A PROPHET INEQUALITY FOR L^2-MARTINGALES

ADAM OSĘKOWSKI

Abstract. Let $k = (k_n)_{n\geq 1}$ be the sequence given by the conditions $k_1 = 0$ and $k_{n+1} = (1 + k_n^2)/2$, $n \geq 1$. We prove that for any L^2-martingale $X = (X_1, X_2, \ldots, X_n)$ we have

$$\mathbb{E} \max_{1\leq k\leq n} X_k \leq \sup_\tau \mathbb{E}X_\tau + k_n \max_{1\leq k\leq n} \sqrt{\operatorname{Var} X_k},$$

where the supremum on the right is taken over all stopping times τ of X which are bounded by n. Furthermore, it is shown that for each n, the constant k_n is the best possible.

1. Introduction

Let $X = (X_1, X_2, \ldots, X_n)$ be a sequence of random variables on some common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $X_n^* = \max_{1\leq k\leq n} X_k$ denote the one-sided maximal function of X. Furthermore, let $M_n = \mathbb{E}X_n^*$ and $V_n = \sup_\tau \mathbb{E}X_\tau$, where the latter supremum is taken over all stopping times τ of X (i.e., all τ adapted to the natural filtration of X). In the literature, comparisons between the numbers M_n and V_n (under various additional assumptions on X) have been called "prophet inequalities". This is due to the natural identification of M_n with the optimal expected return of a prophet or a player endowed with complete foresight; on the other hand, V_n can be regarded as an optimal expected return of a player who knows only past and present, but not the future. Prophet inequalities play a distinguished role in the theory of optimal stopping and have been studied intensively by many mathematicians. The literature on the subject is very large, and it would be impossible to review it here; thus we will content ourselves with a few examples. For example, the ratio prophet inequality of Krengel, Sucheston and Garling (see [8]) asserts that

$$M_n \leq 2V_n$$

for all sequences X of independent nonnegative random variables. Another result in this direction is that of Hill and Kertz [4], which states that

$$M_n - V_n \leq \frac{b-a}{4},$$

for all sequences X of independent random variables all taking values in a given finite interval $[a, b]$.

The motivation for the results of this note comes from the paper of Kennedy and Kertz [7], which contains the study of the following prophet inequality for independent random variables with finite variances.

2010 *Mathematics Subject Classification.* Primary: 60G40, 60E15. Secondary: 60G42.
Key words and phrases. Prophet inequality, optimal stopping, martingale.
Research partially supported by the NCN grant DEC-2012/05/B/ST1/00412.

Theorem 1.1. *For each $n \geq 1$, there exists a finite constant c_n such that for any sequence $X = (X_1, X_2, \ldots, X_n)$ of independent square-integrable random variables we have*

$$M_n \leq V_n + c_n\sqrt{n-1} \max_{1\leq k\leq n} \sqrt{\operatorname{Var} X_k}. \tag{1.1}$$

Actually, Kennedy and Kertz show the above bound with $c_n = 1/2$ and prove that $\liminf_{n\to\infty} c_n \geq \sqrt{\ln 2 - 1/2} = 0.439485\ldots$. Thus, due to the factor $\sqrt{n-1}$, we see that if n is large, the prophet can have a large advantage over the player. However, if one assumes additionally that X is a martingale, the situation is entirely different. Here is the second main result of [7] (see also [3] for a related statement).

Theorem 1.2. *For each $n \geq 1$, let k_n be the minimal constant such that for any martingale $X = (X_1, X_2, \ldots, X_n)$ we have*

$$M_n \leq V_n + k_n \max_{1\leq k\leq n} \sqrt{\operatorname{Var} X_k}. \tag{1.2}$$

Then the sequence k_1, k_2, $\ldots$ increases to 1.

The purpose of this paper is to determine the explicit formula for the constants k_n introduced in the above theorem. Quite interestingly, the definition of these numbers turns out to involve the logistic map, a simple recurrence relation which appears in many places in mathematics. Unfortunately, the recurrence cannot be given a compact formula, but we also provide a good approximation. Here is our main result.

Theorem 1.3. *The sequence $k = (k_n)_{n\geq 1}$ is given by the conditions $k_1 = 0$ and*

$$k_{n+1} = (1 + k_n^2)/2, \qquad n \geq 1.$$

Furthermore, we have $k_n = 1 - O(n^{-1})$; more precisely, $\frac{n}{2}(1 - k_n)$ increases to 1.

This should be compared to the following statement proved by Hill and Kertz [5]: for any martingale $X = (X_1, X_2, \ldots, X_n)$ with values in $[0, 1]$ we have the sharp bound

$$M_n \leq V_n + \left(\frac{n-1}{n}\right)^n.$$

Moreover, for $n \geq 2$, the constants (and hence also the extremal martingales) associated with this inequality and those arising in (1.2) are different.

This theorem will be established in the next section. Our approach rests on the theory of moments (see e.g. [1] and [6]).

2. Proof of Theorem 1.3

It will be convenient to work with the sequence $a = (a_n)_{n\geq 1}$ given by $a_1 = 0$ and the recurrence $a_{n+1} = a_n^2 + 1/4$ for $n \geq 1$. Note that the assertion of Theorem 1.3 can be equivalently stated as $k_n = 2a_n$ for $n \geq 1$. For the sake of clarity, we have decided to split this section into three separate parts.

2.1. **Proof of $k_n \leq 2a_n$.** To study the prophet inequality (1.1), let us introduce a family of certain special functions. Namely, for any nonnegative integer n and any real numbers $t \leq s$, define

$$\phi_n(s,t) = \inf \mathbb{E}\big[X_n^2 - (X_n^* \vee s)\big],$$

where the infimum is taken over all martingales $X = (X_1, X_2, \ldots, X_n)$ satisfying $X_1 = t$ almost surely. Clearly, we have $\phi_1(s,t) = t^2 - s$ and

$$\phi_2(s,t) = \inf \Big\{ \mathbb{E}\big[|t+X|^2 - \big((t+X) \vee s\big)\big] \,:\, \mathbb{E}X = 0 \Big\}. \tag{2.1}$$

Furthermore, conditioning on X_2, we get the recurrence

$$\phi_{n+1}(s,t) = \inf \Big\{ \mathbb{E}\phi_n((t+X) \vee s, t+X) \,:\, \mathbb{E}X = 0 \Big\} \tag{2.2}$$

for $n = 2, 3, \ldots$. Both (2.1) and (2.2) involve evaluating $\inf \mathbb{E}h(X)$ over all random variables X with $\mathbb{E}X = 0$, where h is a given function. This is a standard problem of the theory of moments and can be solved graphically as follows. The required infimum is given by the height, at location $x = 0$, of the lower boundary of the convex hull of the graph of h. See [1], [2] and [6] for more on the subject.

Keeping the above observations in mind, we turn to the explicit formula for ϕ_n.

Theorem 2.1. *For any $n = 1, 2, \ldots$ and any $t \leq s$ we have*

$$\phi_n(s,t) = \begin{cases} (2s - 2a_n)t - s - (s-a_n)^2 & \text{if } s - t < a_n, \\ t^2 - s & \text{if } s - t \geq a_n. \end{cases} \tag{2.3}$$

Proof. We proceed using induction. If $n = 1$, then the identity (2.3) holds true, since $a_1 = 0$. Suppose that (2.3) is valid for some nonnegative n and let us try to compute ϕ_{n+1} with the use of (2.2). To this end, introduce the function $h : \mathbb{R} \to \mathbb{R}$ by $h(x) = \phi_n((t+x) \vee s, t+x)$. A direct computation shows that h is given by

$$h(x) = \begin{cases} (t+x)^2 - s & \text{if } x \leq s - t - a_n, \\ (2s - 2a_n)(t+x) - (s-a_n)^2 - s & \text{if } s - t - a_n < x \leq s - t, \\ (t+x)^2 - (t+x) - a_n^2 & \text{if } x > s - t. \end{cases}$$

Let us describe the convex hull of the graph of h. We easily check that

- h is continuous,
- h is convex and of class C^1 on each of the intervals $(-\infty, s-t)$, $(s-t, \infty)$,
- its one-sided derivatives at $x = s - t$ satisfy $h'(s-t-) \geq h'(s-t+)$,
- h is linear on $(s - t - a_n, s - t)$.

Thus, we need to find a common tangent line to the parabolas $\gamma_1 : x \mapsto (t+x)^2 - s$ and $\gamma_2 : x \mapsto (t+x)^2 - (t+x) - a_n^2$. A little calculation gives that this line is

$$\{(x,y) : y = (2s - 2a_{n+1})(t+x) - s - (s - a_{n+1})^2\},$$

and the tangency points are

$$\begin{aligned} (x_1, \gamma_1(x)) &= \big(s - t - a_{n+1}, \gamma_1(s - t - a_{n+1})\big), \\ (x_2, \gamma_2(x_2)) &= \big(s - t - a_{n+1} + 1/2, \gamma_2(s - t - a_{n+1} + 1/2)\big). \end{aligned}$$

Therefore, the lower boundary Γ of the convex hull of the graph of h is the union of Γ_1, Γ_2 and Γ_3, where

$$\begin{aligned}
\Gamma_1 &= \{(x,y) : x \le x_1,\, y = (t+x)^2 - s\},\\
\Gamma_2 &= \{(x,y) : x_1 < x \le x_2,\, y = (2s-2a_{n+1})(t+x) - s - (s-a_{n+1})^2\},\\
\Gamma_3 &= \{(x,y) : x > x_2,\, y = (t+x)^2 - (t+x) - a_n^2\}.
\end{aligned}$$

Thus, by the above graphical interpretation of ϕ_{n+1} and the fact that x_2 is nonnegative, we obtain that $\phi_{n+1}(s,t) = t^2 - s$ if $x \le x_1$ and $\phi_{n+1}(s,t) = (2s-2a_{n+1})t - s - (s-a_{n+1})^2$ if $x > x_1$. This is precisely the claim. □

Now we will prove that for any L^2-bounded martingale $X = (X_1, X_2, \dots, X_n)$,

$$\mathbb{E}X_n^* \le \mathbb{E}X_1 + 2a_n\sqrt{\operatorname{Var} X_n}. \tag{2.4}$$

This is precisely (1.2) with the constant $2a_n$, because of the martingale property of X. To show the bound, let us introduce the modified centered martingale $\tilde{X} = (X_1 - \mathbb{E}X_1, X_2 - \mathbb{E}X_1, \dots, X_n - \mathbb{E}X_1)$. Applying the definition of ϕ_n conditionally with respect to $\tilde{X}_1$, we get

$$\mathbb{E}\big(\tilde{X}_n^2 - \tilde{X}_n^*\big) \ge \mathbb{E}\phi_n(\tilde{X}_1, \tilde{X}_1) = \mathbb{E}\tilde{X}_1^2 - \mathbb{E}\tilde{X}_1 - a_n^2.$$

However, $\tilde{X}_1$ has expectation 0, so the latter expression is not smaller than $-a_n^2$. Consequently,

$$\mathbb{E}X_n^* - \mathbb{E}X_1 = \mathbb{E}\tilde{X}_n^* \le \mathbb{E}\tilde{X}_n^2 + a_n^2 = \operatorname{Var} X_n + a_n^2.$$

Applying this inequality to the rescaled martingale X/λ (where λ is a fixed positive constant), we obtain

$$\mathbb{E}X_n^* - \mathbb{E}X_1 \le \lambda^{-1}\operatorname{Var} X_n + \lambda a_n^2.$$

The right-hand side, as a function of λ, attains its minimum for the choice $\lambda = \big(\operatorname{Var} X_n\big)^{1/2}/a_n$. Plugging this value of λ above, we obtain the desired estimate (2.4).

2.2. **Proof of $k_n \ge 2a_n$.** Let us now describe the examples which yield the sharpness of (1.2). Let $n \ge 2$ be a fixed integer and consider the sequence $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ of independent mean-zero random variables with the distributions uniquely determined by $\mathbb{P}(\xi_1 = 0) = 1$ and $\xi_k \in \{-a_{n+2-k}, 1/2 - a_{n+2-k}\}$ for $k = 2, 3, \dots, n$. Introduce the stopping time $\tau = \inf\{k : \xi_k < 0\}$ (with the convention $\inf\emptyset = \infty$) and the sequence $X = (X_1, X_2, \dots, X_n)$ given by

$$X_k = \xi_1 + \xi_2 + \ldots + \xi_{\tau\wedge k}.$$

Since the variables $\xi_1, \xi_2, \dots, \xi_n$ are independent and centered, Doob's optional sampling theorem implies that X is a mean-zero martingale. Directly from the definition, we derive that for $k = 2, 3, \dots, n$,

$$\begin{aligned}
&\mathbb{P}\left(X_n = \frac{n-k}{2} - a_n - a_{n-1} - \dots - a_k\right)\\
&= \mathbb{P}\left(X_n^* = \frac{n-k}{2} - a_n - a_{n-1} - \dots - a_{k+1}\right)\\
&= \mathbb{P}(\tau = n+2-k)\\
&= \mathbb{P}(\xi_2 \geq 0,\, \xi_3 \geq 0,\, \dots,\, \xi_{n+1-k} \geq 0,\, \xi_{n+2-k} < 0)\\
&= 2^{n-k} a_n a_{n-1} \dots a_{k+1}(1 - 2a_{k-1})
\end{aligned}$$

(if $k = n$, the expression in the second line is understood to be $\mathbb{P}(X_n^* = 0)$) and, similarly,

$$\begin{aligned}
&\mathbb{P}\left(X_n = \frac{n-1}{2} - a_n - a_{n-1} - \dots - a_2\right)\\
&= \mathbb{P}\left(X_n^* = \frac{n-1}{2} - a_n - a_{n-1} - \dots - a_2\right)\\
&= \mathbb{P}(\xi_k \geq 0 \text{ for all } k)\\
&= 2^{n-1} a_n a_{n-1} \dots a_2.
\end{aligned}$$

Therefore, we derive that

$$\begin{aligned}
\mathbb{E}X_n^* = &\sum_{k=2}^{n} \left(\frac{n-k}{2} - a_n - a_{n-1} - \dots - a_{k+1}\right) \cdot 2^{n-k} a_n a_{n-1} \dots a_{k+1}(1 - 2a_{k-1})\\
&+ \left(\frac{n-1}{2} - a_n - a_{n-1} - \dots - a_2\right) \cdot 2^{n-1} a_n a_{n-1} \dots a_2
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{Var} X_n = &\sum_{k=2}^{n} \left(\frac{n-k}{2} - a_n - a_{n-1} - \dots - a_k\right)^2 \cdot 2^{n-k} a_n a_{n-1} \dots a_{k+1}(1 - 2a_{k-1})\\
&+ \left(\frac{n-1}{2} - a_n - a_{n-1} - \dots - a_2\right)^2 \cdot 2^{n-1} a_n a_{n-1} \dots a_2.
\end{aligned}$$

Some tedious, but rather straightforward calculations show the following two recurrence relations for $\mathbb{E}X_n^*$ and $\operatorname{Var} X_n$:

$$\mathbb{E}X_{n+1}^* = 2a_{n+1}(\mathbb{E}X_n^* - a_{n+1} + 1/2) \tag{2.5}$$

and

$$\operatorname{Var} X_{n+1} = 2a_{n+1} \operatorname{Var} X_n + a_{n+1}(1 - 2a_{n+1})/2. \tag{2.6}$$

These two equations imply that $\mathbb{E}X_n^* = 2a_n^2$ and $\operatorname{Var} X_n = a_n^2$ for all n. This is clear for $n = 1$, and for larger n we use an easy induction. Consequently, $\mathbb{E}X_n^* = 2a_n\big(\operatorname{Var} X_n\big)^{1/2}$ and we are done.

2.3. **Asymptotics.** We turn to the final part of Theorem 1.3, concerning the approximation of the size of k_n. It will be more convenient to work with the sequence $b = (b_n)_{n\geq 1}$ given by $b_n = (1 - k_n)/2$; note that $b_1 = 1/2$ and $b_{n+1} = b_n - b_n^2$ for $n \geq 1$. We will be done if we establish the following.

Lemma 2.2. *The sequence b satisfies the following.*
(i) We have $b_n \leq 1/(n+1)$ for all $n \geq 1$.
(ii) The sequence $(nb_n)_{n\geq 1}$ increases to 1.

Proof. (i) This follows by straightforward induction.
(ii) By the recursion defining b and the condition (i) proved above, we have

$$\frac{(n+1)b_{n+1}}{nb_n} = \frac{(n+1)(1-b_n)}{n} \geq 1.$$

Therefore, the sequence $(nb_n)_{n\geq 1}$ is nondecreasing and bounded by 1, so it converges to a certain limit $g \in [0,1]$. In particular, we get that $b_n \leq g/n$ for all n. Thus, for any nonnegative integers n, m such that $n > m$, we have

$$\begin{aligned} b_n = b_{n-1}(1-b_{n-1}) &= b_{n-2}(1-b_{n-2})(1-b_{n-1}) = \dots \\ &= b_m(1-b_m)(1-b_{m+1})\dots(1-b_{n-1}) \\ &\geq b_m\left(1-\frac{g}{m}\right)\left(1-\frac{g}{m+1}\right)\dots\left(1-\frac{g}{n-1}\right). \end{aligned}$$

Suppose that $g < 1$ and pick a number $\eta \in (g,1)$. Then, if k is sufficiently large, we have $1 - g/k \geq \exp(-\eta/k)$ and hence, if the above numbers m, n are large enough, we get

$$b_n \geq \exp\left(-\eta\left(\frac{1}{m}+\frac{1}{m+1}+\dots+\frac{1}{n-1}\right)\right) \geq \exp(-\eta\ln(n-1))\cdot c$$

for some constant $c > 0$ depending on m and η, but not on n. This implies

$$nb_n \geq \frac{nc}{(n-1)^\eta} \xrightarrow{n\to\infty} \infty,$$

a contradiction. Thus $g = 1$ and the proof is complete. □

References

[1] D. C. Cox, *The best constant in Burkholder's weak-L^1 inequality for the martingale square function*, Proc. Amer. Math. Soc. **85** (1982), pp. 427–433.

[2] D. C. Cox, *Some sharp martingale inequalities related to Doob's inequality*, Inequalities in Statistics and Probability, IMS Lecture Notes-Monograph Series **5** (1984), pp. 78–83.

[3] L. E. Dubins and G. Schwarz, *A sharp inequality for sub-martingales and stopping times*, Soc. Math. de France, Astérisque **157-158** (1988), pp. 129–145.

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

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

[6] J. H. B. Kemperman, *The general moment problem, a geometric approach*, Ann. Math. Statist. **39** (1968), pp. 93–122.

[7] D. P. Kennedy and R. P. Kertz, *A prophet inequality for independent random variables with finite variances*, J. Appl. Probab. **34** (1997), pp. 945–958.

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

Department of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland

E-mail address: ados@mimuw.edu.pl