Prophet regions for independent $[0, 1]$-valued random variables with random discounting

Pieter C. Allaart
University of North Texas *

November 16, 2004

Abstract

Let $X_1, X_2 \dots$ and $B_1, B_2 \dots$ be mutually independent $[0, 1]$-valued random variables, with $\mathrm{E}B_j = \beta > 0$ for all j. Let $Y_j = B_1 \cdots B_{j-1} X_j$ for $j \geq 1$. A complete comparison is made between the optimal stopping value $V(Y_1, \dots, Y_n) := \sup\{\mathrm{E}\,Y_\tau : \tau \text{ is a stopping rule for } Y_1, \dots, Y_n\}$ and $\mathrm{E}(\max_{1 \leq j \leq n} Y_j)$. It is shown that the set of ordered pairs $\{(x, y) : x = V(Y_1, \dots, Y_n), y = \mathrm{E}(\max_{1 \leq j \leq n} Y_j)$ for some sequence $Y_1, \dots, Y_n$ obtained as above$\}$ is precisely the set $\{(x, y) : 0 \leq x \leq 1, x \leq y \leq \Psi_{n,\beta}(x)\}$, where $\Psi_{n,\beta}(x) = [(1-\beta)n + 2\beta]x - \beta^{-(n-2)}x^2$ if $x \leq \beta^{n-1}$, and $\Psi_{n,\beta}(x) = \min_{j \geq 1}\{(1-\beta)jx + \beta^j\}$ otherwise. Sharp difference and ratio prophet inequalities are derived from this result, and an analogous comparison for infinite sequences is obtained.

1 Introduction

Suppose independent $[0, 1]$-valued random variables $X_1, \dots, X_n$ are observed sequentially, and the payoff for stopping at time j is $Y_j = B_1 \cdots B_{j-1} X_j$, where $B_1, B_2, \dots$ are $[0, 1]$-valued random variables with a common mean $\beta > 0$ that are independent of each other and of the sequence $\{X_j\}$. This paper aims to compare the values $\mathrm{E}(\max\{Y_1, \dots, Y_n\})$ and $V(Y_1, \dots, Y_n) := \sup\{\mathrm{E}\,Y_\tau : \tau \text{ is a stopping rule}\}$. Such comparisons have been called *prophet inequalities* in view of the natural interpretation of $\mathrm{E}(\max\{Y_1, \dots, Y_n\})$ as the optimal expected return of a prophet, or player with complete foresight. Prophet inequalities for independent random variables without discounting were first given by Krengel and Sucheston [10]. With $X_1, \dots, X_n$ as

*Mathematics Department, P.O. Box 311430, Denton, TX 76203-1430; e-mail: allaart@unt.edu

AMS 2000 subject classification: Primary 60G40, 60E15; Secondary 62L15.

Key words and phrases: Optimal stopping value, prophet inequality, inequalities for stochastic processes, conjugate duality.

above, they showed that $\mathrm{E}(\max\{X_1,\ldots,X_n\}) \le 2V(X_1,\ldots,X_n)$ and this bound is best possible. Simpler proofs of this and related inequalities were given by Hill and Kertz [6, 7], and Hill [5] subsequently generalized these results by showing that the *prophet region*

$$\{(x,y) : x = V(X_1,\ldots,X_n) \text{ and } y = \mathrm{E}(\max\{X_1,\ldots,X_n\}) \text{ for some sequence } X_1,\ldots,X_n \text{ of independent } [0,1]\text{-valued r.v.'s}\}$$

is exactly the set $\{(x,y) : 0 \le x \le 1, x \le y \le 2x - x^2\}$. Boshuizen [2] extended Hill's result to independent random variables with a constant discount factor β. He showed that for $0 < \beta \le 1$, the set

$$\{(x,y) : x = V(X_1,\beta X_2\ldots,\beta^{n-1}X_n) \text{ and } y = \mathrm{E}(\max\{X_1,\beta X_2,\ldots,\beta^{n-1}X_n\}) \text{ for some sequence } X_1,\ldots,X_n \text{ of independent } [0,1]\text{-valued r.v.'s}\}$$

is precisely the set $\{(x,y) : 0 \le x \le 1, x \le y \le \Phi_\beta(x)\}$, where

$$\Phi_\beta(x) = \begin{cases} 2x - x^2/\beta, & \text{if } x \le 1 - \sqrt{1-\beta} \\ 1 - (2/\beta)(\sqrt{1-\beta} - (1-\beta))(1-x), & \text{if } x > 1 - \sqrt{1-\beta}. \end{cases}$$

More recently, Saint-Mont [11, 12] has given several sharp inequalities for random variables with simultaneous costs and discountings. However, in these papers too the discount factors are deterministic, and indeed the present article seems to be the first work to address the prophet problem with random discounting.

This paper is organized as follows. Section 2 states the main result, a prophet region for a finite sequence of $[0,1]$-valued random variables with random discounting. As Figure 1 below shows, this region is considerably larger than Boshuizen's prophet region for the case of a constant discount factor β. From the prophet region, several sharp difference and ratio prophet inequalities are derived, some of which offer surprising discontinuities as $\beta \uparrow 1$. Taking limits of the prophet region as $n \to \infty$ yields a prophet region for the infinite-horizon case.

The proof of the main result is developed in sections 3 and 4. At the heart of the argument lies a reduction to fewer random variables in all cases except one particularly well-structured case. This reduction is carried out in section 3. The proof is completed in section 4 using conjugate duality. Section 5 briefly discusses a number of natural extensions.

2 Prophet regions

Throughout this paper, the following notation will be used. For subsets A of the underlying probability space, $I(A)$ denotes the indicator function of A. For real numbers x and y, $x \vee y$ denotes the maximum of x and y, and $x^+ := \max\{x, 0\}$. The symbol $\lfloor x \rfloor$ represents the greatest integer less than or equal to x, and $\lceil x \rceil$ is the least integer greater than or equal to x.

In this paper, it will always be assumed that $X_1, X_2, \ldots$ are independent $[0,1]$-valued random variables, and $B_1, B_2, \ldots$ are $[0,1]$-valued random variables having a common mean $\beta > 0$ that are independent of each other and of the sequence $\{X_j\}$. Given such sequences of random variables, define

$$Y_j = B_1 \cdots B_{j-1} X_j, \qquad j = 1, 2, \ldots, \tag{1}$$

where an empty product is taken to be 1.

For $j \in \mathbb{N}$, $\mathcal{F}_j$ is the σ-algebra generated by $\{X_1, \ldots, X_j\} \cup \{B_1, \ldots, B_{j-1}\}$, and $\mathcal{F}_0$ is the trivial σ-algebra. For a collection of random variables $\{X_s : s \in S\}$, $\operatorname{ess\,sup}\{X_s : s \in S\}$ denotes the essential supremum of the collection. (See pages 8-9 of [3] for the definition and existence of essential supremum.)

Let T be the set of all stopping rules with respect to the filtration $\{\mathcal{F}_j\}$. The *value* $V(Y_1, Y_2, \ldots)$ of $Y_1, Y_2, \ldots$ is defined by $V(Y_1, Y_2, \ldots) = \sup\{\mathrm{E}\, Y_t : t \in T\}$, the value of $Y_1, \ldots, Y_n$ is $V(Y_1, \ldots, Y_n) = \sup\{\mathrm{E}\, Y_t : t \in T, t \le n\}$, and the *conditional value* of $Y_m, Y_{m+1}, \ldots, Y_n$ given $\mathcal{F}_j$ is $V(Y_m, Y_{m+1}, \ldots, Y_n | \mathcal{F}_j) = \operatorname{ess\,sup}\{\mathrm{E}(Y_t | \mathcal{F}_j) : t \in T, m \le t \le n\}$.

Definition 2.1 *For $n \ge 1$ and $0 \le x \le 1$, let*

$$\Psi_{n,\beta}(x) = \begin{cases} [(1-\beta)n + 2\beta]x - \beta^{-(n-2)}x^2, & \text{if } x \le \beta^{n-1}, \\ (1-\beta)jx + \beta^j, & \text{if } \beta^j \le x \le \beta^{j-1},\ 1 \le j \le n-1. \end{cases}$$

Theorem 2.2 (Prophet region for finite sequences). *The set of points*

> $\{(x,y) : x = V(Y_1, \ldots, Y_n)$ *and* $y = \mathrm{E}(Y_1 \vee \cdots \vee Y_n)$ *for some sequence* $Y_1, \ldots, Y_n$ *of the form* (1)$\}$

is precisely the set $\{(x,y) : 0 \le x \le 1, x \le y \le \Psi_{n,\beta}(x)\}$.

The prophet region is illustrated in Figure 1 for small values of n. For the sake of comparison, the figure also includes Boshuizen's bound for independent random variables with a fixed discount factor.

The proof of Theorem 2.2 will be developed in the next two sections. Note that Hill's classical prophet region for independent $[0,1]$-valued random variables $X_1, \ldots, X_n$ (see [5]) follows from Theorem 2.2 by setting $\beta = 1$.

Corollary 2.3 (Difference prophet inequalities).

(i) For fixed n and fixed β,

$$\mathrm{E}(Y_1 \vee \cdots \vee Y_n) - V(Y_1, \ldots, Y_n) \le \begin{cases} \beta^{n-2}\{(1-\beta)(n-1) + \beta\}^2/4, & \beta \ge 1 - 1/n, \\ (1-\beta)\lfloor 1/(1-\beta) \rfloor \beta^{\lfloor 1/(1-\beta) \rfloor}, & \text{otherwise.} \end{cases}$$

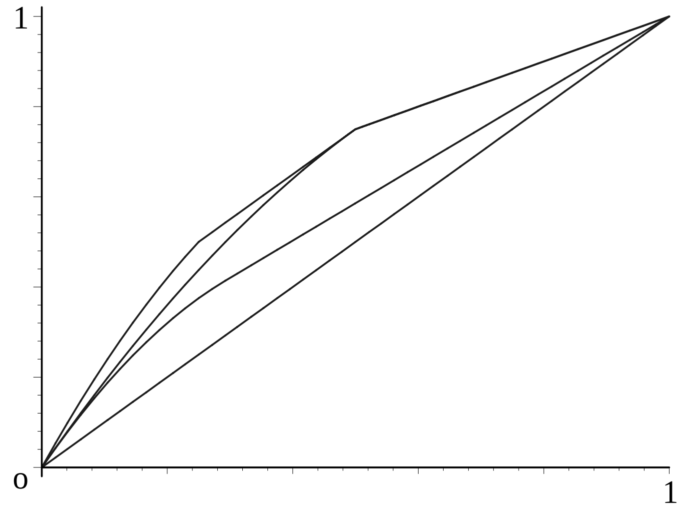

Figure 1: From the bottom, (i) the diagonal $y = x$, (ii) Boshuizen's bound $y = \Phi_{1/2}(x)$, (iii) $y = \Psi_{2,1/2}(x)$, (iv) $y = \Psi_{3,1/2}(x)$. Note that the functions $\Psi_{2,1/2}$ and $\Psi_{3,1/2}$ coincide on $1/2 \le x \le 1$.

(ii) For fixed β and for every n,

$$\mathrm{E}(Y_1 \vee \cdots \vee Y_n) - V(Y_1, \ldots, Y_n) \le \begin{cases} 1/4, & \beta = 1, \\ (1-\beta)\lfloor 1/(1-\beta)\rfloor \beta^{\lfloor 1/(1-\beta)\rfloor}, & \beta < 1. \end{cases}$$

(iii) For fixed n and every β,

$$\mathrm{E}(Y_1 \vee \cdots \vee Y_n) - V(Y_1, \ldots, Y_n) \le (1 - 1/n)^n.$$

(iv) For every n and every β,

$$\mathrm{E}(Y_1 \vee \cdots \vee Y_n) - V(Y_1, \ldots, Y_n) < 1/e.$$

All bounds are sharp, and the bounds in (i)-(iii) are attained.

It will be shown in the next section that the extremal case in Theorem 2.2 is when $B_1, B_2, \ldots$ are i.i.d. taking the values 0 and 1 only. Thus, Corollary 2.3 (iii) also gives the best possible bound for the case where $B_1, B_2, \ldots$ are i.i.d. (without a requirement on the mean).

Observe the discontinuity in the bound (ii): it is not hard to verify that $(1 - \beta)\lfloor 1/(1-\beta)\rfloor \beta^{\lfloor 1/(1-\beta)\rfloor} \to 1/e$ as $\beta \uparrow 1$, while the bound for $\beta = 1$ is $1/4$.

Corollary 2.4 (Ratio prophet inequalities).

(i) For fixed n and fixed β,

$$\mathrm{E}(Y_1 \vee \cdots \vee Y_n) < \{(1-\beta)n + 2\beta\} V(Y_1, \ldots, Y_n).$$

(ii) For fixed n and every β,

$$\mathrm{E}(Y_1 \vee \cdots \vee Y_n) < nV(Y_1, \ldots, Y_n).$$

Both bounds are sharp.

Taking limits in Theorem 2.2 as $n \to \infty$ gives the following result.

Theorem 2.5 (Propet region for infinite sequences). *The set of points (x, y) such that $x = V(Y_1, Y_2, \ldots)$ and $y = \mathrm{E}(\sup_n Y_n)$ for some sequence $Y_1, Y_2, \ldots$ of the form* (1) *is precisely the set $\{(x, y) : 0 \le x \le 1, x \le y \le \Psi_\beta(x)\}$, where*

$$\Psi_\beta(x) = \begin{cases} 2x - x^2, & \text{if } \beta = 1, \\ \min\limits_{j \ge 1} \{(1-\beta)jx + \beta^j\}, & \text{if } \beta < 1 \end{cases}$$

Again, the function $\beta \mapsto \Psi_\beta(x)$ is discontinuous at $\beta = 1$: if $\beta < 1$ we can write

$$\Psi_\beta(x) = (1-\beta)x\lceil \log_\beta x \rceil + \beta^{\lceil \log_\beta x \rceil},$$

from which it follows that $\lim_{\beta \uparrow 1} \Psi_\beta(x) = x - x \ln x$. Interestingly, this function was found in [8] as the upper boundary of a prophet region for infinite sequences of arbitrarily dependent $[0, 1]$-valued random variables.

As a consequence of Theorem 2.5, the difference inequalities (ii) and (iv) of Corollary 2.3 hold (and remain sharp) when $\mathrm{E}(Y_1 \vee \cdots \vee Y_n)$ and $V(Y_1, \ldots, Y_n)$ are replaced by $\mathrm{E}(\sup_n Y_n)$ and $V(Y_1, Y_2, \ldots)$, respectively.

3 Reduction to sequences with a simple structure

The proof of Theorem 2.2 consists of several steps. First, it is shown that in the extremal case, $B_1, B_2, \ldots$ may be assumed to be Bernoulli variables. The next step is to show that it is sufficient to consider random variables $X_1, \ldots, X_n$ such that $X_j \in \{a_j, 1\}$, where $(a_j : j = 1, \ldots, n)$ is a sequence of nonnegative numbers with $a_n = 0$. The main part of the proof consists in showing that the critical case is when the numbers $a_1, \ldots, a_{n-1}$ are nondecreasing. (See Proposition 3.7 below.) Finally, the upper boundary function $\Psi_{n,\beta}(x)$ is derived for this extremal case using the technique of conjugate duality.

The first lemma shows that $V(Y_1, \ldots, Y_n)$ depends on $B_1, \ldots, B_{n-1}$ only through their mean β.

Lemma 3.1
$$V(Y_1,\dots,Y_n)=V(X_1,\beta X_2,\dots,\beta^{n-1}X_n).$$

Proof. Follows easily by backward induction (see [3]) using the recursive relation

$$V(Y_j,\dots,Y_n|\mathcal{F}_{j-1})=\mathrm{E}[Y_j\vee V(Y_{j+1},\dots,Y_n|\mathcal{F}_j)|\mathcal{F}_{j-1}],\qquad 1\le j\le n,$$

and the mutual independence of $B_1,B_2,\dots$ and $X_1,X_2,\dots$. $\square$

The following definition is taken from [6].

Definition 3.2 *Given a $[0,1]$-valued random variable X and constants $0\le a<b\le 1$ let X_a^b, the* balayage *of X on $[a,b]$, denote a random variable such that $X_a^b=X$ if $X\notin[a,b]$, $X_a^b=a$ with probability $(b-a)^{-1}\,\mathrm{E}[(b-X)I(a\le X\le b)]$, and $X_a^b=b$ otherwise.*

Lemma 3.3 *Let $\hat{B}_1,\dots,\hat{B}_{n-1}$ be i.i.d. $\{0,1\}$-valued and independent of $X_1,\dots,X_n$ with $\mathrm{E}\,\hat{B}_1=\beta$, and let $\hat{Y}_j=\hat{B}_1\cdots\hat{B}_{j-1}X_j$ for $j\ge 1$. Then*

(i) $V(\hat{Y}_1,\dots,\hat{Y}_n)=V(Y_1,\dots,Y_n)$, *and*

(ii) $\mathrm{E}(\hat{Y}_1\vee\cdots\vee\hat{Y}_n)\ge\mathrm{E}(Y_1\vee\cdots\vee Y_n)$.

Proof. Statement (i) is an immediate consequence of Lemma 3.1, and (ii) follows easily from Lemma 2.2 of [6], using the fact that $\hat{B}_j=(B_j)_0^1$ for all $j=1,\dots,n-1$. $\square$

In view of Lemma 3.3, we can and will assume for the remainder of this section that $B_1,\dots,B_{n-1}$ take the values 0 and 1 only. It is then useful to define the random variable

$$N=\min\{1\le j\le n-1:B_j=0\}\qquad(\text{or}=n\text{ if no such }j\text{ exists}).$$

The relationship (1) can now be written as $Y_j=X_jI(N\ge j)$, and the value $V(Y_1,\dots,Y_n)$ can be interpreted as the optimal expected return of a player who sees a random number of observations $X_1,\dots,X_N$, where the reward for stopping after time N is zero. Observe that N has a truncated geometric distribution, and $\mathrm{P}(N\ge j+1|N\ge j)=\beta$ for $j=1,\dots,n-1$. *For the remainder of this paper, N will be a fixed random variable independent of $X_1,\dots,X_n$, having the truncated geometric distribution as described above.*

For prophet-like inequalities when the distribution of N is unrestricted, the reader is referred to the paper [1].

Definition 3.4 *For $j=1,\dots,n$, let $v_j:=\sup_{\tau\ge j}\mathrm{E}(Y_\tau|N\ge j-1)$. Let $v_{n+1}:=0$.*

Observe that $v_1 = \mathrm{E}(X_1 \vee v_2)$, while for $2 \leq j \leq n$,

$$v_j = \mathrm{P}(N \geq j | N \geq j-1)\, \mathrm{E}(X_j \vee v_{j+1}) = \beta\, \mathrm{E}(X_j \vee v_{j+1}).$$

Lemma 3.5 *Fix $j \in \{1, \dots, n\}$, let $\tilde{X}_j = (X_j)^1_{v_{j+1}} \vee v_{j+1}$ be independent of N, and let $\tilde{X}_i = X_i$ for $i \neq j$. Define $\tilde{Y}_i = \tilde{X}_i I(N \geq i)$ for $i = 1, \dots, n$. Then*

(i) $V(\tilde{Y}_1, \dots, \tilde{Y}_n) = V(Y_1, \dots, Y_n)$, *and*

(ii) $\mathrm{E}(\tilde{Y}_1 \vee \cdots \vee \tilde{Y}_n) \geq \mathrm{E}(Y_1 \vee \cdots \vee Y_n)$.

Proof. Statement (ii) follows from Lemma 2.2 of [6] and monotonicity. To see (i) in the case $j = 1$, observe that $V(\tilde{Y}_1, \dots, \tilde{Y}_n) = \mathrm{E}(\tilde{X}_1 \vee v_2) = \mathrm{E}(X_1 \vee v_2) = V(Y_1, \dots, Y_n)$. For the case $j \geq 2$, note that

$$\sup_{\tau \geq j} \mathrm{E}(\tilde{Y}_\tau | N \geq j-1) = \beta\, \mathrm{E}(\tilde{X}_j \vee v_{j+1}) = \beta\, \mathrm{E}(X_j \vee v_{j+1}) = v_j.$$

It thus follows inductively that $\sup_{\tau \geq i} \mathrm{E}(\tilde{Y}_\tau | N \geq i-1) = v_i$ for $i = 1, \dots, j$, and hence $V(\tilde{Y}_1, \dots, \tilde{Y}_n) = v_1 = V(Y_1, \dots, Y_n)$. $\Box$

In view of Lemmas 3.3 and 3.5, it is now possible to phrase the problem of maximizing $\mathrm{E}(Y_1 \vee \cdots \vee Y_n)$ for a given value of $V(Y_1, \dots, Y_n)$ as a constrained optimization problem with finitely many variables. However, since the expression for $\mathrm{E}(Y_1 \vee \cdots \vee Y_n)$ depends on the relative ordering of the variables $v_2, \dots, v_n$, a direct analytical solution of this optimization problem appears to be a daunting task. Instead, an entirely probabilistic argument will be given in Proposition 3.7 below to show that the critical case is when $v_2 \leq v_3 \leq \cdots \leq v_n$, leading to a manageable expression for $\mathrm{E}(Y_1 \vee \cdots \vee Y_n)$. The argument uses the following relationship.

Lemma 3.6 *Suppose $X_1, \dots, X_n$ satisfy $X_j \in \{v_{j+1}, 1\}$ for each $j = 1, \dots, n$. Then*

$$v_j = \mathrm{P}(Y_j \vee \cdots \vee Y_n = 1 | N \geq j-1), \qquad j = 1, \dots, n.$$

Proof. At each time j, it is optimal to stop if $Y_j = 1$, but we are indifferent between stopping and continuing if $Y_j = v_{j+1}$ (since this is the optimal expected return given the present situation if we continue), and also if $Y_j = 0$ (since we already missed the last observation, or the value of future observations is zero). Thus among the stopping rules that take at least j observations, the rule $\tau_1 = \min\{i \geq j : Y_i = 1\}$ (or $= n$ if no such i exists) is optimal on the event $\{N \geq j-1\}$, and

$$v_j = \mathrm{E}(Y_{\tau_1} | N \geq j-1) = \mathrm{P}(Y_j \vee \cdots \vee Y_n = 1 | N \geq j-1),$$

since $Y_{\tau_1} = 0$ on the event $\{Y_j \vee \cdots \vee Y_n < 1\}$ in view of the fact that $X_n \in \{0, 1\}$. $\Box$

Proposition 3.7 *Suppose $X_j \in \{v_{j+1}, 1\}$ for every j, and suppose there exist indices j_1 and j_2 with $1 < j_1 < j_2 \le n$ such that $v_{j_1} > v_{j_2}$. Then there exist an integer $m < n$ and random variables $X'_1, \ldots, X'_m$ independent of N, such that the random variables $Y'_j = X'_j I(N \ge j)$ $(j = 1, \ldots, n)$ satisfy:*

(i) $V(Y'_1, \ldots, Y'_m) = V(Y_1, \ldots, Y_n)$, *and*

(ii) $\mathrm{E}(Y'_1 \vee \cdots \vee Y'_m) \ge \mathrm{E}(Y_1 \vee \cdots \vee Y_n)$.

Proof. For brevity, let P_j and E_j denote conditional probability and expectation operators given that $N \ge j$. The construction of the sequence $X'_1, \ldots, X'_m$ depends on two cases.

Case 1. Suppose first that $\max\{v_2, \ldots, v_n\} > v_n$. Fix an integer $2 \le m < n$ such that $v_m = \max\{v_2, \ldots, v_n\}$. Define $X'_1 = X_1, \ldots, X'_{m-1} = X_{m-1}$, and $X'_m =$ the $\{0,1\}$-valued r.v., independent of N, with $\mathrm{E}\, X'_m = \mathrm{E}\, X_m$. Let $Y'_j = X'_j I(N \ge j)$ for $j = 1, \ldots, m$. Finally, define $v'_j = \sup_{\tau \ge j} \mathrm{E}(Y'_\tau | N \ge j-1)$. Statement (i) follows since $v'_m = \beta\, \mathrm{E}\, X'_m = \beta\, \mathrm{E}\, X_m = v_m$, so by backward induction $v'_j = v_j$ for $j = 1, \ldots, m-1$. To see (ii), observe that

$$
\begin{aligned}
\mathrm{E}_{m-1}(v_m \vee Y'_m - v_m \vee Y_m \vee \cdots \vee Y_n) &= \mathrm{E}_{m-1}[(Y'_m - v_m)^+ - (Y_m \vee \cdots \vee Y_n - v_m)^+] \\
&= (1 - v_m)[\mathrm{P}_{m-1}(Y'_m = 1) - \mathrm{P}_{m-1}(Y_m \vee \cdots \vee Y_n = 1)] \\
&= (1 - v_m)(v'_m - v_m) = 0,
\end{aligned}
$$

where the third equality follows using Lemma 3.6. Thus

$$
\begin{aligned}
&\mathrm{E}(Y'_1 \vee \cdots \vee Y'_m) - \mathrm{E}(Y_1 \vee \cdots \vee Y_n) \\
&\quad = \mathrm{E}[(Y_1 \vee \cdots \vee Y_{m-1} \vee Y'_m - Y_1 \vee \cdots \vee Y_n) I(Y_1 \vee \cdots \vee Y_{m-1} < 1, N \ge m-1)] \\
&\quad = \mathrm{E}_{m-1}(v_m \vee Y'_m - v_m \vee Y_m \vee \cdots \vee Y_n)\, \mathrm{P}(N \ge m-1, X_1 \vee \cdots \vee X_{m-1} < 1) \\
&\quad = 0.
\end{aligned}
$$

Case 2. If $\max\{v_2, \ldots, v_n\} = v_n$, then there exist integers k and l with $2 \le k \le l-2$ and $l \le n$ such that

$$
\max\{v_2, \ldots, v_{l-1}\} = v_k \le v_l \le v_{l+1} \le \cdots \le v_n. \tag{2}
$$

The basic idea is to replace the random variables $X_k, \ldots, X_{l-1}$ with a single random variable X'_k whose expectation equals that of X_k, but whose values are the values of X_{l-1}. This is possible since $\beta\, \mathrm{E}\, X_k = v_k \ge v_{l-1} = \beta\, \mathrm{E}\, X_{l-1} \ge \beta v_l$, and so $\mathrm{E}\, X_k \ge v_l$.

Let $m := n - (l - k - 1)$, and define $X'_1, \ldots, X'_m$ by $X'_1 = X_1, \ldots, X'_{k-1} = X_{k-1}$, $X'_k =$ the $\{v_l, 1\}$-valued r.v., independent of N, with $\mathrm{E}\, X'_k = \mathrm{E}\, X_k$, and $X'_j = X_{l-k+j-1}$ for $j = k+1, \ldots, m$. Define Y'_j and v'_j as in Case 1. To check that (i) is satisfied, note that by construction, $v'_{k+1} = v_l$, and thus

$$
v'_k = \beta\, \mathrm{E}(X'_k \vee v'_{k+1}) = \beta\, \mathrm{E}(X'_k \vee v_l) = \beta\, \mathrm{E}\, X'_k = \beta\, \mathrm{E}\, X_k = v_k.
$$

As in Case 1, it follows that $V(Y_1', \ldots, Y_m') = V(Y_1, \ldots, Y_n)$.

The verification of property (ii) requires more work. Note first that the conditions on X_k' imply that

$$\mathrm{P}(X_k' = v_l) = \frac{1 - v_{k+1}}{1 - v_l}\,\mathrm{P}(X_k = v_{k+1}). \tag{3}$$

We can write

$$\begin{aligned}
&\mathrm{E}(Y_1' \vee \cdots \vee Y_m') - \mathrm{E}(Y_1 \vee \cdots \vee Y_n) \\
&\quad = \mathrm{E}[(Y_1' \vee \cdots \vee Y_m' - Y_1 \vee \cdots \vee Y_n) I(N \geq k, Y_1 \vee \cdots \vee Y_{k-1} < 1)] \\
&\quad = \mathrm{E}[(v_k \vee Y_k' \vee \cdots \vee Y_m' - v_k \vee Y_k \vee \cdots \vee Y_n) I(N \geq k, X_1 \vee \cdots \vee X_{k-1} < 1)] \\
&\quad = \mathrm{E}_k[(1 - v_k \vee Y_k \vee \cdots \vee Y_n) - (1 - v_k \vee Y_k' \vee \cdots \vee Y_m')] \\
&\qquad\qquad \times \mathrm{P}(N \geq k, X_1 \vee \cdots \vee X_{k-1} < 1).
\end{aligned} \tag{4}$$

To show that the last expectation is nonnegative, let

$$p = \mathrm{P}_k(N \leq l-2, Y_k \vee \cdots \vee Y_{l-2} < 1),$$

and calculate

$$\begin{aligned}
&\mathrm{E}_k(1 - v_k \vee Y_k \vee \cdots \vee Y_n) \\
&\quad = p(1 - v_k) + \mathrm{P}_k(N \geq l-1, X_k \vee \cdots \vee X_{l-2} < 1)\,\mathrm{E}_{l-1}(1 - Y_{l-1} \vee \cdots \vee Y_n) \\
&\quad = p(1 - v_k) + (1 - v_l)\,\mathrm{P}_k(N \geq l-1, X_k \vee \cdots \vee X_{l-1} < 1) \\
&\qquad\qquad \times \frac{\mathrm{E}_{l-1}(1 - Y_{l-1} \vee \cdots \vee Y_n)}{\mathrm{P}(X_{l-1} < 1)(1 - v_l)} \\
&\quad = p(1 - v_k) + \mathrm{P}_k(N \geq l-1, Y_k \vee \cdots \vee Y_n < 1) \frac{\mathrm{E}_{l-1}(1 - Y_{l-1} \vee \cdots \vee Y_n)}{\mathrm{P}_{l-1}(Y_{l-1} \vee \cdots \vee Y_n < 1)},
\end{aligned} \tag{5}$$

where the last equality follows from Lemma 3.6. Next, let $N' = \min\{N, m\}$, and use the truncated geometric distribution of N to obtain

$$\begin{aligned}
\mathrm{E}_k(1 - v_k \vee Y_k' \vee \cdots \vee Y_m') &= \sum_{i=k}^{m} \mathrm{P}_k(N' = i)\,\mathrm{E}(1 - X_k' \vee \cdots \vee X_i') \\
&= \sum_{j=l-1}^{n} \mathrm{P}_k(N' = j - l + k + 1)\,\mathrm{E}(1 - X_k' \vee X_l \vee \cdots \vee X_j) \\
&= \sum_{j=l-1}^{n} \mathrm{P}_{l-1}(N = j)\,\mathrm{E}(1 - X_k' \vee X_l \vee \cdots \vee X_j).
\end{aligned}$$

Since X_k' and X_{l-1} are both $\{v_l, 1\}$-valued,

$$\mathrm{E}(1 - X_k' \vee X_l \vee \cdots \vee X_j) = \frac{\mathrm{P}(X_k' < 1)}{\mathrm{P}(X_{l-1} < 1)}\,\mathrm{E}(1 - X_{l-1} \vee X_l \vee \cdots \vee X_j),$$

and so

$$\begin{aligned}
\mathrm{E}_k(1 - v_k \vee Y_k' \vee \cdots \vee Y_m') &= \frac{\mathrm{P}(X_k' < 1)}{\mathrm{P}(X_{l-1} < 1)} \mathrm{E}_{l-1}(1 - Y_{l-1} \vee \cdots \vee Y_n) \\
&= (1 - v_{k+1}) \mathrm{P}(X_k < 1) \frac{\mathrm{E}_{l-1}(1 - Y_{l-1} \vee \cdots \vee Y_n)}{\mathrm{P}(X_{l-1} < 1)(1 - v_l)} \\
&= \mathrm{P}_k(Y_k \vee \cdots \vee Y_n < 1) \frac{\mathrm{E}_{l-1}(1 - Y_{l-1} \vee \cdots \vee Y_n)}{\mathrm{P}_{l-1}(Y_{l-1} \vee \cdots \vee Y_n < 1)}. \qquad (6)
\end{aligned}$$

Here the second equality follows from (3), and the last equality follows from Lemma 3.6. Subtracting (6) from (5) gives

$$\begin{aligned}
&\mathrm{E}_k(1 - v_k \vee Y_k \vee \cdots \vee Y_n) - \mathrm{E}_k(1 - v_k \vee Y_k' \vee \cdots \vee Y_m') \\
&\qquad = p \left[1 - v_k - \frac{\mathrm{E}_{l-1}(1 - Y_{l-1} \vee \cdots \vee Y_n)}{\mathrm{P}_{l-1}(Y_{l-1} \vee \cdots \vee Y_n < 1)} \right] \\
&\qquad = p \left[1 - v_k - \mathrm{E}_{l-1}(1 - Y_{l-1} \vee \cdots \vee Y_n | Y_{l-1} \vee \cdots \vee Y_n < 1) \right] \\
&\qquad \geq 0,
\end{aligned}$$

since the condition $N \geq l-1$ implies that $1 - Y_{l-1} \vee \cdots \vee Y_n \leq 1 - X_{l-1} \leq 1 - v_l \leq 1 - v_k$. This, together with the development (4), yields (ii). □

4 Proof of the main theorem

In order to avoid a constrained optimization problem, the method of conjugate duality will be used. Recall that a real-valued function g defined on an interval I is concave if $g(\lambda a + (1-\lambda)b) \geq \lambda g(a) + (1-\lambda) g(b)$ for all $a, b \in I$ and all $\lambda, 0 < \lambda < 1$. As a reference for the following definition and lemma see Chapter 4 of Stoer and Witzgall [13].

Definition 4.1 *Let g be a real-valued function defined on an interval $I \subset \mathbb{R}$. The concave conjugate function g^* of g is defined on the set $I^* = \{\gamma \in \mathbb{R} : \inf_{x \in I}[x\gamma - g(x)] > -\infty\}$ by $g^*(\gamma) = \inf_{x \in I}[x\gamma - g(x)]$.*

Lemma 4.2 *Let g be a concave function defined on an interval $I \subset \mathbb{R}$. Then*

(i) g^* is a concave function and I^* is an interval; and

(ii) $(g^*)^* = g$ and $(I^*)^* = I$ if the hypograph of g, $\{(r, x) \in \mathbb{R} \times I : r \leq g(x)\}$, is closed.

Definition 4.3 *The prophet region $\mathcal{R}_n$ is the set of ordered pairs $\{(x, y) : x = V(Y_1, \ldots, Y_n),\ y = \mathrm{E}(Y_1 \vee \cdots \vee Y_n), Y_j = X_j I(N \geq j)$, for some sequence of independent r.v.'s $X_1, \ldots, X_n\}$. The upper boundary function Γ_n is defined on $[0, 1]$ by $\Gamma_n(x) = \sup\{y : (x, y) \in \mathcal{R}_n\}$.*

Since it is not *a priori* clear that the function Γ_n is concave, mixtures of independent random variables are needed. Following Kertz [9], p. 94, say that a sequence of random variables $X_1, \ldots, X_n$ is a *mixture of independent r.v.'s* under a σ-algebra $\mathcal{G}$ if $X_1, \ldots, X_n$ are conditionally independent given $\mathcal{G}$; that is, $\mathrm{P}(X_1 \in B_1, \ldots, X_n \in B_n | \mathcal{G}) = \mathrm{P}(X_1 \in B_1 | \mathcal{G}) \cdots \mathrm{P}(X_n \in B_n | \mathcal{G})$ for all Borel sets $B_1, \ldots, B_n$ of $\mathbb{R}$. It is assumed that the randomization for a mixture takes place before the values of the random variables are obtained.

Definition 4.4 *Let $X_1, \ldots, X_n$ be a mixture of independent r.v.'s under $\mathcal{G}$, independent of N, and let $Y_j = X_j I(N \geq j)$ for $j = 1, \ldots, n$. A stop rule for $Y_1, \ldots, Y_n$ is a r.v. τ taking values in $\{1, \ldots, n\}$ such that $\{\tau = j\} \in \sigma(\mathcal{G}, Y_1, \ldots, Y_j)$ for each $j = 1, \ldots, n$. The value $V(Y_1, \ldots, Y_n)$ is defined by $V(Y_1, \ldots, Y_n) = \sup\{\mathrm{E}\, Y_\tau : \tau$ is a stop rule for $Y_1, \ldots, Y_n\}$.*

Definition 4.5 *The function Φ_n is defined on $[0,1]$ by $\Phi_n(x) = \sup\{\mathrm{E}(Y_1 \vee \cdots \vee Y_n) : V(Y_1, \ldots, Y_n) = x, Y_j = X_j I(N \geq j)$,* for some mixture $X_1, \ldots, X_n$ of independent r.v.'s that are independent of $N\}$.

It is clear from the above definitions that $\Gamma_n \leq \Phi_n$.

Lemma 4.6 *(i) Φ_n is a concave function with closed hypograph; (ii) $(\Phi_n^*)^* = \Phi_n$; and (iii) $\Phi_n^* = \Gamma_n^*$.*

Proof. Statement (i) follows as in Lemma 4.5 of [9]; (ii) follows from (i) and Lemma 4.2; and (iii) follows as in Proposition 4.6 of [9]. □

Proposition 4.7 *Let $\gamma \in \mathbb{R}$. Then*

$$\Gamma_n^*(\gamma) = \begin{cases} \gamma - 1, & \gamma \leq 1 - \beta, \\ \beta^{j-1}[\gamma - \beta - (1-\beta)j], & (1-\beta)(j-1) \leq \gamma \leq (1-\beta)j, \quad 2 \leq j \leq n, \\ -\beta^{n-2}[\gamma - 2\beta - (1-\beta)n]^2/4, & (1-\beta)n \leq \gamma \leq (1-\beta)n + 2\beta, \\ 0, & (1-\beta)n + 2\beta \leq \gamma. \end{cases}$$

Proof. By Lemma 3.5 and Proposition 3.7, we may assume that $X_j \in \{v_{j+1}, 1\}$ for $j = 1, \ldots, n-1$, $X_n \in \{0, 1\}$, and $v_2 \leq v_3 \leq \cdots \leq v_n$. For brevity, write $V = V(Y_1, \ldots, Y_n)$, $M = \mathrm{E}(Y_1 \vee \cdots \vee Y_n)$, and $q_j = \mathrm{P}(X_j < 1)$, $j = 1, \ldots, n$. The above assumptions and the truncated geometric distribution of N lead to the expression

$$\begin{aligned} M &= 1 - \sum_{j=1}^{n-1} (1-\beta)\beta^{j-1} q_1 \cdots q_j (1 - v_{j+1}) - \beta^{n-1} q_1 \cdots q_n (1 - v_n) \\ &= 1 - \sum_{j=1}^{n-2} (1-\beta)\beta^{j-1} q_1 \cdots q_j (1 - v_{j+1}) - \beta^{n-2} q_1 \cdots q_{n-1} (1 - v_n)^2, \end{aligned}$$

where the second equality follows since $v_n = \beta\,\mathrm{E}\,X_n = \beta(1-q_n)$. Similarly, $V = v_1 = \mathrm{E}\,X_1 = 1 - q_1(1-v_2)$, and so

$$\begin{aligned}\gamma V - M = \gamma\{1 - q_1(1-v_2)\} + \sum_{j=1}^{n-2}(1-\beta)\beta^{j-1}q_1\cdots q_j(1-v_{j+1}) \\ + \beta^{n-2}q_1\cdots q_{n-1}(1-v_n)^2 - 1.\end{aligned} \tag{7}$$

For each k with $2 \le k < n$, the relationship $v_k = \beta\,\mathrm{E}\,X_k = \beta\{1 - q_k(1-v_{k+1})\}$ implies that v_k depends linearly on each of the variables $q_k, \dots, q_{n-1}$, and does not depend on $q_1, \dots, q_{k-1}$. Hence $\gamma V - M$ depends linearly on each q_j, $j = 1, \dots, n-1$. In order to minimize $\gamma V - M$ it therefore suffices to consider values of $q_1, \dots, q_{n-1}$ in $\{0, 1\}$. There are now two cases.

Case 1. $q_k = 0$ for some $k < n$, and $q_j = 1$ for all $j < k$. Then $v_j = \beta v_{j+1}$ for $j = 2, \dots, k-1$, and $v_k = \beta$. Hence (7) simplifies to

$$\gamma V - M = \beta^{k-1}[\gamma - \beta - (1-\beta)k]. \tag{8}$$

Case 2. $q_j = 1$ for all $j < n$. Then

$$\gamma V - M = \beta^{n-2}v_n[\gamma + v_n - 2\beta - (1-\beta)n]. \tag{9}$$

This must be minimized over $v_n \in [0, \beta]$. A closer inspection reveals that the minimum value of (9) is attained at $v_n = v_n^*$, where

$$v_n^* = \begin{cases} \beta, & \gamma \le (1-\beta)n, \\ \beta - (\gamma - (1-\beta)n)/2, & (1-\beta)n \le \gamma \le (1-\beta)n + 2\beta, \\ 0, & (1-\beta)n + 2\beta \le \gamma. \end{cases} \tag{10}$$

If $v_n = \beta$, then (9) simplifies to $\gamma V - M = \beta^{n-1}[\gamma - \beta - (1-\beta)n]$, extending (8) to the value $k = n$. A comparison of the functions $f_k(\gamma) := \beta^{k-1}[\gamma - \beta - (1-\beta)k]$ shows that $f_{k+1}(\gamma) \le f_k(\gamma)$ if and only if $\gamma \ge (1-\beta)k$. Thus for $\gamma \le (1-\beta)n$, $\Gamma_n^*(\gamma) = \min_{1\le k\le n} f_k(\gamma) = f_{k_0}(\gamma)$, where k_0 is the largest integer k such that $\gamma \ge (1-\beta)k$, or $k_0 = 1$ if no such k exists.

On the other hand, if $\gamma \ge (1-\beta)n$ it may be checked that

$$-\beta^{n-2}[\gamma - 2\beta - (1-\beta)n]^2/4 \le f_n(\gamma), \tag{11}$$

since the two sides of this inequality are equal when $\gamma = (1-\beta)n$, and the derivative of the left side with respect to γ is no greater than the derivative of the right side. Since $f_n(\gamma) \le f_j(\gamma)$ for all $j < n$, it follows from (10) and (11) that $\Gamma_n^*(\gamma) = -\beta^{n-2}[\gamma - 2\beta - (1-\beta)n]^2/4$ when $(1-\beta)n \le \gamma \le (1-\beta)n + 2\beta$. Finally, if $\gamma \ge (1-\beta)n + 2\beta$, then $v_n^* = 0$ and $f_j(\gamma) \ge f_n(\gamma) > 0$ for all j, and hence $\Gamma_n^*(\gamma) = 0$, from (9). $\square$

Proposition 4.8 $\Phi_n = (\Gamma_n^*)^* = \Psi_{n,\beta}$.

Proof. That $\Phi_n = (\Gamma_n^*)^*$ follows from Lemma 4.6. To compute $(\Gamma_n^*)^*$, fix $x \in [0,1]$, and define a function $f(\gamma) := f_x(\gamma) := x\gamma - \Gamma_n^*(\gamma)$. Observe from Proposition 4.7 that f is convex, continuous, and piecewise differentiable with derivative

$$f'(\gamma) = \begin{cases} x-1, & \gamma < 1-\beta, \\ x-\beta^{j-1}, & (1-\beta)(j-1) < \gamma < (1-\beta)j, \\ & \qquad\qquad\qquad 2 \le j \le n, \\ x-\beta^{n-2}[(1-\beta)n+2\beta-\gamma]/2, & (1-\beta)n < \gamma < (1-\beta)n+2\beta, \\ x, & (1-\beta)n+2\beta < \gamma. \end{cases}$$

Thus if $x \le \beta^{n-1}$, f attains its minimum value at the point $\gamma^* = (1-\beta)n + 2\beta - 2\beta^{-(n-2)}x$, and $\Phi_n(x) = f(\gamma^*) = [(1-\beta)n+2\beta]x - \beta^{-(n-2)}x^2$. On the other hand, if $x > \beta^{n-1}$, then the minimum value of f occurs at the point $\gamma^* = (1-\beta)j^*$, where $j^* = \max\{j : x \le \beta^{j-1}\}$. Hence $\Phi_n(x) = f(\gamma^*) = (1-\beta)j^*x + b^{j^*}$. In both cases, $\Phi_n(x) = \Psi_{n,\beta}(x)$. $\Box$

Proof of Theorem 2.2. Propositions 3.3 and 4.8 imply that if $x = V(Y_1,\dots,Y_n)$ for some sequence $Y_1,\dots,Y_n$ satisfying (1), then $\mathrm{E}(Y_1 \vee \dots \vee Y_n) \le \Gamma_n(x) \le \Phi_n(x) = \Psi_{n,\beta}(x)$. Conversely, let (x,y) be any point with $0 \le x \le 1$ and $x \le y \le \Psi_{n,\beta}(x)$. Define sequences of random variables $X_1^{(1)},\dots,X_n^{(1)}$ and $X_1^{(2)},\dots,X_n^{(2)}$ by $X_1^{(1)} \equiv x$, $X_2^{(1)} \equiv 0,\dots,X_n^{(1)} \equiv 0$, $X_j^{(2)} \equiv \min\{x\beta^{-(j-1)},1\}$ for $j = 1,\dots,n-1$, and $X_n^{(2)} =$ the $\{0,1\}$-valued r.v with $\mathrm{E}\,X_n^{(2)} = \min\{x\beta^{-(n-1)},1\}$. Define $Y_j^{(i)} = X_j^{(i)} I(N \ge j)$ for $i = 1,2$ and $j = 1,\dots,n$. Then $V(Y_1^{(1)},\dots,Y_n^{(1)}) = V(Y_1^{(2)},\dots,Y_n^{(2)}) = \mathrm{E}(Y_1^{(1)} \vee \dots \vee Y_n^{(1)}) = x$, and $\mathrm{E}(Y_1^{(2)} \vee \dots \vee Y_n^{(2)}) = \Psi_{n,\beta}(x)$.

For $0 \le t \le 1$ and $1 \le j \le n$, let $X_j(t) = (1-t)X_j^{(1)} + tX_j^{(2)}$ and $Y_j(t) = X_j(t)I(N \ge j)$. Note that $X_j^{(i)} \ge \beta\,\mathrm{E}\,X_{j+1}^{(i)}$ for $i = 1,2$ and all j, and therefore $X_j(t) \ge \beta\,\mathrm{E}\,X_{j+1}(t)$ for all t and all j. It follows that $V(Y_1(t),\dots,Y_n(t)) = \mathrm{E}\,X_1(t) = x$. Finally, since $\mathrm{E}(Y_1(t) \vee \dots \vee Y_n(t))$ varies continuously from x at $t = 0$ to $\Psi_{n,\beta}(x)$ at $t = 1$, there is some $t_0 \in [0,1]$ such that $\mathrm{E}(Y_1(t_0) \vee \dots \vee Y_n(t_0)) = y$. $\Box$

5 Extensions and examples

1. *Scaling.* It is straightforward to extend the result of Theorem 2.2 to random variables $X_1,\dots,X_n$ taking values in an interval $[0,b]$ for $b > 0$: the prophet region simply gets scaled by a factor b. More precisely, the set of points (x,y) such that $x = V(Y_1,\dots,Y_n)$ and $y = \mathrm{E}(Y_1 \vee \dots \vee Y_n)$ for some sequence $Y_1,\dots,Y_n$ of the form (1), where $X_1,\dots,X_n \in [0,b]$, is exactly the set $\{(x,y) : 0 \le x \le b, x \le y \le b\Psi_{n,\beta}(x/b)\}$. The results of Corollary 2.3 and Theorem 2.5 are extended analogously.

On the other hand, there seems to be no obvious generalization to arbitrary intervals $[a, b]$, even when $a > 0$. The standard transformation by translation and scaling (e.g. [6, 9]) does not apply here, due to the multiplicative nature of the definition (1). An attempt to derive the prophet region from scratch appears to lead to tedious calculations, and this task is not pursued here.

2. *Fixed discount distributions.* A much more difficult problem arises when the discount factors $B_1, B_2, \ldots$ must come from *given* distributions with mean β. If $n = 2$ however, only one discount factor B is involved, and a sharp difference inequality can be established as follows.

Theorem 5.1 *Let X_1, X_2, and B be independent $[0,1]$-valued r.v.'s with $\mathrm{E}B = \beta$, and let $\phi(u) = \mathrm{E}(B-u)^+$ for $0 \le u \le 1$. Then*

$$\mathrm{E}(X_1 \vee BX_2) - V(X_1, BX_2) \le \sup_{x \in [0,1]} x\phi(\beta x),$$

and this bound is attained.

Proof. Using balayage and backward induction we may assume that, in the extremal case, $X_1 \equiv \mathrm{E}(BX_2) = \beta\,\mathrm{E}\,X_2$, and $X_2 \in \{0, 1\}$. Then

$$\begin{aligned}\mathrm{E}(X_1 \vee BX_2) - V(X_1, BX_2) &= \mathrm{E}(\beta\,\mathrm{E}\,X_2 \vee BX_2) - \beta\,\mathrm{E}\,X_2 = \mathrm{E}(BX_2 - \beta\,\mathrm{E}\,X_2)^+ \\ &= \mathrm{E}(B - \beta\,\mathrm{E}\,X_2)^+\,\mathrm{P}(X_2 = 1) = \phi(\beta\,\mathrm{E}\,X_2)\,\mathrm{E}\,X_2.\end{aligned}$$

Since $\mathrm{E}\,X_2$ can take on any value in $[0, 1]$, the theorem follows. □

The value $\sup_{x\in[0,1]} x\phi(\beta x)$ is easy to calculate for many common distributions of B. As an example, consider the uniform distribution for B. Then $\beta = 1/2$ and $\phi(u) = (1-u)^2/2$, and so $\sup_{x\in[0,1]} x\phi(\beta x) = \max_{x\in[0,1]}((x/2)(1-x/2)^2) = 4/27$.

Example 5.2 Suppose X_1, X_2 and B are independent, all having the uniform distribution on $(0, 1)$, so that $\beta = \frac{1}{2}$. Then

$$V(X_1, BX_2) = \mathrm{E}(X_1 \vee \beta\,\mathrm{E}\,X_2) = \mathrm{E}(X_1 \vee \tfrac{1}{4}) = \frac{17}{32},$$

and the value of $\mathrm{E}(X_1 \vee BX_2)$ is obtained as follows: by the independence of X_2 and B, $\mathrm{P}(BX_2 \le x) = x(1 - \ln x)$ for $0 < x \le 1$, and hence,

$$\mathrm{E}(X_1 \vee BX_2) = \int_0^1 \mathrm{P}(X_1 \vee BX_2 \ge x)\,dx = \int_0^1 [1 - x^2(1 - \ln x)]\,dx = \frac{5}{9}.$$

Thus, the prophet's advantage is $7/288 \approx 0.0243$, which is about one-sixth the bound from Theorem 5.1 for a uniform B.

3. *Discounts greater than one.* While discount factors in real life are usually between 0 and 1, it is conceivable (i.e. in times of deflation) that the discount factors are sometimes greater than 1. How do the prophet inequalities in this paper change in such cases? While refraining from a full study, we illustrate this difference for the special case where $n = 2$ and B_1 takes values in $[0, 2]$.

Proposition 5.3 *For all independent $[0, 1]$-valued random variables X_1 and X_2, and for every $[0, 2]$-valued random variable B with mean β which is independent of X_1 and X_2,*

$$\mathrm{E}(X_1 \vee BX_2) - V(X_1, BX_2) \le \begin{cases} \beta - \beta^2/2, & \beta \le 1, \\ 1/2, & \beta > 1. \end{cases}$$

and this bound is attained.

Proof. As before, we may assume that $B \in \{0, 2\}$ and $X_2 \in \{0, 1\}$. Consider two cases:

Case 1. If $\beta\,\mathrm{E}\,X_2 > 1$, then certainly $\beta > 1$, and

$$\begin{aligned} \mathrm{E}(X_1 \vee BX_2) - V(X_1, BX_2) &\le \mathrm{E}(1 \vee BX_2) - \beta\,\mathrm{E}\,X_2 = \mathrm{E}(1 - BX_2)^+ \\ &= 1 - \mathrm{P}(B = 2, X_2 = 1) = 1 - (\beta/2)\,\mathrm{E}\,X_2 \le 1/2. \end{aligned}$$

Case 2. If $\beta\,\mathrm{E}\,X_2 \le 1$, then we may assume additionally that $X_1 \in \{\beta\,\mathrm{E}\,X_2, 1\}$. Thus,

$$\begin{aligned} \mathrm{E}(X_1 \vee BX_2) - V(X_1, BX_2) &= \mathrm{E}(BX_2 - X_1)^+ \le \mathrm{E}(BX_2 - \beta\,\mathrm{E}\,X_2)^+ \\ &= (2 - \beta\,\mathrm{E}\,X_2)\,\mathrm{P}(B = 2, X_2 = 1) = (\beta/2)\,\mathrm{E}\,X_2(2 - \beta\,\mathrm{E}\,X_2). \end{aligned}$$

This last expression is always at most $1/2$, and is bounded above by $\beta - \beta^2/2$ when $\beta \le 1$, since $\mathrm{E}X_2 \le 1$.

The bound is attained by taking $B \in \{0, 2\}$ with mean β, and if $\beta \le 1$, by taking $X_1 \equiv \beta$ and $X_2 \equiv 1$; and if $\beta > 1$, by taking $X_1 \equiv 1$ and $X_2 \in \{0, 1\}$ with mean $1/\beta$. $\square$

It is interesting to compare the above bound with the corresponding bound from Corollary 2.3 (i): the latter evaluates to $\beta - \beta^2$ if $\beta \le 1/2$, and to $1/4$ if $\beta > 1/2$.

4. *Discounts with restricted variances.* From the point of view of applications, it seems worthwhile to consider models that restrict the fluctuations in the sequence $B_1, B_2 \ldots$. One way to do this is to let the B_j be independent with a given mean, and variances uniformly bounded by some fixed number σ^2. Exact prophet regions may be difficult to obtain for this model. But in the case $n = 2$, a fairly simple difference prophet inequality can be derived.

Theorem 5.4 *Let $0 < \beta \leq 1$ and $\sigma^2 \leq \beta(1-\beta)$. For all independent $[0,1]$-valued random variables X_1, X_2 and B, where B has mean β and variance bounded above by σ^2,*

$$\mathrm{E}(X_1 \vee BX_2) - V(X_1, BX_2) \leq \begin{cases} (\beta^2+\sigma^2)/4\beta, & \text{if } \sigma \leq \beta, \\ \beta\sigma^2/(\beta^2+\sigma^2), & \text{if } \sigma > \beta. \end{cases}$$

This bound is attained.

Proof. The key to the proof is the following result due to Heijnen and Goovaerts [4]. For any $[0,1]$-valued random variable B with mean β and variance σ^2, and for any number $a \in [0,1]$, the largest possible value of $\mathrm{E}(B \vee a)$ is $\psi(a)$, defined by

$$\psi(a) = \begin{cases} \beta + \frac{\sigma^2 a}{\beta^2+\sigma^2}, & a \leq \frac{\beta^2+\sigma^2}{2\beta}, \\ \frac{1}{2}[a + \beta + \sqrt{(a-\beta)^2+\sigma^2}], & \frac{\beta^2+\sigma^2}{2\beta} \leq a \leq \frac{1-\beta^2-\sigma^2}{2(1-\beta)}, \\ \frac{\sigma^2+(1-\beta^2)a}{\sigma^2+(1-\beta)^2}, & a \geq \frac{1-\beta^2-\sigma^2}{2(1-\beta)}. \end{cases} \tag{12}$$

By doing a balayage of B if necessary, we may assume that the variance of B is exactly σ^2. For brevity, write $V := V(X_1, BX_2)$, and $M := \mathrm{E}(X_1 \vee BX_2)$. As before, we have $V = \mathrm{E}(X_1 \vee \beta \,\mathrm{E}\, X_2)$, and, regardless of the distribution of B, the extremal case is when $X_1 \equiv \beta \,\mathrm{E}\, X_2$, and $X_2 \in \{0,1\}$. This implies that $V = \beta \,\mathrm{E}\, X_2$, and

$$\begin{aligned} M &= \mathrm{E}(\beta \,\mathrm{E}\, X_2 \vee B)\,\mathrm{P}(X_2 = 1) + \beta \,\mathrm{E}\, X_2 \,\mathrm{P}(X_2 = 0) \\ &\leq \psi(\beta \,\mathrm{E}\, X_2)\,\mathrm{E}\, X_2 + \beta \,\mathrm{E}\, X_2(1 - \mathrm{E}\, X_2). \end{aligned}$$

It follows that

$$M - V \leq \frac{1}{\beta}[a\psi(a) - a^2], \qquad \text{where } a = \beta \,\mathrm{E}\, X_2.$$

Consider the function $f(a) := a\psi(a) - a^2$. On the interval $0 < a < (\beta^2+\sigma^2)/2\beta$, we have

$$f'(a) = \beta - \frac{2\beta^2 a}{\beta^2+\sigma^2} > 0.$$

On the interval $(\beta^2+\sigma^2)/2\beta < a < (1-\beta^2-\sigma^2)/2(1-\beta)$, we have

$$f'(a) = \frac{1}{2}\left[\frac{(a-\beta)(2a-\beta)+\sigma^2}{\sqrt{(a-\beta)^2+\sigma^2}} - (2a-\beta)\right].$$

With some algebra, this can be seen to be nonpositive exactly when $a \geq (\beta^2+\sigma^2)/2\beta$.

Finally, on the interval $(1-\beta^2-\sigma^2)/2(1-\beta) < a < 1$, we have

$$f'(a) = \frac{\sigma^2}{\sigma^2+(1-\beta)^2}(1-2a) < 0,$$

since the inequality $\sigma^2 \leq \beta(1-\beta)$ implies that $a > 1/2$.

It follows that $f(a)$ is maximized at $a^* := (\beta^2 + \sigma^2)/2\beta$. However, a cannot exceed β, and so the cases $a^* \leq \beta$ and $a^* > \beta$ have to be considered separately. Note that $a^* \leq \beta$ if and only if $\sigma \leq \beta$. Thus, if $\sigma \leq \beta$,

$$M - V \leq \frac{f(a^*)}{\beta} = \frac{a^*}{\beta}[\psi(a^*) - a^*] = \frac{\beta^2 + \sigma^2}{4\beta},$$

while if $\sigma > \beta$,

$$M - V \leq \frac{f(\beta)}{\beta} = \psi(\beta) - \beta = \frac{\beta\sigma^2}{\beta^2 + \sigma^2}.$$

This proves the inequality in the theorem.

What do the extremal distributions look like? If $\sigma \leq \beta$, take $X_1 \equiv a^* = (\beta^2 + \sigma^2)/2\beta$; take $X_2 \in \{0,1\}$ with $\mathrm{P}(X_2 = 1) = a^*/\beta = (\beta^2 + \sigma^2)/2\beta^2$; and take for B the unique distribution with mean β and variance σ^2 on the points $\{0, \beta + \sigma^2/\beta\}$. If $\sigma > \beta$, take $X_1 \equiv \beta$, $X_2 \equiv 1$, and B as above. □

Getting the upper boundary function of the prophet region appears to require tedious calculations, since for fixed V, the maximum value of M is sometimes attained for a in the interior of the interval $[(\beta^2 + \sigma^2)/2\beta, (1 - \beta^2 - \sigma^2)/2(1-\beta)]$. The critical point is then the solution to a quartic equation. These calculations are omitted here.

It is also noted that the method of proof used above does not in any obvious way generalize to more random variables. Such generalization would require additional resources.

Acknowledgement

The author wishes to thank the referee for raising a question concerning discount factors outside the unit interval. This question led to Proposition 5.3.

References

[1] Allaart, P. C. (2004). An application of prophet regions to optimal stopping with a random number of observations. *Optimization* **53**, no. 4, 331-338.

[2] Boshuizen, F. A. (1991). Prophet region for independent random variables with a discount factor. *J. Multivariate Anal.* **37**, 76-84.

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

[4] Heijnen, B. and Goovaerts, M. J. (1989). Best upper bounds on risks altered by deductibles under incomplete information. *Scand. Actuarial J.* **1989**, 23-46.

[5] Hill, T. P. (1983). Prophet inequalities and order selection in optimal stopping problems. *Proc. Amer. Math. Soc.* **88**, 131-137.

[6] 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.

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

[8] Hill, T. P. and Kertz, R. P. (1983). Stop rule inequalities for uniformly bounded sequences of random variables. *Trans. Amer. Math. Soc.* **278**, 197-207.

[9] 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.

[10] 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.

[11] Saint-Mont, U. (1999) Prophet regions for independent random variables with simultaneous costs and discountings. *Sequential Anal.* **18**, no. 2, 119-132.

[12] Saint-Mont, U. (1999) Prophet regions for i.i.d. random variables with simultaneous costs and discountings. *Statistics and Decisions* **17**, 185-203.

[13] Stoer, J. and Witzgall, C. (1970). *Convexity and optimization in finite dimensions I.* Springer-Verlag, New York.