Prophet inequalities for i.i.d. random variables with random arrival times.

Pieter C. Allaart
University of North Texas *

October 3, 2005

Abstract

Suppose $X_1, X_2, \ldots$ are i.i.d. nonnegative random variables with finite expectation, and for each k, X_k is observed at the k-th arrival time S_k of a Poisson process with unit rate which is independent of the sequence $\{X_k\}$. Let $t > 0$ be a finite time horizon. Several comparisons are made between the expected maximum $M(t) := \mathrm{E}[\max_{k\geq 1} X_k \, \mathrm{I}(S_k \leq t)]$ and the optimal stopping value $V(t) := \sup_{\tau\in\mathcal{T}} \mathrm{E}[X_\tau \, \mathrm{I}(S_\tau \leq t)]$, where $\mathcal{T}$ is the set of all $\mathbb{N}$-valued random variables τ such that $\{\tau = i\}$ is measurable with respect to the σ-algebra generated by $X_1, \ldots, X_i$ and $S_1, \ldots, S_i$. For instance, it is shown that $M(t)/V(t) \leq 1 + \alpha_0$, where $\alpha_0 \doteq 0.34149$ is the unique value of α such that $\int_0^1 (y - y\ln y + \alpha)^{-1} \, dy = 1$; and this bound is asymptotically sharp as $t \to \infty$. Another result is that $M(t)/V(t) < 2 - (1 - e^{-t})/t$, and this bound is asymptotically sharp as $t \downarrow 0$. Analogous upper bounds for the difference $M(t) - V(t)$ are also given, under the additional assumption that the X_k are $[0, 1]$-valued.

AMS 2000 subject classification: 60G40, 62L15.

Key words and phrases: Optimal stopping rule, Poisson process, Prophet inequality.

* **Author's address**: Mathematics Department, University of North Texas, P.O. Box 311430, Denton, TX 76203-1430; e-mail: allaart@unt.edu

1 Introduction

S. Karlin [4] studied the following problem. An item is for sale, and independent, identically distributed price offers $X_1, X_2, \ldots$ arrive according to a Poisson process with rate λ. The item must be sold by a fixed time $t > 0$, or it becomes worthless. Thus, if S_i denotes the arrival time of the i-th offer ($i \in \mathbb{N}$), the optimal expected return is given by

$$V(t) := \sup_{\tau \in \mathcal{T}} \mathrm{E}[X_\tau \, \mathrm{I}(S_\tau \leq t)], \tag{1}$$

where $\mathcal{T}$ is the set of all $\mathbb{N}$-valued random variables (stopping rules) τ such that $\{\tau = i\}$ is measurable with respect to the σ-algebra generated by $X_1, \ldots, X_i$ and $S_1, \ldots, S_i$. Assuming the price offers are independent of the arrival process, Karlin showed that $V(t)$ is the unique solution of the initial value problem

$$V' = \lambda \, \mathrm{E}(X_1 - V)^+, \qquad V(0) = 0, \tag{2}$$

and the optimal policy is to accept the first offer whose value X satisfies $X > V(\tau)$, where τ is the amount of time remaining when the offer arrives. Sakaguchi [8] gives explicit solutions of (2) for several common distributions of X_1.

The purpose of this paper is to compare $V(t)$ with the expected maximum

$$M(t) := \mathrm{E}(\max\{X_1, \ldots, X_{N(t)}\}),$$

where $N(t)$ denotes the number of arrivals up to time t. (The maximum of an empty set is taken to be zero.) In particular, reasonably tight upper bounds are given for the ratio $M(t)/V(t)$ and the difference $M(t) - V(t)$, the latter under the additional assumption that the price offers are uniformly bounded. The work is inspired by analogous comparisons in the discrete-time setting, which are known in the literature as *prophet inequalities.* Let $Z_1, \ldots, Z_n$ be independent nonnegative random variables with finite expectations, and define the quantities $V_n := V(Z_1, \ldots, Z_n) := \sup\{\mathrm{E}\, Z_\tau : \tau \text{ is a stopping rule for } Z_1, \ldots, Z_n\}$, and $M_n := \mathrm{E}(\max\{Z_1, \ldots, Z_n\})$. Krengel and Sucheston [6] proved that $M_n \leq 2V_n$, with strict inequality if $Z_1, \ldots, Z_n$ are not all identically equal to zero. Hill and Kertz [2] showed that if the nonnegativity condition is replaced with the condition that the Z_i take values in an interval $[a, b]$, then $M_n - V_n \leq (b-a)/4$. While the constants 2 and 1/4 are best-possible in general, sharper bounds hold if the Z_i are known to be identically distributed. This was shown by Hill and Kertz [3], who proved the following two theorems.

Theorem A (Hill and Kertz [3]). *There exist constants* $\{a_n\}$ *with* $1.1 < a_n < 1.6$*,* $n \geq 2$*, such that if* $Z_1, \ldots, Z_n$ *are i.i.d. nonnegative random variables, then* $M_n \leq a_n V_n$*. This bound is sharp, and holds with strict inequality if* Z_1 *is not identically equal to zero.*

Theorem B (Hill and Kertz [3]). *There exist constants* $\{b_n\}$ *with* $0 < b_n < 1/4$*,* $n \geq 2$*, such that if* $Z_1, \ldots, Z_n$ *are i.i.d. nonnegative random variables taking values in* $[a, b]$*, then* $M_n - V_n \leq b_n(b - a)$*. This bound is attained.*

The precise definitions of the constants a_n and b_n, along with some sample values, are given in Section 2.

Comparisons of the above kind have been called "prophet inequalities" because of the natural interpretation of M_n as the optimal expected return of a player with complete foresight into the future. This paper develops analogous inequalities for a continuous-time model, where observations arrive according to a Poisson process.

No claim is made here as to having "best-possible" bounds; indeed, it seems unlikely that such bounds can be found explicitly (other than by a lucky guess of the extremal distributions), since many of the tools that have worked well in the discrete-time setting (backward and forward induction, balayage, etc.) are inadequate here in view of the continuous time parameter. In particular (and contrary to the discrete-time case), there is no mechanism to immediately reduce the problem to random variables taking only finitely many values.

Throughout the paper it is assumed that $\lambda = 1$. However, the results can easily be restated for arbitrary $\lambda > 0$ by a simple time scaling, replacing 't' with 'λt'. In fact, all of the results are valid (with only trivial modifications) for non-homogeneous Poisson processes as well. The random variables $X, X_1, X_2, \ldots$ are assumed to be nonnegative and i.i.d., with finite expectation, and not identically equal to zero.

Four inequalities are proved: two for the ratio $M(t)/V(t)$, and two for the difference $M(t) - V(t)$, the latter under the additional assumption that X is bounded. While all of the bounds hold for every $t > 0$, one pair of bounds is reasonably good for large values of t (and is asymptotically sharp as $t \to \infty$), while the other pair of bounds is fairly sharp for small values of t (and is asymptotically sharp as $t \downarrow 0$). The bounds for "large" t in Section 2 use Theorems A and B above, and in fact the results are that $M(t)/V(t) \leq \lim_{n\to\infty} a_n \doteq 1.341$, and if X is $[0,1]$-valued, then $M(t) - V(t) \leq \limsup_{n\to\infty} b_n$ (the value of which is not known at present). Obviously, these bounds are far from sharp when t is small, since the prophet's advantage disappears as the expected number of observations approaches zero. This is made precise by the inequalities for "small" t in Section 3. For example, Theorem 3.2 shows that $M(t)/V(t) < 2 - (1 - e^{-t})/t$. The proofs of these inequalities make use of pure threshold rules. The idea is that when the expected number of observations is small, threshold rules should not perform much worse than the optimal rule.

In the discrete-time setting, threshold stopping rules were used by Samuel-Cahn [9, 10], who proved that the advantage of a prophet over a mortal player without foresight does not increase (in the extremal case) when the mortal is limited to the use of pure threshold rules. Rinott and Samuel-Cahn [7] used threshold rules to prove that $M_n \leq 2V_n$ even for a class of negatively dependent random variables.

Just how close the bounds of Sections 2 and 3 are to being sharp is explored in Section 4, which uses two-point and three-point distributions for X to obtain lower bounds on the best-possible ratio and difference constants. Section 5 gives examples in which the price offers arrive according to a discrete-time renewal process. The first example concerns the binomial process, for which it is shown that the expected maximum and stop rule supremum satisfy the bounds of Theorems A and B. The second example shows that the bounds derived in Section 2 fail to hold for general renewal processes.

The following notation is used throughout. For real numbers x and y, $x \vee y$ denotes the maximum of x and y. For constants $a < b$, the collection of all random variables taking values in $[a, b]$ is denoted by $\mathcal{X}_{[a,b]}$. When it is desirable to emphasize the dependence of $M(t)$ and $V(t)$ on the distribution of X, they will be written as $M(t; X)$ and $V(t; X)$, respectively. Finally, for $s > 0$, define $X_s^* := \max\{X_1, \ldots, X_{N(s)}\}$.

2 Long-range prophet inequalities

This section develops upper bounds for the ratio and difference of $M(t)$ and $V(t)$ which are fairly sharp when t is large. The following simple inequality will be helpful.

Lemma 2.1 *For all $s > 0$ and all $c \geq 0$,*

$$\mathrm{E}(X_s^* - c)^+ \leq s\, \mathrm{E}(X - c)^+.$$

Proof. We have

$$\begin{aligned} \mathrm{E}(X_s^* - c)^+ &= \sum_{n=1}^{\infty} \mathrm{E}(X_1 \vee \cdots \vee X_n - c)^+\, \mathrm{P}(N(s) = n) \\ &\leq \sum_{n=1}^{\infty} n\, \mathrm{E}(X - c)^+\, \mathrm{P}(N(s) = n) \\ &= \mathrm{E}\, N(s) \cdot \mathrm{E}(X - c)^+ = s\, \mathrm{E}(X - c)^+. \quad \square \end{aligned}$$

Theorem 2.2 *For all $t > 0$,*

$$M(t) \leq (1 + \alpha_0) V(t),$$

where $\alpha_0 \doteq 0.34149$ is the unique value of α such that $\int_0^1 (y - y \ln y + \alpha)^{-1}\, dy = 1$.

Proof. Fix $n \geq 2$, and let $\delta := t/n$. Consider an intermediate player, to be called the "partial prophet", who has limited foresight in the sense that he can see, at the beginning of each time interval $I_i := ((i-1)\delta, i\delta]$, $i = 1, \ldots, n$, all of the observations (if any) arriving in that interval. For $i = 1, \ldots, n$, let Z_i denote the largest of the observations arriving in the interval I_i (or $Z_i = 0$ if no observations arrive in that interval). It is a routine exercise to verify that $Z_1, \ldots, Z_n$ are i.i.d.

Let $v_j := V(Z_{j+1}, \ldots, Z_n)$, for $j = 0, 1, \ldots, n-1$, and let $v_n := 0$. Since the partial prophet sequentially observes $Z_1, \ldots, Z_n$, his optimal expected return is simply v_0, and, by backward induction [1, p. 50], his optimal rule is to stop in the first time interval I_i for which $Z_i \geq v_i$, and to accept the largest observation, Z_i, in that interval.

Now consider the following stopping rule for the gambler:

> Accept the first observation X_j such that, if X_j arrives in the time interval I_i, then $X_j \geq v_i$.

Let $V_\delta(t)$ denote the expected return from this stopping rule. Define $Z_i' = Z_i \, \mathrm{I}(Z_i \geq v_i)$, and let $X_{i,1}', X_{i,2}', \ldots$ denote the successive values of those X_j arriving after time $(i-1)\delta$ for which $X_j \geq v_i$. Let N_i' be the number of such observations (with $X_j \geq v_i$) that arrive in the interval I_i. Observe that N_i' has a Poisson distribution with mean $\delta \, \mathrm{P}(X \geq v_i)$, and $Z_i' = \max\{X_{i,1}', \ldots, X_{i,N_i'}'\}$. Thus, Lemma 2.1 applied to the sequence $X_{i,1}', X_{i,2}', \ldots$ yields that

$$\mathrm{E} Z_i' \leq \delta \, \mathrm{P}(X \geq v_i) \, \mathrm{E} \, X_{i,1}'. \tag{3}$$

Note that

$$v_0 = \sum_{i=1}^{n} \mathrm{P}(Z_1 < v_1, \ldots, Z_{i-1} < v_{i-1}) \, \mathrm{E}[Z_i \, \mathrm{I}(Z_i \geq v_i)], \tag{4}$$

and

$$V_\delta(t) = \sum_{i=1}^{n} \mathrm{P}(Z_1 < v_1, \ldots, Z_{i-1} < v_{i-1}) \, \mathrm{E}[X_{i,1}' \, \mathrm{I}(Z_i \geq v_i)]. \tag{5}$$

Using (3), we obtain that

$$\begin{aligned}
\mathrm{E}[X_{i,1}' \, \mathrm{I}(Z_i \geq v_i)] &= \mathrm{E} \, X_{i,1}' \, \mathrm{P}(N_i' \geq 1) \geq \mathrm{E} \, Z_i' \cdot \frac{\mathrm{P}(N_i' \geq 1)}{\delta \, \mathrm{P}(X \geq v_i)} \\
&= \mathrm{E} \, Z_i' \cdot \frac{1 - e^{-\delta \, \mathrm{P}(X \geq v_i)}}{\delta \mathrm{P}(X \geq v_i)} \geq \mathrm{E} \, Z_i' \cdot (1 - e^{-\delta})/\delta \\
&= \mathrm{E}[Z_i \, \mathrm{I}(Z_i \geq v_i)](1 - e^{-\delta})/\delta,
\end{aligned}$$

where the second inequality follows since the function $(1 - e^{-\delta p})/p$ is decreasing in p. Substituting this result into (5) and comparing with (4) yields the conclusion

$$v_0 \leq \frac{\delta}{1 - e^{-\delta}} V_\delta(t) \leq \frac{t/n}{1 - e^{-t/n}} V(t). \tag{6}$$

By Theorem A,

$$M(t) = \mathrm{E}(X_1 \vee \cdots \vee X_{N(t)}) = \mathrm{E}(Z_1 \vee \cdots \vee Z_n) < a_n v_0,$$

and, hence,

$$M(t) < a_n \frac{t/n}{1 - e^{-t/n}} V(t).$$

It was shown by Kertz [5] that $\lim_{n \to \infty} a_n = 1 + \alpha_0$. Since n was arbitrary, the theorem follows upon letting $n \to \infty$. $\square$

Theorem 2.3 *Assume that $X_1, X_2, \ldots$ are $[0,1]$-valued. Then for all $t > 0$,*

$$M(t) - V(t) \leq \limsup_{n\to\infty} b_n.$$

Proof. We use the notation from the proof of Theorem 2.2. Since $v_0 \leq \mathrm{P}(N(t) \geq 1) = 1 - e^{-t}$, Theorem B and (6) imply that

$$M(t) - V(t) \leq b_n + (1 - e^{-t}) \left[1 - \frac{n(1 - e^{-t/n})}{t} \right], \qquad \text{for all } n \geq 2. \tag{7}$$

Letting $n \to \infty$ completes the proof. $\square$

Remark 2.4 The value of $\limsup b_n$ does not seem to be known at present. From the numerical values of b_n given in Table 1 ($b_n = \beta_n$), it seems likely that $\limsup_{n\to\infty} b_n \approx .1113$. However, in the absence of an exact value for this limit, equation (7) gives good practical upper bounds by taking $n = 10^6$. For instance, if $t = 1000$ the right hand side of (7) evaluates to .11176, and this is a rigorous bound.

Remark 2.5 The results of Theorems 2.2 and 2.3 hold equally for any non-homogeneous Poisson process with rate function $\lambda(x)$, $x > 0$, provided $\lambda(x)$ is bounded on bounded intervals. The only modification required in the proof of Theorem 2.2 concerns the choice of the n intervals that partition the interval $(0, t]$. Note that the intervals can always be chosen so that the number of arrivals in each interval is Poisson with the *same* parameter $\mu = (1/n) \int_0^t \lambda(x)\,dx$. With this new partition is again associated an i.i.d. sequence of random variables $Z_1, \ldots, Z_n$, and the rest of the proof goes through with μ in place of δ.

How sharp are the bounds in Theorems 2.2 and 2.3? Some insight may be gained by analyzing the extremal distributions from Hill and Kertz [3]. The following notation is taken from that paper. For $n > 1$ and $w, x \in [0, \infty)$, let $\phi_n(w, x) = (n/(n-1))w^{(n-1)/n} + x/(n-1)$. For $\alpha \in [0, \infty)$, define $\eta_{0,n}(\alpha) = \phi_n(0, \alpha)$, and inductively, $\eta_{j,n}(\alpha) = \phi_n(\eta_{j-1,n}(\alpha), \alpha)$ for $j \geq 1$. In their Propositions 3.4 and 3.8, Hill and Kertz show that

(i) there is a unique $\alpha_n \in (0,1)$ such that $\eta_{n-1,n}(\alpha_n) = 1$; and

(ii) there is a unique $\beta_n \in (0,1)$ such that $(n-1)[\eta_{n,n}(\beta_n) - \eta_{n-1,n}(\beta_n)] = 1$.

n	α_n	β_n	n	α_n	β_n
2	.17157	.06250	20	.32055	.10561
3	.22138	.07761	40	.33081	.10839
4	.24811	.08539	60	.33432	.10933
5	.26496	.09020	80	.33609	.10981
6	.27659	.09348	100	.33716	.11010
7	.28513	.09586	1000	.34105	.11114
8	.29166	.09768	10^4	.34144	.11125
9	.29683	.09911	10^5	.34148	.11126
10	.30101	.10027	10^6	.34149	.11126

Table 1: Selected values of α_n and β_n.

Kertz [5, Lemma 6.2(b)] proves further that $\lim_{n\to\infty} \alpha_n = \alpha_0$, with α_0 as defined in the statement of Theorem 2.2. Table 1 gives some sample values of α_n and β_n.

Now let $a_n := 1 + \alpha_n$, and $b_n := \beta_n$. These are the precise definitions of the constants appearing in Theorems A and B.

Proposition 4.4 of Hill and Kertz [3] gives the ε-extremal distributions for Theorem A. They satisfy $\mathrm{P}(Z_1 = 0) = (\eta_{0,n}(\alpha_n))^{1/n} = (\alpha_n/(n-1))^{1/n}$. Observe that by the construction of the random variables $\{Z_i\}$ in the proof of Theorem 2.2, Z_1 has an atom at zero of size at least $\mathrm{P}(N(\delta) = 0) = e^{-\delta} = e^{-t/n}$, but every distribution on $[0, \infty)$ satisfying this condition can be obtained by a suitable choice of the distribution of X. It follows that, as long as

$$t \geq \log((n-1)/\alpha_n),$$

then for every $\varepsilon > 0$ there exists a random variable X and corresponding i.i.d. random variables $Z_1, \ldots, Z_n$ such that

$$M(t;X) = \mathrm{E}(Z_1 \vee \cdots \vee Z_n) > (a_n - \varepsilon)V(Z_1, \ldots, Z_n) \geq (a_n - \varepsilon)V(t;X).$$

For example, if $t \geq \log(99/\alpha_{100}) \doteq 5.683$, then $\sup_X M(t;X)/V(t;X) \geq \alpha_{100} \doteq 1.337$.

Similarly, the extremal distribution for Theorem B (Proposition 5.3 of Hill and Kertz [3]) has $\mathrm{P}(Z_1 = 0) = (\eta_{0,n}(\beta_n))^{1/n}$, so as long as $t \geq \log((n-1)/\beta_n)$, there exists a $[0,1]$-valued random variable X such that $M(t;X) - V(t;X) \geq b_n$. For example, if $t \geq \log(99/\beta_{100}) \doteq 6.802$, then $\sup_{X \in \mathcal{X}_{[0,1]}} (M(t;X) - V(t;X)) \geq \beta_{100} \doteq 0.110$.

Remark 2.6 If X takes values in an interval $[a,b]$ with $0 \leq a < b$, then $M(t) - V(t) \leq b \cdot \limsup_{n\to\infty} b_n$. That the bound is not $(b-a) \limsup_{n\to\infty} b_n$ can be explained by the fact that the random variables Z_i in the proof of Theorem 2.2 are not $[a,b]$-valued but $[0,b]$-valued. Furthermore, the sharpness considerations above no longer apply if a is not very close to zero, since the extremal distributions for Theorem B typically give positive mass to small values.

3 Short-range prophet inequalities

The inequalities obtained in the previous section can be improved considerably when t (and with it the expected number of observations) is small. In this section, we consider pure threshold rules of the form $\tau(c) = \inf\{n : X_n \geq c\}$. The rationale is that for a small expected number of observations the best threshold rule should perform almost as well as the optimal rule. The advantage of using threshold rules is that explicit expressions are available for the expected return. For $c \geq 0$, let $W_c(t)$ denote the expected return from the rule that accepts the first observation whose value is greater than or equal to c. Thus,

$$W_c(t) := W_c(t;X) := \mathrm{E}\left[X_{\tau(c)}\, \mathrm{I}(\tau(c) \leq N(t))\right].$$

The value of $W_c(t)$ is given by

$$W_c(t) = \left[1 - e^{-t\,\mathrm{P}(X \geq c)}\right] \mathrm{E}(X | X \geq c). \tag{8}$$

This can be seen as follows. Let X' be a random variable whose distribution is the same as the conditional distribution of X given $X \geq c$. Note that observations whose value is at least c arrive according to a Poisson process with rate $\lambda' := \mathrm{P}(X \geq c)$. Thus,

$$\begin{aligned} W_c(t) &= \mathrm{E}[X'_1\, \mathrm{I}(N'(t) \geq 1)] = (1 - e^{-\lambda' t})\, \mathrm{E}\, X' \\ &= \left[1 - e^{-t\,\mathrm{P}(X \geq c)}\right] \mathrm{E}(X | X \geq c), \end{aligned}$$

where $N'(t)$ is the number of arrivals up to time t in a Poisson process with rate λ'.

The next lemma is the key to the results in this section.

Lemma 3.1 *For $\gamma > 0$, the function*

$$f(x) := (1 - e^{-x})(1 + \gamma/x), \qquad x > 0$$

does not have a local minimum on $(0, \infty)$.

Proof. Since

$$f'(x) = e^{-x}(1 + \gamma/x) - (1 - e^{-x})\gamma/x^2$$

and

$$f''(x) = -e^{-x}(1 + \gamma/x + 2\gamma/x^2) + 2(1 - e^{-x})\gamma/x^3,$$

it follows that

$$f'(x) + f''(x) = (\gamma/x^3)\{2 - x - (2 + x)e^{-x}\}. \tag{9}$$

Let $g(x) = 2 - x - (2 + x)e^{-x}$. Then $g(0) = g'(0) = 0$, and $g''(x) < 0$ for all $x > 0$. Thus, (9) implies that $f'(x) + f''(x) < 0$ for all $x > 0$. But then there cannot exist a point $x_0 > 0$ such that $f'(x_0) = 0$ and $f''(x_0) \geq 0$. Since f is smooth, the lemma follows. $\Box$

Theorem 3.2 *For all $t > 0$,*

$$\frac{M(t)}{V(t)} < 2 - \frac{1 - e^{-t}}{t}. \tag{10}$$

Proof. The idea is to identify a suitable threshold c such that (10) holds with $W_c(t)$ in place of $V(t)$. For any $c \geq 0$,

$$\begin{aligned} \mathrm{E}[X_t^* \,\mathrm{I}(X_t^* < c)] &= \mathrm{E}[X_t^* \,\mathrm{I}(X_t^* < c, N(t) \geq 1)] \\ &\leq c\,\mathrm{P}(X_t^* < c, N(t) \geq 1) \\ &= c\{\mathrm{P}(X_t^* < c) - P(N(t) = 0)\} \\ &= c\,\mathrm{P}(X_t^* < c) - ce^{-t}, \end{aligned}$$

and

$$\mathrm{E}[X_t^* \,\mathrm{I}(X_t^* \geq c)] = c\,\mathrm{P}(X_t^* \geq c) + \mathrm{E}(X_t^* - c)^+.$$

Adding these expressions and applying Lemma 2.1 gives

$$\begin{aligned} \mathrm{E}X_t^* &\leq c(1 - e^{-t}) + \mathrm{E}(X_t^* - c)^+ \\ &\leq c(1 - e^{-t}) + t\,\mathrm{E}(X - c)^+. \end{aligned} \tag{11}$$

On the other hand, (8) can be written as

$$W_c(t) = \left[1 - e^{-t\,\mathrm{P}(X \geq c)}\right] \left\{ c + \frac{\mathrm{E}(X-c)^+}{\mathrm{P}(X \geq c)} \right\}. \tag{12}$$

Now for any $\alpha > 0$, there exists a unique number $c := c_\alpha$ such that

$$\mathrm{E}(X-c)^+ = c\alpha. \tag{13}$$

For this value of c, (11) and (12) combine to give

$$\frac{M(t)}{W_c(t)} \leq \frac{1 - e^{-t} + t\alpha}{(1-e^{-tp})(1+\alpha/p)}, \tag{14}$$

where $p := \mathrm{P}(X \geq c)$. Let $d(p)$ denote the denominator in the right hand side of (14). Using Lemma 3.1 with $x = tp$ and $\gamma = t\alpha$ we find that $d(p)$ is smallest either when $p = 0^+$ or when $p = 1$. Thus, noting that $\lim_{p \to 0} d(p) = t\alpha$,

$$\frac{M(t)}{W_c(t)} \leq \max \left\{ \frac{1 - e^{-t} + t\alpha}{t\alpha}, \frac{1 - e^{-t} + t\alpha}{(1-e^{-t})(1+\alpha)} \right\}. \tag{15}$$

The first term in the maximum is decreasing in α, and the second is increasing in α. Hence, the right hand side of (15) is minimized when $t\alpha = (1-e^{-t})(1+\alpha)$; that is, when

$$\alpha = \alpha^* := \frac{1-e^{-t}}{t + e^{-t} - 1}.$$

(Note that $\alpha^* > 0$, since $e^{-t} > 1 - t$.) For $\alpha = \alpha^*$, the maximum in (15) reduces to the right hand side of (10). Finally, the inequality is strict since (13) and $\alpha^* > 0$ imply that $\mathrm{P}(X \geq c_{\alpha^*}) > 0$, giving strict inequality in the proof of Lemma 2.1 (and hence, in (11)). □

Remark 3.3 The bound of Theorem 3.2 is sharp if the mortal is limited to the use of pure threshold rules. To see this, let $0 < p < 1$, $\varepsilon = \{1 - e^{-pt} - p(1-e^{-t})\}/(1-p)(1-e^{-t})$, and let X have the distribution $\mathrm{P}(X = 1) = p = 1 - \mathrm{P}(X = \varepsilon)$. Then there are only two essentially different threshold rules: $\tau(\varepsilon)$ and $\tau(1)$. By the choice of ε,

$$W_\varepsilon(t) = (1-e^{-t})\{p + \varepsilon(1-p)\} = 1 - e^{-tp} = W_1(t),$$

using (8). Furthermore, it is not difficult to compute that

$$M(t) = 1 - (1-\varepsilon)e^{-tp} - \varepsilon e^{-t}.$$

Thus,

$$\begin{aligned}\frac{M(t)}{\sup_c W_c(t)} &= \frac{M(t)}{W_1(t)} = 1 + \frac{\varepsilon(e^{-tp} - e^{-t})}{1 - e^{-tp}} \\ &= 1 + \frac{e^{-tp} - e^{-t}}{(1-p)(1-e^{-t})}\left(1 - \frac{p(1-e^{-t})}{1-e^{-tp}}\right) \\ &\to 2 - \frac{1-e^{-t}}{t} \qquad \text{as } p \downarrow 0.\end{aligned}$$

As a consequence, any attempt to improve on Theorem 3.2 will require the use of stopping rules which are more sophisticated than pure threshold rules.

The next result is a bound on the difference $M(t) - V(t)$. Assume that the random variables $X, X_1, X_2, \ldots$ are i.i.d. and $[a, b]$-valued, where $0 \leq a < b$. For any threshold c, define $D_c(t; X) := M(t; X) - W_c(t; X)$. For $t > 0$, let $h_t : [0, 1] \to \mathbb{R}$ be the function

$$h_t(x) = 1 - e^{-tx} - (1 - e^{-t})x.$$

Define $\gamma(t) := t/(1 - e^{-t})$, and $\beta(t) := 1 - \{1 + \log \gamma(t)\}/\gamma(t)$. Routine calculus shows that $\beta(t) = \max_{0 \leq x \leq 1} h_t(x)$, and the maximum is attained at $x = (\log \gamma(t))/t$.

The next theorem presents a universal value c^* which minimizes the largest possible difference $D_c(t; X)$ for $X \in \mathcal{X}_{[a,b]}$. Thus, the threshold c^* would be minimax if the distribution of X were completely unknown.

Theorem 3.4 *For all* $t > 0$*,*

$$\inf_c \sup_{X \in \mathcal{X}_{[a,b]}} D_c(t; X) = [b - \max\{a, c^*\}]\beta(t),$$

where $c^* = b\beta(t)/\{\beta(t) + 1 - e^{-t}\}$*. Moreover, the infimum is attained by the choice* $c = c^*$*.*

The proof of Theorem 3.4 uses the concept of balayage. Given an $[a, b]$-valued random variable X and constants $a \leq c < d \leq b$, let X_c^d denote a random variable such that $X_c^d = X$ if $X \notin [c, d]$, $X_c^d = c$ with probability $(d - c)^{-1}\,\mathrm{E}[(d - X)\,\mathrm{I}(c \leq X \leq d)]$, and $X_c^d = d$ otherwise. It follows immediately that $\mathrm{E}X_c^d = \mathrm{E}\,X$, $\mathrm{E}(X_c^d | X_c^d \geq c) = \mathrm{E}(X | X \geq c)$, and $\mathrm{P}(X_c^d \geq c) = \mathrm{P}(X \geq c)$. Moreover, Lemma 2.2 of Hill and Kertz [2] implies that if Y is a random variable independent of both X and X_c^d, then $\mathrm{E}(X_c^d \vee Y) \geq \mathrm{E}(X \vee Y)$.

Proof of Theorem 3.4. Let $c = c^*$, and assume first that $c^* > a$. Let X be any $[a,b]$-valued random variable, and define $\hat{X} = X_c^b$. By (8), $W_c(t;\hat{X}) = W_c(t;X)$, and by the last remark in the previous paragraph, $M(t;\hat{X}) \geq M(t;X)$. Therefore, $D_c(t;\hat{X}) \geq D_c(t;X)$. So, by replacing X with $\hat{X}$ if necessary, we may assume that $\mathrm{P}(c < X < b) = 0$.

Next, define $p := \mathrm{P}(X = b)$, and $r := \mathrm{P}(X \geq c)$. For $n \geq 1$,

$$
\begin{aligned}
\mathrm{E}(X_1 \vee \cdots \vee X_n) &\leq c\,\mathrm{P}(X_1 \vee \cdots \vee X_n \leq c) + b\,\mathrm{P}(X_1 \vee \cdots \vee X_n = b) \\
&= c(1-p)^n + b\{1-(1-p)^n\} = b - (b-c)(1-p)^n.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
M(t;X) &\leq \sum_{n=1}^{\infty} \{b - (b-c)(1-p)^n\}\,\mathrm{P}(N(t) = n) \\
&= (b-c)(1-e^{-tp}) + c(1-e^{-t}). \qquad (16)
\end{aligned}
$$

On the other hand, by (8),

$$
W_c(t;X) = (1-e^{-tr})\left(c + \frac{(b-c)p}{r}\right).
$$

Regarding p as fixed and r as variable, this last expression is minimized either when $r = p$ or $r = 1$, in view of Lemma 3.1. It follows that

$$
W_c(t;X) \geq \min\{(1-e^{-t})(c+(b-c)p), b(1-e^{-tp})\}.
$$

Subtracting from (16) and rearranging terms, we obtain that

$$
\begin{aligned}
D_c(t;X) &\leq \max\{(b-c)h_t(p), c(e^{-tp} - e^{-t})\} \\
&\leq \max\left\{(b-c)\beta(t), c(1-e^{-t})\right\}.
\end{aligned}
$$

The two terms inside the maximum are equal when $c = c^*$, and so

$$
D_{c^*}(t;X) \leq (b-c^*)\beta(t).
$$

Suppose next that $c^* \leq a$. Then clearly $D_{c^*}(t;X) = D_a(t;X)$, and the preceding argument (with $r = \mathrm{P}(X \geq a) = 1$) yields that

$$
D_{c^*}(t;X) = D_a(t;X) \leq (b-a)\beta(t).
$$

Conversely, for any $c \geq 0$ and $\varepsilon > 0$ the distribution of X can be chosen so that $D_c(t;X) \geq [b - \max\{a, c^*\}]\beta(t) - \varepsilon$:

(i) If $c \leq a$, take $X \in \{a, b\}$ with $\mathrm{P}(X = b) = (\log \gamma(t))/t$. Then $D_c(t;X) = (b-a)\beta(t)$.

(ii) If $a < c < c^*$, take $X \in \{c, b\}$ with $\mathrm{P}(X = b) = (\log \gamma(t))/t$. Then $D_c(t;X) = (b-c)\beta(t) > (b-c^*)\beta(t)$.

(iii) If $c \geq c^*$ and $c > a$, take $X \equiv c - \varepsilon$ (where ε is sufficiently small so that $c - \varepsilon > a$). Then $D_c(t;X) = (c-\varepsilon)(1-e^{-t}) \geq c^*(1-e^{-t}) - \varepsilon = (b-c^*)\beta(t) - \varepsilon$.

Thus, the choice c^* is minimax, and the theorem follows. □

Corollary 3.5 *If X is $[a, b]$-valued with $0 \leq a < b$, then for all $t > 0$,*

$$M(t;X) - V(t;X) \leq \min\left\{(b-a)\beta(t), \frac{b\beta(t)(1-e^{-t})}{\beta(t) + 1 - e^{-t}}\right\}. \tag{17}$$

Remark 3.6 For a non-homogeneous Poisson process with rate function $\lambda(x)$, the bounds corresponding to Theorems 3.2 and 3.4 are obtained by replacing t with $\mu(t) := \int_0^t \lambda(x)\, dx$.

4 How sharp are the bounds?

In this section, the ratio and difference of $M(t;X)$ and $V(t;X)$ are examined for random variables X taking only finitely many values, say $a_1 < a_2 < \cdots < a_n$, where $a_1 \geq 0$. Put $a_0 = 0$. For $i = 0, 1, \ldots, n$, define $r_i := \mathrm{P}(X \geq a_i)$, $\mu_i := \mathrm{E}(X - a_i)^+$, and $E_i := \mathrm{E}(X | X \geq a_i)$. Observe that $\mu_n = 0$, and recursively, for $k = n, n-1, \ldots, 1$,

$$\mu_{k-1} = \mathrm{E}(X - a_k)^+ + r_k(a_k - a_{k-1}) = \mu_k + r_k(a_k - a_{k-1}).$$

A moment's reflection reveals that there are critical times $0 < t_1^* < t_2^* < \cdots < t_{n-1}^* < \infty$ such that the optimal rule is to accept an observation with value a_i with time τ remaining if and only if $\tau \leq t_i^*$ or $i = n$. Set $t_0^* = 0$, and $t_n^* = \infty$. For $1 \leq k \leq n$ and $t \geq t_{k-1}^*$, let $V_k(t)$ denote the expected return, with time t remaining, from the rule:

Accept a_i with time τ remaining if and only if $\tau \leq t_i^*$ or $i \geq k$.

Clearly, it is optimal to accept a_k with time t remaining if and only if $a_k \geq V_k(t)$. Thus, t_k^* is the unique value of $t \geq t_{k-1}^*$ such that $V_k(t) = a_k$. For $k = 1$, we have

$$V_1(t) = (1 - e^{-t})E_1, \qquad t \geq 0,$$

so that

$$t_1^* = -\log(1 - (a_1/E_1)) = -\log(\mu_1/\mu_0) = \log(\mu_0/\mu_1).$$

And, inductively for $k = 2, \ldots, n-1$ and $t \geq t_{k-1}^*$,

$$\begin{aligned} V_k(t) &= \left(1 - e^{-r_k(t - t_{k-1}^*)}\right) E_k + e^{-r_k(t-t_{k-1}^*)} V_{k-1}(t_{k-1}^*) \\ &= E_k - (E_k - a_{k-1}) e^{-r_k(t - t_{k-1}^*)}. \end{aligned}$$

Thus,

$$e^{-r_k(t_k^* - t_{k-1}^*)} = \frac{E_k - a_k}{E_k - a_{k-1}} = \frac{r_k(E_k - a_k)}{r_k(E_k - a_{k-1})} = \frac{\mu_k}{\mu_{k-1}},$$

and so

$$t_k^* = t_{k-1}^* + (1/r_k)\log(\mu_{k-1}/\mu_k), \qquad k = 2, \ldots, n-1.$$

Finally, when $t_{k-1}^* \leq t \leq t_k^*$ $(k = 1, 2, \ldots, n)$,

$$V(t) = V_k(t) = E_k - (E_k - a_{k-1}) e^{-r_k(t - t_{k-1}^*)}. \tag{18}$$

On the other hand, the prophet's value is easily computed to be

$$M(t) = \sum_{i=1}^{n} (a_i - a_{i-1}) \left(1 - e^{-r_i t}\right), \qquad \text{for all } t \geq 0. \tag{19}$$

Example 4.1 Let $n = 2$, and put $a_1 = 1$ and $a_2 = K$, where K is large. Let a be a positive real number such that

$$\log(1 + t/a) < t, \tag{20}$$

and let $r_2 = a/(tK)$. We will examine the ratio $R(t) = M(t)/V(t)$ as $K \to \infty$. First, by (19),

$$M(t) = 1 - e^{-t} + (K-1)\left(1 - e^{-a/K}\right) \to 1 - e^{-t} + a, \qquad \text{as } K \to \infty. \tag{21}$$

Next, $\mu_1 = (a/tK)(K-1) \to a/t$, and $\mu_0 = \mu_1 + 1 \to (a/t) + 1$, so that

$$t_1^* = \log(\mu_0/\mu_1) \to \log(1 + t/a), \qquad K \to \infty.$$

It follows that $t_1^* < t$ for sufficiently large K, so by (18),

$$\begin{aligned} V(t) &= E_2 - (E_2 - a_1) e^{-r_2(t - t_1^*)} \\ &= (K-1)\left(1 - e^{-a(t - t_1^*)/tK}\right) + 1 \\ &\to a - (a/t)\log(1 + t/a) + 1, \qquad K \to \infty. \end{aligned} \tag{22}$$

Together, (21) and (22) yield that

$$R(t) \to \frac{a+1-e^{-t}}{a+1-(a/t)\log(1+t/a)}, \qquad \text{as } K \to \infty. \tag{23}$$

In particular, if $a = 1$, then (20) is met for every $t > 0$, and (23) becomes

$$R(t) \to \frac{2-e^{-t}}{2-\log(1+t)/t}. \tag{24}$$

Numerical experimentation suggests that, for the range $0 < t < 2$, this ratio is close to the maximum ratio over all two-valued random variables.

Example 4.2 Let $n = 3$, and put $a_i = K^{i-1}$ for $i = 1, 2, 3$, where K is again assumed to be large. Let a and b be positive real numbers such that $a < t$ and

$$\log(1 + a/b) < a, \tag{25}$$

and let $r_2 = a/t$, and $r_3 = b/(tK)$. Then

$$\begin{aligned} M(t) &= 1 - e^{-t} + (K-1)\left(1-e^{-a}\right) + K(K-1)\left(1-e^{-b/K}\right) \\ &\sim K\left(1-e^{-a}+b\right) \qquad \text{as } K \to \infty. \end{aligned}$$

Next, $\mu_2 = (b/tK)K(K-1) = b(K-1)/t$, $\mu_1 = \mu_2 + (a/t)(K-1) = (a+b)(K-1)/t$, and $\mu_0 = \mu_1 + 1$. Hence, $t_1^* = \log(\mu_0/\mu_1) = \log(1 + 1/\mu_1) \to 0$, and so

$$t_2^* = t_1^* + (t/a)\log(1 + a/b) \to (t/a)\log(1 + a/b).$$

It follows that $t_2^* < t$ when K is sufficiently large, and then

$$\begin{aligned} V(t) &= K(K-1)\left(1-e^{-b(t-t_2^*)/tK}\right) + K \\ &\sim K\left[b - (b/a)\log(1+a/b) + 1\right]. \end{aligned}$$

Thus,

$$R(t) \to \frac{1+b-e^{-a}}{1+b-(b/a)\log(1+a/b)} \qquad \text{as } K \to \infty.$$

In particular, if $a = 2$ and $b = 1$, then (25) is satisfied, and

$$R(t) \to \frac{2-e^{-2}}{2-(\log 3)/2} \doteq 1.28536, \qquad \text{for } t > 2.$$

Note that this is the same value obtained in (24) for $t = 2$. However, by admitting three-point distributions this ratio can be achieved for *any* $t \geq 2$.

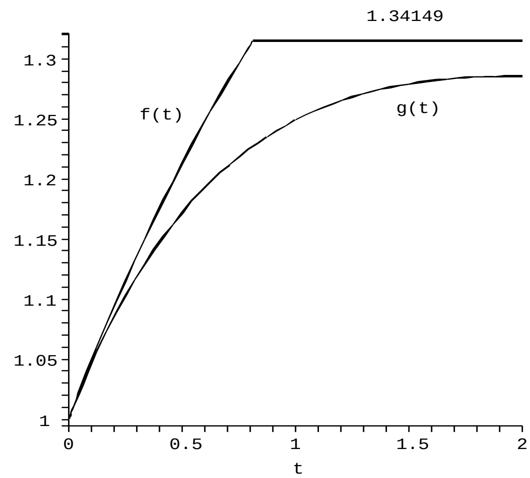

Figure 1: The theoretical best-possible ratio bound is between the two curves $g(t)$ and $\min\{f(t), 1.34149\}$, with $f(t) = 2 - (1 - e^{-t})/t$, and $g(t) = (2 - e^{-t})/\{2 - \log(1 + t)/t\}$.

Observe from Table 1 that the smallest n for which $a_n > 1.28536$ is $n = 8$. A ratio arbitrarily close to $a_8 = 1.29166$ can be obtained when $t \geq \log(7/\alpha_8) \doteq 3.1781$. For smaller values of t, however, Examples 4.1 and 4.2 provide larger ratios than the method discussed at the end of Section 2.

Example 4.2 shows that the bound of Theorem 3.2 is asymptotically sharp as $t \downarrow 0$ in the following sense. Let $f(t) = 2 - (1 - e^{-t})/t$, and $g(t) = (2 - e^{-t})/\{2 - \log(1 + t)/t\}$. That is, $g(t)$ is the right hand side of (24). By Example 4.2, the theoretical best-possible ratio bound is between $g(t)$ and $f(t)$. Straightforward calculations show that

$$f(t) - g(t) = O(t^2) \qquad \text{as } t \downarrow 0.$$

This relationship is illustrated in Figure 1, which also shows the uniform ratio bound from Theorem 2.2.

A similar comparison can be made for the difference bound of Corollary 3.5. Let $\hat{f}(t) = \beta(t)(1 - e^{-t})/(\beta(t) + 1 - e^{-t})$. That is, $\hat{f}(t)$ is the right hand side of (17) for $a = 0, b = 1$. Let X have the distribution given by $\mathrm{P}(X = 1) = 1/(e^t + 1) = 1 - \mathrm{P}(X = (1 - e^{-t})/2)$. Then $t_1^* = t$ exactly, and equations (18) and (19) yield

$$M(t) - V(t) = \frac{1}{2}\left[(1 + e^{-t})(1 - \exp\{-t/(e^t + 1)\}) - e^{-t}(1 - e^{-t})\right]. \tag{26}$$

Let $\hat{g}(t)$ denote the right hand side of (26). A straightforward calculation shows that

$$\hat{f}(t) - \hat{g}(t) = O(t^3) \qquad \text{as } t \downarrow 0.$$

Thus, the bound of Corollary 3.5 is asymptotically quite sharp as $t \downarrow 0$.

5 Other renewal processes: examples in discrete time

The main purpose of this section is to show that the conclusions of Theorems 2.2 and 2.3 may fail if the Poisson process governing the arrivals of observations is replaced by an arbitrary renewal process. The general setup is as follows. Let $T_1, T_2, \dots$ be i.i.d. random variables taking values in the positive integers, and assume that $X, X_1, X_2, \dots$ are i.i.d. nonnegative random variables, independent of the T_i. Call the random times $S_k = T_1 + \cdots + T_k$ ($k \in \mathbb{N}$) the *renewal times*, and assume that for each k, the random variable X_k is observed at time S_k. Put $S_0 = 0$. For $n \in \mathbb{N}$, let $N_n = \max\{k : S_k \le n\}$. In other words, N_n is the number of observations that arrive by time n. As before, we wish to compare the values $M_n := \mathrm{E}(\max\{X_1, \dots, X_{N_n}\})$ and $V_n := \sup_{\tau \in \mathcal{T}} \mathrm{E}[X_\tau \, \mathrm{I}(S_\tau \le n)]$, where $\mathcal{T}$ is the set of all $\mathbb{N}$-valued random variables (stopping rules) τ such that $\{\tau = i\}$ is measurable with respect to the σ-algebra generated by $X_1, \dots, X_i$ and $S_1, \dots, S_i$.

The problem will be easier to analyze if we represent it as follows. For each $j \in \mathbb{N}$, define

$$Y_j = \begin{cases} X_k, & \text{if } j = S_k \quad (k = 1, 2, \dots), \\ 0, & \text{otherwise.} \end{cases}$$

It is not difficult to see that $M_n = \mathrm{E}(Y_1 \vee \cdots \vee Y_n)$, and $V_n = \sup\{\mathrm{E}\, Y_\tau : \tau \text{ is a stopping rule for } Y_1, \dots, Y_n\}$. Thus, the problem is reduced to that of stopping an ordinary sequence of random variables, and standard methods can be applied to solve it.

Observe that the Y_j are, in general, neither independent nor identically distributed. However, there is one important exception.

Example 5.1 Let $0 < p < 1$, and assume that $\mathrm{P}(T_1 = k) = (1-p)^{k-1} p$ for $k = 1, 2, \dots$. This yields the *binomial process*, which has the property that the events $\{j$ *is a renewal time*$\}$, $j \in \mathbb{N}$, are mutually independent and have probability p. Since the X_i are i.i.d., this implies that the Y_j are i.i.d. with common distribution $pF + (1-p)\delta_{\{0\}}$, where F is the distribution of X, and $\delta_{\{0\}}$ denotes Dirac measure at zero. It follows from Theorem A and the above representations that $M_n \le a_n V_n$. Similarly, if X is $[0,1]$-valued then so is Y_1, and Theorem B implies that $M_n - V_n \le b_n$. The sharpness of these inequalities depends on the value of p: the first bound is sharp if $1 - p \le (\eta_{0,n}(\alpha_n))^{1/n}$, that is, if $p \ge (\alpha_n/(n-1))^{1/n}$. Likewise, the second bound is attained if $p \ge (\beta_n/(n-1))^{1/n}$. It is not clear how sharp the bounds are when p is smaller than the indicated values.

The next example shows that the best possible ratio and difference bounds are, in general, strictly greater than a_n and b_n.

Example 5.2 Fix $n \in \mathbb{N}$. Let $0 < p < 1$, and assume that $\mathrm{P}(T_1 = 1) = p = 1 - \mathrm{P}(T_1 = n)$. We compute V_n by backward induction. For $i = 1, 2, \ldots, n$, let γ_i denote the supremum, over all stopping times τ such that $i \le \tau \le n$, of $\mathrm{E}[Y_\tau | \, i \textit{ is a renewal time}]$. Then $\gamma_n = \mathrm{E}\, X$, and $\gamma_i = \mathrm{E}(X \vee p\gamma_{i+1})$ for $i = 1, 2, \ldots, n-1$, since if i is a renewal moment, then the next renewal moment is either $i+1$ or $i+n$, and $i+n$ is beyond the time horizon. Finally,

$$V_n = p\gamma_1 + (1-p)\gamma_n. \tag{27}$$

Now let X have a distribution on two points ε and 1, where $0 < \varepsilon < 1$, and the probability $\pi := \mathrm{P}(X = 1)$ is chosen so that

$$p\,\mathrm{E}\, X = \varepsilon. \tag{28}$$

It follows immediately that $\gamma_i = \mathrm{E}\, X$ for $i = 1, \ldots, n$, and hence, by (27), $V_n = \mathrm{E}\, X$. On the other hand,

$$\begin{aligned} M_n &= \mathrm{E}\, X_1 + \sum_{k=2}^{n} \mathrm{E}(X_2 \vee \cdots \vee X_k - X_1)^+ \, \mathrm{P}(N_n = k) \\ &= \mathrm{E}\, X + \sum_{k=2}^{n-1} p^k (1-p)(1-\varepsilon)(1-\pi)\{1 - (1-\pi)^{k-1}\} \\ &\qquad\qquad\qquad + p^n (1-\varepsilon)(1-\pi)\{1 - (1-\pi)^n\} \\ &= \mathrm{E}\, X + (1-\varepsilon)p\pi \left[\frac{1 - \{p(1-\pi)\}^n}{1 - p(1-\pi)} - 1 \right], \end{aligned}$$

where the last equality follows after routine simplification. Since $\mathrm{E}X = \pi + (1-\pi)\varepsilon$, (28) implies that $\varepsilon = p\pi/\{1-p(1-\pi)\}$, and $1-\varepsilon = (1-p)/\{1-p(1-\pi)\}$. Thus, we obtain the expressions

$$D_n := M_n - V_n = \frac{p(1-p)\pi}{1-p(1-\pi)}\left[\frac{1-\{p(1-\pi)\}^n}{1-p(1-\pi)} - 1\right]$$

and

$$R_n := \frac{M_n}{V_n} = 1 + p(1-p)\left[\frac{1-\{p(1-\pi)\}^n}{1-p(1-\pi)} - 1\right].$$

Now as $\pi \downarrow 0$, R_n increases to $1+p^2-p^{n+1}$, which is maximized for $p = (2/(n+1))^{1/(n-1)}$. It follows that R_n can be arbitrarily close to

$$c_n := 1 + \left(\frac{2}{n+1}\right)^{2/(n-1)}\left(\frac{n-1}{n+1}\right).$$

Observe that $\lim_{n\to\infty} c_n = 2$. Thus, the conclusion of Theorem 2.2 fails to hold for this case when n is sufficiently large. (In fact, $c_5 \doteq 1.3849 > 1+\alpha_0$.)

As for the difference D_n, note that

$$\lim_{n\to\infty} D_n = \frac{p^2(1-p)\pi(1-\pi)}{\{1-p(1-\pi)\}^2}. \tag{29}$$

For fixed p, this is maximized at $\pi = (1-p)/(2-p)$, and substituting this into (29) yields that $\lim_{n\to\infty} D_n = p^2/4$. Thus, if we choose n sufficiently large, p sufficiently close to 1, and $\pi = (1-p)/(2-p)$, then D_n will be arbitrarily close to $1/4$.

Note that it is not known whether $\limsup_{n\to\infty} b_n < 1/4$, though Table 1 suggests this should be the case. If this is true, then the conclusion of Theorem 2.3 too fails to hold for this example.

The last example raises an interesting question: do there exist a renewal process (in discrete or continuous time) and a distribution for X such that $M(t;X)/V(t;X) > 2$, or (if X is $[0,1]$-valued) $M(t;X) - V(t;X) > 1/4$? If not, why do these classical constants for the independent case appear as upper bounds in a problem concerning i.i.d. random variables? These questions will be addressed in a future paper.

References

[1] Chow, Y. S., Robbins, H., and Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping*, Houghton Mifflin, Boston.

[2] Hill, T. P. and Kertz, R. P. (1981). Additive comparisons of stop rule and supremum expectations of uniformly bounded independent random variables. *Proc. Amer. Math. Soc.* **83**, 582-585.

[3] Hill, T.P., Kertz, R.P. (1982). Comparisons of stop rule and supremum expectations of i.i.d. random variables. *Ann. Probab.* **10**, no. 2, 336-345.

[4] Karlin, S. (1962). Stochastic models and optimal policy for selling an asset. *Studies in applied probability and management science*, 148-158. Stanford Univ. Press.

[5] Kertz, R. P. (1986). Stop rule and supremum expectations of i.i.d. random variables: a complete comparison by conjugate duality. *J. Multivariate Anal.* **19**, 88-112.

[6] Krengel, U. and Sucheston, L. (1978). On semiamarts, amarts, and processes with finite value. *Probability on Banach Spaces*, Ed. by J. Kuelbs, Marcel Dekker, New York.

[7] Rinott, Y. and Samuel-Cahn, E. (1987). Comparisons of optimal stopping values and prophet inequalities for negatively dependent random variables. *Ann. Statist.* **15**, no. 4, 1482-1490.

[8] Sakaguchi, M. (1976). Optimal stopping problems for randomly arriving offers. *Math. Japonicae* **21**, 201-217.

[9] Samuel-Cahn, E. (1984). Comparison of threshold stop rules and maximum for independent nonnegative random variables. *Ann. Probab.* **12**, no. 4, 1213-1216.

[10] Samuel-Cahn, E. (1988). Prophet inequalities for threshold rules for independent bounded random variables. *Statistical decision theory and related topics* IV, Vol. 2 (West Lafayette, Ind., 1986), 177-182, Springer, New York.