satisfies $P(X_n' = 1) = X_{n-1} = 1 - P(X_n' = 0)$; note that $X_0', \ldots, X_n'$ is a martingale with $E(\max_{j \leqslant n} X_j') \geqslant E(\max_{j \leqslant n} X_j)$. Let $X_j' = X_n'$ for all $j > n$, let

$$t_1 = \min\{k \geqslant 1 : X_k' = 0 \text{ or } X_k' \geqslant X_0'\},$$

and define $\hat{X}_0 = X_0'$ and $\hat{X}_1 = X_{t_1}'$. Similarly define $\hat{X}_2, \ldots, \hat{X}_n$ (e.g. $\hat{X}_2 = X_{t_2}'$, where $t_2 = \min\{k > t_1 : X_k' = 0 \text{ or } X_k' \geqslant X_{t_1}'\}$). The process $\hat{X}_0, \ldots, \hat{X}_n$ is a martingale (since $t_j < \infty$ a.e. for all j) with $P(\hat{X}_{m+1} \geqslant \hat{X}_m) + P(\hat{X}_{m+1} = 0) = 1$, for all $m = 0, \ldots, n-1$, and satisfying $E(\max_{j \leqslant n} \hat{X}_j) = E(\max_{j \leqslant n} X_j')$. □

3. Prophet inequalities for finite sequences. In this section are given the main result (Theorem 3.2) and resulting inequalities for finite sequences of random variables taking values in [0, 1]. Fix $n \geqslant 1$.

DEFINITION 3.1. C_n denotes the closed, convex set in $\mathbf{R}^2$ given by

$$C_n = \{(x, y) : x \leqslant y \leqslant x(1 + n(1 - x^{1/n})); 0 \leqslant x \leqslant 1\}.$$

(Note that C_1 is the set appearing in (1).)

THEOREM 3.2. *The set of ordered pairs $\{(x, y) : x = V(X_0, \ldots, X_n)$ and $y = E(\max_{j \leqslant n} X_j)$ for some $X_0, \ldots, X_n\}$ is precisely the set C_n.*

PROOF. (For alternative proofs of the martingale claims in the following argument the reader may refer to **[6]**.) First it will be shown that $\{(x, y) : x = V(X_0, \ldots, X_n)$ and $y = E(\max_{j \leqslant n} X_n)$ for some $X_0, \ldots, X_n\}$ is a subset of C_n. Fix any process $X_0, \ldots, X_n$ taking values in [0, 1], and notice that by Propositions 2.2 (letting $X_j = X_n$ for all $j > n$) and 2.6 it suffices to show

(4) if $X_0, \ldots, X_n$ is a martingale with $P(X_{m+1} \geqslant X_m) + P(X_{m+1} = 0) = 1$ for all $m = 0, \ldots, n-1$, then $E(\max_{j \leqslant n} X_j) \leqslant EX_0(1 + n(1 - (EX_0)^{1/n}))$.

By Jensen's inequality, (4) will follow once it is shown that

(5) if $(X_0, \mathcal{G}_0), \ldots, (X_k, \mathcal{G}_k)$ is a martingale with $P(X_{m+1} \geqslant X_m) + P(X_{m+1} = 0) = 1$ for all $m = 0, \ldots, k-1$, then $E(X_0 \vee \cdots \vee X_k \mid \mathcal{G}_0) \leqslant X_0 + kX_0(1 - X_0^{1/k})$ a.e.

The proof of (5) will be by induction on k. First note that, by the martingale property, $X_j = 0$ a.e. on $\{X_0 = 0\}$ for all $j = 1, \ldots, n$, so it remains only to show (5) holds a.e. on $\{X_0 > 0\}$.

For $k = 1$, check that

$$\begin{aligned} E(X_0 \vee X_1 \mid \mathcal{G}_0) &= E\big(X_0 I_{(X_1=0)} + X_1 I_{(X_1 \geq X_0)} \mid \mathcal{G}_0\big) \\ &= X_0 + E\big((X_1 - X_0) I_{(X_0 \leq X_1)} \mid \mathcal{G}_0\big) \\ &\leq X_0 + E\big((1 - X_0) X_1 \mid \mathcal{G}_0\big) = X_0 + (1 - X_0) E(X_1 \mid \mathcal{G}_0) \\ &= X_0 + X_0(1 - X_0) \quad \text{a.e. on } \{X_0 > 0\}. \end{aligned}$$

Assume (5) is true for $k = m$, and show it is true for $k = m + 1$ as follows: calculate

$$\begin{aligned} E(X_0 \vee \cdots \vee X_{m+1} \mid \mathcal{G}_0) &= E\big(X_0 I_{(X_1=0)} + (X_1 \vee \cdots \vee X_{m+1}) I_{(X_0 \leq X_1)} \mid \mathcal{G}_0\big) \\ &= X_0 P(X_1 = 0 \mid \mathcal{G}_0) + E\big(E(X_1 \vee \cdots \vee X_{m+1} \mid \mathcal{G}_1) I_{(X_0 \leq X_1)} \mid \mathcal{G}_0\big) \\ &\leq X_0 P(X_1 = 0 \mid \mathcal{G}_0) + E\big(\big(X_1 + m X_1 (1 - X_1^{1/m})\big) I_{(X_0 \leq X_1)} \mid \mathcal{G}_0\big) \\ &= X_0 + E\big(\big[(m+1) - m X^{1/m} - (X_0/X_1)\big] I_{(X_0 \leq X_1)} X_1 \mid \mathcal{G}_0\big) \\ &\leq X_0 + E\big((m+1)(1 - X_0^{1/(m+1)}) I_{(X_0 \leq X_1)} X_1 \mid \mathcal{G}_0\big) \\ &= X_0 + (m+1)(1 - X_0^{1/(m+1)}) E(X_1 \mid \mathcal{G}_0) \\ &= X_0 + (m+1) X_0 (1 - X_0^{1/(m+1)}) \quad \text{a.e. on } \{X_0 > 0\}, \end{aligned}$$

where the first inequality follows from the induction hypothesis, and the second inequality by maximizing the function $f(x) = (m+1) - m x^{1/m} - w/x$ for $x > w > 0$. This establishes (5), and the "$\subseteq C_n$" part of the proof.

It remains to show that for each point $(x, y) \in C_n$, there is a process $X_0, X_1, \ldots, X_n$ taking values in $[0, 1]$ and satisfying $x = V(X_0, \ldots, X_n)$ and $y = E(\max_{j \leq n} X_j)$. This follows immediately from the following proposition, which identifies a particularly simple and well-structured class of extremal processes for C_n. □

PROPOSITION 3.3. *For every point $(x, y) \in C_n$, there is a sequence of random variables $X_0, \ldots, X_n$, each taking at most two values, which is both Markov and a martingale, and which satisfies $V(X_0, \ldots, X_n) = x$ and $E(\max_{j \leq n} X_j) = y$.*

PROOF. For $(x, y) = (0, 0)$ or $(1, 1)$, take $X_0 = \cdots = X_n = x$ a.e. Let $(x, y) \in C_n \backslash \{(0,0), (1,1)\}$, and define the process $X_0', \ldots, X_n'$ by $X_0' \equiv x$ and

$$P\big(X_m' = x^{(n-m)/n} \mid \mathcal{F}_{m-1}'\big) = x^{1/n} = 1 - P\big(X_m' = 0 \mid \mathcal{F}_{m-1}'\big) \quad \text{for } m = 1, \ldots, n.$$

Then $X_0', \ldots, X_n'$ is both Markov and a martingale (so $V(X_0', \ldots, X_n') = x$) and satisfies $E(\max_{j \leq n} X_j') = x(1 + n(1 - x^{1/n}))$. (This example essentially appears in [**6**], and is included here for ease of reference.)

Now let $a = (y - x)/(nx(1 - x^{1/n}))$, and $b = 1 - a$, and define $X_0, \ldots, X_n$ by $X_m = a X_m' + bx$ for $m = 0, \ldots, n$. Then $X_0, \ldots, X_n$ is again Markov and a martingale with $V(X_0, \ldots, X_n) = x$ and $E(\max_{j \leq n} X_j) = y$. □

The next three inequalities follow immediately from Theorem 3.2.

COROLLARY 3.4 (DUBINS AND PITMAN [6]). *Let $X_0, \dots, X_n$ be a martingale taking values in* [0, 1]. *Then*

$$E\Big(\max_{j \le n} X_j\Big) \le EX_0\Big(1 + n\big(1 - (EX_0)^{1/n}\big)\Big). \tag{6}$$

COROLLARY. 3.5. *Let $X_0, \dots, X_n$ be any random variables taking values in* [0, 1]. *Then*

$$E\Big(\max_{j \le n} X_j\Big) \le V(X_0, \dots, X_n) + (n/(n+1))^{n+1} \tag{7}$$

and

$$E\Big(\max_{j \le n} X_j\Big) < (n+1)V(X_0, \dots, X_n) \quad \text{if } P\Big(\max_{j \le n} X_j > 0\Big) > 0. \tag{8}$$

The following two results are immediate consequences of Theorem 3.2 and Proposition 3.3.

THEOREM 3.6. *The set of ordered pairs $\{(x, y)\colon x = EX_0$ and $y = E(\max_{j \le n} X_j)$ for some martingale $X_0, \dots, X_n\}$ is precisely the set C_n.*

THEOREM 3.7. *The set of ordered pairs $\{(x, y)\colon x = V(X_0, \dots, X_n)$ and $y = E(\max_{j \le n} X_j)$ for some Markov process $X_0, \dots, X_n\}$ is precisely the set C_n.*

REMARKS. Inequalities (6) and (7) are attained: for (6), choose $X_0', \dots, X_n'$ as in the proof of Proposition 3.3 (with $y = x(1 + n(1 - x^{1/n}))$; for (7), require further that $x = (n/(n+1))^n$). By considering $X_0', \dots, X_n'$ with x sufficiently close to zero, (8) can be seen to be sharp. The weak inequality version of (8) and the process $X_0', \dots, X_n'$ of Proposition 3.3 have appeared on p. 514 of Blackwell and Dubins [**1**] and in Proposition 1 of Hill and Kertz [**8**].

For the collection of random variables $X_0, \dots, X_n$ taking values in $[a, b]$, $-\infty < a < b < \infty$, the set of ordered pairs $\{(x, y)\colon x = V(X_0, \dots, X_n)$ and $y = E(\max_{j \le n} X_j)$ for some $X_0, \dots, X_n\}$ is precisely the set

$$\Big\{(x, y)\colon x \le y \le x + n(x-a)\big(1 - ((x-a)/(b-a))^{1/n}\big);\ a \le x \le b\Big\}$$

(similarly for Markov processes and martingales taking values in $[a, b]$).

4. Prophet inequalities for infinite sequences. In this section the analogous results for infinite sequences of random variables taking values in [0, 1] are developed.

DEFINITION 4.1. C denotes the convex set in $\mathbf{R}^2$ given by $C = \{(x, y)\colon x \le y < x - x \ln x; 0 < x < 1\} \cup \{(0,0), (1,1)\}$. (In terms of the sets C_n of Definition 3.1, $C = \bigcup_{n=1}^{\infty} C_n$.)

THEOREM 4.2. *The set of ordered pairs $\{(x, y)\colon x = V(X_0, X_1, \dots)$ and $y = E(\sup_n X_n)$ for some process $X_0, X_1, \dots\}$ is precisely the set C.*

PROOF. Since $V(X_0, X_1, \dots) = \lim_{n \to \infty} V(X_0, \dots, X_n)$ and $E(\sup_n X_n) = \lim_{n \to \infty} E(\max_{j \le n} X_j)$, it follows immediately from Theorem 3.2 that

$$\begin{aligned} C &\subseteq \Big\{(x, y)\colon x = V(X_0, X_1, \dots) \text{ and } y = E\Big(\sup_n X_n\Big) \text{ for some } X_0, X_1, \dots\Big\} \\ &\subseteq \{(x, y)\colon x \le y \le x - x \ln x; 0 < x < 1\} \cup \{(0,0), (1,1)\}. \end{aligned} \tag{9}$$

The remainder of the proof consists solely of showing that the inequality in question is strict, that is, for every sequence $X_0, X_1, \ldots$ with $V(X_0, X_1, \ldots) = x \in (0, 1)$ and $y = E(\sup_n X_n)$ it follows that $y < x - x \ln x$.

Fix $X_0, X_1, \ldots$ with $V(X_0, X_1, \ldots) = x \in (0, 1)$. The proof will be complete once it is shown that there exists a sequence $\hat{X}_0, \hat{X}_1, \ldots$ with $V(\hat{X}_0, \hat{X}_1, \ldots) = x$ but with $E(\sup_n \hat{X}_n) > E(\sup_n X_n)$.

First, notice that for every martingale $X_0', X_1', \ldots,$

$$E\Big(\sup_n X_n'\Big) \leqslant E(X_0')(1 - \ln EX_0'); \tag{10}$$

$$E\Big(\sup_n X_n' \mid X_0'\Big) \leqslant X_0'(1 - \ln X_0') \quad \text{a.e. on } \{X_0' > 0\}; \tag{11}$$

and if X_0' is not a.e. constant, then there is a $\delta > 0$ with

$$E\Big(\sup_n X_n'\Big) \leqslant E(X_0'(1 - \ln X_0')) \leqslant EX_0'(1 - \ln EX_0') - \delta, \tag{12}$$

where (10) follows from (9) since $V(X_0', X_1', \ldots) = EX_0'$, (11) follows from (10) (using regular conditional distributions, if necessary, as in §4.3 of [**2**]), and (12) follows from (11) and the strict convexity of the function $f(x) = x - x \ln x$.

Now, it may be assumed (from Proposition (2.2)) that the given sequence $X_0, X_1, \ldots$ is a martingale, and even that $X_0 \equiv x$ (otherwise consider the sequence $X_{-1} \equiv x, X_0, X_1, \ldots$).

Case 1. $P(X_1 = x/\alpha) = \alpha = 1 - P(X_1 = 0)$ for some $0 < x \leqslant \alpha < 1$.

First it will be shown that

$$E\Big(\sup_n X_n\Big) \leqslant x(2 - \ln x - \alpha + \ln \alpha). \tag{13}$$

To establish (13), let $x_1 = x/\alpha$, and calculate

$$\begin{aligned} E\Big(\sup_n X_n\Big) &= x(1 - \alpha) + \int_{X_1 = x_1} E(X_1 \vee X_2 \vee \cdots \mid X_1)\, dP \\ &\leqslant x(1 - \alpha) + \int_{X_1 = x_1} X_1(1 - \ln X_1)\, dP \\ &= x(1 - \alpha) + \alpha(x_1 - x_1 \ln x_1) = x(2 - \ln x - \alpha + \ln \alpha), \end{aligned}$$

where the inequality follows from (11) since $X_1, X_2, \ldots$ is a martingale.

Next, fix $\varepsilon > 0$ with $\varepsilon < x(-1 - \ln \alpha + \alpha)/2$. Since $\ln y - y$ increases to -1 as y increases to 1, there exists $\hat{\alpha} \in (\alpha, 1)$ with

$$\ln \hat{\alpha} - \hat{\alpha} > \ln \alpha - \alpha + \varepsilon/x. \tag{14}$$

From (9), for $\hat{x}_1 = x/\hat{\alpha}$, there exists a martingale $\tilde{X}_1 \equiv \hat{x}_1, \tilde{X}_2, \ldots$ satisfying $E(\sup_{n \geqslant 1} \tilde{X}_n) \geqslant \hat{x}_1(1 - \ln \hat{x}_1) - \varepsilon$. Define $\hat{X}_0, \hat{X}_1, \ldots$ by $\hat{X}_0 \equiv x$; $P(\hat{X}_1 = \hat{x}_1) = \hat{\alpha} = 1 - P(\hat{X}_1 = 0)$; and $P(\hat{X}_j = 0 \mid \hat{X}_1 = 0) = 1 = P(\hat{X}_j = \tilde{X}_j \mid \hat{X}_1 = \hat{x}_1)$ a.e. for $j > 1$. The process $\hat{X}_0, \hat{X}_1, \ldots$ is a martingale satisfying

$$E\Big(\sup_n \hat{X}_n\Big) \geqslant x(2 - \ln x - \hat{\alpha} + \ln \hat{\alpha}) - \varepsilon, \tag{15}$$

since

$$\begin{aligned} E\Big(\sup_n \hat{X}_n\Big) &= x(1-\hat{\alpha}) + \int_{\hat{X}_1=\hat{x}_1} E\big(\hat{X}_1 \vee \hat{X}_2 \vee \cdots \mid \hat{X}_1\big)\, dP \\ &= x(1-\hat{\alpha}) + \hat{\alpha} E\big(\tilde{X}_1 \vee \tilde{X}_2 \vee \cdots\big) \\ &\geqslant x(1-\hat{\alpha}) + \hat{\alpha}(\hat{x}_1 - \hat{x}_1 \ln \hat{x}_1 - \varepsilon) \\ &\geqslant x(2 - \ln x - \hat{\alpha} + \ln \hat{\alpha}) - \varepsilon, \end{aligned}$$

where the first inequality follows from the construction of $\hat{X}_1, \hat{X}_2, \ldots$ and the second inequality follows because $\hat{x}_1 = x/\hat{\alpha}$ and $0 < \hat{\alpha} < 1$.

Now by (13), (14) and (15) we have

$$\begin{aligned} E\Big(\sup_n X_n\Big) &\leqslant x(2 - \ln x - \alpha + \ln \alpha) \\ &< x(2 - \ln x - \hat{\alpha} + \ln \hat{\alpha} - \varepsilon/x) \leqslant E\Big(\sup_n \hat{X}_n\Big), \end{aligned}$$

and Case 1 is completed.

General case (*Reduction to Case* 1). By an argument similar to that in Proposition 2.6, it may be assumed that, for some $\alpha \in (0, 1)$, $P(X_1 \geqslant x) = \alpha = 1 - P(X_1 = 0)$. It will be shown that there is a martingale $\hat{X}_0 \equiv x$, $\hat{X}_1, \ldots$ with $P(\hat{X}_1 = x/\alpha) = \alpha = 1 - P(\hat{X}_1 = 0)$ and satisfying $E(\sup_n \hat{X}_n) \geqslant E(\sup_n X_n)$, thereby reducing the general case to Case 1.

Assume X_1 is not constant a.e. on $\{X_1 \geqslant x\}$, otherwise Case 1 applies (except in the degenerate case $X_1 = x$ a.e., which is solved by deleting X_1 and considering X_0, $X_2, X_3, \ldots$).

For some $\delta > 0$,

$$\begin{aligned} E\Big(\sup_n X_n\Big) &= x(1-\alpha) + \int_{X_1 \geqslant x} E\big(X_1 \vee X_2 \vee \cdots \mid X_1\big)\, dP \\ &\leqslant x(1-\alpha) + \int_{X_1 \geqslant x} X_1(1 - \ln X_1)\, dP \\ &= x(1-\alpha) + \alpha \int_{X_1 \geqslant x} X_1(1 - \ln X_1)\, d(P/\alpha) \\ &\leqslant x(1-\alpha) + \alpha(x_1 - x_1 \ln x_1 - \delta) \\ &\leqslant x(1-\alpha) + \alpha(x_1 - x_1 \ln x_1) - \delta, \end{aligned} \tag{16}$$

where $x_1 = \alpha^{-1} \int_{X_1 \geqslant x} X_1 = x/\alpha$, and where the first inequality follows from (11), and the second inequality from the strict concavity of the function $x - x \ln x$. From (9) and Proposition 3.3 there is a martingale $\hat{X}_0 \equiv x$, $\hat{X}_1, \ldots$ with $P(\hat{X}_1 = x_1) = \alpha = 1 - P(\hat{X}_1 = 0)$ and satisfying $E(\hat{X}_1 \vee \hat{X}_2 \vee \cdots \mid \hat{X}_1 = x_1) \geqslant x_1 - x_1 \ln x_1 - (\delta/2\alpha)$.Thus

$$\begin{aligned} E\Big(\sup_n \hat{X}_n\Big) &= x(1-\alpha) + \int_{\hat{X}_1 = x_1} E\big(\hat{X}_1 \vee \hat{X}_2 \vee \cdots \mid \hat{X}_1 = x_1\big)\, dP \\ &\geqslant x(1-\alpha) + \alpha(x_1 - x_1 \ln x_1 - (\delta/2\alpha)) \geqslant E\Big(\sup_n X_n\Big), \end{aligned}$$

where the second inequality follows from (16). This completes the reduction to Case 1, and the proof. □

COROLLARY 4.3 (CF. [**6**, p. 227]). *Let $X_0, X_1, \ldots$ be a martingale taking values in* $[0, 1]$ *with* $0 < EX_0 < 1$. *Then*

$$E\left(\sup_n X_n\right) < EX_0(1 - \ln EX_0). \tag{17}$$

REMARK. Under the conditions of Corollary 4.3 with $EX_0 = x$, the weak expectation inequality $E(\sup_n X_n) \leqslant x - x \ln x$ can also be proved from the weak distributional inequality $P(\sup_n X_n \geqslant y) \leqslant x/y$ for $y \geqslant x$ [**5**, p. 314] as follows:

$$E\left(\sup_n X_n\right) = \int_0^\infty P\left(\sup_n X_n \geqslant \lambda\right) d\lambda \leqslant \int_0^x 1\, d\lambda + \int_x^1 x/\lambda\, d\lambda = x - x \ln x.$$

COROLLARY 4.4. *Let $X_0, X_1, \ldots$ be any random variables taking values in* $[0, 1]$. *Then*

$$E\left(\sup_n X_n\right) < V(X_0, X_1, \ldots) + e^{-1}. \tag{18}$$

PROOF. Note that $y = x + e^{-1}$ is tangent to the curve $y = x - x \ln x$ at $x = e^{-1}$; apply Theorem 4.2. □

PROPOSITION 4.5. *For each $(x, y) \in C$, there is a sequence of random variables $X_0, X_1, \ldots$, each taking at most two values, which is both Markov and a martingale, and which satisfies $EX_0 = x$ and $E(\sup_n X_n) = y$.*

PROOF. Apply Proposition 3.3 and Theorem 4.2 □

The following two theorems follow immediately from Theorem 4.2 and Proposition 4.5.

THEOREM 4.6. *The set of ordered pairs $\{(x, y)\colon x = EX_0$ and $y = E(\sup_n X_n)$ for some martingale $X_0, X_1, \ldots\}$ is precisely the set C.*

THEOREM 4.7. *The set of ordered pairs $\{(x, y)\colon x = V(X_0, X_1, \ldots)$ and $y = E(\sup_n X_n)$ for some Markov process $X_0, X_1, \ldots\}$ is precisely the set C.*

REMARKS. Inequalities (17) and (18) are sharp. To come arbitrarily close to equality in (17), let $x = EX_0$ and $y_\varepsilon = EX_0(1 - EX_0) - \varepsilon$ and choose the process $X_0, X_1, \ldots$ associated with the point $(x, y_\varepsilon) \in C$ as in Proposition 3.3; for (18) require further that $x = e^{-1}$.

For the collection of infinite sequences $X_0, X_1, \ldots$ taking values in $[a, b]$, $-\infty < a < b < \infty$, the set of ordered pairs $\{(x, y)\colon x = V(X_0, X_1, \ldots)$ and $y = E(\sup_n X_n)$ for some $X_0, X_1, \ldots\}$ is precisely the set

$$\{(x, y)\colon x \leqslant y < x + (x - a) \ln((x - a)/(b - a)); a < x < b\} \cup \{(a, a), (b, b)\}.$$

5. Applications to order selection in optimal stopping problems. In this section the results of the two previous sections are used to solve a nonprophet problem, in this case to determine the advantage a player may obtain by rearranging the order of

observation of a given sequence. Recall that the assumption that all random variables take values in [0, 1] is still in effect. Fix $n \geqslant 1$.

Let $\mathcal{P}_n$ denote the permutations of $\{0, 1, \ldots, n\}$, and $\mathcal{P}$ the permutations of $\{0, 1, 2, \ldots\}$.

DEFINITION 5.1. For any sequence of random variables $X_0, X_1, \ldots,$

$$U(\{X_0, X_1, \ldots, X_n\}) = \max\{V(X_{\pi(0)}, \ldots, X_{\pi(n)}) : \pi \in \mathcal{P}_n\}, \text{ and}$$

$$U(\{X_0, X_1, \ldots\}) = \sup\{V(X_{\pi(0)}, X_{\pi(1)}, \ldots) : \pi \in \mathcal{P}\}.$$

$U(\{X_0, X_1, \ldots\})$ represents the optimal expected gain of a player free to select (deterministically) the order of observation of the collection of random variables $\{X_0, X_1, \ldots\}$, and free to stop whenever he pleases. As is shown in Theorem 3.11 of [7], $U(\{X_0, X_1, \ldots\})$ even equals $W(\{X_0, X_1, \ldots\})$, the optimal expected gain of a player free to select the order of observation randomly (as a function of past observations only).

The main results of this section, Theorem 5.2 and 5.3, give complete solutions to the question of the advantage a player may obtain by rearranging the order of observations of given uniformly bounded finite, and infinite, sequences of random variables. Recall C_n and C in Definitions 3.1 and 4.1.

THEOREM 5.2. *The set of ordered pairs* $\{(x, y)\colon x = V(X_0, \ldots, X_n)$ *and* $y = U(\{X_0, \ldots, X_n\})$ *for some process* $X_0, X_1, \ldots, X_n\}$ *is precisely the set* C_n.

THEOREM 5.3. *The set of ordered pairs* $\{(x, y)\colon x = V(X_0, X_1, \ldots)$ *and* $y = U(\{X_0, X_1, \ldots\})$ *for some process* $X_0, X_1, \ldots\}$ *is precisely the set* C.

PROOFS OF THEOREMS 5.2 AND 5.3. For Theorem 5.2, first notice that $V(X_0, X_1, \ldots, X_n) \leqslant U(\{X_0, X_1, \ldots, X_n\}) \leqslant E(\max_{j \leqslant n} X_j)$. Next, it is shown that the random variables $X_m = aX'_m + bx$, $m = 0, \ldots, n$, of Proposition 3.3 satisfy $V(X_n, X_{n-1}, \ldots, X_1, X_0) = U(\{X_0, \ldots, X_n\}) = E(\max_{j \leqslant n} X_j)$. The random variables X'_m, $m = 0, \ldots, n$, of Proposition 3.3 satisfy

$$x + (n-1)x(1 - x^{1/n}) = E\Big(\max_{j \leqslant n} X'_j\Big) \geqslant U(\{X'_0, \ldots, X'_n\})$$

$$\geqslant V(X'_n, X'_{n-1}, \ldots, X'_1, X'_0) = E(X'_{n-t}) = x + (n-1)x(1 - x^{1/n}),$$

where t is the stop rule for $X'_{n-j}, j = 0, \ldots, n$, given by $t = \min\{j\colon X'_{n-j} \neq 0\}$. Thus, $V(X_n, \ldots, X_0) = aV(X'_n, \ldots, X'_0) + bx = aE(\max_{j \leqslant n} X'_j) + bx = E(\max_{j \leqslant n} X_j)$, and $V(X_n, \ldots, X_0) = U(\{X_0, \ldots, X_n\}) = E(\max_{j \leqslant n} X_j)$. This concludes the proof of Theorem 5.2; the proof of Theorem 5.3 is similar. □

The analogs of the corollaries and theorems for Markov sequences and martingales found at the ends of §§3 and 4 also follow easily, as do the corresponding results for random variables taking values in $[a, b]$. The analog of (18), for example, is

$$U(\{X_0, X_1, \ldots\}) < V(X_0, X_1, \ldots) + e^{-1}. \tag{19}$$

A probabilistic interpretation of (19) is that a player should never pay more than e^{-1} for the privilege of rearranging the order of observations of a given sequence of random variables (taking values in [0, 1]).

Another similar application of prophet inequalities to a nonprophet problem, that of determination of the advantage of using nonmeasurable stop rules (see [**11**]), is also possible.

REFERENCES

1. D. Blackwell and L. E. Dubins, *A converse to the dominated convergence theorem*, Illinois J. Math. **7** (1963), 508–514.

2. L. Breiman, *Probability*, Addison-Wesley, Reading, Mass, 1968.

3. A. Brunel and U. Krengel, *Parier avec un prophete dans le cas d'un processus sous-additif*, C. R. Acad. Sci. Paris Ser. A–B **288** (1979), 57–60.

4. Y. Chow, H. Robbins and D. Siegmund, *Great expectations: the theory of optimal stopping*, Houghton Mifflin, Boston, Mass., 1971.

5. J. L. Doob, *Stochastic processes*, Wiley, New York, 1953.

6. L. E. Dubins and J. Pitman, *A maximal inequality for skew fields*, Z. Wahrsch. Verw. Gebiete **52** (1980), 219–227.

7. T. Hill, *Prophet inequalities and order selection in optimal stopping problems*, Proc. Amer. Math. Soc. **88** (1983), 131–137.

8. T. Hill and R. Kertz, *Ratio comparisons of supremum and stop rule expectations*, Z. Wahrsch. Verw. Gebiete **56** (1981), 283–285.

9. _____, *Additive comparisons of stop rule and supremum expectations of uniformly bounded independent random variables*, Proc. Amer. Math. Soc. **83** (1981), 582–585.

10. _____, *Comparisons of stop rule and supremum expectations of i.i.d. random variables*, Ann. Probab. **10** (1982), 336–345.

11. T. Hill and V. Pestien, *The advantage of using non-measurable stop rules*, Ann. Probab. (to appear).

12. U. Krengel and L. Sucheston, *Semiamarts and finite values*, Bull. Amer. Math. Soc. **83** (1977), 745–747.

13. _____, *On semiamarts, amarts, and processes with finite value*, Adv. in Prob. Related Topics **4** (1978), 197–266.

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA, GEORGIA 30332