Ratio Prophet Inequalities when the Mortal has Several Choices*†

David Assaf, Larry Goldstein and Ester Samuel-Cahn
Hebrew University, University of Southern California and Hebrew University

Abbreviated Title: Prophet Inequalities

April 30, 2001

Abstract

Let X_i be non-negative, independent random variables with finite expectation, 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 of $E[\max_{i=1,\ldots,k} X_{t_i}]$. For $n \geq k$ the optimal return is $V_k^n(X_1, \ldots, X_n) = \sup E[\max_{i=1,\ldots,k} X_{t_i}]$ where the supremum is over all stopping rules $t_1, \ldots, t_k$ such that $P(t_i \leq n) = 1$. We show that for a sequence of constants g_k which can be evaluated recursively, the inequality $EX_n^* < g_k V_k^n(X_1, \ldots, X_n)$ holds for all such $X_1, \ldots, X_n$ and all $n \geq k$; $g_1 = 2$, $g_2 = 1+e^{-1} = 1.3678\ldots$, $g_3 = 1+e^{1-e} = 1.1793\ldots$, $g_4 = 1.0979\ldots$ and $g_5 = 1.0567\ldots$. Similar results hold for infinite sequences $X_1, X_2, \ldots$. Since with five choices the mortal is thus guaranteed over 94% of the prophet's value, more than five choices may not be practical.

1 Introduction and Summary

The classical ratio "prophet inequality" states that for nonnegative independent random variables with known distribution and finite expectation, $X_1, \ldots, X_n, n \geq 2$, the inequality

$$E(X_n^*) < 2V(X_1, \ldots, X_n) \tag{1}$$

holds, where $X_n^* = \max(X_1, \ldots, X_n) = X_1 \vee \ldots \vee X_n$, $V(X_1, \ldots, X_n) = \max_{t \in T_n} E(X_t)$, and T_n is the collection of all stopping rules based on $X_1, \ldots, X_n$. (A stopping rule t is in T_n if the event $\{t = k\}$ depends only on $X_1, \ldots, X_k$ and possibly some external randomization, and $P(t \leq n) = 1$). Inequality (1) extends non-strictly to infinite sequences of random variables, with maximum replaced by supremum, provided $E(\sup X_i) < \infty$, where the rules are required to satisfy $P(t < \infty) = 1$. Inequality (1) cannot hold with a smaller constant

*AMS 2000 subject classifications. Primary 60G40; secondary 60E15.

†Key words and phrases: optimal stopping rules, ratio prophet inequality, best choice multiple stopping options.

replacing 2, and thus 2 is known as a "best bound". See e.g. Hill and Kertz (1981), and some earlier references mentioned there. The term "prophet inequality" stems from the fact that EX_n^* may be considered the return to a "prophet" who has complete foresight and can thus choose the best (largest) observation, while $V(X_1, \ldots, X_n)$ is the value obtained by a "mortal" (henceforth called "statistician"), who must decide whether to stop or not as the sequence unfolds, with no possibility of recalling any passed up observations.

In the present paper we are considering a situation where the statistician is given k, $k \leq n$, opportunities to choose variables by means of k stopping rules. The return is defined as the expected value of the *largest* of the k choices. As an example, the case $k = 2$ may correspond to a situation in which you put your first selected item (perhaps a house or a job offer) "on hold" as a guaranteed fallback value. You then proceed sequentially to select a second item (which should be of greater value than the first unless it is the last one) and finish by taking the better of the two items selected.

Multiple stopping rules, in a general setting, are studied by Stadje (1985). In connection with prophet inequalities they are studied by Kennedy (1987). Kennedy considers the case where the statistician receives the expected value of the sum of his k choices. When the payoff is the expected value of his maximal choice, as described above, the problem is studied in Assaf and Samuel-Cahn (2000). They show that there exist simple k-choice rules for the statistician, called "threshold rules", with values $W_k^n(X_1, \ldots, X_n)$, such that for any independent $X_i \geq 0$ the inequality

$$E(X_n^*) < \left(\frac{k+1}{k}\right) W_k^n(X_1, \ldots, X_n) \tag{2}$$

holds. Since threshold rules are usually not optimal,

$$W_k^n(X_1, \ldots, X_n) \leq V_k^n(X_1, \ldots, X_n),$$

where $V_k^n(X_1, \ldots, X_n)$ is the optimal k-choice value. Hence, by (2),

$$E(X_n^*) < \left(\frac{k+1}{k}\right) V_k^n(X_1, \ldots, X_n). \tag{3}$$

It turns out that, except when $k = 1$, the constant $(k+1)/k$ is not the best constant in this inequality. In the present paper we prove Theorem 1.1, which provides a sequence of improved constants.

We assume henceforth that all random variables in the stopping sequences we consider have known distributions and are independent, non-negative with finite expectation, and not identically zero.

Theorem 1.1 *For $k = 1, 2, \ldots$, let $g_k = g_k(0)$ where the functions $g_k(x)$ are defined recursively by (8). Then for all $n \geq k$ and any $X_1, \ldots, X_n$,*

$$E(X_n^*) < g_k V_k^n(X_1, \ldots, X_n). \tag{4}$$

The first six values of the g_k sequence are $g_1 = 2$, $g_2 = 1 + e^{-1} = 1.3678\ldots$, $g_3 = 1 + e^{1-e} = 1.1793\ldots$, $g_4 = 1.0979\ldots$, $g_5 = 1.0567\ldots$, $g_6 = 1.0341\ldots$.

For $X_1, X_2, \ldots$, an infinite sequence of such variables with value $V_k^\infty(X_1, X_2 \ldots)$, the inequality

$$E(\sup{}_{i=1,2,\ldots} X_i) \leq g_k V_k^\infty(X_1, X_2, \ldots) \tag{5}$$

holds provided the left hand side of (5) is finite.

That Theorem 1.1 gives considerable improvement over (3) is supported by numerical results and Assertion 3.1, which proves that for $k \geq 2$, $g_k(0) < (k+1)/k$. However, except for $k = 1$, no claim about having a best bound is made here. We prove Theorem 1.1 by induction on n for each fixed k, and by solving a differential equation, as explained in Section 3. In principle, once the result (4) for some k is known, it is a simple matter to obtain (at least numerically) the result (4) for $k+1$. For practical situations, it seems that no more than five choices would be of much real interest, since with five choices in the worst case scenario, the statistician is already guaranteed over 94% of the value of the prophet, for any n.

In our proofs we need the following generalization of (1), which is also of interest in its own right.

Theorem 1.2 *For $n \geq 2$ and $x = P(X_n^* = 0) < 1$,*

$$EX_n^* < (2-x)V_1^n(X_1, \ldots, X_n). \tag{6}$$

In the infinite case, with $x = P(\sup_{i=1,2,\ldots} X_i = 0)$,

$$E(\sup{}_{i=1,2,\ldots} X_i) \leq (2-x)V_1^\infty(X_1, X_2, \ldots). \tag{7}$$

The expression $2-x$ is a best bound. (For $x = 1$, (6) holds with equality.)

Similar to the generalization of (1) to (6) we have a generalization of Theorem 1.1 to 1.3; this requires the following definition. For $0 \leq y < 1$, let

$$\begin{aligned} u_1(y) &= 0, \quad \text{and define for } k \geq 1, \\ u_{k+1}(y) &= -\int_y^1 e^{-u_k(u)} du, \quad h_k(y) = e^{u_k(y)}, \quad \text{and} \quad g_k(y) = h_k(y) + 1 - y. \end{aligned} \tag{8}$$

Theorem 1.3 *The functions g_k are strictly decreasing. If $n \geq k$ and $x = P(X_n^* = 0) < 1$, then*

$$EX_n^* < g_k(x)V_k^n(X_1, \ldots, X_n). \tag{9}$$

In particular, for $0 \leq y < 1$ we have

$$g_1(y) = 2 - y, \tag{10}$$
$$g_2(y) = e^{-(1-y)} + 1 - y, \tag{11}$$
$$g_3(y) = \exp\{1 - e^{1-y}\} + 1 - y, \quad \textit{and} \tag{12}$$
$$g_4(y) = \exp\{e^{-1}[Ei(1) - Ei(e^{1-y})]\} + 1 - y \tag{13}$$

where

$$Ei(y) = \oint_{-\infty}^{y} \frac{e^z}{z} dz, \quad y > 0. \tag{14}$$

Similar statements to (9) hold non-strictly for the infinite case by taking limits.

Since the functions $g_k(y)$ are decreasing, Theorem 1.1 follows from Theorem 1.3. The reason the functions $g_k(y)$ are given explicitly only for $k = 1, 2, 3, 4$ is that further functions can be obtained only through numerical evaluation.

The paper is organized as follows. In Section 2 we introduce some preliminary notions, and prove Lemmas which will simplify our later derivations. In Section 3 we prove Theorem 1.2, yielding the $k = 1$ case of Theorem 1.3, identifying $g_1(y) = 2 - y$, as well as Theorem 1.3, using a basic inequality relating the values EX_n^* and V_k^n through $g_{k+1}(y)$, obtained as the solution of a differential equation based on $g_k(y)$.

2 Preliminaries

In the following, we make the non-triviality

Assumption 2.1 *The value $V_k^{n-1}(X_2, \dots, X_n)$ cannot be attained with less than k choices. That is,*

$$V_k^{n-1}(X_2, \dots, X_n) > V_{k-1}^{n-1}(X_2, \dots, X_n).$$

We shall also need the following

Definition 2.1 *Let $X_2, \dots, X_n$ be given, and $k < n$. The value $b_k = b_k(X_2, \dots, X_n)$ is called the indifference value for the k-choice problem if one is indifferent between (i) picking b_k as a first choice and being left with $k-1$ choices among $X_2, \dots, X_n$, and (ii) not choosing b_k and having k choices among $X_2, \dots, X_n$. Thus,*

$$V_k^n(b_k, X_2, \dots, X_n) = V_k^{n-1}(X_2, \dots, X_n) = V_{k-1}^{n-1}(X_2, \dots, X_n \vee b_k). \tag{15}$$

The requirement that $k < n$ in the definition of an indifference value is needed, since for $k \geq n$ the trivial relation $V_k^n(X_1, \dots, X_n) = EX_n^*$ holds.

Assumption 2.1 has the following important consequence.

Proposition 2.1 *The function*

$$\phi(z) = V_{k-1}^{n-1}(X_2, \dots, X_n \vee z) \tag{16}$$

is strictly increasing in z for $z \in [c, \infty)$ for any $c \geq 0$ such that

$$P(\max\{X_2, \dots, X_n\} \leq c) > 0. \tag{17}$$

In particular, under Assumption 2.1, $\phi(z)$ is strictly increasing in z for $z \in [b_k, \infty)$, and the indifference value b_k is unique and positive.

Proof: Let $z \geq c$. By (17), $P(\max\{X_2, \ldots, X_n\} \leq z) > 0$, and there is positive probability that the best $k-1$ choice rule for $(X_2, \ldots, X_n \vee z)$ will choose z. With $z < y$, let $\tilde{V}_{k-1}^{n-1}(X_2, \ldots, X_n \vee y)$ be the value of applying the optimal $k-1$ choice rule for $(X_2, \ldots, X_n \vee z)$ applied to $(X_2, \ldots, X_n \vee y)$. Hence,

$$\phi(y) = V_{k-1}^{n-1}(X_2, \ldots, X_n \vee y) \geq \tilde{V}_{k-1}^{n-1}(X_2, \ldots, X_n \vee y) > V_{k-1}^{n-1}(X_2, \ldots, X_n \vee z) = \phi(z).$$

Furthermore, $P(\max\{X_2, \ldots, X_n\} \leq b_k) > 0$. If not, then for some $j \geq 2$ we must have $P(X_j > b_k) = 1$. But in that case one would use one of the k choices to pick X_j rather than to pick $X_1 = b_k$, contradicting the definition of b_k as an indifference value. Hence, b_k is unique, as if b and b^* are both indifference values, with say $b^* < b$, from (15) and (16) it would follow that $\phi(b) = \phi(b^*)$, contradicting the strict monotonicity of ϕ in $[b^*, \infty)$.

To see that b_k is positive, note that $b_k = 0$ would, by use of (15), contradict Assumption 2.1. □

The interpretation of $b_k(X_2, \ldots, X_n)$ in relation to the optimal k-choice rule for $X_1, \ldots, X_n$ is as follows. When an $X_1 > b_k(X_2, \ldots, X_n)$ is observed, the optimal action is to pick X_1 as a first choice. When $X_1 = b_k(X_2, \ldots, X_n)$ one is indifferent between picking X_1 or not, and if $X_1 < b_k(X_2, \ldots, X_n)$ then X_1 should not be picked.

We introduce the following notation. Let

$$D_k^n(X_1, \ldots, X_n) = EX_n^* - V_k^n(X_1, \ldots, X_n) \quad \text{and} \tag{18}$$

$$R_k^n(X_1, \ldots, X_n) = \frac{EX_n^*}{V_k^n(X_1, \ldots, X_n)}. \tag{19}$$

In the following series of lemmas our aim is to replace the given sequence of random variables $X_1, \ldots, X_n$ by another sequence $\hat{X}_1, \ldots, \hat{X}_n$, say, so that

$$R_k^n(X_1, \ldots, X_n) \leq R_k^n(\hat{X}_1, \ldots, \hat{X}_n). \tag{20}$$

Since

$$R_k^n(X_1, \ldots, X_n) = \frac{D_k^n(X_1, \ldots, X_n)}{V_k^n(X_1, \ldots, X_n)} + 1, \tag{21}$$

to prove (20) it suffices that

$$D_k^n(X_1, \ldots, X_n) \leq D_k^n(\hat{X}_1, \ldots, \hat{X}_n) \quad \text{and} \quad V_k^n(X_1, \ldots X_n) \geq V_k^n(\hat{X}_1, \ldots, \hat{X}_n).$$

Thus our lemmas will be stated in terms of the differences D_k^n and values V_k^n, rather than directly in terms of R_k^n.

Lemma 2.1 *For $k < n$ and any $X_1, X_2, \ldots, X_n$ with $b_k = b_k(X_2, \ldots, X_n)$,*

$$D_k^n(X_1, \ldots, X_n) \leq D_k^n(b_k, X_2, \ldots, X_n) \tag{22}$$

and

$$V_k^n(X_1, \ldots, X_n) \geq V_k^n(b_k, X_2, \ldots, X_n). \tag{23}$$

Proof: Let F be the distribution function of X_1. Clearly

$$E[X_1 \vee \ldots \vee X_n] = \int E[x \vee X_2 \vee \ldots \vee X_n] dF(x),$$

and since the value x of X_1 will be known before a decision whether to pick it or not must be made,

$$V_k^n(X_1, \ldots, X_n) = \int V_k^n(x, X_2, \ldots, X_n) dF(x).$$

It follows that $D_k^n(X_1, \ldots, X_n) = \int D_k^n(x, X_2, \ldots, X_n) dF(x)$, and hence it suffices to show (22) and (23) for $X_1 = x$, where x is any constant.

Case 1: $x \leq b_k$. Then

$$V_k^n(x, X_2, \ldots, X_n) = V_k^{n-1}(X_2, \ldots, X_n) = V_k^n(b_k, X_2, \ldots, X_n).$$

Thus (23) holds, and since $E[x \vee X_2 \vee \ldots \vee X_n] \leq E[b_k \vee \ldots \vee X_n]$, (22) holds.

Case 2: $x > b_k$. Here (23) is trivial. Also, for any $t_2, \ldots, t_k \in T_n$ strictly greater than one,

$$\begin{aligned} E[x \vee X_{t_2} \vee \ldots \vee X_{t_k}] &= E[b_k \vee X_{t_2} \vee \ldots \vee X_{t_k}] + E[x - (b_k \vee X_{t_2} \vee \ldots \vee X_{t_k})]^+ \\ &\geq E[b_k \vee X_{t_2} \vee \ldots \vee X_{t_k}] + E[x - (b_k \vee X_2 \vee \ldots \vee X_n)]^+. \end{aligned} \tag{24}$$

Taking supremum over $t_2, \ldots, t_k$ first on the left and then on the right side of (24) yields

$$V_k^n(x, X_2, \ldots, X_n) \geq V_k^n(b_k, X_2, \ldots, X_n) + E[x - (b_k \vee X_2 \vee \ldots \vee X_n)]^+. \tag{25}$$

On the other hand

$$E[x \vee X_2 \vee \ldots \vee X_n] = E[b_k \vee X_2 \vee \ldots \vee X_n] + E[x - (b_k \vee X_2 \vee \ldots \vee X_n)]^+. \tag{26}$$

Clearly (26) and (25) yield (22) for this case. □

Lemma 2.2 *Let $X_1, \ldots, X_n$ be given, $b_k = b_k(X_2, \ldots, X_n)$ and $P(X_1 = 0) = 1 - \alpha$. Let*

$$\tilde{X}_1 = \begin{cases} 0 & 1-\alpha \\ b_k & \alpha. \end{cases}$$

Then

$$D_k^n(X_1, \ldots, X_n) \leq D_k^n(\tilde{X}_1, X_2, \ldots, X_n) \tag{27}$$

and

$$V_k^n(X_1, \ldots, X_n) \geq V_k^n(\tilde{X}_1, X_2, \ldots, X_n). \tag{28}$$

Proof: Let $\hat{X}_1$ have the conditional distribution of X_1, given $X_1 \neq 0$. Since

$$V_k^n(X_1, \ldots, X_n) = (1-\alpha) V_k^{n-1}(X_2, \ldots, X_n) + \alpha V_k^n(\hat{X}_1, X_2, \ldots, X_n),$$

and

$$D_k^n(X_1, \ldots, X_n) = (1-\alpha) D_k^{n-1}(X_2, \ldots, X_n) + \alpha D_k^n(\hat{X}_1, X_2, \ldots, X_n)$$

the result follows immediately from Lemma 2.1. □

Lemma 2.3 *Let $X_2, \ldots, X_n$ be given, $n > k$, and let $b_k = b_k(X_2, \ldots, X_n)$. Let $\hat{X}_i = X_i I(X_i > b_k), i = 2, \ldots, n$, and let $\hat{b}_k = b_k(\hat{X}_2, \ldots, \hat{X}_n)$. Then*

$$b_k \geq \hat{b}_k. \tag{29}$$

Proof: We have that

$$\begin{aligned} V_{k-1}^{n-1}(\hat{X}_2, \ldots, \hat{X}_n \vee b_k) &= V_{k-1}^{n-1}(X_2, \ldots, X_n \vee b_k) = V_k^{n-1}(X_2, \ldots, X_n) \\ &\geq V_k^{n-1}(\hat{X}_2, \ldots, \hat{X}_n) = V_{k-1}^{n-1}(\hat{X}_2, \ldots, \hat{X}_n \vee \hat{b}_k), \end{aligned}$$

where the inequality is a consequence of $X_i \geq \hat{X}_i$ a.s. Inequality (29) now follows by Proposition 2.1 for $c = 0$. □

Remark 2.1. In spite of Lemma 2.3 it is possible that one set of variables is stochastically smaller than the other, but its indifference number is larger, as the following simple example shows. Let $n = 3,\ k = 2$ and $Y_3,\ X_3$ be identically distributed, with $P(X_3 = 1) = 2/3 = 1 - P(X_3 = 0)$, and let

$$X_2 = \begin{cases} 1 & 1/3 \\ 1/2 & 1/3 \\ 0 & 1/3 \end{cases}, \qquad Y_2 = \begin{cases} 1 & 1/2 \\ 1/2 & 1/6 \\ 0 & 1/3. \end{cases}$$

Clearly $X_2 \overset{st}{<} Y_2$, but it is easily checked that $b_2(X_2, X_3) = 1/4 > b_2(Y_2, Y_3) = 1/6$.

That the above Lemmas can be used together is the content of Lemma 2.4.

Lemma 2.4 *For any $X_1, \ldots, X_n$, $n > k$ such that $P(X_n^* = 0) = x$, $0 \leq x < 1$, there exist $\tilde{X}_1, \ldots, \tilde{X}_n$ and $\tilde{b}_k = b_k(\tilde{X}_2, \ldots, \tilde{X}_n)$ such that*

1. *$P(\tilde{X}_n^* = 0) = x$,*
2. *$\tilde{X}_i = \tilde{X}_i I(\tilde{X}_i > \tilde{b}_k)$ for $i = 2, \ldots, n$,*
3. *$\tilde{X}_1$ takes the values $\tilde{b}_k$ and 0 only, and*
4. $$D_k^n(X_1, \ldots, X_n) \leq D_k^n(\tilde{X}_1, \ldots, \tilde{X}_n) \tag{30}$$
   and
   $$V_k^n(X_1, \ldots, X_n) \geq V_k^n(\tilde{X}_1, \ldots, \tilde{X}_n). \tag{31}$$

Proof: Let $b_k = b_k(X_2, \ldots, X_n)$. By Lemma 2.2 we may without loss of generality assume that $X_1 = 0$ and b_k with probabilities $1 - \alpha$ and α respectively. Let $\hat{X}_i = X_i I(X_i > b_k)$, $i = 2, \ldots, n$ and $\hat{X}_1 = 0$ and b_k with probability $1 - \hat{\alpha}$ and $\hat{\alpha}$ respectively, where $\hat{\alpha}$ as given in (37) is determined so that $P(\hat{X}_n^* = 0) = x$. We shall show that

$$D_k^n(X_1, \ldots, X_n) \leq D_k^n(\hat{X}_1, \ldots, \hat{X}_n) \tag{32}$$

and

$$V_k^n(X_1,\ldots,X_n) \geq V_k^n(\hat{X}_1,\ldots,\hat{X}_n). \tag{33}$$

Let $\hat{b}_k = b_k(\hat{X}_2,\ldots,\hat{X}_n)$. Then by Lemma 2.3, $b_k \geq \hat{b}_k$ and thus it follows that $\hat{X}_i = \hat{X}_i I(\hat{X}_i > \hat{b}_k)$, $i = 2,\ldots,n$. Thus if we set $\tilde{X}_i = \hat{X}_i$ for $i = 2,\ldots,n$ then $\tilde{b}_k = \hat{b}_k$, and 2. holds. Now let $\tilde{X}_1 = 0$ and $\tilde{b}_k$ with probability $1-\hat{\alpha}$ and $\hat{\alpha}$ respectively. Thus 1. and 3. are satisfied. Now (30) and (31) will follow from (32) and (33) together with Lemma 2.2.

Inequality (33) follows since by the definition of b_k and (15)

$$\begin{aligned} & V_{k-1}^{n-1}(\hat{X}_2,\ldots,\hat{X}_n \vee b_k) = V_{k-1}^{n-1}(X_2,\ldots,X_n \vee b_k) \\ & = V_k^{n-1}(X_2,\ldots,X_n) = V_k^n(X_1,\ldots,X_n), \end{aligned} \tag{34}$$

whereas clearly $V_k^{n-1}(\hat{X}_2,\ldots,\hat{X}_n) \leq V_k^{n-1}(X_2,\ldots,X_n)$ and thus

$$\begin{aligned} & V_k^n(\hat{X}_1,\ldots,\hat{X}_n) = \hat{\alpha} V_{k-1}^{n-1}(\hat{X}_2,\ldots,\hat{X}_n \vee b_k) + (1-\hat{\alpha}) V_k^{n-1}(\hat{X}_2,\ldots,\hat{X}_n) \\ & \leq V_k^n(X_1,\ldots,X_n), \end{aligned} \tag{35}$$

which is (33). For any $X_1,\ldots,X_n$ let

$$X^*_{[2,n]} = X_2 \vee \cdots \vee X_n \quad \text{and} \quad \hat{X}^*_{[2,n]} = \hat{X}_2 \vee \cdots \vee \hat{X}_n. \tag{36}$$

Let $r = P(X^*_{[2,n]} = 0)$ and $s = P(0 < X^*_{[2,n]} \leq b_k)$. Then $x = P(X_n^* = 0) = (1-\alpha)r$, and also $x = P(\hat{X}_n^* = 0) = (1-\hat{\alpha})(r+s)$, and thus,

$$\begin{aligned} (1-\hat{\alpha}) &= (1-\alpha)r/(r+s) \quad \text{and so} \\ \hat{\alpha} &= 1-(1-\alpha)r/(r+s). \end{aligned} \tag{37}$$

Thus, using (37)

$$E\hat{X}_n^* = E\hat{X}^*_{[2,n]} + b_k(r+s)\hat{\alpha} = E\hat{X}^*_{[2,n]} + b_k(s+\alpha r), \tag{38}$$

whereas

$$\begin{aligned} EX_n^* &= (1-\alpha)EX^*_{[2,n]} + \alpha E[X^*_{[2,n]} \vee b_k] \\ &= (1-\alpha)EX^*_{[2,n]} + \alpha\{b_k + E[\hat{X}^*_{[2,n]} - b_k]^+\} \\ &= (1-\alpha)EX^*_{[2,n]} + \alpha\{b_k + E\hat{X}^*_{[2,n]} - (1-r-s)b_k\} \\ &= (1-\alpha)EX^*_{[2,n]} + \alpha E\hat{X}^*_{[2,n]} + b_k\alpha(r+s) \\ &\leq (1-\alpha)(E\hat{X}^*_{[2,n]} + sb_k) + \alpha E\hat{X}^*_{[2,n]} + b_k\alpha(r+s) \\ &= E\hat{X}^*_{[2,n]} + b_k(s+\alpha r) \\ &= E\hat{X}_n^*, \end{aligned} \tag{39}$$

by (38). Hence, together with (33), we have (32). □

3 The Differential Equation Approach

We begin this section with the

Proof of Theorem 1.2 We prove Theorem 1.2 by induction on n. For $n = 1$, we have

$$\frac{EX^*}{V_1^1(X)} = 1 < 2 - x = g_1(x), \quad \text{for all } 0 \leq x < 1.$$

With $x = P(X^*_{[2,n]} = 0)$, assume as our induction hypothesis that

$$\frac{EX^*_{[2,n]}}{V_1^{n-1}(X_2, \ldots, X_n)} < 2 - x. \tag{40}$$

Without loss of generality, we may assume the variables are as in Lemma 2.4; letting

$$X_1 = \begin{cases} 0 & 1-\alpha \\ b_1 & \alpha \end{cases}$$

where b_1 is the indifference value, i.e. satisfies $b_1 = V_1^{n-1}(X_2, \ldots, X_n)$, we have

$$EX_n^* = b_1 \alpha x + EX^*_{[2,n]}.$$

Since

$$V_1^n(X_1, \ldots, X_n) = V_1^{n-1}(X_2, \ldots, X_n) = b_1,$$

we have by (40),

$$\begin{aligned} \frac{EX_n^*}{V_1^n(X_1, \ldots, X_n)} &= \frac{b_1 \alpha x + EX^*_{[2,n]}}{b_1} \\ &< \alpha x + 2 - x \\ &= 2 - (1-\alpha)x \\ &= g_1((1-\alpha)x). \end{aligned}$$

But now the induction in complete, since $(1 - \alpha)x = P(X_n^* = 0)$. □

To see that $2 - x$ is the best bound, let $n = 2$, $0 < \mu \leq 1$, and

$$X_1 = \begin{cases} \mu & 1-x \\ 0 & x \end{cases} \tag{41}$$

and let

$$X_2 = \begin{cases} 1 & \mu \\ 0 & 1-\mu. \end{cases} \tag{42}$$

Then $V_1^2(X_1, X_2) = \mu$ and $E(X_2^*) = \mu + (1 - \mu)\mu(1 - x)$ and thus we have

$$\begin{aligned} E(X_2^*)/V_1^2(X_1, X_2) &= 2 - x - \mu(1-x) \quad \text{and} \\ P(X_2^* = 0) &= (1-\mu)x. \end{aligned}$$

Letting $\mu \to 0$ we have $E(X_2^*)/V_1^2(X_1, X_2) \to 2 - x$ while $P(X_2^* = 0) \to x$. Since $0 \le x < 1$ is arbitrary it follows that $2 - x$ cannot be improved upon. □

Note that Theorem 1.2 shows that inequality (43) of the following Lemma 3.1 is satisfied for $k = 1$ by $g_1(y) = 2 - y$.

Lemma 3.1 *Suppose that for a fixed k there exists a function $g_k(y)$ such that for any $n \ge k$ and any $Y_1, \dots, Y_n$ the inequality*

$$EY_n^* < g_k(x) V_k^n(Y_1, \dots, Y_n) \tag{43}$$

holds for $x = P(Y_n^* = 0) < 1$. Then for any $X_2, \dots, X_n$, $n \ge k+1$, with $X_i = X_i I(X_i > a), i = 2, \dots, n$ for some constant $a > 0$, we have that

$$\{(g_k(x) - 1 + x)a + EX_{[2,n]}^*\}/g_k(x) < V_{k+1}^n(a, X_2, \dots, X_n), \tag{44}$$

where $x = P(X_{[2,n]}^* = 0)$.

Proof: Let $Y_i = [X_i - a]^+, i = 2, \dots, n$ and $Y_{[2,n]}^* = Y_2 \vee \dots \vee Y_n$. Note that $EY_{[2,n]}^* = EX_{[2,n]}^* - (1-x)a$. Thus, by (43), since $P(Y_{[2,n]}^* = 0) = P(X_{[2,n]}^* = 0) = x$,

$$\begin{aligned} &V_{k+1}^n(a, X_2, \dots, X_n) \ge a + V_k^{n-1}(Y_2, \dots, Y_n) > a + EY_{[2,n]}^*/g_k(x) \\ &= a + (EX_{[2,n]}^* - (1-x)a)/g_k(x) = \{(g_k(x) - 1 + x)a + EX_{[2,n]}^*\}/g_k(x). \end{aligned} \tag{45}$$
□

We now derive an inequality for $k+1$ choices. By Lemma 2.4 for $n > k+1$ we need only consider random variables such that $X_1 = b_{k+1}$ and 0 with probabilities α and $1 - \alpha$ respectively, and $X_i = X_i I(X_i > b_{k+1})$ where $b_{k+1} = b_{k+1}(X_2, \dots, X_n)$. For short write $V_{k+1}^n = V_{k+1}^n(X_1, \dots, X_n)$. Then

$$V_{k+1}^n = V_{k+1}^n(X_1, \dots, X_n) = V_{k+1}^{n-1}(X_2, \dots, X_n). \tag{46}$$

From (44) with $a = b_{k+1}$ we have

$$b_{k+1} < \frac{g_k(x) V_{k+1}^n - EX_{[2,n]}^*}{g_k(x) - 1 + x}, \tag{47}$$

where $x = P(X_{[2,n]}^* = 0)$.

The following Lemma is the key step in establishing Theorem 1.3.

Lemma 3.2 *Suppose that for a fixed k there exists a function $g_k(x)$ such that for all $n \ge k$ and all $X_1, \dots, X_n$, $EX_n^* < g_k(x) V_k^n(X_1, \dots, X_n)$ for $x = P(X^* = 0)$, $0 \le x < 1$, and let*

$$h_k(x) = g_k(x) - 1 + x. \tag{48}$$

Suppose that a solution h_{k+1} in $[0, 1)$ exists to

$$h_{k+1}'(x) = \frac{h_{k+1}(x)}{h_k(x)}, \tag{49}$$

such that $h'_{k+1}(x)$ *is nondecreasing, and such that*

$$g_{k+1}(x) = h_{k+1}(x) + 1 - x > 1 \quad \text{for all } 0 \le x < 1. \tag{50}$$

Then

$$\begin{aligned} &EX_n^* < g_{k+1}(x) V_{k+1}^n (X_1, \ldots, X_n), \\ &\text{for all } n \ge k+1 \text{ and all } X_1, \ldots, X_n, \text{ where } x = P(X^* = 0). \end{aligned} \tag{51}$$

Proof: Again, by Lemma 2.4, we need only consider random variables such that $X_1 = b_{k+1}$ and 0 with probabilities α and $1-\alpha$ respectively, and $X_i = X_i I(X_i > b_{k+1})$ where $b_{k+1} = b_{k+1}(X_2, \ldots, X_n)$. We proceed by induction on n for fixed $k+1$. For our base case $n = k+1$ the only requirement for (51) to hold is that $g_{k+1}(x) > 1$, for $0 \le x < 1$, which is assumed. Now assume that (51) holds for some $n - 1 \ge k+1$, and consider $X_1, \ldots, X_n$; let $x = P(X^*_{[2,n]} = 0)$. For $n \ge k+2$ we have by use of (47),

$$\begin{aligned} EX_n^* &= \alpha x b_{k+1} + EX^*_{[2,n]} \\ &< \frac{\alpha x (g_k(x) V_{k+1}^n - EX^*_{[2,n]})}{g_k(x) - 1 + x} + EX^*_{[2,n]} \\ &= \frac{\alpha x g_k(x) V_{k+1}^n + EX^*_{[2,n]}(g_k(x) - 1 + (1-\alpha)x)}{g_k(x) - 1 + x}. \end{aligned}$$

The induction assumption and (46) yield that

$$EX^*_{[2,n]} < g_{k+1}(x) V_{k+1}^{n-1} = g_{k+1}(x) V_{k+1}^n, \tag{52}$$

hence,

$$\begin{aligned} EX_n^* &< \frac{\alpha x g_k(x) V_{k+1}^n + g_{k+1}(x) V_{k+1}^n (g_k(x) - 1 + (1-\alpha)x)}{g_k(x) - 1 + x} \\ &= \left\{ \frac{\alpha x [g_k(x) - g_{k+1}(x)]}{g_k(x) - 1 + x} + g_{k+1}(x) \right\} V_{k+1}^n. \end{aligned} \tag{53}$$

Our induction will be complete if we can show that for any $0 \le x < 1$ and any $0 < \alpha \le 1$ the value in the curly bracket on the right hand side of (53) is less than or equal to $g_{k+1}(x - \alpha x)$, since $P(X_n^* = 0) = (1-\alpha)x = x - \alpha x$. Rearranging terms, it suffices to show

$$\frac{g_{k+1}(x) - g_{k+1}(x - \alpha x)}{\alpha x} \le \frac{g_{k+1}(x) - g_k(x)}{g_k(x) - 1 + x}. \tag{54}$$

We can simplify the approach somewhat by rewriting (54) in terms of the functions h_k and h_{k+1} using (48),

$$\frac{h_{k+1}(x) - h_{k+1}(x - \alpha x)}{\alpha x} \le \frac{h_{k+1}(x)}{h_k(x)}. \tag{55}$$

But by the mean value theorem, the value of the left hand side of (55) is $h'_{k+1}(x - \theta x)$ for some $0 < \theta < \alpha$, and hence, since by our assumption $h'_{k+1}(x)$ is nondecreasing,

$$\frac{h_{k+1}(x) - h_{k+1}(x - \alpha x)}{\alpha x} = h'_{k+1}(x - \theta x) \le h'_{k+1}(x) = \frac{h_{k+1}(x)}{h_k(x)}. \qquad \square$$

Proof of Theorem 1.3: We show that the functions defined in (8) satisfy the conditions of Lemma 3.2. First, since h_{k+1} in (8) is positive, it satisfies (49) of Lemma 3.2 if and only if

$$u'_{k+1}(x) = e^{-u_k(x)}. \tag{56}$$

for

$$u_j(x) = \log h_j(x), \quad j = k, k+1. \tag{57}$$

Since we want the smallest solution $g_{k+1}(x)$, we take $h_{k+1}(1) = 1$ and therefore have chosen in (8) the solution for which $u_{k+1}(1) = 0$.

To verify the properties of these functions claimed in Theorem 1.3 we begin by proving that $u'_k e^{u_k} < 1$ for all $k \geq 1$, for the functions u_k defined in (8). The case $k = 1$ for $u_1(x) = 0$ is trivial, and we proceed by induction, assuming the inequality is true for k. Then

$$u'_k(x) < e^{-u_k(x)},$$

and integrating from x to 1 and using that $u_k(1) = 0$ we derive that

$$-u_k(x) < \int_x^1 e^{-u_k(y)} dy$$

or that

$$\exp\left\{-\left(u_k(x) + \int_x^1 e^{-u_k(y)} dy\right)\right\} < 1,$$

which is equivalent to $u'_{k+1} e^{u_{k+1}} < 1$.

We can now verify the claim made in Theorem 1.3 that the functions g_k defined in (8) are strictly decreasing; we have $g'_k < 0$ if and only if $h'_k < 1$, if and only if $u'_k e^{u_k} < 1$.

Next we show that the functions h'_{k+1} are non-decreasing. The inequality $u'_k e^{u_k} < 1$, or $u'_k < e^{-u_k}$ is equivalent to $u'_k < u'_{k+1}$. Hence

$$\frac{h'_k}{h_k} < \frac{h'_{k+1}}{h_{k+1}},$$

which with (49) yields

$$h''_{k+1}(x) = \frac{h'_{k+1} h_k - h_{k+1} h'_k}{h_k^2} > 0,$$

and that h'_{k+1} is increasing.

Next, we need to show that $g_{k+1}(x) > 1$ for $0 \leq x < 1$. Since g_{k+1} is strictly decreasing, for $0 \leq x < 1$ we have

$$g_{k+1}(x) > g_{k+1}(1) = h_{k+1}(1) = e^{u_{k+1}(1)} = 1.$$

Lastly, Theorem 1.2 gives the base step for the induction with $g_1(x) = 2 - x$, and therefore $h_1(x) = 1$, and $u_1(x) = 0$. For $k = 2$ we have

$$u_2(x) = -\int_x^1 1 dy = -(1-x), \quad h_2(x) = e^{-(1-x)},$$

and so

$$g_2(x) = e^{-(1-x)} + 1 - x.$$

Then

$$u_3(x) = -\int_x^1 e^{1-y} dy = 1 - e^{1-x}, \quad h_3(x) = \exp(1 - e^{1-x})$$

and

$$g_3(x) = \exp(1 - e^{1-x}) + 1 - x.$$

Thus

$$u_4(x) = e^{-1} \int_x^1 e^{e^{(1-y)}} dy. \tag{58}$$

For (58) we can make a change of variables so as to use existing tables. Set $e^{(1-y)} = z$. Then (58) can also be written as

$$u_4(x) = -e^{-1} \oint^{e^{(1-x)}} \frac{e^z}{z} \, dz + c. \tag{59}$$

The function

$$Ei(x) = \oint_{-\infty}^{x} \frac{e^x}{z} \, dz, \qquad x > 0 \tag{60}$$

is tabulated, see e.g. Abramowitz and Stegun (1964) Table 5.1. (58) with the requirement $u_4(1) = 0$ can be written as

$$u_4(x) = e^{-1}[Ei(1) - Ei(e^{1-x})]. \qquad \square \tag{61}$$

In particular for $x = 0$ we get $u_4(0) = e^{-1}[Ei(1) - Ei(e)] \approx -2.32337$ and thus $g_4 = g_4(0) = 1.0979$ as in Theorem 1.1.

Further numerical integration yields the values

$$g_5 = 1.0567\ldots \quad , \quad g_6 = 1.0341\ldots$$

We conclude the paper with the proof of Assertion 3.1, showing that the bounds derived here are strictly better than the bounds of Assaf and Samuel-Cahn (2000), for all $k \geq 2$.

Assertion 3.1 *For $k \geq 2$, $g_k(0) < (k+1)/k$.*

Proof: The assertion is equivalent to

$$h_k(0) < \frac{1}{k}, \quad k = 2, 3, \ldots. \tag{62}$$

By Theorem 1.3, $h_2(0) = e^{-1} < 1/2$, thus (62) holds for $k = 2$. We proceed by induction. Showing (62) for $k + 1$ is equivalent to

$$\log(k+1) < -u_{k+1}(0). \tag{63}$$

Now

$$-u_{k+1}(0) = \int_0^1 e^{-u_k(x)} dx = \int_0^1 \frac{1}{h_k(x)} dx.$$

We shall show

$$h_k(x) \le w_k(x) = (1 + (k-1)x)/k, \quad \text{for } 0 \le x \le 1. \tag{64}$$

Then it follows that

$$-u_{k+1}(0) = \int_0^1 \frac{1}{h_k(x)} dx \ge \int_0^1 \frac{k}{1 + (k-1)x} dx = \frac{k \log k}{k-1} > \log(k+1). \tag{65}$$

To see the last inequality in (65), note that it is equivalent to

$$\frac{\log k}{k-1} > \log(1 + \frac{1}{k}), \tag{66}$$

and since $1/k > \log(1 + 1/k)$, (66) clearly holds for all $k \ge 2$.

It remains to show (64). Consider the difference

$$m_k(x) = w_k(x) - h_k(x).$$

We must show that

$$m_k(x) \ge 0 \quad \text{for } 0 \le x \le 1. \tag{67}$$

By the induction hypotheses, $m_k(0) > 1/k - 1/k = 0$, and clearly $m_k(1) = 1 - 1 = 0$.

We have shown in the proof of Theorem 1.3 that $h_k'(x)$ is increasing in x for $0 \le x \le 1$. Thus $h_k(x)$ is convex, and since $w_k(x)$ is linear, $m_k(x)$ is concave. Since a concave function taking non-negative values at the endpoints of an interval must be non-negative on that interval, (67) follows. □

References

Abramowitz, M. and Stegun, I.A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables.* Department of Commerce. National Bureau of Standards, Washington D.C.

Assaf, D. and Samuel-Cahn, E. (2000). Simple ratio prophet inequalities for a mortal with multiple choices. *Journal of Applied Probablity*, **37**.

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

Kennedy, D.P. (1987). Prophet type inequalities for multi-choice optimal stopping. *Stochastic Proc. and their Applications.*, **24**, 77-88.

Stadje, W. (1985). On multiple stopping rules. *Optimization*, **16**, 401-418.