

Wald for non-stopping times: The rewards of impatient prophets

Alexander E. Holroyd* Yuval Peres† Jeffrey E. Steif‡

Abstract

Let X_1, X_2, \dots be independent identically distributed nonnegative random variables. Wald's identity states that the random sum $S_T := X_1 + \dots + X_T$ has expectation $\mathbb{E}T \cdot \mathbb{E}X_1$ provided T is a stopping time. We prove here that for any $1 < \alpha \leq 2$, if T is an arbitrary nonnegative random variable, then S_T has finite expectation provided that X_1 has finite α -moment and T has finite $1/(\alpha - 1)$ -moment. We also prove a variant in which T is assumed to have a finite exponential moment. These moment conditions are sharp in the sense that for any i.i.d. sequence X_i violating them, there is a T satisfying the given condition for which S_T (and, in fact, X_T) has infinite expectation.

An interpretation is given in terms of a prophet being more rewarded than a gambler when a certain *impatience* restriction is imposed.

Keywords: Wald's identity; stopping time; moment condition; prophet inequality.

AMS MSC 2010: 60G50.

Submitted to ECP on 14 June 2014, final version accepted on 11 November 2014.

1 Introduction

Let X_1, X_2, \dots be independent identically distributed (i.i.d.) nonnegative random variables, and let T be a nonnegative integer-valued random variable. Write $S_n = \sum_{i=1}^n X_i$ and $X = X_1$. Wald's identity [14] states that if T is a *stopping time* (which is to say that for each n , the event $\{T = n\}$ lies in the σ -field generated by X_1, \dots, X_n), then

$$\mathbb{E}S_T = \mathbb{E}T \cdot \mathbb{E}X. \tag{1.1}$$

In particular, if X and T have finite mean then so does S_T .

It is natural to ask whether similar conclusions can be obtained if we drop the requirement that T be a stopping time. It is too much to hope that the equality (1.1) still holds. (For example, suppose that X_i takes values 0, 1 with equal probabilities, and let T be 1 if $X_2 = 0$ and otherwise 2. Then $\mathbb{E}S_T = 1 \neq \frac{3}{2} \cdot \frac{1}{2} = \mathbb{E}T \cdot \mathbb{E}X$.) However, one may still ask when S_T has finite mean. It turns out that finite means of X and T no longer suffice, but stronger moment conditions do. Our main result gives sharp moment conditions for this conclusion to hold. In addition, when the moment conditions fail, with a suitably chosen T we can arrange that even the final summand X_T has infinite mean. Here is the precise statement. (If $T = 0$ we take by convention $X_T = 0$).

*Microsoft Research, USA. E-mail: holroyd@microsoft.com

†Microsoft Research, USA. E-mail: peres@microsoft.com

‡Chalmers University of Technology and Göteborg University, Sweden. E-mail: steif@math.chalmers.se

Theorem 1.1. *Let X_1, X_2, \dots be i.i.d. nonnegative random variables, and write $S_n := \sum_{i=1}^n X_i$ and $X = X_1$. For each $\alpha \in (1, 2]$, the following are equivalent.*

- (i) $\mathbb{E}X^\alpha < \infty$.
- (ii) *For every nonnegative integer-valued random variable T satisfying $\mathbb{E}T^{1/(\alpha-1)} < \infty$, we have $\mathbb{E}S_T < \infty$.*
- (iii) *For every nonnegative integer-valued random variable T satisfying $\mathbb{E}T^{1/(\alpha-1)} < \infty$, we have $\mathbb{E}X_T < \infty$.*

The special case $\alpha = 2$ of Theorem 1.1 is particularly natural: then the condition on X in (i) is that it have finite variance, and the condition on T in (ii) and (iii) is that it have finite mean. At the other extreme, as $\alpha \downarrow 1$, (ii) and (iii) require successively higher moments of T to be finite. One may ask what happens when T satisfies an even stronger condition such as a finite exponential moment – what condition must we impose on X , if we are to conclude $\mathbb{E}S_T < \infty$? The following provides an answer, in which, moreover, the independence assumption may be relaxed.

Theorem 1.2. *Let X_1, X_2, \dots be i.i.d. nonnegative random variables, and write $S_n := \sum_{i=1}^n X_i$ and $X = X_1$. The following are equivalent.*

- (i) $\mathbb{E}[X(\log X)^+] < \infty$.
- (ii) *For every nonnegative integer-valued random variable T satisfying $\mathbb{E}e^{cT} < \infty$ for some $c > 0$, we have $\mathbb{E}S_T < \infty$.*
- (iii) *For every nonnegative integer-valued random variable T satisfying $\mathbb{E}e^{cT} < \infty$ for some $c > 0$, we have $\mathbb{E}X_T < \infty$.*

Moreover, if X_1, X_2, \dots are assumed identically distributed but not necessarily independent, then (i) and (ii) are equivalent.

On the other hand, in the following variant of Theorem 1.1, dropping independence results in a different moment condition for T .

Proposition 1.3. *Let X be a nonnegative random variable. For each $\alpha \in (1, 2]$, the following are equivalent.*

- (i) $\mathbb{E}X^\alpha < \infty$.
- (ii) *For every nonnegative integer valued random variable T satisfying $\mathbb{E}T^{\alpha/(\alpha-1)} < \infty$, and for any X_1, X_2, \dots identically distributed with X (but not necessarily independent), we have $\mathbb{E}S_T < \infty$.*

In order to prove the implications (iii) \Rightarrow (i) of Theorems 1.1 and 1.2, we will assume that (i) fails, and construct a suitable T for which $\mathbb{E}X_T = \infty$ (and thus also $\mathbb{E}S_T = \infty$). This T will be the *last* time the random sequence is in a certain (time-dependent) deterministic set, i.e.

$$T := \max\{n : X_n \in B_n\}$$

for a suitable sequence of sets B_n . It is interesting to note that, in contrast, no T of the form $\min\{n : X_n \in B_n\}$ could work for this purpose, since such a T is a stopping time, so Wald’s identity applies. In the context of Theorem 1.2, for instance, T will take the form

$$T := \max\{n : X_n \geq e^n\}.$$

The results here bear an interesting relation to so-called prophet inequalities; see [7] for a survey. A central prophet inequality (see [11]) states that if X_1, X_2, \dots are independent (not necessarily identically distributed) nonnegative random variables then

$$\sup_{U \in \mathcal{U}} \mathbb{E}X_U \leq 2 \sup_{S \in \mathcal{S}} \mathbb{E}X_S, \tag{1.2}$$

where \mathcal{U} denotes the set of all positive integer-valued random variables and \mathcal{S} denotes the set of all stopping times. The left side is of course equal to $\mathbb{E} \sup_i X_i$. The factor 2 is sharp. The interpretation is that a prophet and a gambler are presented sequentially with the values X_1, X_2, \dots , and each can stop at any time k and then receive payment X_k . The prophet sees the entire sequence in advance and so can obtain the left side of (1.2) in expectation, while the gambler can only achieve the supremum on the right. Thus (1.2) states that the prophet's advantage is at most a factor of 2.

The inequality (1.2) is uninteresting when (X_i) is an infinite i.i.d. sequence, but for example applying it to $X_i \mathbf{1}[i \leq n]$ (where n is fixed and (X_i) are i.i.d.) gives

$$\sup_{\substack{U \in \mathcal{U}: \\ U \leq n}} \mathbb{E} X_U \leq 2 \sup_{\substack{S \in \mathcal{S}: \\ S \leq n}} \mathbb{E} X_S, \tag{1.3}$$

(and the factor of 2 is again sharp). How does this result change if we replace the condition that U and S are bounded by n with a moment restriction? It turns out that the prophet's advantage can become infinite, in the following sense. Let X_1, X_2, \dots be any i.i.d. nonnegative random variables with mean 1 and infinite variance. By Theorem 1.1, there exists an integer-valued random variable T so that $\mu := \mathbb{E} T < \infty$ but $\mathbb{E} X_T = \infty$. Then we have

$$\sup_{\substack{U \in \mathcal{U}: \\ \mathbb{E} U \leq \mu}} \mathbb{E} X_U = \infty; \quad \sup_{\substack{S \in \mathcal{S}: \\ \mathbb{E} S \leq \mu}} \mathbb{E} X_S \leq \mu.$$

Here the first claim follows by taking $U = T$ and the second claim follows from Wald's identity.

Interpreting impatience as meaning that the time at which we stop has mean at most μ , we see that impatience hurts the gambler much more than the prophet.

Our proof of the implication (i) \Rightarrow (ii) in Theorem 1.1 will rely on a concentration inequality which is due to Hsu and Robbins [8] for the important special case $\alpha = 2$, and a generalization due to Katz [10] for $\alpha < 2$. For expository purposes, we include a proof of the Hsu-Robbins inequality, which is different from the original proof. Thus, we give a complete proof from first principles of Theorem 1.1 in the case $\alpha = 2$. Erdős [3, 4] proved a converse of the Hsu-Robbins result; we will also obtain this converse in the case of nonnegative random variables as a corollary of our results.

Throughout the article we will write $X = X_1$ and $S_n := \sum_{i=1}^n X_i$. If $T = 0$ then we take $X_T = 0$ and $S_T = 0$.

2 The case of exponential tails

In this section we give the proof of Theorem 1.2, which is relatively straightforward. We start with a simple lemma relating X_T and S_T for T of the form that we will use for our counterexamples. The same lemma will be used in the proof of Theorem 1.1.

Lemma 2.1. *Let X_1, X_2, \dots be i.i.d. nonnegative random variables. Let T be defined by*

$$T := \max\{k : X_k \in B_k\}$$

for some sequence of sets B_k for which the above set is a.s. finite, and where we take $T = 0$ and $X_T = 0$ when the set is empty. Then

$$\mathbb{E} S_T = \mathbb{E}[(T - 1)^+] \cdot \mathbb{E} X + \mathbb{E} X_T.$$

Wald for non-stopping times

Proof. Observe that $\mathbf{1}[T = k]$ and S_{k-1} are independent for every $k \geq 1$. Therefore,

$$\begin{aligned} \mathbb{E}S_T &= \mathbb{E} \sum_{k=1}^{\infty} S_k \mathbf{1}[T = k] \\ &= \mathbb{E} \sum_{k=1}^{\infty} (S_{k-1} + X_k) \mathbf{1}[T = k] \\ &= \sum_{k=1}^{\infty} \mathbb{E}S_{k-1} \cdot \mathbb{P}(T = k) + \mathbb{E} \sum_{k=1}^{\infty} X_k \mathbf{1}[T = k] \\ &= \mathbb{E}[(T-1)^+] \cdot \mathbb{E}X + \mathbb{E}X_T. \end{aligned} \quad \square$$

Proof of Theorem 1.2. We first prove that (i) and (ii) are equivalent, assuming only that the X_i are identically distributed (not necessarily independent).

Assume that (i) holds, i.e. $\mathbb{E}[X(\log X)^+] < \infty$, and that T is a nonnegative integer-valued random variable satisfying $\mathbb{E}e^{cT} < \infty$. Observe that $X_k \leq e^{ck} + X_k \mathbf{1}[X_k > e^{ck}]$, so

$$S_T \leq \sum_{k=1}^T e^{ck} + \sum_{k=1}^T X_k \mathbf{1}[X_k \geq e^{ck}].$$

The first sum equals

$$\frac{e^c(e^{cT} - 1)}{e^c - 1}$$

which has finite expectation. The expectation of the second sum is at most

$$\sum_{k=1}^{\infty} \mathbb{E}(X \mathbf{1}[X \geq e^{ck}]) = \mathbb{E} \sum_{k=1}^{\infty} X \mathbf{1}[X \geq e^{ck}] = \mathbb{E} \left(X \left\lfloor \frac{(\log X)^+}{c} \right\rfloor \right) < \infty.$$

Hence $\mathbb{E}S_T < \infty$ as required, giving (ii).

Now assume that (i) fails, i.e. $\mathbb{E}[X(\log X)^+] = \infty$, but (ii) holds (still without assuming independence of the X_i). Taking $T \equiv 1$ in (ii) shows that $\mathbb{E}X < \infty$. Now let

$$T := \max\{k \geq 1 : X_k \geq e^k\}, \quad (2.1)$$

where T is taken to be 0 if the set above is empty and ∞ if it is unbounded. Then

$$\mathbb{P}(T \geq k) \leq \sum_{i=k}^{\infty} \mathbb{P}(X_i \geq e^i) \leq \sum_{i=k}^{\infty} \frac{\mathbb{E}X}{e^i}$$

by Markov's inequality. The last sum is $(\mathbb{E}X)e^{1-k}/(e-1)$, and hence $\mathbb{E}e^{ck} < \infty$ for suitable $c > 0$ (and in particular T is a.s. finite). On the other hand,

$$\mathbb{E}S_T = \mathbb{E} \sum_{k=1}^{\infty} X_k \mathbf{1}[k \leq T] \geq \mathbb{E} \sum_{k=1}^{\infty} X \mathbf{1}[X \geq e^k] = \mathbb{E}(X[(\log X)^+]),$$

which is infinite, contradicting (ii).

Now assume that the X_i are i.i.d. We have already established that (i) and (ii) are equivalent, and (ii) immediately implies (iii) since $S_T \geq X_T$. It therefore suffices to show that (iii) implies (i). Suppose (i) fails and (iii) holds. Taking $T \equiv 1$ in (iii) shows that $\mathbb{E}X < \infty$. Now take the same T as in (2.1). As argued above, $\mathbb{E}S_T = \infty$ and $\mathbb{E}e^{cT} < \infty$ for some $c > 0$ (so $\mathbb{E}T < \infty$). Hence (iii) gives $\mathbb{E}X_T < \infty$. But this contradicts Lemma 2.1. \square

Remark Conditions (i) and (iii) cannot be equivalent if the i.i.d. condition is dropped since if $X_1 = X_2 = X_3 = \dots$, then $X_T = X_1$ for every T and so (iii) just corresponds to X having a first moment.

3 The case $\alpha = 2$ and the Hsu-Robbins Theorem

In this section we prove Theorem 1.1 in the important special case $\alpha = 2$ (so $1/(\alpha - 1) = 1$). We will use the following result of Hsu and Robbins [8]. See [5, §6.11.1] for an alternative proof of this result, arguably simpler than the original proof in [8], and making use of a result of [6]. For expository purposes we give yet another proof, which is self-contained, and based on an argument from [2].

Theorem 3.1 (Hsu and Robbins). *Let X_1, X_2, \dots be i.i.d. random variables with finite mean μ and finite variance. Then for all $\epsilon > 0$,*

$$\sum_{n=1}^{\infty} \mathbb{P}(|S_n - n\mu| \geq n\epsilon) < \infty.$$

Proof. We may assume without loss of generality that $\mu = 0$ and $\mathbb{E}X^2 = 1$. Let $X_n^* := \max\{X_1, \dots, X_n\}$ and $S_n^* := \max\{S_1, \dots, S_n\}$. Observe that for any $h > 0$, the stopping time $\tau_h := \min\{k : S_k \geq h\}$ satisfies

$$\begin{aligned} \mathbb{P}(S_n > 3h, X_n^* \leq h) &\leq \mathbb{P}(\tau_h < n, S_{\tau_h} \leq 2h) \mathbb{P}(S_n > 3h \mid \tau_h < n, S_{\tau_h} \leq 2h) \\ &\leq \mathbb{P}(\tau_h \leq n)^2 \end{aligned} \tag{3.1}$$

where the last step used the strong Markov property at time τ_h . Now Kolmogorov's maximum inequality (see e.g. [9, Lemma 4.15]) implies that

$$\mathbb{P}(\tau_h \leq n) = \mathbb{P}(S_n^* \geq h) \leq \frac{\mathbb{E}S_n^2}{h^2} = \frac{n}{h^2}.$$

Applying this with $h = \epsilon n/3$ we infer from (3.1) that

$$\mathbb{P}\left(S_n > n\epsilon, X_n^* \leq \frac{\epsilon n}{3}\right) \leq \frac{81}{\epsilon^4 n^2}.$$

Moreover, we have

$$\mathbb{P}\left(X_n^* > \frac{\epsilon n}{3}\right) \leq n\mathbb{P}\left(X_1 > \frac{\epsilon n}{3}\right).$$

Combining the last two inequalities give

$$\mathbb{P}(S_n > n\epsilon) \leq \frac{81}{\epsilon^4 n^2} + n\mathbb{P}\left(X_1 > \frac{\epsilon n}{3}\right).$$

The first term on the right is summable in n , and the second term is summable by the assumption of finite variance. Applying the same argument to $-S_n$ completes the proof. \square

We will also need a simple fact of real analysis, a converse to Hölder's inequality, which we state in a probabilistic form. See, e.g., Lemma 6.7 in [12] for a related statement. The proof method is known as the "gliding hump"; see [13] and the references therein.

Lemma 3.2. *Let $p, q > 1$ satisfy $1/p + 1/q = 1$. Assume that a nonnegative random variable X satisfies $\mathbb{E}Xg(X) < \infty$ for every nonnegative function g that satisfies $\mathbb{E}g^q(X) < \infty$. Then $\mathbb{E}X^p < \infty$.*

Proof. Assume $\mathbb{E}X^p = \infty$. Letting $\psi_k := \mathbb{P}(\lfloor X \rfloor = k)$, we have $\sum_{k=1}^{\infty} \psi_k k^p = \mathbb{E}\lfloor X \rfloor^p = \infty$, so we can choose integers $0 = a_0 < a_1 < a_2, \dots$ such that for each $\ell \geq 1$,

$$S_\ell := \sum_{k \in [a_{\ell-1}, a_\ell)} \psi_k k^p \geq 1.$$

Wald for non-stopping times

Denote the interval $[a_{\ell-1}, a_\ell)$ by I_ℓ and let g be defined on $[0, \infty)$ by

$$g(x) := \frac{\lfloor x \rfloor^{p-1}}{\ell S_\ell} \text{ for } x \in I_\ell.$$

Since $(p-1)q = p$, we obtain

$$\mathbb{E}g^q(X) = \sum_{\ell=1}^{\infty} \sum_{k \in I_\ell} \psi_k \frac{k^p}{\ell^q S_\ell^q} = \sum_{\ell=1}^{\infty} \frac{1}{\ell^q S_\ell^{q-1}} < \infty.$$

On the other hand

$$\mathbb{E}Xg(X) \geq \sum_{\ell=1}^{\infty} \sum_{k \in I_\ell} \psi_k \frac{k^p}{\ell S_\ell} = \sum_{\ell=1}^{\infty} \frac{1}{\ell} = \infty. \quad \square$$

We can now proceed with the main proof.

Proof of Theorem 1.1, case $\alpha = 2$. We will first show that (i) and (ii) are equivalent. Assume (i) holds, i.e. $\mathbb{E}X^2 < \infty$, and let T satisfy $\mathbb{E}T < \infty$. We may assume without loss of generality that $\mathbb{E}X = 1$. By the nonnegativity of the X_i , we have

$$\mathbb{P}(S_T \geq 2n) \leq \mathbb{P}(T \geq n) + \mathbb{P}(S_n \geq 2n). \quad (3.2)$$

Since $\mathbb{E}T < \infty$, the first term on the right is summable in n . Since $\mathbb{E}X^2 < \infty$ and $\mathbb{E}X = 1$, Theorem 3.1 with $\epsilon = 1$ implies that the second term is also summable. We conclude that $\mathbb{E}S_T < \infty$.

Now assume (ii). To show that X has finite second moment, using Lemma 3.2 with $p = q = 2$, we need only show that for any nonnegative function g satisfying $\mathbb{E}g^2(X) < \infty$, we have $\mathbb{E}Xg(X) < \infty$. Given such a g , consider the integer valued random variable

$$T_g := \max\{k \geq 1 : g(X_k) \geq k\}, \quad (3.3)$$

where T_g is taken to be 0 if the set is empty or ∞ if it is unbounded. We have

$$\mathbb{E}T_g = \sum_{k=1}^{\infty} \mathbb{P}(T_g \geq k) \leq \sum_{k=1}^{\infty} \sum_{\ell=k}^{\infty} \mathbb{P}(g(X_\ell) \geq \ell) = \sum_{\ell=1}^{\infty} \ell \mathbb{P}(g(X) \geq \ell).$$

Since $\mathbb{E}g^2(X) < \infty$, the last expression is finite, and hence $\mathbb{E}T_g < \infty$. Thus, by assumption (ii), we have $\mathbb{E}S_{T_g} < \infty$. However

$$\begin{aligned} \mathbb{E}S_{T_g} &= \mathbb{E} \sum_{k=1}^{\infty} X_k \mathbf{1}[k \leq T_g] \geq \mathbb{E} \sum_{k=1}^{\infty} X_k \mathbf{1}[g(X_k) \geq k] \\ &= \mathbb{E} \sum_{k=1}^{\infty} X \mathbf{1}[g(X) \geq k] \geq \mathbb{E}X[g(X)], \end{aligned} \quad (3.4)$$

so that $\mathbb{E}X[g(X)] < \infty$, which easily yields $\mathbb{E}Xg(X) < \infty$ as required.

Clearly (ii) implies (iii). Finally, we proceed as in the proof of Theorem 1.2 to show (iii) implies (i). Suppose (i) fails and (iii) holds. Taking $T \equiv 1$ in (iii) shows that $\mathbb{E}X < \infty$. Since $\mathbb{E}X^2 = \infty$, Lemma 3.2 implies the existence of a g with $\mathbb{E}g^2(X) < \infty$ but $\mathbb{E}Xg(X) = \infty$. Let T_g be defined as in (3.3) above, for this g . The argument above shows that $\mathbb{E}S_{T_g} = \infty$ while $\mathbb{E}T_g < \infty$, and so the assumption (iii) gives $\mathbb{E}X_{T_g} < \infty$. However this contradicts Lemma 2.1. \square

We also obtain the following converse of the Hsu-Robbins Theorem due to Erdős.

Corollary 3.3. *Let X_1, X_2, \dots be i.i.d. nonnegative random variables with finite mean μ . Write $S_n = \sum_{i=1}^n X_i$ and $X = X_1$. If, for all $\epsilon > 0$,*

$$\sum_{n=1}^{\infty} \mathbb{P}(|S_n - n\mu| \geq n\epsilon) < \infty,$$

then X has a finite variance.

Proof. Without loss of generality, we can assume that $\mu = 1$. By Theorem 1.1 with $\alpha = 2$, it suffices to show that $\mathbb{E}S_T < \infty$ for all T with finite mean. However, this is immediate from (3.2) – the first term on the right is summable since T has finite mean, and the second term is summable by the assumption of the corollary with $\epsilon = 1$. \square

4 The case of $\alpha < 2$

The proof of Theorem 1.1 in the general case follows very closely the proof for $\alpha = 2$. We need the following replacement of Theorem 3.1 due to Katz [10], whose proof we do not give here. A converse of the results in [10] appears in [1]. We will also use the general case of Lemma 3.2.

Theorem 4.1 (Katz). *Let X_1, X_2, \dots be i.i.d. random variables satisfying $\mathbb{E}|X_1|^t < \infty$ with $t \geq 1$. If $r > t$, then, for all $\epsilon > 0$,*

$$\sum_{n=1}^{\infty} n^{r-2} \mathbb{P}(|S_n| \geq n^{r/t}\epsilon) < \infty.$$

Proof of Theorem 1.1, case $\alpha < 2$. We first prove that (i) implies (ii). Assume that $\mathbb{E}X^\alpha < \infty$, and T is an integer valued random variable with $\mathbb{E}T^{1/(\alpha-1)} < \infty$. Observe that

$$\mathbb{P}(S_T \geq n) \leq \mathbb{P}(T \geq \lceil n^{\alpha-1} \rceil) + \mathbb{P}(S_{\lceil n^{\alpha-1} \rceil} \geq n). \tag{4.1}$$

Since $\mathbb{P}(T \geq \lceil n^{\alpha-1} \rceil) \leq \mathbb{P}(T^{1/(\alpha-1)} \geq n)$, the first term on the right is summable. For the second, we have

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{P}(S_{\lceil n^{\alpha-1} \rceil} \geq n) &\leq \sum_{k=1}^{\infty} \sum_{\substack{n \geq 1: \\ \lceil n^{\alpha-1} \rceil = k}} \mathbb{P}(S_k \geq n) \\ &\leq \sum_{k=1}^{\infty} \sum_{\substack{n \geq 1: \\ \lceil n^{\alpha-1} \rceil = k}} \mathbb{P}(S_k \geq (k-1)^{\frac{1}{\alpha-1}}) \end{aligned}$$

since $\lceil n^{\alpha-1} \rceil = k$ implies that $n \geq (k-1)^{1/(\alpha-1)}$. It is easy to check that there exists C_α such that for all $k \geq 1$,

$$\#\{n : \lceil n^{\alpha-1} \rceil = k\} \leq C_\alpha k^{\frac{2-\alpha}{\alpha-1}}.$$

Hence the last double sum is at most

$$C_\alpha \sum_{k=1}^{\infty} k^{\frac{2-\alpha}{\alpha-1}} \mathbb{P}(S_k \geq (k-1)^{\frac{1}{\alpha-1}}).$$

Now using Theorem 4.1 with $t = \alpha$ and $r = \alpha/(\alpha-1)$ and $\epsilon = \frac{1}{2}$ (and noting that $k^{1/(\alpha-1)}/2 \leq (k-1)^{1/(\alpha-1)}$ for large enough k), we conclude that the above expression is finite. Hence $\mathbb{E}S_T < \infty$, as required.

Next we show that (ii) implies (i). To show that X has a finite α -moment, using Lemma 3.2, it suffices to show that for any nonnegative function g satisfying

$\mathbb{E}g^{\alpha/(\alpha-1)}(X) < \infty$, we have $\mathbb{E}Xg(X) < \infty$. Given such a g , consider as before the integer valued random variable

$$T_g := \max\{k \geq 1 : g(X_k) \geq k\},$$

where T_g is taken to be 0 if the set is empty or ∞ if it is unbounded. Observe that

$$\begin{aligned} \sum_{k=1}^{\infty} k^{\frac{2-\alpha}{\alpha-1}} \mathbb{P}(T_g \geq k) &\leq \sum_{k=1}^{\infty} k^{\frac{2-\alpha}{\alpha-1}} \sum_{\ell=k}^{\infty} \mathbb{P}(g(X_\ell) \geq \ell) \\ &\leq \sum_{\ell=1}^{\infty} \ell^{\frac{1}{\alpha-1}} \mathbb{P}(g(X) \geq \ell). \end{aligned}$$

If $\mathbb{E}g^{\alpha/(\alpha-1)}(X) < \infty$ then the last sum is finite and hence $\mathbb{E}T_g^{1/(\alpha-1)} < \infty$. By assumption (ii) we have $\mathbb{E}S_{T_g} < \infty$. However, as argued in (3.4), $\mathbb{E}S_{T_g} \geq \mathbb{E}X[g(X)]$. Therefore $\mathbb{E}X[g(X)] < \infty$, so $\mathbb{E}Xg(X) < \infty$ as required.

Clearly (ii) implies (iii). Finally, suppose (i) fails and (iii) holds. Taking $T \equiv 1$ in (iii) shows that $\mathbb{E}X < \infty$. Since $\mathbb{E}X^\alpha = \infty$, Lemma 3.2 implies the existence of a g with $\mathbb{E}g^{\alpha/(\alpha-1)}(X) < \infty$ but $\mathbb{E}Xg(X) = \infty$. Then, as before, T_g as defined above gives a contradiction to Lemma 2.1. \square

5 The dependent case

Proof of Proposition 1.3. Assume (i) holds. If $\mathbb{E}T^{\alpha/(\alpha-1)} < \infty$ and X_1, X_2, \dots are as in (ii), then we can write

$$S_T \leq \sum_{k=1}^T k^{\frac{1}{\alpha-1}} + \sum_{k=1}^T X_k \mathbf{1}[X_k \geq k^{\frac{1}{\alpha-1}}].$$

The first sum is at most $T^{\alpha/(\alpha-1)}$ which has finite expectation. The expectation of the second sum is at most

$$\mathbb{E} \sum_{k=1}^{\infty} X \mathbf{1}[X \geq k^{\frac{1}{\alpha-1}}] \leq \mathbb{E}(XX^{\alpha-1}) = \mathbb{E}X^\alpha < \infty.$$

Hence $\mathbb{E}S_T < \infty$, as claimed in (ii).

Now assume (ii) holds. To show that X has finite α -moment, using Lemma 3.2, it is enough to show that for any nonnegative g satisfying $\mathbb{E}g^{\alpha/(\alpha-1)}(X) < \infty$, we have $\mathbb{E}Xg(X) < \infty$. It is easily seen that it suffices to only consider g that are integer valued. Given such a g , let T be $g(X)$ and let all the X_i be equal to X . Then $\mathbb{E}T^{\alpha/(\alpha-1)} < \infty$. By (ii), $\mathbb{E}S_T < \infty$. However, by construction $S_T = Xg(X)$, concluding the proof. \square

References

- [1] Baum, L. E., Katz, M., Convergence rates in the law of large numbers, *Trans. Amer. Math. Soc.* **120**, 1965, 108–123. MR-0198524
- [2] Chow, Y. S. and Teicher, H., Probability theory. Independence, interchangeability, martingales. Second edition. Springer Texts in Statistics. Springer-Verlag, 1988. MR-0953964
- [3] Erdős, P., On a theorem of Hsu and Robbins, *Ann. Math. Statist.* **20**, 1949, 286–291. MR-0030714
- [4] Erdős, P., Remark on my paper “On a theorem of Hsu and Robbins”, *Ann. Math. Statist.* **21**, 1950, 138. MR-0032970
- [5] Gut, A., *Probability: A Graduate Course*. Second Edition. Springer, 2013. MR-2977961

Wald for non-stopping times

- [6] Hoffmann-Jørgensen, J., Sums of independent Banach space valued random variables. *Studia Math.* **LII**, 1974, 159–186. MR-0356155
- [7] Hill, T. P. and Kertz, R. P., A survey of prophet inequalities in optimal stopping theory, Strategies for sequential search and selection in real time, Amer. Math. Soc., *Contemp. Math.*, **125**, 1992, 191–207, MR-1160620
- [8] Hsu, P. L. and Robbins, H., Complete convergence and the law of large numbers, *Proc. Nat. Acad. Sci. U.S.A.* **33**, 1947, 25–31. MR-0019852
- [9] Kallenberg, O., *Foundations of Modern Probability*. Second Edition. Springer, 2002. MR-1876169
- [10] Katz, M., The probability in the tail of a distribution, *Ann. Math. Statist.* **34**, 1963, 312–318. MR-0144369
- [11] Krengel, U. and Sucheston, L., On semiamarts, amarts, and processes with finite value. Probability on Banach spaces, 197–266, *Adv. Probab. Related Topics* **4**, Dekker, 1978. MR-0515432
- [12] Royden, H. L., *Real analysis*. Third edition. Macmillan, 1988. MR-1013117
- [13] Sokal, A. D. A really simple elementary proof of the uniform boundedness theorem. *Amer. Math. Monthly* **118**, 2011, 450–452. MR-2805031
- [14] Wald, A., On cumulative sums of random variables. *Ann. Math. Statist.* **15**, 1944, 283–296. MR-0010927

Acknowledgments. We thank the anonymous referee for helpful comments. Most of this work was carried out when JES was visiting Microsoft Research at Redmond, WA, and he thanks the Theory Group for its hospitality. JES also acknowledges the support of the Swedish Research Council and the Knut and Alice Wallenberg Foundation.