A Sharp Ratio Inequality for Optimal Stopping When Only Record Times are Observed

Pieter C. Allaart
Department of Mathematics, University of North Texas, Denton, Texas, USA

Abstract: Let $X_1, \ldots, X_n$ be independent, identically distributed random variables that are nonnegative and integrable, with known continuous distribution. These random variables are observed sequentially, and the goal is to maximize the expected X value at which one stops. Let V_n denote the optimal expected return of a player who can observe at time j only whether X_j is a relative record ($j = 1, \ldots, n$), and W_n that of a player who observes at time j the actual value of X_j. It is shown that $V_n > a_n W_n$, where $a_n = \max_{1 \le k < n}(k/n) \sum_{j=k}^{n-1} 1/j$, and this inequality is sharp.

Keywords: Optimal stopping rule; Relative record; Secretary problem.

Subject Classifications: 60G40.

Let $X, X_1, \ldots, X_n$ be independent, identically distributed (i.i.d.) random variables that are nonnegative and integrable, with known continuous distribution. Recently, Samuel-Cahn (2007) investigated the problem of maximizing $\mathrm{E}X_\tau$ over those stopping times τ that use only information about the record times of the sequence $X_1, \ldots, X_n$. Precisely, let $Y_1 \equiv 1$, and for $j = 2, \ldots, n$, let $Y_j = 1$ if $X_j > \max\{X_1, \ldots, X_{j-1}\}$, and $Y_j = 0$ otherwise. Define the σ-algebras $\mathcal{G}_j := \sigma(\{Y_1, \ldots, Y_j\})$, $j = 1, \ldots, n$, and let

$$V_n := \sup_{\tau \le n} \mathrm{E}\, X_\tau,$$

the supremum being over all stopping times τ adapted to the filtration $\{\mathcal{G}_j\}$ with $\tau \le n$ a.s.

Samuel-Cahn shows that the optimal rule is among the rules

$$t_n(k) = \min\{j > k : Y_j = 1\} \wedge n, \quad k = 1, \ldots, n-1,$$

and the expected payoff from the rule $t_n(k)$ is

$$V_n(k) := \mathrm{E}\, X_{t_n(k)} = k \left[\sum_{j=k+1}^{n-1} \frac{\mathrm{E}M_j}{j(j-1)} + \frac{\mathrm{E}X}{n-1} \right], \tag{1}$$

where $M_j := \max\{X_1, \ldots, X_j\}$. Thus, $V_n = \max_{1 \le k < n} V_n(k)$.

Samuel-Cahn goes on to describe the asymptotic properties of V_n as $n \to \infty$ for various distributions of X, and to give asymptotic comparisons of V_n with the usual optimal stopping value

$$W_n := \sup_{\tau \le n} \mathrm{E}\, X_\tau,$$

Address correspondence to P. C. Allaart, Department of Mathematics, University of North Texas, 1155 Union Circle #311430, Denton, TX 76203-5017, USA; E-mail: allaart@unt.edu

where the supremum is over all stopping times $\tau \le n$ adapted to the filtration $\{\mathcal{F}_j\}$ defined by $\mathcal{F}_j := \sigma(\{X_1, \ldots, X_j\})$, $j = 1, \ldots, n$.

For *fixed* n, however, (1) immediately yields a sharp inequality comparing V_n and W_n. Let

$$s(k,n) := \sum_{j=k}^{n} \frac{1}{j}, \qquad a_n := \max_{1 \le k < n} (k/n) s(k, n-1).$$

Theorem 1. *For every $n \ge 2$,*

$$V_n > a_n W_n, \tag{2}$$

and the constant a_n is best possible.

Proof. Let $M_0 \equiv 0$. Since X is continuous it is not degenerate, and therefore (as is easy to show), $\mathrm{E} M_{j+1} - \mathrm{E}\, M_j < \mathrm{E}\, M_j - \mathrm{E}\, M_{j-1}$ for all $j \ge 1$. Hence, there is a strictly concave function that interpolates the points $(j, \mathrm{E}\, M_j)$, $j = 0, 1, \ldots, n$, so that

$$\mathrm{E} M_j > \frac{j}{n} \mathrm{E}\, M_n, \quad j = 1, \ldots, n-1.$$

Substituting this into (1) yields

$$V_n > \max_{1 \le k < n} k \left[\sum_{j=k+1}^{n-1} \frac{\mathrm{E}\, M_n}{n(j-1)} + \frac{\mathrm{E}\, M_n}{n(n-1)} \right] = a_n \,\mathrm{E}\, M_n \ge a_n W_n.$$

To see that the bound is sharp, let X initially have distribution

$$\mathrm{P}(X = K) = K^{-1} = 1 - \mathrm{P}(X = 0), \tag{3}$$

where $K > 1$. Note that for this distribution, $W_n = \mathrm{E}\, M_n$. Moreover,

$$\mathrm{E} M_j = K\, \mathrm{P}(M_j = K) = K[1 - (1 - K^{-1})^j] \to j \quad \text{as } K \to \infty,$$

and so, given $\varepsilon > 0$, K may be chosen so large that

$$\mathrm{E} M_j \le \left(\frac{j}{n} + \varepsilon \right) \mathrm{E}\, M_n$$

for $j = 1, \ldots, n$. Hence

$$\begin{aligned} V_n(k) &\le \left[a_n + k \sum_{j=k+1}^{n} \frac{\varepsilon}{j(j-1)} \right] \mathrm{E}\, M_n \\ &= \left[a_n + \left(1 - \frac{k}{n}\right) \varepsilon \right] \mathrm{E}\, M_n < (a_n + \varepsilon)\, \mathrm{E}\, M_n. \end{aligned}$$

Of course, the distribution in (3) is not continuous, but it may be approximated arbitrarily closely by a continuous distribution with respect to for instance the Lévy-Prohorov metric.

Since the functionals $\mathrm{E}M_j$ are continuous with respect to this metric, the sharpness of the bound (2) follows. □

It is interesting to note that the constant a_n is exactly the probability of selecting the best of n items under an optimal strategy in the classical secretary problem or dowry problem. Gilbert and Mosteller (1966, pp. 39-41) gave a large table of values of a_n (for instance, $a_2 = a_3 = 1/2$, $a_4 = 11/24$, $a_5 = 13/30$, etc.) and showed that $\lim_{n\to\infty} a_n = 1/e$, a fact that has since become common knowledge. However, it does not appear to have been demonstrated explicitly in the literature that the a_n's are strictly decreasing from a_3 onward. This intuitively plausible fact depends on the following elegant result from number theory due to Kürschák (1918); see also Pólya and Szegö (1976, Problem VIII.251, p. 154 and pp. 358/59).

Lemma 1. *Let k and n be positive integers with $k < n$. Then $s(k,n)$ is not an integer.*

Proposition 1. *The constants $\{a_n\}$ are strictly decreasing for $n \geq 3$.*

Proof Let

$$\alpha_{n,k} := (k/n)s(k,n-1), \quad k = 1,\dots,n-1,$$

and let k_n be the value of k that maximizes $\alpha_{n,k}$ for fixed n. It follows as in Gilbert and Mosteller (1966, p. 39) that

$$k_n = \min\{k : s(k+1,n-1) < 1\}. \tag{4}$$

From this it is clear that k_{n+1} is equal to either k_n or $k_n + 1$. If $k_{n+1} = k_n$, then

$$\begin{aligned} a_n - a_{n+1} &= \alpha_{n,k_n} - \alpha_{n+1,k_n} \\ &= k_n \left[\frac{s(k_n,n-1)}{n} - \frac{s(k_n,n)}{n+1}\right] \\ &= \frac{k_n}{n(n+1)}\big(s(k_n,n-1) - 1\big) > 0 \end{aligned}$$

by (4), the inequality being strict for $n \geq 3$ in view of Lemma 1. And if $k_{n+1} = k_n + 1$, then

$$\begin{aligned} a_n - a_{n+1} &= \alpha_{n,k_n} - \alpha_{n+1,k_n+1} \\ &= \frac{k_n}{n}s(k_n,n-1) - \frac{k_n+1}{n+1}s(k_n+1,n) \\ &= \frac{n-k_n}{n(n+1)}\big(1 - s(k_n+1,n-1)\big) > 0 \end{aligned}$$

by (4). Thus, a_n is strictly decreasing for $n \geq 3$. □

Remarks. *(a)* It is well known that $k_n \approx n/e$, the approximation being fairly accurate even for small values of n. Thus, the rule "let roughly a proportion $1/e$ of the observations go by, then stop with the next relative record" guarantees an expected payoff at least a_n

times the payoff of a player who can observe the actual X_i's, regardless of the distribution of X.

(b) It is clear from the proof of Theorem 1 that even a "prophet" who can foretell the future can do no better than $1/a_n$ times the value of a player observing only record times. In particular, the advantage of a prophet over such a player is bounded by the factor e.

(c) Hill and Kertz (1982) showed that for i.i.d. random variables, $\mathrm{E}M_n \leq c_n W_n$, where $\{c_n\}$ is a sequence of numbers, believed to be strictly increasing, with known limit $1.34149\cdots$. Seen in this light, the result of this note shows that the opportunity to fully observe the random variables gives a much greater advantage than the gift of foresight.

REFERENCES

Gilbert, J. P. and Mosteller, F. (1966). Recognizing the Maximum of a Sequence. *J. Amer. Statist. Assoc.* 61: 35-73.

Hill, T. P. and Kertz, R. P. (1982). Comparisons of Stop Rule and Supremum Expectations of I.I.D. Random Variables, *Annals of Probability* 10: 336-345.

Kürschák, J. (1918). On the Harmonic Series. (Hungarian). *Mat. Phys. Lapok* 27: 299-300.

Pólya, G. and Szegö, G. (1976). *Problems and Theorems in Analysis*, Vol. II. New York: Springer-Verlag.

Samuel-Cahn, E. (2007). Optimal Stopping for I.I.D. Random Variables Based on the Sequential Information of the Location of Relative Records Only, *Sequential Analysis* 26: 395-401.