The Annals of Probability
1985, Vol. 13, No. 2, 566–571

OPTIMAL STOPPING OF INDEPENDENT RANDOM VARIABLES AND MAXIMIZING PROPHETS

BY D. P. KENNEDY

University of Cambridge

The prophet inequality for a sequence of independent nonnegative random variables shows that the ratio of the mean of the maximum of the sequence to the optimal expected return using stopping times is always bounded by 2; i.e., on average, the proportional advantage of a prophet with complete foresight over a gambler using nonanticipating stopping rules is at most 2. Here, an inequality linking the mean of the sum of the k largest order statistics of the sequence and the optimal expected return is derived. This implies that if the k largest order statistics are close to the maximum in mean then the proportional advantage of the prophet is at most of order $(k+1)/k$. An extension of the additive prophet inequality for uniformly bounded independent random variables is also given.

1. Introduction. Let $\{X_r, r \geq 1\}$ be a sequence of independent random variables on a probability space $(\Omega, \mathscr{F}, P)$ and let $\mathscr{T}$ represent the set of (finite-valued) stopping times with respect to the natural filtration. Krengel and Sucheston [6], [7], showed that if the random variables X_r are nonnegative, then there exists a universal constant, C, such that

$$E(\sup_{r\geq 1} X_r) \leq C \sup_{T\in\mathscr{T}} EX_T. \tag{1.1}$$

Furthermore, Garling (cf. [7]) showed that $C = 2$ and that 2 is the best possible constant in the sense that the upper bound in (1.1) is attained for certain random variables. This and similar inequalities linking the mean of the maximum of the sequence and the optimum of the stopped sequence have come to be called "prophet" inequalities as $E(\sup_{r\geq 1} X_r)$ represents the expected return to a prophet (a player with complete foresight) in a game where the rewards are represented by the sequence $\{X_r, r \geq 1\}$ while $\sup_{T\in\mathscr{T}} EX_T$ is the optimal expected return of a gambler using nonanticipating stopping rules. Thus (1.1) shows that, on average, the gambler does proportionally at least half as well as the prophet.

In a recent series of papers Hill and Kertz ([2], [3], [4], [5]) have considered various extensions and refinements of (1.1). In particular ([4]), they have shown that if the random variables $\{X_r, r \geq 1\}$ are uniformly bounded, taking values in $[a, b]$, then

$$E(\sup_{r\geq 1} X_r) - \sup_{T\in\mathscr{T}} EX_T \leq (b-a)/4, \tag{1.2}$$

and again this is the best possible such inequality.

While (1.1) and (1.2) give tight bounds on how the gambler fares relative to the prophet, there are situations where, for example, the gambler does very much better than half the return of the prophet and in this paper we attempt to pin

Received December 1983; revised May 1984.
AMS 1980 *subject classifications*. Primary 60G40, 62L15.
Key words and phrases. Optimal stopping, prophet inequality, order statistics.

this down by relating the gambler's return to order statistics of the sequence $\{X_r, r \geq 1\}$ other than the maximum. For each $n \geq 1$ and $1 \leq k \leq n$, let $M_{k,n}$ be the kth largest order statistic among $X_1, \cdots, X_n$, i.e.,

$$M_{k,n} = \vee \wedge_{i=1}^{k} X_{r_i}$$

where the maximum extends over all $r_1, \cdots, r_k$ with $1 \leq r_1 < r_2 < \cdots < r_k \leq n$. Let $M_k = M_{k,\infty} = \lim_{n\to\infty} M_{k,n}$ be the kth largest order statistic from the whole sequence $\{X_r, r \geq 1\}$, so $M_1 = \sup_{r\geq 1} X_r$. We will prove the following generalization of (1.1).

THEOREM 1. *If $\{X_r, r \geq 1\}$ is a sequence of independent nonnegative random variables, then for each $k \geq 1$,*

$$E[\textstyle\sum_{i=1}^{k} M_i] \leq (k + 1)\sup_{T\in\mathscr{T}} EX_T, \tag{1.3}$$

and the inequality is the best possible.

What (1.3) shows is that if there are k order statistics which are close to $\sup_{r\geq 1} X_r$ in mean, then the proportional advantage of the prophet is not 2 but $(k + 1)/k$. Of course, this reduces to (1.1) in the case $k = 1$. The analogous result to (1.2) becomes the following.

THEOREM 2. *If $\{X_r, r \geq 1\}$ is a sequence of independent random variables taking values in $[a, b]$ then for each $k \geq 1$,*

$$(1/k)E[\textstyle\sum_{i=1}^{k} M_i] - \sup_{T\in\mathscr{T}} EX_T \leq (b - a)/(k + 1)^{(k+1)/k}. \tag{1.4}$$

Again, we observe that (1.4) reduces to (1.2) in the case $k = 1$, though for $k > 1$ it is not clear whether the bound in (1.4) is tight. Unlike (1.2) the inequality of Theorem 2 may be vacuous in that it is possible that the left-hand side of (1.4) is negative for $k > 1$ (consider the trivial case $X_1 \equiv 1$, $X_r \equiv 0$, $r > 1$). We present the proofs in Sections 2 and 3 and we note that Theorem 1 could be deduced from the proof of Theorem 2 (in particular, from Corollary 3, below) but we present a direct argument as it suggests the proof of the second result.

2. Proof of Theorem 1. First note that for nonnegative real numbers α, β, γ we have

$$\alpha \vee \beta + \gamma \leq \alpha + \beta \vee \gamma, \quad \text{for} \quad \alpha \geq \gamma. \tag{2.1}$$

Assume that $X_r \geq 0$ for each r, and for $n \geq r$, let

$$v_{r,n} = \sup\{EX_T\colon T \in \mathscr{T}, r \leq T \leq n\}$$

with $v_{r,n} = 0$ for $r > n$. Then (cf. Chow, Robbins and Siegmund [1]) $v_{n,n} = EX_n$, and $v_{r,n} = E(X_r \vee v_{r+1,n})$, $r \leq n$. Note that

$$\lim_{n\to\infty} v_{r,n} = \sup\{EX_T\colon T \in \mathscr{T}, r \leq T < \infty\} = v_r,$$

say, with $v_r = E(X_r \vee v_{r+1})$ and $v_1 = \sup_{T\in\mathscr{T}} EX_T$. Observe that $v_{1,n} \geq v_{2,n} \geq \cdots \geq v_{n,n}$, for each n. The proof makes repeated use of (2.1) and the following obvious equality. For any α, and $n \geq k$,

$$\sum_{i=1}^{k} \alpha \vee M_{i,n} = \sum_{i=1}^{k-1} \alpha \vee M_{i,n-1} + \alpha \vee M_{k,n-1} \vee X_n, \tag{2.2}$$

where by convention we take $M_{i,n} = 0$ for $i > n$, and the empty sum to be zero. Now, for fixed $n \geq k$, we will show by backwards induction on $r = n, n-1, \cdots, k$ that

$$E[\sum_{i=1}^{k} M_{i,n}] \leq E[\sum_{i=1}^{k} v_{r+2,n} \vee M_{i,r} + v_{r+1,n}] \tag{2.3}$$

and furthermore, the right-hand side of (2.3) is decreasing in r. Clearly (2.3) is true (with equality) for $r = n$; for $n \geq r > k$, using the fact that $v_{r,n}$ is decreasing in r, we have

$$\begin{aligned} E[\textstyle\sum_{i=1}^{k} v_{r+2,n} \vee M_{i,r} + v_{r+1,n}] & \\ &\leq E[\textstyle\sum_{i=1}^{k} v_{r+1,n} \vee M_{i,r} + v_{r+1,n}] \\ &= E[\textstyle\sum_{i=1}^{k-1} v_{r+1,n} \vee M_{i,r-1} + v_{r+1,n} \vee M_{k,r-1} \vee X_r + v_{r+1,n}], \end{aligned}$$

the second relation coming from (2.2). Now, take $\alpha = v_{r+1,n} \vee M_{k,r-1}$, $\beta = X_r$ and $\gamma = v_{r+1,n}$ in (2.1) and we see that this expression is bounded above by

$$E[\sum_{i=1}^{k} v_{r+1,n} \vee M_{i,r-1} + X_r \vee v_{r+1,n}] = E[\sum_{i=1}^{k} v_{r+1,n} \vee M_{i,r-1} + v_{r,n}],$$

giving the inductive step. Setting $r = k$ in (2.3) and observing that

$$\sum_{i=1}^{k} v_{k+2,n} \vee M_{i,k} = \sum_{i=1}^{k} v_{k+2,n} \vee X_i$$

gives

$$E[\sum_{i=1}^{k} M_{i,n}] \leq E[\sum_{i=1}^{k} v_{k+2,n} \vee X_i + v_{k+1,n}].$$

Letting $n \to \infty$ and using the fact that $v_i \geq E(X_i \vee v_{k+2})$ for $i \leq k$, gives

$$\begin{aligned} E[\textstyle\sum_{i=1}^{k} M_i] &\leq E[\textstyle\sum_{i=1}^{k} v_{k+2} \vee X_i + v_{k+1}] \leq \textstyle\sum_{i=1}^{k+1} v_i \\ &\leq (k+1)v_1 = (k+1)\sup_{T\in\mathscr{T}} EX_T, \end{aligned} \tag{2.4}$$

which establishes (1.3).

To see that the inequality (1.3) is tight, take $X_1 = X_2 = \cdots = X_k = 1$, $X_r = 0$, $r > k+1$ and $X_{k+1} = 1/\delta$ with probability δ and $X_{k+1} = 0$ with probability $1 - \delta$. Then $v_1 = 1$ and $E[\sum_{i=1}^{k} M_i] = k + 1 - \delta \to k + 1$ as $\delta \to 0$.

It should be noted that (1.3) provides an improvement on the inequality (1.1) only if $EM_2 > \frac{1}{2}EM_1$. In the above example the ratio EM_2/EM_1 tends to $\frac{1}{2}$ and so the question is raised as to whether, say in the case $k = 2$, (1.3) is the best bound that may be obtained for $EM_1 < 2EM_2$. The answer is no, as is shown by the following result, the argument for which was suggested by an Associate Editor of a first draft of this work.

PROPOSITION. *If $\{X_r, r \geq 1\}$ is a sequence of independent nonnegative random variables then*

$$EM_1 \leq ((1+\gamma)/2\gamma^2)\sup_{T\in\mathscr{T}} EX_T \tag{2.5}$$

where $\gamma = \gamma(\alpha)$ is the nonnegative root of

$$(1 - \gamma)^3 = (1 - \alpha)(1 + \gamma) \quad \textit{with} \quad \alpha = EM_2/EM_1.$$

PROOF. Suppose that $0 \le \beta \le 1$, $0 \le \gamma \le 1$ satisfy

$$\alpha \ge 1 - (1 - \gamma)^2(1 - \beta). \tag{2.6}$$

Then $P\{M_1 > \beta EM_1\} \ge \gamma$, for if not

$$P\{M_1 \le \beta EM_1\} > 1 - \gamma \tag{2.7}$$

which in turn implies that

$$P\{M_2 \le \beta EM_1 \mid M_1 = X_k, M_1\} \ge P\{M_1 \le \beta EM_1\} > 1 - \gamma;$$

now consider

$$\begin{aligned} E(M_1 - M_2) &> (1 - \gamma)E(M_1 - \beta EM_1)_+ \ge (1 - \gamma)E(M_1 - EM_1)_+ \\ &= (1 - \gamma)E(M_1 - EM_1)_- > (1 - \gamma)^2(1 - \beta)EM_1, \end{aligned}$$

from (2.7).

Hence, $\{1 - (1 - \gamma)^2(1 - \beta)\}EM_1 > EM_2$, contradicting (2.6), and disproving (2.7). By considering $T = \inf\{k: X_k \ge \beta EM_1\}$ we see that $\sup_{T\in\mathscr{T}} EX_T \ge \gamma\beta EM_1$, and maximising $\gamma\beta$ subject to $\alpha \ge 1 - (1 - \gamma)^2(1 - \beta)$ gives (2.5).

Now, as $\alpha = EM_2/EM_1 \to 1$, $\gamma(\alpha) \to 1$ and hence $\sup_{T\in\mathscr{T}} EX_T/EM_1 \to 1$, showing that the lower bound in (1.3) cannot be best possible for all values of α, $1/2 \le \alpha < 1$. For $\alpha = 1$, $M_1 = M_2 =$ constant, a.s. and so $\sup_{T\in\mathscr{T}} EX_T = EM_1$. It would be interesting to obtain a lower bound for the ratio $\sup_{T\in\mathscr{T}} EX_T/EM_1$ in terms of $\alpha = EM_2/EM_1$; a direct numerical calculation shows that the bound provided by (1.3), $(1 + \alpha)/3$, is strictly larger than that given by (2.5), $2\gamma^2(\alpha)/(1 + \gamma(\alpha))$, for all values of $\alpha \le 0.992$.

One might conjecture that for some k we always have $\sup_{T\in\mathscr{T}} EX_T > EM_k$, but the following example shows that this is not true. Take $X_1 = X_2 = \cdots = X_k = 1$; for $k + 1 \le i \le 2k$, let X_i take the values α_i/δ_i and 0 with probabilities δ_i and $1 - \delta_i$ respectively, $1 > \alpha_i > \delta_i > 0$ and let $X_r = 0$ for $r > 2k$. Assuming that $\alpha_{k+1} + \cdots + \alpha_{2k} \le 1$ we have $v_i \le EX_i + \cdots + EX_{2k} \le 1$ for $k + 1 \le i \le 2k$, whence $v_1 = v_k = 1$. But $EM_k = 1 + \{\bigwedge_{k+1}^{2k}(\alpha_i/\delta_i) - 1\}\prod_{k+1}^{2k}\delta_i > 1$.

Examining the proof of Theorem 1, in particular the relation (2.4), might lead one to conjecture that it could be improved to

$$E[\textstyle\sum_{i=1}^{k} M_i] \le \sup\{E[\sum_1^{k+1} X_{T_i}]: T_i \in \mathscr{T}, T_1 < T_2 < \cdots < T_{k+1}\}. \tag{2.8}$$

Again, this is not true. Let $X_i = 1$, $1 \le i \le k$, and suppose that X_{k+1} takes the values 2, 0 with probabilities δ, $1 - \delta$, X_{k+2} takes values $3/\varepsilon$, 0 with probabilities ε, $1 - \varepsilon$, and $X_r \equiv 0$, $r > k + 2$. Provided $1 > \delta \ge 1/2$, it is easily seen that the right-hand side of (2.8) is $k + 2 + 2\delta$ while the left-hand side is $k + \delta + 3 - \varepsilon$ contradicting (2.8) as $\varepsilon \to 0$.

Finally, it has been pointed out by E. Samuel-Cahn that a technique used by her in [8] leads to a simple and elegant proof of Theorem 1. For $b \ge 0$ let $T(b) = \min\{r \le n: X_r \ge b\}$ with $T(b) = n + 1$, if there is no such r. By choosing b^* to be the unique solution to $b = E\sum_{i=1}^{n}(X_i - b)_+$, then the argument in [8] shows that $b^* \le EX_{T(b^*)}$ from which Theorem 1 follows when it is observed that

$$\sum_{i=1}^{k} M_{i,n} \le kb^* + \sum_{i=1}^{n}(X_i - b^*)_+.$$

3. Proof of Theorem 2. By scaling the uniformly bounded random variables $\{X_r,\ r \ge 1\}$, without any loss of generality we may assume that they take values in $[0, 1]$. The argument is similar to that of the previous section except that in place of (2.1) we use the following inequality. For $k \ge 0$ and real numbers $0 \le \gamma \le \alpha_k \le \alpha_{k-1} \le \cdots \le \alpha_0 \le 1$, and $0 \le \beta \le 1$ we have

$$\alpha_k \vee \beta + \gamma - \gamma \prod_{i=1}^{k}[(\alpha_i \vee \beta) \wedge \alpha_{i-1}] \le \alpha_k + \beta \vee \gamma - (\beta \vee \gamma)\prod_{i=0}^{k}\alpha_i, \tag{3.1}$$

where by convention the empty product is 1. This inequality follows easily after observing that for $k \ge 1$

$$\prod_{i=1}^{k}[(\alpha_i \vee \beta) \wedge \alpha_{i-1}] = [(\alpha_k \vee \beta) \wedge \alpha_0]\prod_{i=1}^{k-1}\alpha_i.$$

Now take $M_{0,n} = 1$ and then for $0 \le \alpha \le 1$, $i \ge 1$,

$$\alpha \vee M_{i,n} = (\alpha \vee M_{i,n-1} \vee X_n) \wedge (\alpha \vee M_{i-1,n-1}). \tag{3.2}$$

With the same notation as in the previous section we shall establish, for fixed $n \ge k$, by backwards induction on $r = n, n-1, \cdots, k$ that

$$\begin{aligned} &E[\textstyle\sum_{i=1}^{k} M_{i,n}] \\ &\quad \le E[\textstyle\sum_{i=1}^{k} v_{r+2,n} \vee M_{i,r} + v_{r+1,n} - v_{r+1,n}\prod_{i=(k-n+r+1)_+}^{k}(v_{r+2,n} \vee M_{i,r})], \end{aligned} \tag{3.3}$$

and furthermore, the right-hand side of (3.3) is decreasing in r. This inequality is immediate for $r = n$. Then, for $n \ge r > k$, using (2.2) and (3.2), since $v_{r,n}$ is decreasing in r, we have

$$\begin{aligned} &E[\textstyle\sum_{i=1}^{k} v_{r+2,n} \vee M_{i,r} + v_{r+1,n} - v_{r+1,n}\prod_{i=(k-n+r+1)_+}^{k}(v_{r+2,n} \vee M_{i,r})] \\ &\quad \le E[\textstyle\sum_{i=1}^{k} v_{r+1,n} \vee M_{i,r} + v_{r+1,n} - v_{r+1,n}\prod_{i=(k-n+r+1)_+}^{k}(v_{r+1,n} \vee M_{i,r})] \\ &\quad = E[\textstyle\sum_{i=1}^{k-1} v_{r+1,n} \vee M_{i,r-1} + v_{r+1,n} \vee M_{k,r-1} \vee X_r + v_{r+1,n} \\ &\qquad\quad - v_{r+1,n}\prod_{i=(k-n+r+1)_+}^{k}(v_{r+1,n} \vee M_{i,r-1} \vee X_r) \wedge (v_{r+1,n} \vee M_{i-1,r-1})]. \end{aligned}$$

Now apply (3.1) with $\alpha_i = v_{r+1,n} \vee M_{i,r-1}$, $i = (k-n+r)_+, \cdots, k$; $\alpha_i = 1$ for $i < (k-n+r)_+$; $\beta = X_r$ and $\gamma = v_{r+1,n}$. Using the fact that X_r is independent of $M_{i,r-1}$ we see that this last expression is bounded above by

$$E[\textstyle\sum_{i=1}^{k} v_{r+1,n} \vee M_{i,r-1} + v_{r,n} - v_{r,n}\prod_{i=(k-n+r)_+}^{k}(v_{r+1,n} \vee M_{i,r-1})],$$

which establishes the induction and hence (3.3). Taking $r = k$ in (3.3) and noting that

$$\prod_{i=1}^{k}(\alpha \vee M_{i,k}) = \prod_{i=1}^{k}(\alpha \vee X_i)$$

we see that

$$\begin{aligned} E[\textstyle\sum_{i=1}^{k} M_{i,n}] &\le E[\textstyle\sum_{i=1}^{k} v_{k+2,n} \vee M_{i,k} + v_{k+1,n} - v_{k+1,n}\prod_{i=1}^{k}(v_{k+2,n} \vee M_{i,k})] \\ &= E[\textstyle\sum_{i=1}^{k} X_i \vee v_{k+2,n} + v_{k+1,n} - v_{k+1,n}\prod_{i=1}^{k}(X_i \vee v_{k+2,n})]. \end{aligned}$$

Using the independence of $X_1, \cdots, X_k$ and letting $n \to \infty$ yields

$$\begin{aligned} E[\textstyle\sum_{i=1}^k M_i] &\le \textstyle\sum_{i=1}^k E(X_i \vee v_{k+2}) + v_{k+1} - v_{k+1} \prod_{i=1}^k E(X_i \vee v_{k+2}) \\ &\le \textstyle\sum_{i=1}^{k+1} v_i - \prod_{i=1}^{k+1} v_i \\ &\le (k+1)v_1 - v_1^{k+1}. \end{aligned} \tag{3.4}$$

It follows that

$$\begin{aligned} (1/k)E[\textstyle\sum_{i=1}^k M_i] - \sup_{T\in\mathcal{T}} EX_T &\le (v_1 - v_1^{k+1})/k \\ &\le 1/(k+1)^{(k+1)/k}, \end{aligned}$$

where the final inequality is obtained by maximizing $(x - x^{k+1})/k$ in $0 \le x \le 1$. This completes the proof of Theorem 2. From (3.4) we derive the following result (which in the case $k = 1$ was given by Hill ([2], Theorem 2.3)).

COROLLARY 3. *If $\{X_r, r \ge 1\}$ is a sequence of independent random variables taking values in* $[0, 1]$ *then for each* $k \ge 1$,

$$E[\textstyle\sum_{i=1}^k M_i] \le (k+1)v - v^{k+1}$$

where $v = \sup_{T\in\mathcal{T}} EX_T$.

Acknowledgements. The author is grateful to Saul Jacka for very useful discussions and, in particular, for suggestions leading to the examples of Section 2. Thanks are also due to the Associate Editor for his helpful comments.

REFERENCES

[1] CHOW, Y. S., ROBBINS, H. and SIEGMUND, D. (1971). *Great Expectations: The Theory of Optimal Stopping*. Houghton-Mifflin, Boston, Massachusetts.

[2] HILL, T. P. (1983). Prophet inequalities and order selection in optimal stopping problems. *Proc. Amer. Math. Soc.* **88** 131–137.

[3] HILL, T. P. and KERTZ, R. P. (1981). Ratio comparisons of supremum and stop rule expectations. *Z. Wahrsch. verw. Gebiete* **56** 283–285.

[4] HILL, T. P. and KERTZ, R. P. (1981). Additive comparisons of stop rule and supremum expectations of uniformly bounded independent random variables. *Proc. Amer. Math. Soc.* **83** 582–585.

[5] HILL, T. P. and KERTZ, R. P. (1982). Comparisons of stop rule and supremum expectations of i.i.d. random variables. *Ann. Probab.* **10** 336–345.

[6] KRENGEL, U. and SUCHESTON, L. (1977). Semiamarts and finite values. *Bull. Amer. Math. Soc.* **83** 745–747.

[7] KRENGEL, U. and SUCHESTON, L. (1978). On semiamarts and amarts, and processes with finite value. In *Probability on Banach Spaces* (ed. J. Kuelbs). Marcel Dekker, New York.

[8] SAMUEL-CAHN, E. (1984). Comparison of threshold stop rules and maximum for independent nonnegative random variables. *Ann. Probab.* **12** 1213–1216.

STATISTICAL LABORATORY
UNIVERSITY OF CAMBRIDGE
16 MILL LANE
CAMBRIDGE CB2 1SB
UNITED KINGDOM