The Annals of Probability
1992, Vol. 20, No. 3, 1222–1228

A DIFFERENCE PROPHET INEQUALITY FOR BOUNDED I.I.D. VARIABLES, WITH COST FOR OBSERVATIONS

BY ESTER SAMUEL-CAHN

Hebrew University of Jerusalem

Let X_i be i.i.d. random variables, $0 \le X_i \le 1$ and $c \ge 0$, and let $Y_i = X_i - ic$. It is shown that for all n, all c and all such X_i, $E(\max_{i \ge 1} Y_i) - \sup_t EY_t < e^{-1}$, where t is a stopping rule and e^{-1} is shown to be the best bound for which the inequality holds. Specific bounds are also obtained for fixed n or fixed c. These results are very similar to those obtained by Jones for a similar problem, where $0 \le X_i \le 1$ are independent but not necessarily identically distributed. All results are valid and unchanged also when Y_i is replaced by $Y_i^* = \max_{1 \le j \le i} X_j - ic$.

1. Introduction. In a recent paper, Jones (1990) considers the following "prophet problem." Let X_i be independent, $i = 1, 2, \dots,$ $0 \le X_i \le 1$ and let $c \ge 0$ be a fixed constant. Consider the optimal stopping problem for the sequence $Y_i = X_i - ic$, $i = 1, 2, \dots,$ which corresponds to a reward X_i minus a fixed cost c of sampling, for each observation. Let $V(Y_1, \dots, Y_n) = \sup\{EY_t\colon t \le n,\ t \text{ a stopping rule}\}$. Let $[x]$ denote the largest integer smaller than x. Jones (1990) proves the following very interesting "difference prophet inequality" (see Theorem A):

(a) For $0 < c \le 1$ fixed and all n and $0 \le X_i \le 1$,

$$E\Big(\max_{1 \le i \le n} Y_i\Big) - V(Y_1, \dots, Y_n) \le [1/c]c(1-c)^{[1/c]}. \tag{1.1}$$

(b) For $n \ge 1$ fixed and all $c \ge 0$ and $0 \le X_i \le 1$,

$$E\Big(\max_{1 \le i \le n} Y_i\Big) - V(Y_1, \dots, Y_n) \le (1 - 1/n)^n. \tag{1.2}$$

(c) For all $c \ge 0$, $n \ge 1$ and $0 \le X_i \le 1$,

$$E\Big(\max_{1 \le i \le n} Y_i\Big) - V(Y_1, \dots, Y_n) \le e^{-1}. \tag{1.3}$$

All the preceding bounds are sharp; that is, the constants in the right-hand sides cannot be replaced by any smaller constants. The term "difference prophet inequality" is justified since $E(\max_{1 \le i \le n} Y_i)$ can be interpreted as the expected optimal return to the prophet who can foresee the future and choose the largest Y_i. Jones also shows that the corresponding "ratio prophet inequality" is unbounded.

In the present note we consider the preceding problem when X_i are taken to be i.i.d., $0 \le X_i \le 1$. The case where there is no cost of sampling (i.e., $c = 0$) was one of the first "prophet problems" considered. See Krengel and

Received April 1990; revised February 1991.
AMS 1980 *subject classifications*. Primary 60G40; secondary 60E15.
Key words and phrases. Prophet inequality, optimal stopping, cost of observation.

Sucheston (1978), Hill and Kertz (1981a) and Samuel–Cahn (1984), for the ratio prophet inequality, and Hill and Kertz (1981b) for the difference prophet inequality. The case of i.i.d. X_i is considered in Hill and Kertz (1982) and Kertz (1986). For $c = 0$, the i.i.d. case turns out to be a much more difficult problem than the problem where the assumption identically distributed is dropped. Unlike the case for $c = 0$, when $c > 0$ the i.i.d. problem turns out to be a considerably simpler problem. This is due to the very simple form of the optimal rule and its value in this case, which was obtained by MacQueen and Miller (1960), and Chow and Robbins (1961).

We clearly need consider only $c \le 1$, otherwise both the maximum and the optimal stopping value are obtained for $n = 1$ and the difference is 0. The purpose of the present note is to prove the following:

THEOREM 1. *Let X_i be i.i.d., $0 \le X_1 \le 1$.*
(a) *For $0 < c \le 1$ fixed, and all n and X_1,*

$$E\Big(\max_{1\le i\le n} Y_i\Big) - V(Y_1,\dots,Y_n) \le [1/c]c(1-c)^{[1/c]+1}. \tag{1.4}$$

(b) *For $n \ge 1$ fixed, and all $c \ge 0$ and X_1,*

$$E\Big(\max_{1\le i\le n} Y_i\Big) - V(Y_1,\dots,Y_n) \le (1-1/n)^{n+1}. \tag{1.5}$$

(c) *For all $c \ge 0$, all X_1 and all finite or infinite sequences,*

$$E\Big(\max_{1\le i} Y_i\Big) - V(Y_1,Y_2,\dots) < e^{-1}. \tag{1.6}$$

All bounds are the best possible.

Comparing Theorem 1 with Jones' results (1.1), (1.2) and (1.3), we see that for the problem considered here there is very little difference whether one considers just the independent case or the i.i.d. case. This is quite different from the findings when $c = 0$. The "ratio prophet inequality" in the i.i.d. case is again unbounded.

Letting $c \to 0$ in the right-hand side of (1.4) yields e^{-1}. If we could argue continuity, we would get that, for $0 \le X_i \le 1$ and i.i.d. X_i, $E(\max_{1\le i} X_i) - V(X_1, X_2,\dots) \le e^{-1}$, where this is the best bound. This is, however, false, as follows from the result of Hill and Kertz (1981b). They show that even for nonidentically distributed, independent X_i, $0 \le X_i \le 1$, the preceding difference is bounded by $1/4 < e^{-1}$, and $1/4$ is the best bound and is attainable both in the finite $n \ge 2$ and the infinite case.

2. The proof. We may assume $n \ge 2$; otherwise the theorem is trivial.

Chow and Robbins (1961) consider, among other problems, the optimal stopping problem (for infinite horizon) for the payoff sequence $Y_i^* = \max_{1\le j\le i} X_j - ic$, $i = 1,2,\dots,$ where $X, X_1, X_2,\dots,$ are i.i.d. with finite expectation and $c > 0$. They show [see also Chow, Robbins and Siegmund (1971), page 56(a)] that for this payoff one is in the monotone case, an optimal rule s exists and is given by the simple structure

$$s = \inf\{i\colon X_i \geq \beta\}, \tag{2.1}$$

where β is the unique value for which $E(X-\beta)^+ = c$. (Note that possibly $\beta < 0$.) Also $EY_s^* = \beta$. This beautiful and rather surprising result implies, since $Y_i \leq Y_i^*$, that also for $Y_i = X_i - ic$ the rule s is optimal and $EY_s = \beta$, even though the sequence $\{Y_i\}$ does not correspond to the monotone case. By Exercise 3 on page 60 of Chow, Robbins and Siegmund (1971) it follows that $s_n = \min\{s, n\}$ is optimal for Y_i^* among all stopping rules satisfying $t \leq n$.

One might tend to believe, by an argument similar to that above, that $s_n = \min(s, n)$ is also optimal for Y_i among all $t \leq n$. This is, however, generally false (for $\beta > 0$), as is easily seen by direct backward induction. This was pointed out by a referee, to whom the author is very grateful. (The reason this conclusion fails is that when forced to stop at stage n, the return under Y_n is generally smaller than the return under Y_n^*. Note, however, that when X_i are Bernoulli 0, 1 random variables, the return under Y_{s_n} is exactly equal to the return under $Y_{s_n}^*$.)

We shall still be content using s_n with payoff sequence Y_i, and our first goal will be to maximize the difference $E(\max_{1 \leq i \leq n} Y_i) - EY_{s_n}$.

We show that this difference is always less than or equal to the right-hand sides in (1.4) and (1.5), respectively, but the maximum difference achieves equality to the right-hand sides of (1.4) and (1.5). Actually we show that this equality is achieved for (some special) Bernoulli random variables, taking only the values 0 and 1 with positive probability. For Bernoulli random variables it is, however, immediate that $EY_{s_n}^* = EY_{s_n}$, and hence for these variables $V(Y_1, \dots, Y_n) = EY_{s_n}$. This shows that equality of the right-hand sides and the left-hand sides of (1.4) and (1.5) holds for those variables. Theorem 1 then follows, since generally

$$E\Big(\max_{1 \leq i \leq n} Y_i\Big) - V(Y_1, \dots, Y_n) \leq E\Big(\max_{1 \leq i \leq n} Y_i\Big) - EY_{s_n}.$$

PROPOSITION 2.1.

$$EY_{s_n} = \beta - (1-u)^n\{\beta - E(X|X < \beta)\}, \tag{2.2}$$

where $u = P\{X \geq \beta\}$ and where the value in (2.2) is β when $u = 1$.

PROOF. The result follows quite straightforwardly, using $E(X|X \geq \beta) = c/u + \beta$. $\square$

For fixed X and c, β and u are uniquely determined. We consider first the case $\beta \geq 0$.

Let X_i' be the "dilated" random variable obtained from X_i by letting $X_i = X_i'$ when $X_i < \beta$, but for $X_i \geq \beta$, X_i' assumes the values 1 and β only, with probabilities p and $u - p$, respectively, where

$$p = c/(1 - \beta). \tag{2.3}$$

Let Y_i', $i = 1, \dots, n$, be the Y values defined for the X_i'.

PROPOSITION 2.2. (i) $EY'_{s_n} = EY_{s_n}$.
(ii) $E(\max_{1 \le i \le n} Y_i') \ge E(\max_{1 \le i \le n} Y_i)$. *Hence dilation can only increase the difference* $E(\max_{1 \le i \le n} Y_i) - EY_{s_n}$.

PROOF. Clearly X_i and X_i' have the same β and u values. Thus (i) follows from (2.2). Since $\max_{1 \le i \le n} Y_i$ is a convex function of $X_1, \dots, X_n$ and since the expected value of any convex function can only increase (or remain unchanged) by any dilation, (ii) follows. □

The preceding argument shows that from the point of view of maximizing $E(\max_{1 \le i \le n} Y_i)$ one need consider Bernoulli random variables only, since any X_i satisfying $0 \le X_i \le 1$ can be dilated to take the values 0 and 1 only. We shall show that for given β, in order to minimize (2.2) one again need consider only Bernoulli random variables.

PROPOSITION 2.3. *Fix* $\beta \ge 0$. *Then* EY_{s_n} *is minimized for* $X_i' = 1$ *and* 0 *with probabilities* p *and* $1 - p$, *where* p *is given in* (2.3).

PROOF. For $\beta = 0$ the statement is clear. For $\beta > 0$ and fixed $u = P(X \ge \beta)$, clearly (2.2) is minimized by taking $E(X|X < \beta) = 0$; that is, $P(0 < X < \beta) = 0$. But now (2.2) is minimized if and only if u is minimized; that is, we must take $u = p$. But then $P\{X_i' = \beta\} = u - p = 0$, and the proposition follows. □

Henceforward we shall take X_i to be Bernoulli. Fix c and let $r = \sup\{i: 1 - ic > -c\}$; that is,

$$r = 1 + [1/c], \tag{2.4}$$

where $[x]$ is as before. Note that $Y_i \le 0$ for all $i > r$. Let $D_n = E(\max_{1 \le i \le n} Y_i) - EY_{s_n}$.

PROPOSITION 2.4. *For all* $n \ge r$, *and* c *and* p *fixed*,

$$D_n \le D_r = (1 - p)^r (r - 1)c. \tag{2.5}$$

PROOF. For $n \ge r$,

$$\begin{aligned} E\Big(\max_{1 \le i \le n} Y_i\Big) = E\Big\{\sup_{i \ge 1} Y_i\Big\} &= \{1 - (1 - p)^r\} \\ &\quad - c \sum_{i=1}^{r} ip(1 - p)^{i-1} - c(1 - p)^r \\ &= \beta + (1 - p)^r \{c(r - 1) - \beta\}, \end{aligned} \tag{2.6}$$

where the right-hand side is obtained using

$$\sum_{i=1}^{r} ip(1-p)^{i-1} = \{1-(1-p)^{r+1}\}/p - (r+1)(1-p)^r$$

and $\beta = 1 - c/p$ [by (2.3)]. Using (2.6) and (2.2) and the fact that Y_{s_n} is nondecreasing in n yields (2.5). □

PROPOSITION 2.5. *For c fixed, and all p and n,*

$$D_n \le D_r = [1/c]c(1-c)^{[1/c]+1}. \tag{2.7}$$

PROOF. For c fixed, r is fixed. For $n \ge r$ it follows that (2.5) is maximized when p is minimized, which, by (2.3) means that one must take $\beta = 0$ and $p = c$. But then the value of the right-hand side of (2.5) equals the right-hand side in (2.7).

For $n < r$, one has, similarly to (2.6),

$$E\Big(\max_{1\le i\le n} Y_i\Big) = \beta + (1-p)^n\{c(n-1) - \beta\} \tag{2.8}$$

and hence $D_n = (1-p)^n c(n-1)$. This is 0 when $n = 1$, and for $n > 1$ one has $D_{n+1}/D_n > 1$ iff $p < 1/n$; that is, iff $1/p > n$. But $1/p = 1/c > [1/c] = r - 1 \ge n$, so $D_n < D_r$ for all $n < r$; that is, (2.7) holds. □

Note that (2.7) yields (1.4) once we show that the case $\beta < 0$ is of no concern. Also note that equality holds in (1.4) when $n = [1/c] + 1$ and $X_i = 1$ and 0 with probabilities c and $1 - c$.

The same argument can be used to show equality in (1.5). For $p = c$ we can rewrite (2.5) and the previous argument as

$$D_n \le D_r = (1-p)^r p(r-1) \quad \text{for all } n. \tag{2.9}$$

For r fixed, we can now choose p. By (2.4), $r - 1 < 1/p \le r$; that is, $1/r \le p < 1/(r-1)$. But $(1-p)^r p$ is increasing for $0 \le p \le 1(r+1)$ and decreasing for $1/(r+1) \le p \le 1$. In the permissible interval of p values, (2.9) is therefore maximal for $p = 1/r$, for which the right-hand side of (2.9) becomes $D_r = (1 - 1/r)^{r+1}$, which is the right-hand side of (1.5) for $n = r$. Since $(1 - 1/n)^{n+1}$ increases to e^{-1} and since $E\{\sup_{i\ge 1} Y_i\} = E\{\max_{1\le i\le n} Y_i\}$ for n sufficiently large, while by (2.2), $V(Y_1, \dots, Y_n) \le \beta = V(Y_1, Y_2, \dots)$, (1.6) follows.

To complete the proof of the theorem it remains to consider the case $\beta < 0$. Again we need consider only Bernoulli variables, but now, by (2.1), we have $p - c = \beta$ and $s \equiv 1$. Thus $EY_{s_n} = \beta$ for all n. Thus for all $n \le r$,

$$E\Big\{\max_{1\le i\le n} Y_i\Big\} = \beta/p + (1-p)^n\{c(n-1) - \beta/p\}$$

and for $n \geq r$ we have

$$E\Big\{\sup_{i \geq 1} Y_i\Big\} = E\Big\{\max_{1 \leq i \leq n} Y_i\Big\} = E\Big\{\max_{1 \leq i \leq r} Y_i\Big\}.$$

Thus for $n \leq r$,

$$D_n = \beta(1-p)/p + (1-p)^n\{(p-\beta)(n-1) - \beta/p\} \tag{2.10}$$

and $D_n = D_r$ for $r \leq n$; $D_n = 0$ for $n = 1$. It thus suffices to consider $n \leq r$. For fixed p one can write $D_n = \beta A_n + B_n$, where $A_n > 0$ and B_n do not involve β. Hence D_n is increasing in β; that is, for $\beta \leq 0$, D_n is maximal when $\beta = 0$. But the case $\beta = 0$ was covered by our previous arguments. [Note, however, that for fixed p, changing β will change r through the relationships $p = \beta + c$ and (2.4). Since c decreases with increasing β, $r = r(c)$ can only increase. Thus if n satisfied $n \leq r$ for the original c, (2.10) will remain unchanged by decreasing c.] The proof of Theorem 1 is thus complete. □

Our proof shows that for $X = 1$ and 0 with probability $1/n$ and $1 - 1/n$, respectively, and $c = 1/n$, one has equality in (1.5). It is of interest to note that Jones' (1990) example, which achieves equality in (1.2) (see Example 4.1), is $X_1 = 0$ and $X_2 = \cdots = X_n =$ the previous X and $c = 1/n$. Hence the difference between the i.i.d. case and the more general case is really very small!

To see that the ratio $E\{\max_{1 \leq i \leq n} Y_i\}/V(Y_1, \dots, Y_n)$ is unbounded for all $n > 1$, one can look at (2.8) over β for $\beta > 0$ and fixed c, and let $\beta \to 0$ ($p \to c$). The ratio then tends to infinity.

Remarks. (a) It follows from Hill and Kertz (1983) [see (5) of Kertz (1986)] that the value in the right-hand side of (1.2) [i.e., $(1 - 1/n)^n$] is also the best bound of the difference $E(\max_{1 \leq i \leq n} X_i) - V(X_1, \dots, X_n)$, where $0 \leq X_i \leq 1$ and X_i can have *any kind of dependence*. It is curious that this bound should be the same as for independent random variables, with an added cost $c \geq 0$ per observation.

(b) For i.i.d. X_i, $0 \leq X_i \leq 1$ and $c = 0$, Hill and Kertz (1983), and also Kertz (1986), show that $E(\max_{1 \leq i \leq n} X_i) - V(X_1, \dots, X_n) \leq b_n$ where the constants b_n are best bounds and can be computed recursively. By the remark at the end of the introduction, clearly $b_n \leq 1/4$ for all n.

(c) Since $\{\max_{1 \leq i \leq n} Y_i\} = \{\max_{1 \leq i \leq n} Y_i^*\}$ and since generally $V(Y_1^*, \dots, Y_n^*) \geq V(Y_1, \dots, Y_n)$ it follows that all parts of Theorem 1 hold also with Y_i replaced by Y_i^*. That the bounds are still the best possible follows by our proof. A similar statement is correct also for Jones' theorem as can be seen by his Example 4.1.

Acknowledgment. I wish to express my appreciation and gratitude to Professor Ted Hill for making possible a short visit to Georgia Technical Institute and for his hospitality. During this visit I became acquainted with Jones' work.

REFERENCES

CHOW, Y. S. and ROBBINS, H. (1961). A martingale system theorem and applications. *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* **1** 93–104. Univ. California Press, Berkeley.

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

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

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

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

HILL, T. and KERTZ, R. (1983). Stop rule inequalities for uniformly bounded sequences of random variables. *Trans. Amer. Math. Soc.* **278** 197–207.

JONES, M. L. (1990). Prophet inequalities for cost of observation stopping problems. *J. Multivariate Anal.* To appear.

KERTZ, R. (1986). Stop rule and supremum expectations of i.i.d. random variables: A complete comparison by conjugate duality. *J. Multivariate Anal.* **19** 88–112.

KRENGEL, U. and SUCHESTON, L. (1978). On semiamarts, amarts, and processes with finite value. *Adv. in Probab.* **4** 197–266. Dekker, New York.

MACQUEEN, J. B. and MILLER, R. G., JR. (1960). Optimal persistence policies. *Oper. Res.* **8** 362–380.

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

DEPARTMENT OF STATISTICS
HEBREW UNIVERSITY
JERUSALEM 91905
ISRAEL