A Survey of Prophet Inequalities
in Optimal Stopping Theory

THEODORE P. HILL AND ROBERT P. KERTZ

ABSTRACT. This paper surveys the origin and development of what has come to be known as "prophet inequalities" in optimal stopping theory. Included is a review of all published work to date on these problems, including extensions and variations, descriptions and examples of the main proof techniques, and a list of a number of basic open problems.

1. Introduction

The main purpose of this paper is to provide a brief survey of what has come to be known as "prophet inequalities" or "prophet problems" in the theory of optimal stopping. This survey includes summaries of the basic results, subsequent extensions and variations of these results, main proof tools and techniques (with concrete examples), and a list of open problems.

Although the term "prophet" has been used in other mathematical and probabilistic contexts, the expression "prophet inequality" in optimal stopping theory is generally associated with the following problem. Given a class $\mathcal{C}$ of sequences of integrable random variables $\vec{X} = (X_1, X_2, \dots)$, find universal inequalities valid for all $\vec{X}$ in $\mathcal{C}$ which compare the expected supremum of the sequence with the optimal stopping value of the sequence. That is, if M denotes the expected supremum

$$M = M(\vec{X}) = E\left[\sup_n X_n\right]$$

and V denotes the optimal–stopping value (over the set $\mathcal{T} = \mathcal{T}(\vec{X})$ of stop rules for $\vec{X}$)

$$V = V(\vec{X}) = \sup_{t \in \mathcal{T}} EX_t$$

then a *prophet inequality* for a class $\mathcal{C}$ is an inequality in M and V which is valid for all sequences $\vec{X}$ in $\mathcal{C}$. The term "prophet" arises from the optimal–stopping interpretation of M, which is the optimal expected return of a player endowed with complete foresight, who observes the sequence $X_1, X_2, \ldots$ and may stop whenever he pleases, thereby incurring a reward equal to the variable at the time of stopping. With complete foresight (or inside information, or equivalently, the ability to return to previously observed values), such a player obviously stops always with the largest values, and wins on the average M, which is at least as large as the optimal return V of the non-prophet player (i.e., $M \geq V$).

Although there have been many comparisons of M and V for *fixed* distributions $\vec{X}$, apparently the first universal inequality for a large natural class of random variables is the following now-classical result of Krengel, Sucheston and Garling [**49, 50**] which has directly or indirectly inspired most of the results mentioned in this paper.

If $X_1, X_2, \ldots$ are independent and ≥ 0, then

(1.1) $$M \leq 2V$$

and the bound "2" is sharp.

In other words, if $\mathcal{C}$ is the class of sequences of independent non–negative random variables, then

(1.1′) $$\sup_{\vec{X} \in \mathcal{C}} \frac{M(\vec{X})}{V(\vec{X})} = 2.$$

This result is both surprising and elegant; it says that a player with complete foresight may never win more, on the average, than twice that of an ordinary gambler when sequentially observing and stopping along a sequence of independent nonnegative random variables. Its discovery, first with the *multiplicative* upper bound "4" by Krengel and Sucheston [**50**], has inspired a number of similar inequalities such as the *additive* and, *regional* inequalities (1.2) and (1.3) below.

If $X_1, X_2, \ldots$ are independent with values in $[0,1]$, then (cf. [**33, 26**])

(1.2) $$M - V \leq \frac{1}{4}$$

and

(1.3) $$M \leq 2V - V^2.$$

In fact from (1.3) follows easily that the *region* (often called the prophet region) for the class $\mathcal{C}$ of sequences of independent $[0,1]$ valued random variables is precisely the convex region in the plane given by $V \leq M \leq 2V - V^2$. Note that

inequalities (1.1) and (1.2) both follow easily from (1.3). Analogous inequalities for a variety of other classes $\mathcal{C}$ (e.g., arbitrarily-dependent and uniformly bounded, i.i.d., averages of independent r.v.'s, exchangeable r.v.'s etc.) as well as for a variety of other stopping options (e.g., stopping with partial recall, stopping several times, using only threshold stopping rules, etc.) have been found in recent years, and will be summarized in Sections 2 and 3 below.

Applications of prophet inequalities have been mainly to other problems in optimal stopping theory. The classical inequality (1.1) was orginally discovered by Krengel and Sucheston [50] in their study of semiamarts (processes $\vec{X}$ satisfying $\sup_t EX_t < \infty$), and (1.1) implied a representation theorem for semiamarts, namely,

> a nonnegative independent sequence $X_1, X_2, \ldots$ is a semiamart if and only if $E(\sup_n X_n) < \infty$.

An application of (1.1) for *order-selection* problems is the following (cf. Hill [26]). Let $W(\{X_1, X_2, \ldots\})$ denote the value of the *unordered collection* of random variables $\{X_1, X_2, \ldots\}$ to a player who is free to select the order of observation of the variables as well as the stop rule, that is, $W = W(\{X_1, X_2, \ldots\}) = \sup\{V(X_{\pi(1)}, X_{\pi(2)}, \ldots) : \pi$ is a permutation of the positive integers$\}$. Then the value of a sequence of nonnegative independent r.v.'s to a player free to select both order *and* stop rule is at most twice that of a player who is only free to select the stop rule (and must observe the sequence in a predetermined order); i.e., $W \leq 2V$. This inequality follows easily from (1.1) since clearly $W \leq M$, so

$$(1.4) \qquad \text{if } X_1, X_2, \ldots \text{ are independent and } \geq 0, \text{ then } W \leq 2V.$$

The same example to show that (1.1) is sharp also shows that (1.4) is sharp: taking $X_1 \equiv 1$; $X_2 = \epsilon^{-1}$ with probability ϵ, and $= 0$ otherwise shows that $V = 1$ and $W = M = 2 - \epsilon$. Inequalities analogous to (1.4) for *nonmeasurable stopping* and *stopping with partial recall* have been obtained using (1.1) in [37] and [28]. For several other applications of prophet inequalities, the reader is referred to Hill [28] and Kertz [45].

2. Prophet Inequalities for Different Classes of Random Vectors

For the class $\mathcal{C}_n$ of random vectors $\vec{X} = (X_1, \ldots, X_n)$ consisting of independent r.v.'s $X_1, \ldots, X_n$ taking values in $[0,1]$, the prophet inequalities of (1.1), (1.2) and (1.3) indicate the advantage of the prophet over the gambler. For which classes of random vectors does the prophet have a greater advantage over the gambler than with this class? For which classes of random vectors is the prophet's advantage greatest? On the other hand, for which classes of random vectors does the prophet have a smaller advantage over the gambler than with this class?

If no boundedness assumptions are imposed on the random variables, then the prophet can have arbitrarily large (proportionate) advantage for two independent

r.v.'s, as the following example of [33] indicates:

(2.1)   for any $M > 0$, let $X_1 \equiv 1$ and define X_2 by

$$P(X_2 = 2M) = \frac{1}{2} = P(X_2 = -2M); \text{ then } V(X_1, X_2) = 1$$

$$\text{and } E(X_1 \vee X_2) = M + \left(\frac{1}{2}\right)$$

However, if uniformly bounded r.v.'s are considered, the class comparison becomes more meaningful. If arbitrarily-dependent r.v.'s are considered, then the prophet's greatest advantage over the gambler is given by the following sharp inequalities [35]:

for nonnegative r.v.'s $X_1, \ldots, X_n$,

(2.2) $M \leq nV$;

for r.v.'s $X_1, \ldots, X_n$ taking values in [0,1],

(2.3) $M - V \leq ((n-1)/n)^n$, and

(2.4) $M \leq V - (n-1)V(V^{1/(n-1)} - 1)$;

and for r.v.'s $X_1, X_2, \ldots$ taking values in [0,1]

(2.5) $M - V < e^{-1}$, and

(2.6) $M \leq V - V \ln V$.

For what type of distributions does the prophet do best possible? These extremal distributions have been found [35]; for example, for $0 < x < 1$, $V = x$ and $M = x - (n-1)x(x^{1/(n-1)} - 1)$ for process $X_1, \ldots, X_n$ defined by $X_1 \equiv x$, and for $m = 2, \ldots, n$,

if $X_{m-1} = x^{(n-m+1)/(n-1)}$, then

$$P(X_m = x^{(n-m)/(n-1)} | X_1, \ldots, X_{m-1}) = x^{1/(n-1)} =$$
$$1 - P(X_m = 0 | X_1, \ldots, X_{m-1})$$

and if $X_{m-1} = 0$, then

$$P(X_m = 0 | X_1, \ldots, X_{m-1}) = 1.$$

This process is a martingale. Indeed, the inequalities (2.2)–(2.6) are sharp for martingales, and for Markov chains. For details on these results, see [8, 20, 35, 45]. Recently, it has been observed [38, 39] that for the class of sequences $X_j = Y_j - cj$, $j = 1, 2, \ldots$ where $Y_1, Y_2, \ldots$ are independent r.v.'s taking values in [0,1] and $c > 0$, the prophet's 'proportionate' advantage is also $+\infty$ and the prophet's 'difference' advantage is also 'e^{-1}'. In this case, the 'difference'

inequality (2.5) is already sharp for the subclass in which $Y_1, Y_2, \ldots$ are i.i.d. r.v.'s [58].

Variations on these classes of sequences give classes with advantage to the prophet strictly between the independent r.v. case and the martingale case. For example, for $p > 1$ and any martingale $Y_1, Y_2, \ldots,$

$$(2.7) \qquad E(\sup_{i \geq 1} |Y_i|^p) \leq (p/(p-1))^p \sup_{i \geq 1} E(|Y_i|^p)$$

(with $(p/(p-1))^p \uparrow \infty$ as $p \downarrow 1$, and $(p/(p-1))^p \downarrow e$ as $p \uparrow \infty$), so the universal 'proportionate' constant for this case (with fixed $p > 1$) is finite and > 2. Inequality (2.7) is Doob's inequality ([17, p. 317]); sharpness of Doob's inequality was shown in [13, 15, 19]. The finite sequence cases and other prophet inequalities in this setting are given in [13, 15, 19, 25]. Prophet inequalities for other classes of submartingales which are formed from martingales are given in [19, 23, 25, 46]. The constant 'e' also appears as the universal 'proportionate' constant (i.e., $M \leq eV$) for the class of positive parts of sums of i.i.d. r.v.'s having strictly negative mean, as was shown by Darling, Liggett and Taylor [16] (prior to the results (1.1) of Krengel and Sucheston).

There are now several different proofs of the prophet inequalities (1.1), (1.2), and (1.3) for the class of independent r.v.'s [4, 26, 32, 33, 41, 50], and also a verification of which distributions can be extremal [48]. There are other classes of distributions which give some prophet inequalities with universal constants which are close to those of the independent case. These include a class of discounted independent r.v.'s [7]; classes of averages and weighted sums of independent r.v.'s [11, 27, 49, 50]; a class of finite sequences of exchangeable r.v.'s [21, 22]; a class of negatively dependent r.v.'s [53, 54, 57]; and the class of positive parts of sums of i.i.d. r.v.'s having mean zero [47].

If the class of r.v.'s is not only independent, but also identically distributed, then the prophet's advantage decreases. This was shown in [34, 43], where specific universal constants and boundary curves for the prophet regions are given via implicit equations which allow numerical calculations, with lower values than in the independent case.

3. VARIATIONS ON THE PROPHET–GAMBLER COMPARISON

Changes in the stopping options, changes in the underlying process structure, and changes in the optimization criteria have led to other natural prophet inequalities which give insights into the original prophet inequalities (1.1), (1.2), and (1.3).

Samuel-Cahn [55, 57] has shown that the prophet inequalities (1.1), (1.2), and (1.3) hold even if the gambler is restricted to choose from a smaller collection of stop rules. If one defines a pure threshold stop rule $t(c)$ by $t(c) = \min\{1 \leq j < n : X_j \geq c\}$ if this set is nonempty, and $= n$ otherwise, for some constant

c, and denotes $V_0 = V_0(X_1, \ldots, X_n) = \sup_{c>0} E(X_{t(c)})$, then by using a median constant m of the distribution of $\max_{j \leq n} X_j$, she showed that for the class of nonnegative independent r.v.'s $X_1, \ldots, X_n$,

$$(3.1) \qquad M \leq 2E(X_{t(m)}) \leq 2V_0 \leq 2V.$$

This gives an additional verification of the prophet inequality (1.1) and shows that, for this class of random vectors, limiting the gambler to use of pure threshold stop rules does not improve the prophet's advantage (in the worst-case scenario; of course for certain fixed sequences, such a restriction *does* improve the prophet's advantage). This threshold-stop rule approach has also led to other prophet inequalities for this class of sequences and for classes of negatively dependent r.v.'s [5, 53-57].

Hill and others [9, 24, 26, 28, 35] have shown that gamblers with various types of more general stopping options can do as well (in the worst-case situation) as a prophet does against a gambler with the usual stopping options. For example, if

$W_1 = \sup\{EX_t : t \text{ is a possibly nonmeasurable stop rule, } t \leq n\}$, and

$W_2 = \max\{V(X_{\pi(1)}, \ldots, X_{\pi(n)}) : \pi \text{ is any permutation of } \{1, \ldots, n\}\}$,

then for the class of nonnegative independent r.v.'s $X_1, \ldots, X_n$,

$$(3.2) \qquad W_i \leq 2V \text{ for } i = 1, 2,$$

and in both cases the inequality is sharp. For other stopping options which yield the same type of comparison inequality as (3.2), see [9, 26, 28].

The gambler also does better against the prophet if the prophet receives an average of s choices or if the gambler is given several choices. For example, Kennedy [41, 42] has shown that if M_s is the sth largest order statistic in $X_1, X_2, \ldots$, then for the class of nonnegative, independent r.v.'s $X_1, X_2, \ldots$,

$$(3.3) \qquad E(s^{-1}(M_1 + \cdots + M_s)) \leq (1 + \frac{1}{s})V;$$

and

$$(3.4) \qquad M \leq C_r \cdot \sup\{E(X_{\tau_1} + \cdots + X_{\tau_r}) : \text{ stop rules } \tau_1 < \cdots < \tau_r\},$$

where $2 \geq C_r \downarrow 1$. For other prophet inequalities of this type, see [44].

A combination of changes in the underlying process structures and changes in the choice mechanism have led to prophet inequalities for transforms of processes [45, 51, 60] and for parallel processes [3, 6, 29]. For an example of a prophet inequality in the transform setting let $X_0, X_1, \ldots$ be integrable r.v.'s and $U_1, U_2, \ldots$ be r.v.'s taking values in [0,1], and define the transform of X by U as $U * X = (U * X)_n = X_0 + \sum_{i=1}^{n} U_i(X_i - X_{i-1})$. Krengel and Sucheston [51]

have shown that if $X_0, \ldots, X_n$ are nonnegative r.v.'s with $E(X_i | X_{i-1}) = EX_i$ for $i = 1, \ldots, n$, and $E(X_0 \wedge EX_0) \le E(X_n \wedge EX_n)$, then

(3.5)
$$\sup\{E(U * X) : U_1, \ldots, U_n\} \le 3\sup\{E(U * X) : U_i \in \sigma(X_{i-1}) \text{ for } i = 1, \ldots, n\},$$
and the bound '3' is sharp.

As an example of a prophet inequality in a parallel process setting, let $\vec{X}_1, \ldots, \vec{X}_n$ be n independent sequences of independent $[0,1]$-valued random variables, with $\vec{X}_i = (X_{i1}, X_{i2}, \ldots)$ for $i = 1, \ldots, n$, and let

$$\widehat{V} = \widehat{V}(\vec{X}_1, \ldots, \vec{X}_n) := \max_{1 \le i \le n} V(\vec{X}_i);$$

then Hill and Kennedy [29] have shown that

(3.6)
$$E(\sup_{1 \le i \le n; 1 \le j} X_{ij}) \le 1 - (1 - \widehat{V})^{n+1}$$

and the upper bound is attained.

Prophet inequalities over classes of random vectors have also been proved under other reward criteria, in a game–theoretic formulation and minimax criteria [31, 59], based on stochastic orderings of distributions [19, 23, 25, 46, 53, 54], and with other reward functions [30, 31]. For an example in the stochastic order setting, let μ be any probability measure (p.m.) on $\mathbb{R}$ with $\int |x| d\mu(x) < \infty$, then the Hardy and Littlewood p.m. μ^* is the least upper bound in the stochastic order of the set

$$\{\nu : \text{ there is a martingale } (X_t)_{0 \le t \le 1} \text{ satisfying, } X_1 \text{ has p.m. } \mu$$
$$\text{and } \sup_{0 \le t \le 1} X_t \text{ has p.m. } \nu\}.$$

This result has been applied to give expectation–based prophet inequalities in [19, 23, 25, 46].

4. TECHNIQUES AND METHODS OF PROOF

The purpose of this section is to briefly describe some of the various analytical, probabilistic and algebraic tools which have proved useful in establishing prophet inequalities similar to (1.1)–(1.3).

Classical optimal-stopping theory. Many of the basic ideas in optimal stopping (see especially Chow, Robbins and Siegmund [12]) are used repeatedly throughout the study of prophet inequalities. For example, the basic backward induction principle used to explicitly calculate V says that, given the process has not been stopped before time j, it is optimal to stop at time j if and only if X_j is at least as much as the conditional value $V(X_{j+1}, \ldots | X_1, \ldots, X_j)$ from time $j+1$ on, given $X_1, \ldots, X_j$, where $V(X_{j+1}, \ldots | X_1, \ldots, X_j) = \text{ess sup}\{E(X_t | X_1, \ldots,$

X_j) : $t > j$}. That is, letting $\mathcal{F}_j$ denote the sigma algebra generated by $X_1, \ldots, X_j$,

(4.1) $V(X_j, X_{j+1}, \ldots | \mathcal{F}_j) = X_j \vee V(X_{j+1}, X_{j+2}, \ldots | \mathcal{F}_j)$ a.e. for all $j \geq 1$.

Hence, if the $\{X_j\}$ are *independent*, then the value $V(X_1, \ldots, X_n)$ of a finite sequence is v_1, where $v_n \leq v_{n-1} \leq \cdots \leq v_1$ are defined inductively by $v_n = EX_n$, and $v_j = E(X_j \vee v_{j+1})$. An *optimal stop rule* t^* (i.e., $EX_{t^*} = V(X_1, \ldots, X_n)$) in this case is

$$t^* = \min\{j < n : X_j \geq v_{j+1}\} \text{ if this set is nonempty,}$$
$$\text{and } = n \text{ otherwise,}$$

and so to establish inequalities of the form (1.1)–(1.3) for classes of independent r.v.'s the supremum over stop rules in the definition of V can be replaced by the relatively simple explicit expression for EX_{t^*}.

In some cases, an even simpler optimal stop rule can be shown to exist. For example, in establishing a prophet inequality for bounded i.i.d. r.v.'s with cost c of observation, Samuel-Cahn made powerful use of the Chow, Robbins and Siegmund result that for this class of r.v.'s there is always an optimal rule s^* of the form

$$s^* = \inf\{i : X_i \geq \beta\},$$

where β is the unique value for which

$$E[X - \beta]^+ = c.$$

Dilation and convexity. In establishing an inequality of the form (1.1)–(1.3), it is often useful to restrict a given class $\mathcal{C}$ to a much smaller class $\widetilde{\mathcal{C}}$ containing the extremal distributions. For this purpose, dilations (or balayage) arguments have been very useful, via the following lemma [33].

DEFINITION. If X is any integrable r.v., and $\infty < a < b < \infty$, $(X)_a^b$ is a r.v. satisfying $(X)_a^b = X$ if $X \notin [a, b]$; $= a$ with probability $(b - a)^{-1} \int_{X \in [a,b]} (b - X)$; and $= b$ with probability $(b - a)^{-1} \int_{X \in [a,b]} (X - a)$. (Such an $(X)_a^b$ is a "maximal balayage" of X on the interval $[a, b]$.)

LEMMA 4.1. *(i)* $EX = E[(X)_a^b]$. *(ii) If Y is any r.v. independent of both X and $(X)_a^b$, then*

$$E(X \vee Y) \leq E((X)_a^b \vee Y).$$

This lemma, which is a special case of the fact that X is a balayage of Y if and only if X is convexly dominated by Y (which is also equivalent to the pair (Y, X) being martingalizable), can be used to establish inequalities of the form (1.1)–(1.3) as follows. If constants $a < b$ can be found so that $V(X_1, \ldots, (X_j)_a^b, \ldots) = V(X_1, \ldots, X_j, \ldots)$, then the independent sequence $X_1, \ldots, (X_j)_a^b, \ldots$ is "more

extremal" than the independent sequence $X_1, \ldots, X_j, \ldots$, since V for both sequences is the same, but by Lemma 4.1 (ii),

$$M(X_1, \ldots, (X_j)_a^b, \ldots) \geq M(X_1, \ldots, X_j, \ldots).$$

The simplest example is for a finite sequence $X_1, \ldots, X_n$ of independent r.v.'s taking values in $[a, b]$. Since (by backward induction above), V depends on X_n only through its expectation EX_n, replacing X_n by the two-valued r.v. $(X_n)_a^b$ results in a more extremal distribution. Similarly, X_{n-1} can be replaced by the three-valued r.v. $\left((X_{n-1})_a^{EX_n}\right)_{EX_n}^b$, and so on. In the case $\mathcal{C}$ is the class of i.i.d. r.v.'s this idea was used in [34] to show that the extremal distribution had support on at most $n + 1$ points.

Induction. Backward induction is, as mentioned above, repeatedly used to calculate V for finite sequences. In addition, for some inequalities of the form (1.1)–(1.3) both forward induction *and* backward induction are used simultaneously. For example, these techniques can be used as follows to show (cf. [4, 41, 53]) that for $X_i \geq 0$, not all identically 0,

(4.2)
$$E \left[\max_{1 \leq i \leq n} X_i - V_n \right]^+ < V_n$$

where $V_n = V(X_1, \ldots, X_n)$. By eliminating the +, this yields (1.1) with strict inequality. For $n = 1$ (4.2) states $E[X_1 - EX_1]^+ < EX_1$, which is obvious for $X_1 \geq 0$, not identically 0. Set $V(X_2, \ldots, X_n) = V_{n-1}$, and assume the induction hypothesis

$$E \left[\max_{2 \leq i \leq n} X_i - V_{n-1} \right]^+ \leq V_{n-1}$$

(with strict inequality if not all X_i identically 0). Now

$$E \left[\max_{1 \leq i \leq n} X_i - V_n \right]^+ \leq E \left[\max_{1 \leq i \leq n} X_i - V_{n-1} \right]^+ \leq E[X_1 - V_{n-1}]^+$$

$$+ E \left[\max_{2 \leq i \leq n} X_i - V_{n-1} \right]^+ \leq E[X_1 - V_{n-1}]^+ + V_{n-1} = V_n$$

and strict inequality follows if either X_1 or one of $X_2, \ldots, X_n$ is not identically 0. The last inequality follows from (4.1).

Conditioning. Another way to restrict a given class $\mathcal{C}$ to a smaller class $\tilde{\mathcal{C}}$ is through conditioning. For example, conditioning on X_1 yields

$$M = \int M(x_1, X_2, \ldots) dP_{X_1}(x_1)$$

and

$$V = \int V(x_1, X_2, \ldots) dP_{X_1}(x_1),$$

so for many classes $\mathcal{C}$, X_1 may be replaced by a (worst-case) constant x_1 resulting in a more extremal distribution. In the case of an inequality like (1.2), this says

$$M(x_1, X_2, \dots) - V(x_1, X_2, \dots) \geq M(X_1, X_2, \dots) - V(X_1, X_2, \dots).$$

Observe that to obtain M, any of the r.v.'s X_j may be conditioned on, whereas for V, in general it is not true that V is the conditional expectation of V given X_j for $j > 1$, since given that X_j is a constant x_j "gives information about the future." In some cases in a proper setting one can condition on "interior" X_j's in order to reduce $\mathcal{C}$ (cf. [27]). It should also be observed that for other natural $\mathcal{C}$, such as the class of uniformly bounded i.i.d. r.v.'s, conditioning on X_1, although valid, does not result in a new sequence which is still in the class, and the reduction is worthless.

Constrained optimization. The general problem of establishing an inequality like (1.1) is by definition a constrained optimization problem, in this case

$$\begin{array}{ll}
\text{maximize} & M(X_1, X_2, \dots)/V \\
\text{subject to:} & V(X_1, X_2, \dots) = V \\
& X_1, X_2, \dots \text{ independent} \\
& 0 \leq X_j \leq 1 \text{ for all } j.
\end{array}$$

In general, both the constraints and objective function are unwieldly in this generality, but in many cases, using backward induction, dilation (via Lemma 4.1) etc. one may reduce this setup to a more tractable one, say, of only a finite number of variables. For example, in establishing the finite horizon analog of (1.1) for nonnegative i.i.d. r.v.'s (cf. [32]), dilation and backward induction were used to first show that the extremal distributions have support on at most $n + 1$ points, in this case on the points $0 = x_0 < x_1 < \cdots < x_n$, where $x_j = V(X_1, \dots, X_j)$ for $j = 1, \dots, n$. Then, since the partial-moment relations force constraints on the probabilities $p_j = P(X_1 = x_j)$ (e.g., $x_1 = \sum_{j=1}^n p_j x_j$, and $x_2 = \sum_{j=1}^n p_j x_j + x_1 p_0$), the whole problem can be reduced to maximizing a rational function of n variables (the $\{p_j\}$'s) over a certain constraint set.

Conjugate duality. A powerful technique especially useful in establishing "region" inequalities such as (1.3) is a tool used by Cox and Kertz [15] and Kertz [43] which reduces the original constrained maximization problem to an *unconstrained* minimization problem which is generally much easier to solve. The idea is essentially as follows. For a concave function f defined on an interval I in $\mathbb{R}$, define the *conjugate function* f^* on $D^* = \{\gamma \in \mathbb{R} : \inf_{v \in I}[v\gamma - f(v)] > -\infty\}$ by

$$f^*(\gamma) = \inf_{v \in I}[v\gamma - f(v)].$$

Next, in the context of (1.3), let $\Gamma(v) = \sup\{m : m = M(X_1, \dots)$ and $v = V(X_1, \dots)$ for some sequence of independent r.v.'s $X_1, X_2, \dots$ taking values

in $[0,1]\}$; that is, $\Gamma(v)$ is the upper boundary of the "prophet region" for this particular choice of $\mathcal{C}$. Next, verify that on $I = [0,1]$, Γ is a concave, continuous function with closed hypograph, from which general theory (Young's inequality, etc.) of conjugate functions implies that $(\Gamma^*)^* = \Gamma$, so Γ determines Γ^* and vice versa. Then, since

$$\Gamma^*(\gamma) = \inf\{\gamma V(X_1, X_2, \ldots) - M(X_1, X_2, \ldots) : X_1, X_2, \ldots$$
are independent r.v.'s taking values in $[0,1]\}$,

the problem is now reduced to the *unconstrained* (over $\mathcal{C}$) minimization problem of determining Γ^*, which can often be solved using dilations, mixtures, etc. Finally, determining Γ from Γ^* completes the solution.

Dynamic programming and verification lemmas. To prove analogs of (1.1)–(1.3) for certain classes $\mathcal{C}$, such as uniformly bounded martingales, techniques and results from dynamic programming and gambling theory have also proved useful. For example, to demonstrate that

(4.3)   for all martingales $X_1, \ldots, X_n$ taking values in $[0,1]$,

$$M \leq V + (n-1)V(1 - V^{1/(n-1)})$$

Dubins and Pitman [20] used the following verification lemma from abstract gambling theory due to Dubins and Freedman [18].

LEMMA 4.2. *If (Γ, F, u) is any measurable gambling problem, and Q is a real–valued Σ–measurable function on F satisfying*

   (i) $u(f) \leq Q(f)$ *for each $f \in F$; and*
   (ii) $\int Q(\bar{f})\gamma(d\bar{f}) \leq Q(f)$ *for every $\gamma \in \Gamma(f)$, each $f \in F$,*

then for each $f \in F$, any Γ process starting at f satisfies $E[u(f_t)] \leq Q(f)$ for every $\{\mathcal{F}_n\}$ stop rule t.

To prove (4.3) using Lemma 4.2, take: $F = \{f = (x, y, r) : 0 \leq x \leq y \leq 1, r = 0, 1, 2, \ldots\}$; $u(f) = u(x, y, r) = y$; $\Gamma(f) = \{\gamma_X : \gamma_X$ is the distribution of $(X, X \vee y, r - 1)$ where X is a r.v. with $0 \leq X \leq 1$ and $EX \leq x\}$ if $r = 1, 2, \ldots$, and $\Gamma(f) = \{\delta(f)\}$ if $r = 0$; and $Q(f) = Q(x, y, r) = y + rx(1 - y^{1/r})$, so $Q(x, x, n-1) = x + (n-1)x(1 - x^{1/(n-1)})$. Then (4.3) follows by showing that (i) and (ii) of Lemma 4.2 hold for all $r \geq 1$, using convexity and "coalescing" techniques.

Algebraic inequalities. Several purely algebraic inequalities have also proved useful in establishing prophet inequalities. For example, Kennedy [41] used the inequality

$$\alpha \vee \beta + \gamma \leq \alpha + \beta \vee \gamma \text{ for all } \alpha \geq \gamma \geq 0 \text{ and all } \beta \geq 0$$

and backward induction to establish a generalization of (1.1) for order statistics, namely,

if $X_1, X_2, \ldots$ are independent and ≥ 0, then for each $k \geq 1$,

$$E\left(\sum_{j=1}^{k} X_{[j]}\right) \leq (k+1)V,$$

where $X_{[j]}$ denotes the jth order statistic of $X_1, X_2, \ldots$ (so $X_{[1]} = \sup_{n \geq 1} X_n$, etc.).

Similarly, algebraic inequalities such as

$$(\alpha + \beta - \alpha\beta) \vee \gamma \leq \alpha + (1 - \alpha)(\beta \vee \gamma) \text{ for all } \alpha, \beta, \gamma \in [0, 1]$$

were used in [29] to establish prophet inequalities for parallel processes.

Moment theory. In establishing prophet inequalities for martingale–based processes, basic results from classical moment theory have proved useful. In Cox and Kertz [15], the result from moment theory

if h is a bounded Borel function, then $\phi(t) := \inf\{E[h(X)] : X$ is an integrable r.v. with $EX = t\}$ is the height at location $x = t$ of the lower boundary of the convex hull of the graph of h

was combined with conditioning arguments, convexity, and a conjugate function formulation to obtain prophet regions for p^{th} absolute moments of martingales. In a similar problem, the above moment result was applied in [45] repeatedly to a maximization function

$$\phi_n(s, t) = \sup\left\{E\left(\max_{1 \leq j \leq n}\left\{s, t + \sum_{i=1}^{j} Y_i\right\}\right) : Y_1, \ldots, Y_n\right.$$

is a martingale difference sequence with

$$\left. EY_1 = 0 \text{ and } 0 \leq t + \sum_{i=1}^{j} Y_i \leq 1 \text{ a.e. for each } 1 \leq j \leq n\right\}$$

to conclude that for any martingale $X_1, \ldots, X_n$ taking values in $[0,1]$,

$$E(\max_{1 \leq j \leq n} X_j | X_1 = x) \leq x + (n - 1)(1 - x^{1/(n-1)})x,$$

and (4.3) then follows by Jensen's inequality and the concavity of $g(x) = x(1 - x^{1/(n-1)})$.

5. OPEN PROBLEMS

A number of basic and interesting prophet questions remain open. In the following inequalities, the sharp universal constants $\{k_i\}$ and functions $\{\phi_i\}$ are unknown (although some bounds, e.g., $k_1 \leq 2$, are known in special cases). Recall that $V = V(\vec{X})$ is the optimal value of $\vec{X}$ to a player free to select the stop rule only, $W = W(\vec{X})$ is the optimal value to a player free to select both

the order of observation and the stop rule, and $M = M(\vec{X})$ the optimal value to a player with complete information or foresight, e.g., a prophet, so $V \leq W \leq M$.

Q1. If $Y_1, Y_2, \ldots$ are i.i.d. r.v.'s taking values in $[0,1]$, and $X_j = (Y_1 + \cdots + Y_j)/j$, then

$$M \leq k_1 V;$$
$$M - V \leq k_2;$$
$$M \leq \phi_1(V);$$
$$M \leq k_3 W;$$
$$M - W \leq k_4; \text{ and}$$
$$M \leq \phi_2(W).$$

Q2. If $X_1, \ldots, X_n$ are exchangeable r.v.'s taking values in $[0,1]$, then

$$M \leq k_5 V;$$
$$M - V \leq k_6;$$
$$M \leq \phi_3(V);$$
$$M \leq k_7 W;$$
$$M - W \leq k_8; \text{ and}$$
$$M \leq \phi_4(W).$$

Q3. What is the largest natural class $\mathcal{C}$ of random variables for which $M \leq 2V$? For example, $\mathcal{C}$ contains sequences of independent and more generally, negatively dependent, nonnegative random variables, and sequences of averages of independent nonnegative random variables, but is there a large natural $\mathcal{C}$ including all the known results for which it is true? Similarly, what is the largest natural class $\mathcal{C}$ for which $M - V \leq 1/4$, or $M \leq 2V - V^2$?

Q4. In several cases involving "costs" or other limits, surprising discontinuities in the bounds exist. For example, the extremal bounds as functions of cost c in Jones [38, 39] and Samuel-Cahn [58] are discontinuous at $c = 0$, and the finite horizon result of Klass [47] is discontinuous in the limit as is seen in Darling, Liggett and Taylor's result [16].

Can these counter-intuitive discontinuities be explained?

Q5. What are the universal bounds analogous to (1.1)–(1.3) for comparison of the value V with the "threshold-value" $V_T \leq V$ of a sequence to a player forced to use only threshold stop rules (e.g., what are the best universal constants so that if $X_1, X_2, \ldots$ are independent and ≥ 0, then $V \leq k_9 V_t$, and $V - V_T \leq k_{10}$, etc.)?

In addition, most of the natural analogs of the above questions for order statistics as well as averages and sums of order statistics, also remain open, as do the questions for multi-choice, or multiple-stopping options.

References

1. Badewitz, T., *Prophetenregionen für Folgen von unabhängigen Zufallsvariablen*, Diplomarbeit, University of Göttingen (1989).

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

3. Boshuizen, F., *Prophet inequalities for parallel processes: a note on a proof of T. P. Hill and D. P. Kennedy*, Preprint, Free University of Amsterdam (1987).

4. _____, *Simple proofs for some prophet inequalities*, Preprint, Free University of Amsterdam (1988).

5. _____, *Comparisons of threshold stop rule and supremum expectations for negatively dependent and independent random vectors*, Stoch. Anal. Appl. (1991), (to appear).

6. _____, *Optimal stopping and switching in a matrix of random variables and some related inequalities*, Report WS-341, Free University of Amsterdam (1988).

7. _____, *Prophet inequality for independent random variables with a discount factor*, J. Multivariate Anal. (1991), (to appear).

8. _____, *Optimal stopping theory and prophet inequalities for stochastic processes*, Preprint, Free University of Amsterdam (1989).

9. _____, *Prophet regions for look-ahead stopping rules for bounded random variables*, Stoch. Anal. Appl. **7** (1989), 261–271.

10. _____, *Multivariate prophet inequalities for negatively dependent random variables*, This Proceedings.

11. Brunel, A. and Krengel, U., *Parier avec un prophete dans le cas d'un process sous-additif*, C.R. Acad. Sci., Series A **288** (1979), 57–60.

12. Chow, Y. S., Robbins, H., and Siegmund, D., *Great Expectations: The Theory of Optimal Stopping*, Houghton Mifflin Company, Boston, (1971).

13. Cox, D. C., *Some sharp inequalities related to Doob's inequality*, Inequalities in Statistics and Probability, IMS Lecture Notes - Monograph Series **4**, (1983), pp. 78–83.

14. Cox, D. C. and Kertz, R. P., *Common strict character of some sharp infinite-sequence martingale inequalities*, Stoch. Proc. Appl. **20** (1985), 169–179.

15. _____, *Prophet regions and sharp inequalities for p^{th} absolute moments of martingales*, J. Mult. Anal. **18** (1986), 242–273.

16. Darling, D. A., Liggett, T. and Taylor, H., *Optimal stopping for partial sums*, Ann. Math. Stat. **43** (1972), 1363–1368.

17. Doob, J. L., *Stochastic Processes*, John Wiley and Sons, New York, (1953).

18. Dubins, L. E. and Freedman, D., *A sharper form of the Borel-Cantelli lemma and the strong law*, Ann. Math. Stat. **36** (1965), 800–807.

19. Dubins, L. E. and Gilat, D., *On the distribution of maxima of martingales*, Proc. Am. Math. Soc. **68** (1978), 337–338.

20. Dubins, L. E. and Pitman, J., *A maximal inequality for skew fields*, Z. Wahr. verw. Geb. **52** (1980), 219–227.

21. Elton, J., *Continuity properties of optimal stopping value*, Proc. Am. Math. Soc. **105** (1989), 736–746.

22. Elton, J. and Kertz, R. P., *Comparison of stop rule and maximum expectations for finite sequences of exchangeable random variables*, Stoch. Anal. Appl. (1991), (to appear).

23. Gilat, D., *The best bound in the L log L inequality of Hardy and Littlewood and its martingale counterpart*, Proc. Amer. Math. Soc. **97** (1986), 429–436.

24. _____, *A prophet inequality with order selection for two independent random variables*, Preprint (1985).

25. _____, *On the ratio of the expected maximum of a martingale and the L_p-norm of its last term*, Isr. J. Math. **63** (1988), 270–280.

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

27. _____, *Prophet inequalities for averages of independent non-negative random variables*, Math. Zeit. **192** (1986), 427–436.

28. _____, *Expectation inequalities associated with prophet problems*, Stoch. Anal. Appl. **5** (1987), 299–310.

29. Hill, T. P. and Kennedy, D. P., *Prophet inequalities for parallel processes*, J. Mult. Anal. **31** (1989), 236–243.

30. _____, *Optimal stopping problems with generalized objective functions*, J. Appl. Prob. (1991), (to appear).

31. _____, *Sharp inequalities for optimal stopping based on ranks*, Ann. Applied Prob. (1991), (to appear).

32. Hill, T. P. and Kertz, R. P., *Ratio comparisons of supremum and stop rule expectations*, Z. Wahr. verw. Geb. **56** (1981), 283–285.

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

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

35. _____, *Stop rule inequalities for uniformly bounded sequences of random variables*, Trans. Am. Math. Soc. **278** (1983), 197–207.

36. Hill, T. P. and Krengel, U., *A prophet inequality related to the secretary problem*, This Proceedings (1990).

37. Hill, T. P. and Pestien, V., *The advantage of using non-measurable stop rules*, Ann. Prob. **11** (1983), 442–450.

38. Jones, M., *Universal Constants in Optimal Stopping Theory*, Ph.D. Thesis, Georgia Institute of Technology, Atlanta (1989).

39. _____, *Prophet inequalities for cost of observation stopping problems*, J. Mult. Anal. **34** (1990), 238-253.

40. Kennedy, D. P., *An extension of the prophet inequality*, Lecture Notes in Control and Information Sciences **61** (1984), 104–110 in "Filtering and Control of Random Processes," Ed. by H. Korezlioglu, G. Mazziotto and J. Szpirglas, Springer-Verlag, Berlin.

41. _____, *Optimal stopping of independent random variables and maximizing prophets*, Ann. Prob. **13** (1985), 566–571.

42. _____, *Prophet-type inequalities for multi-choice optimal stopping*, Stoch. Proc. Appl. **24** (1987), 77–88.

43. Kertz, R. P., *Stop rule and supremum expectations of i.i.d. random variables: a complete comparison by conjugate duality*, J. Mult. Anal. **19** (1986), 88–112.

44. _____, *Comparison of optimal value and constrained maxima expectations for independent random variables*, Adv. Appl. Prob. **18** (1986), 311–340.

45. _____, *Prophet problems in optimal stopping: results, techniques and variations*, Preprint (1987).

46. Kertz, R. P. and Rösler, U., *Martingales with given maxima and terminal distributions*, Isr. J. Math. **69** (1990), 173-192.

47. Klass, M., *Maximizing $E \max_{1 \leq k \leq n} S_k^+ / E S_n^+$: a prophet inequality for sums of i.i.d. mean zero random variates*, Ann. Prob. **17** (1989), 1243–1247.

48. Kohlruss, G., *Extremal distributions for the prophet region in the independent case*, Preprint, Angewandte Mathematik und Informatik, University of Münster (1989).

49. Krengel, U. and Sucheston, L., *Semiamarts and finite values*, Bull. Am. Math. Soc. **83** (1977), 745-747.

50. _____, *On semiamarts, amarts, and processes with finite value*, Probability on Banach Spaces, Ed. by J. Kuelbs, Marcel Dekker, New York, 1978.

51. _____, *Prophet compared to the gambler: an inequality for transforms of processes*, Ann. Prob. **15** (1987), 1593-1599.

52. _____, *Stopping rules and tactics indexed by a directed set*, J. Multivariate Analysis, No. 2 **11** (1981), 199–229.

53. Rinott, Y. and Samuel-Cahn, E., *Comparisons of optimal stopping values and prophet inequalities for negatively dependent random variables*, Ann. Stat. **15** (1987), 1482–1490.

54. _____, *Orderings of optimal stopping values and prophet inequalities for certain multivariate distributions*, J. Multivariate Anal. (1991), (to appear).

55. Samuel-Cahn, E., *Comparisons of threshold stop rules and maximum for independent nonnegative random variables*, Ann. Prob. **12** (1984), 1213–1216.

56. _____, *Prophet inequalities for threshold rules for independent random variables*, Fourth Purdue Symposium on Statistics, Decision Theory and Related Topics, Ed. by J. Berger and S. Gupta, Springer-Verlag, Berlin, 1987.

57. _____, *Prophet inequalities for bounded negatively-dependent random variables*, Preprint, Hebrew University (1989).

58. _____, *A difference prophet inequality for bounded i.i.d. variables, with cost for observations*, Preprint, Hebrew University (1990).

59. Schmitz, N., *Games against a prophet*, This Proceedings (1990).

60. Sucheston, L. private communication, excerpts from forthcoming book by Edgar and Sucheston (1989).

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

E-mail: ma201rk@gitvml.bitnet