האוניברסיטה העברית בירושלים

THE HEBREW UNIVERSITY OF JERUSALEM

WHY IS ONE CHOICE DIFFERENT?

by

DAVID ASSAF and ESTER SAMUEL-CAHN

Discussion Paper # 335 September 2003

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

Why is One Choice Different?

by

David Assaf and Ester Samuel-Cahn
The Hebrew University of Jerusalem

Dedicated to Herman Chernoff on the Occasion of his Eightieth Birthday

AMS 2000 subject classification. Primary 60G40; secondary 60E15.

Supported by the Israel Science Foundation Grant number 879/01

Key words and phrases: Optimal stopping rules, ratio prophet inequalities, multiple stopping options.

Abstract

Let X_i be nonnegative independent random variables with finite expectations and $X_n^* = \max\{X_1, \ldots, X_n\}$. The value EX_n^* is what can be obtained by a "prophet". A "mortal" on the other hand, may use $k \geq 1$ stopping rules $t_1, \ldots, t_k$ yielding a return $E[\max_{i=1,\ldots,k} X_{t_i}]$. For $n \geq k$ the optimal return is $V_k^n(X_1, \ldots, X_n) = supE[\max_{i=1,\ldots,k} X_{t_i}]$ where the supremum is over all stopping rules which stop by time n. The well known "prophet inequality" states that for all such X_i's and one choice $EX_n^* < 2V_1^n(X_1, \ldots, X_n)$ and the constant "2" cannot be improved on for any $n \geq 2$. In contrast we show that for $k = 2$ the best constant d satisfying $EX_n^* < dV_2^n(X_1, \ldots, X_n)$ for all such X_i's depends on n. On the way we obtain constants c_k such that $EX_{k+1}^* < c_k V_k^{k+1}(X_1, \ldots, X_{k+1})$.

1 Introduction and summary

The classical "ratio prophet inequality" states that for nonnegative independent random variables, not all identically zero, with known distributions and finite expectations, the inequality

$$EX_n^* < 2V_1^n(X_1, \ldots, X_n) \tag{1}$$

holds for all $n \geq 2$, where $X_n^* = \max\{X_1, \ldots, X_n\} = X_1 \vee \cdots \vee X_n$, and EX_n^* is the return to a prophet who can foresee the entire future, while $V_1^n(X_1, \ldots, X_n) = \sup_{t \in T_n} EX_t$ is the optimal return to a mortal who employs an optimal stopping rule. Here T_n denotes the collection of all stopping rules for $X_1, \ldots, X_n$, which stop no later than by time n. Inequality (1) extends nonstrictly also to infinite sequences of random variables, and stopping rules which satisfy $P(t < \infty) = 1$, provided $E(\sup X_i) < \infty$. See, for example, Hill and Kertz (1981), and some earlier references mentioned there, as well as Samuel-Cahn (1984).

The constant "2" in (1) is a "best bound", i.e. cannot be replaced by any smaller constant, as the following well known example shows.

Example 1. Let $n = 2$, $X_1 = \alpha$ $X_2 = 1$ and 0 with probabilities α and $1 - \alpha$, respectively, where $0 < \alpha < 1$. Then $V_1^2(X_1, X_2) = \alpha$ and $EX_2^* = 2\alpha - \alpha^2$. Thus $EX_2^*/V_1^2(X_1, X_2) = 2 - \alpha \to 2$ as $\alpha \to 0$.

The above example shows that "2" is a best bound for any n, since clearly a "best bound" can only increase with n, as one can always take additional

X's to be identically zero, to attain a bound obtained for a smaller value of n.

In a recent paper, Assaf, Goldstein and Samuel-Cahn (2002), (henceforth AGS), a situation where the mortal has several choices is considered. Let k be the number of choices, and $n \geq k$. When the mortal uses the k stopping rules $1 \leq t_1 \leq \quad \leq t_k \leq n$ his expected return is $E[X_{t_1} \vee \cdots \vee X_{t_k}]$, i.e. the expected value of the *maximal* of the k values chosen. Here clearly later choices may/will depend on the values chosen earlier. Let $n \geq k$ and

$$V_k^n(X_1, \ldots, X_n) = \sup_{1 \leq t_1 \leq \cdots \leq t_k \leq n} E[X_{t_1} \vee \cdots \vee X_{t_k}]$$

denote the optimal k-choice value, $k = 1, 2, \ldots$. In AGS the inequality (1) is generalized and prophet inequalities are obtained for this situation, under the same assumptions on the X_i's, as mentioned above. In particular they show: There exist constants g_k such that for any $n \geq k$ and any non-negative $X_1, \ldots, X_n$ with finite expectations, not all identically zero, the inequalities

$$EX_n^* < g_k V_k^n(X_1, \ldots, X_n) \tag{2}$$

hold. The values of g_k are given explicitly for $k = 1, \ldots 6$. In particular $g_1 = 2,\ g_2 = 1 + e^{-1} = 1.3678\ldots$ and $g_3 = 1 + e^{1-e} = 1.1793\ldots$

The purpose of the present note is to show that, unlike the situation for $k = 1$, the best bounds for more than one choice is n-dependent. In particular we show, for $k = 2$ that

$$EX_3^* < 1.25 V_2^3(X_1, X_2, X_3) \tag{3}$$

and 1.25 is a best bound for $k = 2, n = 3$.

We give an example with $k = 2, n = 4$, to show that the bound for this case is larger, and hence the bound depends on n. More generally, for $n = k + 1$ we obtain values c_k such that $EX_{k+1}^* < c_k V_k^{k+1}(X_1, \ldots, X_{k+1})$, where $c_k < g_k$. (Note, however, that no claim about best bound was made in AGS, regarding g_k holding for all n, except for $k = 1$. The question whether the g_k's are best bounds holding for all n is thus still open.)

The fact that best bounds may be n-dependent in some cases is not new. As an example, in the class of i.i.d. non-negative X_i's, and a single choice, the n-dependence is shown in Hill and Kertz (1982). It is new, however, in the present context of general independent non-negative X_i's for the case of $k > 1$ choices. (This is in contrast to the one choice case discussed above).

2 An inequality, and examples

The results of AGS are actually more general than inequality (2). Theorem 1.3 there states that for any non-negative X_i's with finite expectations satisfying $P(X_n^* = 0) = x$, where $0 \le x < 1$, the ratio prophet inequalities

$$EX_n^* < g_k(x) V_k^n(X_1, \dots, X_n) \tag{4}$$

hold, for $k = 1, 2, \dots$ and $n \ge k$. The functions $g_k(x)$ are defined inductively and are monotone decreasing. The first three functions are

$$\begin{aligned} g_1(x) &= 2 - x \\ g_2(x) &= e^{-(1-x)} + 1 - x \\ g_3(x) &= \exp\{1 - e^{1-x}\} + 1 - x \end{aligned} \tag{5}$$

For $k = 1$ inequality (4) yields a best possible bound for all values of x.

Let $R_k^n(X_1, \dots, X_n) = EX_n^* / V_k^n(X_1, \dots, X_n)$, and note that $\sup R_k^n(X_1, \dots, X_n)$ over all $X_1, \dots, X_n$ is the best bound for k choices and n observations.

With $g_k(x)$ as in (5), for $k = 2, 3, \dots$ let

$$c_k = 1 + \sup_{0 \le p < 1} p[(g_{k-1}(p) - 1)/(g_{k-1}(p) - 1 + p)]. \tag{6}$$

Our main result is the following

Theorem. *For any $k = 2, 3, \dots$ and any independent non-negative X_i's with finite expectations, not identically zero*

$$R_k^{k+1}(X_1, \dots, X_{k+1}) < c_k \tag{7}$$

and for $k = 2$ the value of $c_2 = 5/4$ is a best bound.

The values c_3 and c_4 can be obtained numerically, and are $c_3 = 1.1189\dots$ attained for $p = .2852\dots$ and $c_4 = 1.0646\dots$ attained for $p = .1709\dots$. These values should be compared with the values g_k of (2), in particular $g_2 = 1.3678\dots$, $g_3 = 1.1793\dots$ and $g_4 = 1.0979\dots$ respectively.

We restate some definitions and a lemma from AGS, needed in the proof. We first make the "nontriviality assumption" for $n > k$ regarding $X_2, \dots, X_n$ and $k \ge 2$ which assumes that the value $V_k^{n-1}(X_2, \dots, X_n)$ cannot be attained with less than k choices.

Definition. Let $X_2, \ldots, X_n$ be given and $1 \leq k < n$. The value $b_k = b_k(X_2, \ldots, X_n)$ is called the *indifference value* for the k choice problem if, when $X_1 \equiv b_k$, one is indifferent between (i) picking b_k as a first choice, and being left with $k-1$ choices among $X_2, \ldots, X_n$, and (ii) not picking b_k and having k choices among $X_2, \ldots, X_n$. Here, for $k = 1$, the value of a no-choice option is 0. Clearly for general X_1 an optimal policy will pick X_1 if $X_1 > b_k$, be indifferent between picking it or not, when $X_1 = b_k$, and not pick X_1 when $X_1 < b_k$.

It is shown in AGS that, under the nontriviality assumption, b_k is uniquely defined and positive.

We restate Lemma 2.4 of AGS with a slight change of notation.

Lemma 1. *For any independent non-negative $Y_1, \ldots, Y_n$ with finite expectations such that $P(Y_n^* = 0) = x$, $0 \leq x < 1$, there exist independent non-negative $X_1, \ldots, X_n$ having finite expectations with $b_k = b_k(X_2, \ldots, X_n)$ such that*

(i) $P(X_n^* = 0) = x$

(ii) $X_i = X_i I(X_i > b_k)$ *for* $i = 2, \ldots, n$

(iii) X_1 *takes values* 0 *and* b_k *only*

(iv) $R_k^n(Y_1, \ldots, Y_n) \leq R_k^n(X_1, \ldots, X_n)$.

In what follows we therefore may, and shall, assume that the X_i's are as in Lemma 1. Let $X_{[2,n]}^* = \max\{X_2, \ldots, X_n\}$. Note that $p = P(X_{[2,n]}^* = 0) > 0$ since if just for some $i \geq 2$ one would have $P(X_i = 0) = 0$, (ii) would imply $P(X_i > b_k) = 1$ contradicting the fact that b_k is the indifference value.

Now

$$V_k^n(b_k, X_2, \ldots, X_n) = V_k^{n-1}(X_2, \ldots, X_n) = b_k + V_{k-1}^{n-1}([X_2 - b_k]^+, \ldots, [X_n - b_k]^+) \tag{8}$$

where the first equality follows from the definition, and the rightmost equality follows since if b_k is picked as a first choice, the optimal continuation is to maximize the residual value, i.e. the value for the sequence $[X_2 - b_k]^+, \ldots, [X_n - b_k]^+$, with the remaining $k-1$ choices, since b_k is already guaranteed.

Lemma 2. *With X_i's as in Lemma 1, $i = 1, \ldots, n$ and $n > k \geq 2$,*

$$b_k < [(g_{k-1}(p) - 1) \,/\, (g_{k-1}(p) - 1 + p)] EX^*_{[2,n]} \tag{9}$$

where $p = P(X^*_{[2,n]} = 0)$.

Proof. From (8) and (4) it follows that

$$\begin{aligned} EX^*_{[2,n]} &\geq V_k^{n-1}(X_2, \ldots, X_n) = b_k + V_{k-1}^{n-1}\left([X_2 - b_k]^+, \ldots, [X_n - b_k]^+\right) \\ &> b_k + \frac{1}{g_{k-1}(p)}\left[E([X_2 - b_k]^+ \vee \cdots \vee [X_n - b_k])^+)\right] \\ &= b_k + \frac{1}{g_{k-1}(p)}\left[EX^*_{[2,n]} - (1-p)b_k\right] \end{aligned}$$

Now (9) follows by rearranging the relevant terms. □

Proof of Theorem.

Note that with X_i's as in Lemma 1 for $i = 2, \ldots, n$ the ratio $R_k^n(X_1, \ldots, X_n)$ will be maximal for $X_1 \equiv b_k$, since $V_k^n(X_1, \ldots, X_n)$ remains unchanged as long as X_1 satisfies (iii) of Lemma 1, and EX^*_n increases when $P(X_1 = b_k) = 1$.

Then, by (9)

$$\begin{aligned} EX^*_n &= pb_k + EX^*_{[2,n]} \\ &< [1 + p\,(g_{k-1}(p) - 1) \,/\, (g_{k-1}(p) - 1 + p)]\, EX^*_{[2,n]} \end{aligned} \tag{10}$$

Also, for $n = k+1$ we have, by (8), that $EX^*_{[2,k+1]} = V_k^{k+1}(X_1, \ldots, X_{k+1})$, and (7) follows.

Since $g_1(p) = 2 - p$ the square bracket in (10) is $[1 + p(1-p)] \leq 5/4$, with equality for $p = 1/2$, and hence $c_2 = 5/4$.

To see that 5/4 is a best bound, consider the following

Example 2.

Let $0 < \alpha < 1/2$

$$X_1 \equiv \alpha \qquad X_2 = \begin{cases} 2\alpha & \text{with prob. } \frac{1}{2} \\ 0 & \text{with prob. } \frac{1}{2} \end{cases} \qquad X_3 = \begin{cases} 1 & \text{with prob. } \alpha \\ 0 & \text{with prob. } 1 - \alpha \end{cases}$$

It is easily verified that here

$$V_2^3(X_1, X_2, X_3) = \alpha(2 - \alpha/2) \text{ and } EX_3^* = \alpha(5 - 3\alpha)/2.$$

Thus $R_2^3(X_1, X_2, X_3) = (5 - 3\alpha)/(4 - \alpha) \to 5/4$ as $\alpha \to 0$.

□

To show that the bound for $k = 2$ is n-dependent, it suffices to show an example of $X_1, \ldots, X_4$ for which $R_2^4(X_1, \ldots, X_4) > 5/4$. The following is such an example.

Example 3. Let

$$X_1 = .00112352 \qquad X_2 = \begin{cases} .00229297 & \text{with prob. } .449 \\ 0 & \text{with prob. } .551 \end{cases}$$

$$X_3 = \begin{cases} .00329067 & \text{with prob. } .146 \\ 0 & \text{with prob. } .854 \end{cases} \qquad X_4 = \begin{cases} 1 & \text{with prob. } .001 \\ 0 & \text{with prob. } .999 \end{cases}$$

The prophet value here is $EX_4^* = .002886456$ while the value to the statistician is .00229297, yielding the ratio 1.2588.

Remark. For $k \geq 3$ we do not believe that the values c_k of the theorem are best bounds for $R_k^{k+1}(X_1, \ldots, X_{k+1})$. For example, for $k = 3$ we believe that the best bound for $R_3^4(X_1, \ldots, X_4)$ is $1 + (5\sqrt{5} - 11)/2 = 1.0901\ldots$, while $c_3 = 1.1189$.

References

[1] Assaf, D., Goldstein, L. and Samuel-Cahn, E. (2002). Ratio prophet inequalities when the mortal has several choices. *Ann. Applied Probability*, **12**, 972–984.

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

[3] Hill, T.P. and Kertz, R.P. (1982). Comparisons of stop rule and supremum expectations of i.i.d. random variables. *Ann. Probability*, **10**, 336–345.

[4] Samuel-Cahn, E. (1984). Comparison of threshold stop rules and maximum for independent nonnegative random variables. *Ann. Probability*, **12**, 1213–1216.