Stochastic Inequalities
IMS Lecture Notes – Monograph Series
Volume 22 (1993)

OPTIMAL STOPPING VALUES AND PROPHET INEQUALITIES FOR SOME DEPENDENT RANDOM VARIABLES

By YOSEF RINOTT[1] and ESTER SAMUEL-CAHN

University of California, San Diego and Hebrew University

This paper concerns results on comparisons of stopping values, and prophet inequalities for dependent random variables. We describe general results for negatively dependent random variables, and some examples for the case of positive dependence.

1. Introduction

Let $\mathbf{Z} = (Z_1, \ldots, Z_n)$ be a finite sequence of random variables, having a known distribution, and such that $E|Z_i| < \infty$. As usual, a random variable t taking values in $\{1, 2, \ldots\}$ is said to be a *stopping rule* for $\mathbf{Z}$ if the event $\{t = i\}$ is determined by $Z_1, \ldots, Z_i$, $i = 1, 2, \ldots$, and $P(t \leq n) = 1$. (Infinite sequences and unbounded stopping rules have been studied by the methods described below, with minor technical modifications. For simplicity we consider only finite sequences in this paper.) The *optimal stopping value* corresponding to $\mathbf{Z}$ is defined by $V(\mathbf{Z}) = \sup_t EZ_t$, where the supremum is taken over all stopping rules for $\mathbf{Z}$. $V(\mathbf{Z})$ can be regarded as the best expected value attainable by a statistician who is restricted to stopping on the basis of observations which have already been taken. On the other hand, if one could decide when to stop on the basis of complete information about the whole sequence, including future observations, the relevant value would be EZ^*, where $Z^* = \max(Z_1, \ldots, Z_n)$. The quantity EZ^* is thus the value for a *prophet* who can foresee future observations. Clearly $V(\mathbf{Z}) \leq EZ^*$. Inequalities of the type

$$EZ^* \leq cV(\mathbf{Z}), \tag{1}$$

for $\mathbf{Z}$ in some collection of finite sequences, with constant c depending only on this subclass, are called *ratio prophet inequalities*. For a recent survey on such inequalities, with history and bibliography, see Hill and Kertz (1992).

We shall be interested mainly in two problems:

---

[1]Work supported in part by National Science Foundation Grant DMS-9001274.

AMS 1991 *subject classifications.* Primary 60G40; secondary 60E15, 62H05.

Key words and phrases. Negative dependence, positive dependence, random replacement schemes.

1. Determine sequences $\mathbf{X}$ and $\mathbf{Y}$ of dependent random variables for which the optimal stopping values comparison

$$V(\mathbf{X}) \leq V(\mathbf{Y}) \tag{2}$$

   is valid.

2. Obtain prophet inequalities for collections of dependent sequences.

Qualitatively, if the Y's tend to be larger than the X's then one may expect (2) to hold. However, this is not obvious in the presence of dependence, where the possibility of prediction of future values aids in obtaining a high optimal stopping value for the statistician. Thus, $\mathbf{X} \leq_{st} \mathbf{Y}$ (meaning $Eh(\mathbf{X}) \leq Eh(\mathbf{Y})$ for any nondecreasing function h defined on $\mathbf{R}^n$) does not necessarily imply $V(\mathbf{X}) \leq V(\mathbf{Y})$. For example consider

$$(X_1, X_2) = \begin{cases} (2,10) & \text{w.p. } 1/2 \\ (0,-10) & \text{w.p. } 1/2 \end{cases} \tag{3}$$

and (Y_1, Y_2) independent with $P(Y_1 = 2) = 1$, $P(Y_2 = 10) = P(Y_2 = -10) = 1/2$. Then $\mathbf{X} \leq_{st} \mathbf{Y}$, but $V(\mathbf{X}) = (1/2)\cdot 10 + (1/2)\cdot 0 = 5$, whereas $V(\mathbf{Y}) = 2$. While dependence may work to increase the value through prediction, it also affects the value (both for the statistician and the prophet) directly. For example, for the prophet value, it is well known that EZ^* for independent Z's would be smaller than EZ^* for the same marginal Z's satisfying suitable negative dependence conditions. As we shall show, this also applies to the optimal stopping value. Thus it is natural to expect (2) to hold, for example, when the Y's are in some sense more negatively dependent than the X's.

A good portion of this paper contains a survey and reorganization of previous work of the authors on value comparisons and prophet inequalities for dependent random variables. In the next section we shall bring results from Rinott and Samuel-Cahn (1987) on value comparisons for negatively dependent random variables. In Section 3 we discuss examples of such comparisons under positive dependence. It should be clear from the above discussion that in this case one should anticipate difficulties, because, while the dependence tends to increase the value through prediction, the positive nature of the dependence works to decrease the prophet's value or the statistician's optimal stopping value. Section 5 concerns random replacement schemes. We discuss some results and a conjecture which appeared in Rinott and Samuel-Cahn (1991) and some further partial results on the conjecture. Finally, in Section 5, we reorganize and unify results from our aforementioned two papers, on prophet inequalities for certain classes of dependent random variables. Prophet inequalities for other classes are given in Hill and Kertz (1992).

2. Value Comparisons Under Negative Dependence

DEFINITION The random variables $Z_1, \ldots, Z_n$ are said to be *Negatively lower orthant dependent in sequence* (NLODS) if

$$(4) \qquad P(Z_i < a_i | Z_1 < a_1, \ldots, Z_{i-1} < a_{i-1}) \leq P(Z_i < a_i)$$

for $i = 2, 3, \ldots, n$, and all constants $a_1, \ldots, a_n$ for which the conditional probability in (4) is defined.

It is easy to see that condition (4) is weaker than most of the well known conditions of negative dependence. Thus if $Z_1, \ldots, Z_n$ are *Negatively associated* (NA), i.e., $\mathrm{cov}\{f_1(Z_i, i \in A_1), f_2(Z_j, j \in A_2)\} \leq 0$, for any pair of *disjoint* subsets A_1, A_2 of $\{1, \ldots, n\}$ and any nondecreasing functions f_1, f_2, then they are also NLODS. Likewise, if $Z_1, \ldots, Z_n$ are *Negatively dependent in sequence* (NDS), meaning that $Z_1, \ldots, Z_{i-1} | Z_i = a_i$ is decreasing stochastically in a_i, or if $Z_1, \ldots, Z_n$ are *Conditionally decreasing in sequence* (CDS), i.e., $Z_i | Z_1 = a_1, \ldots, Z_{i-1} = a_{i-1}$ is decreasing stochastically in $a_1, \ldots, a_{i-1}$, for $i = 2, 3, \ldots, n$, then they are NLODS. On the other hand it is easy to see that (4) implies *Negative lower orthant dependence*, that is, $P(Z_1 < a_1, \ldots, Z_n < a_n) \leq \prod_{i=1}^n P(Z_i < a_i)$.

Examples of distributions satisfying (4) include the multinomial, multivariate normal with negative correlations, and permutation distributions, including sampling without replacement, all of which are NA. See Joag-Dev and Proschan (1983) for further details.

THEOREM 2.1 (Rinott and Samuel-Cahn (1987)) *Let $Y_1, \ldots, Y_n$ be NLODS random variables, and let $X_1, \ldots, X_n$ be independent random variables such that for each i, X_i and Y_i have the same marginal distribution, $i = 1, \ldots, n$. Then $V(\mathbf{X}) \leq V(\mathbf{Y})$.*

PROOF Given a sequence of random variables $\mathbf{Z} = (Z_1, \ldots, Z_n)$, and a vector $\mathbf{c} = (c_1, \ldots, c_n)$, with $c_n = -\infty$, and possibly $c_i = -\infty$ for some $i < n$, define the stopping rule $t(\mathbf{c}) = \min\{i \leq n : Z_i \geq c_i\}$. Since $c_n = -\infty$, we have $t(\mathbf{c}) \leq n$. Then for $i > 1$,

$$(5) \; Z_{t(\mathbf{c})} = c_1 + [Z_1 - c_1]^+ + \sum_{i=2}^{n} \{c_i - c_{i-1} + [Z_i - c_i]^+\} \cdot I(t(\mathbf{c}) > i-1),$$

where $\{c_i - c_{i-1} + [Z_i - c_i]^+\} = Z_i - c_{i-1}$ if $c_i = -\infty$.

Recall (or see, e.g., Chow, Robbins and Siegmund (1971, Theorem 3.2)) that for independent $X_1, \ldots, X_n$, the optimal stopping rule is of the form $t(\mathbf{c}^*)$, with

$$(6) \quad c^*_{i-1} = E(X_i \vee c^*_i) = c^*_i + E[X_i - c^*_i]^+, \; i = 2, \ldots, n, \quad c^*_n = -\infty.$$

One can see directly, or from (5), that

$$V(\mathbf{X}) = EX_{t(\mathbf{c}^*)} = c_1^* + E[X_1 - c_1^*]^+. \tag{7}$$

The constants of (6), which are optimal for the sequence $X_1, \ldots, X_n$, need not be optimal for the sequence $Y_1, \ldots, Y_n$, and therefore

$$V(\mathbf{Y}) \geq EY_{t(\mathbf{c}^*)}. \tag{8}$$

Next note that for $Y_1, \ldots, Y_n$ which are NLODS, see (4), we have

$$E\{h(Y_i) \cdot I(Y_1 < a_1, \ldots, Y_{i-1} < a_{i-1})\} \geq$$
$$E\{h(Y_i)\} \cdot E\{I(Y_1 < a_1, \ldots, Y_{i-1} < a_{i-1})\}$$

for any nondecreasing function h. In particular, since $I(t(\mathbf{c}) > i-1) = I(Y_1 < c_1, \ldots, Y_{i-1} < c_{i-1})$, we have

$$E\{(c_i - c_{i-1} + [Y_i - c_i]^+) \cdot I(t(\mathbf{c}) > i-1)\} \geq \tag{9}$$
$$(c_i - c_{i-1} + E[Y_i - c_i]^+) \cdot EI(t(\mathbf{c}) > i-1).$$

To prove the theorem combine (8) with (5) applied to $Y_1, \ldots, Y_n$, and (9), to obtain

$$V(\mathbf{Y}) \geq c_1^* + E[Y_1 - c_1^*]^+ + \sum_{i=2}^{n} \{c_i^* - c_{i-1}^* + E[Y_i - c_i^*]^+\} \cdot EI(t(\mathbf{c}^*) > i-1). \tag{10}$$

Because X_i and Y_i have the same marginal distributions, we can replace $E[Y_i - c_i^*]^+$ by $E[X_i - c_i^*]^+$. Then, by (6), the r.h.s. of (10) reduces to $c_1^* + E[X_1 - c_1^*]^+ = V(\mathbf{X})$, the last equality following from (7), and the proof is complete. □

Theorem 2.1 generalizes the next result due to O'Brien (1983). Our attempts to generalize this result in a different direction are described in Section 4 on random replacement schemes.

COROLLARY 2.1 *Let $(I_1, \ldots, I_n)$ and $(J_1, \ldots, J_n)$ denote random sampling with and without replacement, respectively, from $\{1, \ldots, N\}$, $n \leq N$. Let $X_k = r_k(I_k)$ and $Y_k = r_k(J_k)$, where for all $k = 1, \ldots, n$, $r_k(i) \leq r_k(j)$ if $1 \leq i < j \leq N$. Then $V(\mathbf{X}) \leq V(\mathbf{Y})$.*

PROOF This follows from the fact that $X_1, \ldots, X_n$ are independent, while $Y_1, \ldots, Y_n$ are NA, hence NLODS, and Theorem 2.1 applies. □

In Rinott and Samuel-Cahn (1991), we consider (among other things) the following problem. Given independent $Z_1, \ldots, Z_n$, let

$$V(s) = V(\mathbf{Z} | \sum_{i=1}^{n} Z_i = s)$$

denote the optimal stopping value with respect to observations having the conditional distribution of $\mathbf{Z}$ given $\sum_{i=1}^{n} Z_i = s$. In trying to understand the interaction between dependence and stochastic ordering, and how they affect optimal stopping values, it is natural to seek conditions under which $V(s)$ increases as a function of s. We found that if each of the Z_i's has a log-concave density, or probability function, then indeed $V(s)$ is increasing. In this case the observations are also NA, see Joag-Dev and Proschan (1983).

3. Value Comparisons Under Positive Dependence

In view of Theorem 2.1 and previous discussions, it appears natural to look for structures of positively dependent random variables $(X_1, \ldots, X_n)$, such that for independent $(Y_1, \ldots, Y_n)$ with X_i and Y_i having the same marginal distribution for each $i = 1, \ldots, n$, we have,

$$V(\mathbf{X}) \leq V(\mathbf{Y}). \tag{11}$$

Association in the sense of Esary, Proschan and Walkup (1967) is an example of a well-known strong condition of positive dependence. The variables $X_1, \ldots, X_n$ are said to be associated if $\text{cov}(f_1(X_1, \ldots, X_n), f_2(X_1, \ldots, X_n)) \geq 0$ for any pair of nondecreasing functions f_1 and f_2. While in Theorem 2.1, a suitable (and rather weak) notion of negative dependence was sufficient for the value comparison, this is not the case for comparisons under positive dependence. For example, the variables (X_1, X_2) of (3) are easily shown to be associated. However, if we set Y_i to be independent having the same marginal distribution as X_i, $i = 1, 2$, then $V(\mathbf{Y}) = 1 < V(\mathbf{X}) = 5$.

With the lack, so far, of general results of comparisons of the type (11) under positive dependence, we shall settle for a few examples. In the first three examples it is easy to see that the X's are associated.

EXAMPLE 3.1 *Let Z_i be independent random variables, and let $0 \leq \alpha_i \leq 1$ be constants. Set $X_1 = Z_1$ and $X_i = \alpha_i X_{i-1} + (1 - \alpha_i) Z_i$, $i = 2, \ldots, n$. Let $(Y_1, \ldots, Y_n)$ be independent random variables with Y_i having the same marginal distribution as X_i, $i = 1, \ldots, n$. Then*

$$V(\mathbf{X}) \leq V(\mathbf{Y}).$$

In the special case that $\alpha_i = (i-1)/i$, we obtain the averages $X_i = \frac{1}{i} \sum_{j=1}^{i} Z_j$.

PROOF The proof is by induction on n. For $n = 2$, set $a = EZ_2$. We have

$$V(X_1, X_2) = E\{X_1 \vee [\alpha_2 X_1 + (1 - \alpha_2) E Z_2]\}$$

$$
\begin{aligned}
&= E\{X_1 I(X_1 \geq a)\} + \alpha_2 E\{X_1 I(X_1 < a)\} \\
&\quad +(1-\alpha_2) a P(X_1 < a) \\
&\leq E\{X_1 I(X_1 \geq a)\} + \alpha_2 E X_1 \cdot P(X_1 < a) \\
&\quad +(1-\alpha_2) a P(X_1 < a) \\
&= E\{X_1 I(X_1 \geq a)\} + E X_2 \cdot P(X_1 < a) \\
&\leq \sup_b \{ E\{X_1 I(X_1 \geq b)\} + E X_2 \cdot P(X_1 < b) \} \\
&= E\{X_1 \vee E X_2\} = E\{Y_1 \vee E Y_2\} = V(Y_1, Y_2).
\end{aligned}
$$

Now set $V^{(n-1)}(X_1) = \sup_{2 \leq t \leq n} E(X_t | X_1)$, and $\tilde{V}^{(n-1)}(Y_1) = \sup_{2 \leq t \leq n} E(Y_t | Y_1)$, which actually does not depend on Y_1. Note that the Markov structure of the X's implies $EV^{(n-1)}(X_1) = V(X_2, \ldots, X_n)$, and clearly $E\tilde{V}^{(n-1)}(Y_1) = V(Y_2, \ldots, Y_n)$. The induction hypothesis can be expressed in the form

$$EV^{(n-1)}(X_1) = V(X_2, \ldots, X_n) \leq V(Y_2, \ldots, Y_n) = E\tilde{V}^{(n-1)}(Y_1).$$

Note that $V^{(n-1)}(X_1)$ is a nondecreasing function of X_1. From the structure of the sequence $(X_1, \ldots, X_n)$ with $0 \leq \alpha_i \leq 1$, it is not hard to see that there exists a value $-\infty \leq c \leq \infty$ such that $X_1 \geq V^{(n-1)}(X_1)$ if and only if $X_1 \geq c$. We obtain

$$
\begin{aligned}
V(X_1, \ldots, X_n) &= E\{X_1 \vee V^{(n-1)}(X_1)\} \\
&= E\{X_1 I(X_1 \geq c)\} + E\{V^{(n-1)}(X_1) I(X_1 < c)\} \\
&\leq E\{X_1 I(X_1 \geq c)\} + EV^{(n-1)}(X_1) \cdot P(X_1 < c) \\
&\leq E\{X_1 I(X_1 \geq c)\} + E\tilde{V}^{(n-1)}(Y_1) \cdot P(X_1 < c) \\
&\leq \sup_b \{ E\{X_1 I(X_1 \geq b)\} + E\tilde{V}^{(n-1)}(Y_1) \cdot P(X_1 < b) \} \\
&= E\{X_1 \vee E\tilde{V}^{(n-1)}(Y_1)\} = E\{Y_1 \vee E\tilde{V}^{(n-1)}(Y_1)\} \\
&= V(Y_1, \ldots, Y_n),
\end{aligned}
$$

where the first inequality follows from the monotonicity of $V^{(n-1)}(X_1)$, the second inequality follows from the induction hypothesis, and the last equality follows by the independence of the Y_i's. □

The next example generalizes Example 3.1. Note that a real valued Markov chain can always be represented in the form $X_1 = Z_1$, $X_i = f_i(X_{i-1}, Z_i)$ with independent Z's. Note also that if the functions $f_i(x, z)$ are increasing in x, then the sequence $X_1, \ldots, X_n$ is *Conditionally Increasing in Sequence*, i.e., $P(X_{i+1} > x \,|\, X_1 = x_1, \ldots, X_i = x_i)$ is nondecreasing in $x_1, \ldots, x_i$, for all x and $i = 1, \ldots, n-1$. This implies that $X_1, \ldots, X_n$ are associated. For details see Barlow and Proschan (1975, Theorem 4.7).

EXAMPLE 3.2 *Let Z_i be independent random variables, and let $X_1, \dots, X_n$ have the Markov structure $X_1 = Z_1$, $X_i = f_i(X_{i-1}, Z_i)$, where Z_i are independent, and $f_i(x,z)$ are functions satisfying $0 \le \frac{\partial f_i}{\partial x} \le 1$, $i = 2, \dots, n$ (or $0 \le f_i(x', z) - f_i(x,z) \le x' - x$ for any $x' > x$ in the nondifferentiable case). Let $(Y_1, \dots, Y_n)$ be independent random variables with Y_i having the same marginal distribution as X_i, $i = 1, \dots, n$. Then*

$$V(\mathbf{X}) \le V(\mathbf{Y}).$$

PROOF As in the proof of Example 3.1 it suffices to show that

1. $V^{(n-1)}(X_1) = \sup_{2 \le t \le n} E(X_t | X_1)$, is nondecreasing in X_1,
2. there exists a value $-\infty \le c \le \infty$ such that $X_1 \ge V^{(n-1)}(X_1)$ if and only if $X_1 \ge c$.

1. is readily shown by induction using the monotonicity of f_i in x. We prove 2. by showing that for $X_1' > X_1$, $V^{(n-1)}(X_1') - V^{(n-1)}(X_1) \le X_1' - X_1$. For $n = 2$, this follows readily from $V^{(1)}(X_1) = E[f_2(X_1, Z_2) | X_1] = \int f_2(X_1, z) dF(z)$, and $\frac{\partial f_2}{\partial x} \le 1$, where F denotes the distribution of Z_2.

The proof now requires induction; we prefer to demonstrate the case $n = 3$, and leave the details of the induction to the reader. For $n = 3$ we have $V^{(2)}(X_1) = E\{[f_2(X_1, Z_2) \vee E(f_3(X_2, Z_3) | X_2)] | X_1\}$, and $f_3(X_2, Z_3) = f_3(f_2(X_1, Z_2), Z_3)$. We have $E(f_3(X_2, Z_3) | X_2) = h(X_2)$, say, where (by arguments as above), $X_2' > X_2$ implies $h(X_2') - h(X_2) \le X_2' - X_2$. Define $g(X_1, Z_2) = h(X_2) = h(f_2(X_1, Z_2))$. Then, replacing f_2 by f for brevity, we have $V^{(2)}(X_1) = \int [f(X_1, z) \vee g(X_1, z)] dF(z)$ where F denotes the cdf of Z_2, $0 \le \frac{\partial f}{\partial x} \le 1$, and if g is differentiable $\frac{\partial g}{\partial x} \le 1$ (by the chain rule), and in any case $X_1' > X_1$ implies $g(X_1', z) - g(X_1, z) \le X_1' - X_1$. Note that $f(X_1', z) \vee g(X_1', z) - f(X_1, z) \vee g(X_1, z) \le [f(X_1', z) - f(X_1, z)] \vee [g(X_1', z) - g(X_1, z)]$, so that for $X_1' > X_1$

$$f(X_1', z) \vee g(X_1', z) - f(X_1, z) \vee g(X_1, z) \le X_1' - X_1,$$

and substituting Z_2 for z, and taking expectations, we obtain for $n = 3$, $V^{(n-1)}(X_1') - V^{(n-1)}(X_1) \le X_1' - X_1$. □

EXAMPLE 3.3 *Let $Z_0, Z_1, \dots, Z_n$ be independent random variables and let $X_i = Z_0 + Z_i$, $i = 1, \dots, n$. Let $Y_1, \dots, Y_n$ be independent random variables with Y_i having the same marginal distribution as X_i, $i = 1, \dots, n$. Then*

$$V(\mathbf{X}) \le V(\mathbf{Y}).$$

PROOF It is not hard to verify the relations

$$V(\mathbf{X}) \le EZ_0 + V(Z_1, \dots, Z_n) \le V(\mathbf{Y}).$$

The first inequality above is left to the reader. In order to prove the second, set $Y_i = Z_{0i}+Z_i$, where all the Z's are independent, and Z_{0i} is distributed like Z_0, $i = 1,\dots,n$. We then have $V(\mathbf{Y}) = Ef(Z_{01},\dots,Z_{0n-1},Z_1,\dots,Z_{n-1})$ where f is a convex function (which depends on the constant $E(Z_0 + Z_n)$). For example, for $n = 3$ we have $V(\mathbf{Y}) = E\{(Z_{01} + Z_1) \vee E[(Z_{02} + Z_2) \vee E(Z_{03} + Z_3)]\}$. By Jensen's inequality and the independence of the Z's, we obtain a lower bound to the latter expression by replacing the variables Z_{0i} by their expectation EZ_0. The lower bound thus obtained is readily seen to equal $EZ_0 + V(Z_1,\dots,Z_n)$. □

EXAMPLE 3.4 *Let $X_1,\dots,X_n$ be a martingale, and let $Y_1,\dots,Y_n$ be independent random variables with Y_i having the same marginal distribution as X_i, $i = 1,\dots,n$. Then*

$$V(\mathbf{X}) \leq V(\mathbf{Y}).$$

PROOF Simply note that $V(\mathbf{X}) = EX_1 = EY_1 \leq V(\mathbf{Y})$. □

Note that being a martingale, $X_1,\dots,X_n$ are nonnegatively correlated, but need not be associated.

4. Value Comparisons for Random Replacement Schemes

Random replacement schemes were introduced by Karlin (1974). Consider sampling from a finite population, say $\mathcal{N} = \{1,\dots,N\}$; when the ith observation is taken, it is returned to the population with some probability, say π_i, independently of observation values, and removed with probability $1 - \pi_i$. The observations are taken at each step at random, that is, with equal probability for every number present in the population. Clearly sampling with and without replacement are special cases, and one might look for a hierarchy of comparisons, or ordering, generalizing the comparison in Corollary 2.1.

We now define random replacement schemes more formally. Set $\pi = (\pi_1,\dots,\pi_{n-1})$ with $0 \leq \pi_i \leq 1$, and let U_i be independent Bernoulli variables, $P(U_i = 1) = \pi_i$, $i = 1,\dots,n-1$. Consider an urn (or population) containing the values $\{1,\dots,N\}$. Select a value J_1 at random from the urn; return it if $U_1 = 1$, and remove it from the urn if $U_1 = 0$. Now select J_2 at random from the resulting urn, and return it if and only if $U_2 = 1$. Continue in this manner until a sample $(J_1,\dots,J_n)$ is obtained. Now define $X_k = r_k(J_k)$, where the real valued functions $r_k(i)$, $i \in \{1,\dots,N\}$, $k = 1,\dots,n$ are monotone nondecreasing in i for each k. This monotonicity will always be assumed in the sequel. Other conditions on $r_k(i)$ will appear later. The functions $r_k(i)$ may be seen as the reward for drawing the value i at step k, and at each step the rewards increase with the value drawn. Define the optimal stopping value to be

$$V_\pi^{(n)} = \sup_t EX_t,$$

where the supremum is taken over stopping rules with respect to the fields $\mathcal{F}_k(J_1, \ldots, J_k, U_1, \ldots, U_k)$. This means that the content of the urn at the time of the (possible) next draw is always known.

In these terms, Corollary 2.1 can be recast in the form:

$$V_{\mathbf{1}}^{(n)} \le V_{\mathbf{0}}^{(n)}, \tag{12}$$

where $\mathbf{1}$ ($\mathbf{0}$) denotes the $n-1$-vector of 1's (0's).

The following generalization of (12) holds.

THEOREM 4.1 (Rinott and Samuel-Cahn (1991)). *For any* π *and all* $n \le N$,

$$V_\pi^{(n)} \le V_{\mathbf{0}}^{(n)}. \tag{13}$$

We shall review the proof of this theorem at the end of this section.

One may conjecture that (12) can also be generalized to:

$$V_{\mathbf{1}}^{(n)} \le V_\pi^{(n)}. \tag{14}$$

However this is not true in general. For $N = n = 3$, $r_1(\cdot) \equiv 0$, $r_2(1) = 0$, $r_2(2) = r_2(3) = 3$, $r_3(1) = r_3(2) = 0$, $r_3(3) = 4$, we have, $V_{11}^{(3)} = 22/9 > V_{01}^{(3)} = 21/9$. It is possible that (14) holds if the functions $r_k(i)$ do not depend on k, or perhaps also when they are decreasing in k, for each i, i.e., values are discounted in time of observation.

Note that in general the sequence $\mathbf{X}$ obtained in random replacement schemes is not NLODS, even when $r_k(i)$ does not depend on k. For example, if $n = N = 3$, and $r_k(i) = i, \quad i,k = 1,2,3$, and $\pi_1 = 0, \pi_2 = 1$, then it is easily seen that $P(X_3 < 3 | X_2 < 3) = 3/4 > P(X_3 < 3) = 2/3$. Thus, (14) cannot be derived from Theorem 2.1. For $n = 2$, (14) is easy:

LEMMA 4.1 *For* $n = 2$, (14) *holds.*

NOTATION Define $V_{\pi_1,\ldots,\pi_{k-1}}^{(k)}(\mathcal{M})$ to be the optimal stopping value when initially the urn contains the elements of an ordered set $\mathcal{M}$ where $|\mathcal{M}| \ge k$, at most k draws are allowed, and the replacement probabilities are $\pi_1, \ldots, \pi_{k-1}$. The functions r_k are suppressed in this notation. We may use $V_{\pi_1,\ldots,\pi_{k-1}}^{(k)}(\mathcal{M})$ with $r_1, \ldots, r_k$, and also with $r_2, \ldots, r_{k+1}$. We shall comment on this point when the latter case occurs, although the notation should be clear from the context.

PROOF OF LEMMA 4.1 For $n = 2$, we have

$$\begin{aligned} V_{\pi_1}^{(2)} = V_{\pi_1}^{(2)}(\mathcal{N}) &= \pi_1 E\{X_1 \vee V^{(1)}(\mathcal{N})\} \\ &\quad +(1-\pi_1)E\{X_1 \vee V^{(1)}(\mathcal{N} - \{J_1\})\} \\ &= \pi_1 V_1^{(2)}(\mathcal{N}) + (1-\pi_1)V_0^{(2)}(\mathcal{N}) \geq V_1^{(2)}(\mathcal{N}), \end{aligned} \tag{15}$$

where the final inequality follows from $V_1^{(2)}(\mathcal{N}) \leq V_0^{(2)}(\mathcal{N})$, which is a simple case of (12). Here, $V^{(1)}$ was used with respect to the function r_2. □

For $n > 3$, we are unable to prove (14) even in the seemingly simple case of $r_k(i)$ not depending on k. However, if it is true, then the following lemma could be a step in the right direction. It simplifies (14), which involves a random replacement scheme on the r.h.s., to a comparison between two deterministic schemes: complete replacement, and removal of the first draw followed by complete replacement. The case of $n = 3$ of (14), with some restrictions on r_k, will be derived from this lemma later.

LEMMA 4.2 *Fix $m \geq 3$. $V_1^{(n)} \leq V_\pi^{(n)}$ for all $3 \leq n \leq m$ and N satisfying $n \leq N$, if and only if*

$$V_{1,\dots,1}^{(n)} \leq V_{0,1,\dots,1}^{(n)}, \tag{16}$$

holds for all $3 \leq n \leq m$ and $n \leq N$.

PROOF Clearly, (16) is necessary. To prove sufficiency, we shall make use of the following straightforward generalization of (15):

$$\begin{aligned} V_{\pi_1,\dots,\pi_{n-1}}^{(n)}(\mathcal{N}) &= \pi_1 E\{X_1 \vee V_{\pi_2,\dots,\pi_{n-1}}^{(n-1)}(\mathcal{N})\} \\ &\quad +(1-\pi_1)E\{X_1 \vee V_{\pi_2,\dots,\pi_{n-1}}^{(n-1)}(\mathcal{N} - \{J_1\})\}. \end{aligned} \tag{17}$$

Assuming (16) holds, we now prove $V_1^{(n)} \leq V_\pi^{(n)}$ by induction on n. In the present notation the latter inequality is expressed as

$$V_{1,\dots,1}^{(n)}(\mathcal{N}) \leq V_{\pi_1,\dots,\pi_{n-1}}^{(n)}(\mathcal{N}), \tag{18}$$

which we consider for all n, N such that $n \leq N = |\mathcal{N}|$. The induction hypothesis, see (18), for $n-1$ is

$$V_{1,\dots,1}^{(n-1)}(\mathcal{N}) \leq V_{\pi_2,\dots,\pi_{n-1}}^{(n-1)}(\mathcal{N}),$$

It holds for $n = 3$ $(n - 1 = 2)$ by Lemma 4.1. Applying the induction hypothesis to the r.h.s. of (17) twice, the second time with the population being $\mathcal{N} - \{J_1\}$ instead of $\mathcal{N}$ we obtain the first inequality below:

$$\begin{aligned} V_{\pi_1,\dots,\pi_{n-1}}^{(n)}(\mathcal{N}) &\geq \pi_1 E\{X_1 \vee V_{1,\dots,1}^{(n-1)}(\mathcal{N})\} \\ &\quad +(1-\pi_1)E\{X_1 \vee V_{1,\dots,1}^{(n-1)}(\mathcal{N} - \{J_1\})\} \\ &= \pi_1 V_{1,\dots,1}^{(n)}(\mathcal{N}) + (1-\pi_1)V_{0,1,\dots,1}^{(n)}(\mathcal{N}) \geq V_{1,\dots,1}^{(n)}(\mathcal{N}), \end{aligned} \tag{19}$$

where the equality follows from $E\{X_1 \vee V^{(n-1)}_{1,\dots,1}(\mathcal{N})\} = V^{(n)}_{1,\dots,1}(\mathcal{N})$, and $E\{X_1 \vee V^{(n-1)}_{1,\dots,1}(\mathcal{N}-\{J_1\})\} = V^{(n)}_{0,1,\dots,1}(\mathcal{N})$, and the last inequality follows from (16). In this proof, $V^{(n-1)}$ was always used with respect to $r_2,\dots,r_n$. □

We cannot prove (16) even for $r_k(i)$ not depending on k, but it appears like a more tractable conjecture than (14). A computer search with a variety of functions $r_k(i)$ (not depending on k), and $N \le 15$, did not produce a counterexample to (16). In the case $n = 2$, (14) is already established in Lemma 4.1. For $n = 3$ we have

PROPOSITION 4.1 *Let $n = 3$, $N \ge 3$ and $r_1(i) \ge r_2(i)$, for $i = 1,\dots,N$. Then*

$$V^{(3)}_{1,1} \le V^{(3)}_{0,1}.$$

Clearly, our assumption holds if for all i, $r_k(i)$ is decreasing in k. This is a natural assumption which says that the earlier you observe a certain element i, the higher its value. In other words, there is a cost for time, or for taking more observations. By Lemma 4.2 we conclude that under the conditions of Proposition 4.1, (14) holds for $n = 3$, i.e.,

$$V^{(3)}_{1,1} \le V^{(3)}_{\pi_1,\pi_2}.$$

PROOF OF PROPOSITION 4.1 Define $\bar{r}_k = \frac{1}{N}\sum_{j=1}^{N} r_k(j)$ and

$$\bar{r}_k[i] = \frac{1}{N-1}\sum_{j:i\ne j=1}^{N} r_k(j).$$

Let $A(i) = \frac{1}{N}\sum_{j=1}^{N}\{r_2(j) \vee \bar{r}_3[i]\}$, and $B(i) = \frac{1}{N-1}\sum_{j:i\ne j=1}^{N}\{r_2(j) \vee \bar{r}_3[i]\}$. Note that if for some i, $A(i) > B(i)$, then it is readily seen that $r_2(i) > \bar{r}_3[i]$ and $r_2(i) > A(i)\,(> B(i)\,)$. Since $r_1(i) \ge r_2(i)$, we conclude that $A(i) > B(i)$ implies $r_1(i) > A(i) > B(i)$. It is now easy to see that

$$\frac{1}{N}\sum_{i=1}^{N}\{r_1(i) \vee B(i)\} \ge \frac{1}{N}\sum_{i=1}^{N}\{r_1(i) \vee A(i)\}. \tag{20}$$

Note that the l.h.s. of (20) equals $V^{(3)}_{0,1}$. In order to proceed we now need a simple lemma whose proof is given in Rinott and Samuel-Cahn (1991, Lemma 3.3).

LEMMA 4.3 *Let $h(x), g(x), x \in \mathbb{R}$, be an increasing and a decreasing function, respectively. If X is a random variable such that the expectations below exist, then*

$$E\{h(X) \vee g(X)\} \ge E\{h(X) \vee Eg(X)\}. \tag{21}$$

By Lemma 4.3, the r.h.s. of (20) is $\geq$ than

$$\frac{1}{N}\sum_{j=1}^{N}\{r_1(j) \vee \frac{1}{N}\sum_{i=1}^{N} A(i)\}. \tag{22}$$

Finally, it suffices to show that the r.h.s. of (22) is $\geq$ than $V_{1,1}^{(3)}$. We have,

$$\frac{1}{N}\sum_{i=1}^{N} A(i) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{N}\sum_{i=1}^{N}\{r_2(j) \vee \bar{r}_3[i]\}. \tag{23}$$

Applying Lemma 4.3, or simple convexity, to the inner sum in the r.h.s. of (23) and noting the relation $\frac{1}{N}\sum_{i=1}^{N}\bar{r}_3[i] = \bar{r}_3$ we conclude that the r.h.s. of (23) is $\geq \frac{1}{N}\sum_{j=1}^{N}\{r_2(j) \vee \bar{r}_3\}$. Denote the latter quantity by C. Thus the r.h.s. of (22) is $\geq \frac{1}{N}\sum_{j=1}^{N}\{r_1(j) \vee C\}$, which is exactly $V_{1,1}^{(3)}$ and the proof of Proposition 4.1 is complete. □

For the proof of Theorem 4.1 we shall need a simple lemma whose proof can be found in Rinott and Samuel-Cahn (1991).

LEMMA 4.4 *Let J be a random element of $\mathcal{N}$. Then for any $m \leq N-1$,*

$$V_{\mathbf{0}}^{(m)}(\mathcal{N}) \leq EV_{\mathbf{0}}^{(m)}(\mathcal{N}-\{J\}).$$

In words, removing a random (known) element from the population before sampling, increases the average stopping value for sampling without replacement.

Perhaps this lemma is best explained by an example. For $N=3$, $m=2$, and $r_k(i)=i$, we have $V_{\mathbf{0}}^{(2)}(\{1,2,3\}) = (1/3)\cdot(2+3)/2+(1/3)\cdot 2+(1/3)\cdot 3 = 5/2$, corresponding to the first sampled item being 1,2, or 3, respectively. If prior to sampling, a random element J is removed from $\mathcal{N}=\{1,2,3\}$, we have $EV_{\mathbf{0}}^{(2)}(\{1,2,3\}-\{J\}) = (1/3)\cdot 3+(1/3)\cdot 3+(1/3)\cdot 2 = 8/3$, corresponding to the removed element J being 1,2, or 3, respectively.

PROOF OF THEOREM 4.1 It is easy to see that arguments similar to those given for Lemma 4.2 imply also that in order to prove $V_{\pi}^{(n)} \leq V_{\mathbf{0}}^{(n)}$ it suffices to prove $V_{1,0,\ldots,0}^{(n)} \leq V_{0,\ldots,0}^{(n)}$. In order to prove the latter inequality, note that $V_{\mathbf{0}}^{(n-1)}(\mathcal{N}-\{J_1\})$ is decreasing in J_1, while $X_1 = r_1(J_1)$ is increasing in J_1. Applying Lemma 4.3 and then Lemma 4.4 to obtain the inequalities below, we have

$$\begin{aligned} V_{0,\ldots,0}^{(n)} &= E\{X_1 \vee V_{\mathbf{0}}^{(n-1)}(\mathcal{N}-\{J_1\})\} \\ &\geq E\{X_1 \vee EV_{\mathbf{0}}^{(n-1)}(\mathcal{N}-\{J_1\})\} \\ &\geq E\{X_1 \vee V_{\mathbf{0}}^{(n-1)}(\mathcal{N})\} = V_{1,0,\ldots,0}^{(n)}, \end{aligned}$$

and the proof is complete. In this proof, $V^{(n-1)}$ was used with respect to $r_2, \ldots, r_n$. □

5. Prophet Inequalities

In this section we review certain prophet inequalities for some of the models discussed above. We start with simple technical lemmas.

LEMMA 5.1 *Let $(X_1, \ldots, X_n)$ be either NLODS random variables (including independent random variables), or observations arising under any random replacement scheme of the type described in Section 4, with $r_k(i) \geq r_{k+1}(i)$ for all i, k. Then for any constant c,*

$$E\{[X_k - c]^+ \,|\, X_1 \vee \cdots \vee X_{k-1} < c\} \geq E[X_k - c]^+, \; k = 2, \ldots, n.$$

PROOF For independent random variables the result is obvious (with equality), and the inequality follows easily from the definition of NLODS variables. For random replacement schemes the result follows from the fact that conditionally on any values of $X_1 < c, \ldots X_{k-1} < c$, and any replacement indicators $U_1, \ldots, U_n$, X_k is distributed as $r_k(J)$, where J is drawn from an urn from which some elements I with $r_k(I) < c$ have been removed. Here the monotonicity of $r_k(i)$ in k was used. □

Henceforth we shall consider only nonnegative random variables, and exclude (without further mention) the trivial case that they are all identically zero. For such a sequence $(X_1, \ldots, X_n)$, and $b \geq 0$, let $t(b)$ denote the stopping time: $t(b) = \inf\{k : X_k \geq b\} \wedge n$.

LEMMA 5.2 *Let $(X_1, \ldots, X_n)$ be nonnegative random variables. Let $b > 0$ be the unique constant satisfying $\sum_{k=1}^n E[X_k - b]^+ = b$. Suppose*

$$E\{[X_k - b]^+ \,|\, X_1 \vee \cdots \vee X_{k-1} < b\} \geq E[X_k - b]^+, \; k = 2, \ldots, n. \tag{24}$$

Then

$$b < EX_{t(b)}. \tag{25}$$

PROOF

$$\begin{aligned} EX_{t(b)} &\geq E\{X_{t(b)} I(X_1 \vee \cdots \vee X_n \geq b)\} \\ &= E\{\, bI(X_1 \vee \cdots \vee X_n \geq b) \\ &\qquad + \sum_{k=1}^{n} [X_k - b]^+ I(X_1 \vee \cdots \vee X_{k-1} < b)\,\} \\ &\geq bP(X_1 \vee \cdots \vee X_n \geq b) \end{aligned}$$

$$+\sum_{k=1}^{n} E[X_k - b]^+ P(X_1 \vee \cdots \vee X_{k-1} < b)$$
$$> \ bP(X_1 \vee \cdots \vee X_n \geq b)$$
$$+P(X_1 \vee \cdots \vee X_n < b) \sum_{k=1}^{n} E[X_k - b]^+ = b\,;$$

here the first inequality holds because $X_i \geq 0$, the second inequality follows from (24), and the last inequality from $P(X_1 \vee \cdots \vee X_{k-1} < b) \geq P(X_1 \vee \cdots \vee X_n < b)$, with strict inequality for the first k such that $P(X_k > b) > 0$. □

LEMMA 5.3 *Let $(Y_1, \ldots, Y_n)$ be any nonnegative random variables, and let $b \geq 0$ be the unique constant satisfying $\sum_{k=1}^{n} E[Y_k - b]^+ = b$. Then $E\{Y_1 \vee \cdots \vee Y_n\} \leq 2b$.*

PROOF Simply take expectations on both sides of the simple relation: $Y_1 \vee \cdots \vee Y_n \leq b + \sum_{k=1}^{n} [Y_k - b]^+$. □

THEOREM 5.1 *Let $(X_1, \ldots, X_n)$ be nonnegative random variables which are either NLODS (including independent random variables), or observations arising under any random replacement scheme of the type described in Section 4, with $r_k(i) \geq r_{k+1}(i)$ for all i, k, or any other random variables which satisfy for every constant c,*

$$E\{[X_k - c]^+ \mid X_1 \vee \cdots \vee X_{k-1} < c\} \geq E[X_k - c]^+, \ k = 2, \ldots, n.$$

Then the prophet inequality

$$E\{X_1 \vee \cdots \vee X_n\} < 2V(\mathbf{X}) \tag{26}$$

holds. Moreover, if $(Y_1, \ldots, Y_n)$ are any *nonnegative random variables such that for each i, X_i and Y_i have the same (marginal) distribution, $i = 1, \ldots, n$, then*

$$E\{Y_1 \vee \cdots \vee Y_n\} < 2V(\mathbf{X}).$$

PROOF It clearly suffices to prove the second part of the theorem. Noting that the quantity b defined in Lemmas 5.2 - 5.3 depends on marginal distributions only, and applying Lemmas 5.1 - 5.3 we have,

$$\frac{1}{2} E\{Y_1 \vee \cdots \vee Y_n\} \leq b < EX_{t(b)} \leq V(\mathbf{X}). \tag{27}$$ □

For *independent* random variables the inequality (26) was obtained by Krengel and Sucheston (1978). This latter article provided the inspiration to a large body of results on prophet inequalities. For independent $0 \leq X_k \leq 1$, Hill (1983) sharpened the result to

$$E\{X_1 \vee \cdots \vee X_n\} < 2V(\mathbf{X}) - V(\mathbf{X})^2 .$$

The negative dependence condition of Theorem 5.2 below, which generalizes Hill's result, is stronger than the NLODS condition; however it is weaker than CDS (see Section 2).

THEOREM 5.2 (Samuel-Cahn (1991)) *Let* $0 \leq X_k \leq 1$, $k = 1, \ldots, n$, *and suppose that* $X_1, \ldots, X_n$ *are negatively dependent in the sense that* $P(X_k < a_k | X_1 < a_1, \ldots, X_{k-1} < a_{k-1})$ *is nondecreasing in* $a_1, \ldots, a_{k-1}$, *for all* $k = 2, \ldots, n$. *Then*

$$E\{X_1 \vee \cdots \vee X_n\} < 2V(\mathbf{X}) - V(\mathbf{X})^2 .$$

For positively dependent random variables we quote a result for averages (recall Example 3.1).

THEOREM 5.3 (Hill (1986)) *Let* Z_i *be independent nonnegative random variables, and consider the averages* $X_i = \frac{1}{i}\sum_{j=1}^{i} Z_j$, $i = 1, \ldots, n$. *Then*

$$E\{X_1 \vee \cdots \vee X_n\} < 2V(\mathbf{X}).$$

REFERENCES

BARLOW, R. E. AND PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing Probability Models.* Holt, Rinehart and Winston, New York.

CHOW, Y. S., ROBBINS, H. AND SIEGMUND, D. (1971). *Great Expectations: The Theory of Optimal Stopping.* Houghton Mifflin, Boston, MA.

ESARY, J. D., PROSCHAN, F. AND WALKUP, D. W. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **44** 1466-1474.

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

HILL, T. P. (1986). Prophet inequalities for averages of independent non-negative random variables. *Math. Zeit.* **192** 427-436.

HILL, T. P. AND KERTZ, R. P. (1992). A survey of prophet inequalities in optimal stopping theory. In *Strategies for Sequential Search and Selection in Real-Time.* F.T. Bruss, T.S. Ferguson, and S.M. Samuels, eds. American Mathematical Society, Providence, RI. 191-207.

JOAG-DEV, K. AND PROSCHAN, F. (1983). Negative association of random variables, with applications. *Ann. Statist.* **11** 286-295.

KARLIN, S. (1974). Inequalities for symmetric sampling plans I. *Ann. Statist.* **2** 1065-1094.

KRENGEL, U. AND SUCHESTON, L. (1978). On semiamarts, amarts, and processes with finite value. In *Probability on Banach Spaces,* J. Kuelbs, ed. Marcel Dekker, New York. 197-266.

O'BRIEN, G. L. (1983). Optimal stopping when sampling with and without replacement. *Z. Wahrsch. verw. Gebiete.* **64** 125-128.

RINOTT, Y. AND SAMUEL-CAHN, E. (1987). Comparisons of optimal stopping values and prophet inequalities for negatively dependent random variables. *Ann. Statist.* **15** 1482-1490.

RINOTT, Y. AND SAMUEL-CAHN, E. (1991). Orderings of optimal stopping values and prophet inequalities for certain multivariate distributions. *J. Multivariate Anal.* **37** 104-114.

SAMUEL-CAHN, E. (1991). Prophet inequalities for bounded negatively dependent random variables. *Statist. Prob. Lett.* **12** 213-216.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, SAN DIEGO
LA JOLLA, CA 92093

DEPARTMENT OF STATISTICS
HEBREW UNIVERSITY
JERUSALEM 91905, ISRAEL