MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

Lévy's Zero-one Law in Game-Theoretic Probability

Glenn SHAFER, Vladimir VOVK and Akimichi TAKEMURA

METR 2009–17 May 2009

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.keisu.t.u-tokyo.ac.jp/research/techrep/index.html

Lévy's zero-one law in game-theoretic probability

Glenn Shafer[1,2], Vladimir Vovk[2], and Akimichi Takemura[3]

May, 20009

Abstract

We prove a game-theoretic version of Lévy's zero-one law, and deduce several corollaries from it, including Kolmogorov's zero-one law, the ergodicity of Bernoulli shifts, and a zero-one law for dependent trials. Our secondary goal is to explore the basic definitions of game-theoretic probability theory, with Lévy's zero-one law serving a useful role.

1 Introduction

In this article we continue the investigation of zero-one laws of game-theoretic probability theory started in [7]. Our main result is a game-theoretic version of Lévy's [5] zero-one law, from which we deduce game-theoretic versions of Bártfai and Révész's [1] zero-one law, Kolmogorov's zero-one law ([4], Appendix), and the ergodicity of Bernoulli shifts (see, e.g., [3], Section 8.1, Theorem 1). The last two results have been established in [7], but our proofs are different: we obtain them as easy corollaries of our main result.

We start our exposition in Section 2 by introducing our general prediction protocol and defining the game-theoretic notions of expectation and probability. For more information on game-theoretic probability theory, see, e.g., [6]. In Section 3 we prove Lévy's zero-one law for our prediction protocol, and in Section 4 we derive other zero-one laws as corollaries.

We will be using the standard notation $\mathbb{N} = \{1, 2, \ldots\}$ for the set of all natural numbers and $\mathbb{R} = (-\infty, \infty)$ for the set of all real numbers. Alongside $\mathbb{R}$ we will often consider sets, such as $(-\infty, \infty]$ and $[-\infty, \infty]$, obtained from $\mathbb{R}$ by adding $-\infty$ or ∞ (or both). We set $0 \times \infty := 0$ (but do not assign any default value to $\infty + (-\infty)$). The indicator function of a subset E of a given set X will be denoted $\mathbb{I}_E$; i.e., $\mathbb{I}_E : X \to \mathbb{R}$ takes the value 1 on E and the value 0 outside E. The words such as "positive" and "negative" are to be understood in the wide sense of inequalities $\geq$ and $\leq$ rather than $>$ and $<$.

[1]Department of Accounting and Information Systems, Rutgers Business School—Newark and New Brunswick, 180 University Avenue, Newark, NJ 07102, USA.

[2]Computer Learning Research Centre, Department of Computer Science, Royal Holloway, University of London, Egham, Surrey TW20 0EX, England.

[3]Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

2 Game-theoretic expectation and probability

We consider a general perfect-information game between three players, called Forecaster, Skeptic, and Reality. The game proceeds in discrete time. First we describe the game formally, and then briefly explain the intuition behind the formal description.

Let $\mathcal{X}$ be a fixed set, which we will call the *outcome space*, and let $(-\infty,\infty]^{\mathcal{X}}$ stand for the set of all functions $f:\mathcal{X}\to(-\infty,\infty]$. A function $\mathcal{E}:(-\infty,\infty]^{\mathcal{X}}\to[-\infty,\infty]$ is called a *superexpectation functional* if it satisfies the following axioms:

1. If $f,g\in(-\infty,\infty]^{\mathcal{X}}$ satisfy $f\le g$, then $\mathcal{E}(f)\le\mathcal{E}(g)$.
2. If $f\in(-\infty,\infty]^{\mathcal{X}}$ and $c\in(0,\infty)$, then $\mathcal{E}(cf)=c\mathcal{E}(f)$.
3. If $f,g\in(-\infty,\infty]^{\mathcal{X}}$, then $\mathcal{E}(f+g)\le\mathcal{E}(f)+\mathcal{E}(g)$ (with the right-hand side understood to be ∞ when $\mathcal{E}(f)=\infty$ or $\mathcal{E}(g)=\infty$).
4. For each $c\in(-\infty,\infty]$, $\mathcal{E}(c)=c$, where the c in parentheses is the function in $(-\infty,\infty]^{\mathcal{X}}$ that is identically equal to c.
5. For any sequence of positive functions $f_1,f_2,\ldots$ in $(-\infty,\infty]^{\mathcal{X}}$,

$$\mathcal{E}\left(\sum_{k=1}^{\infty}f_k\right)\le\sum_{k=1}^{\infty}\mathcal{E}\left(f_k\right). \tag{1}$$

Let $\mathbf{E}$ be the set of all superexpectation functionals.

Axiom 4 implies $\mathcal{E}(0)=0$ (so that we can allow $c=0$ in Axiom 2). This, in combination with Axiom 1, implies

$$f\ge0\Longrightarrow\mathcal{E}(f)\ge0. \tag{2}$$

Axioms 3 and 4 imply that

$$\mathcal{E}(f+c)=\mathcal{E}(f)+c \tag{3}$$

for each $c\in\mathbb{R}$ (indeed, $\mathcal{E}(f+c)\le\mathcal{E}(f)+\mathcal{E}(c)=\mathcal{E}(f)+c$ and $\mathcal{E}(f)\le\mathcal{E}(f+c)+\mathcal{E}(-c)=\mathcal{E}(f+c)-c$). From (2) and (3) we can see that

$$\mathcal{E}(f)\le0\Longrightarrow\inf_{x\in\mathcal{X}}f(x)\le0. \tag{4}$$

We will refer to property (4) as *coherence*. Replacing the $=$ in Axiom 2 with $\le$ would lead to an equivalent statement.

Axioms 1–5 are relaxations of the standard properties of the expectation functional: cf., e.g., Axioms 1–5 in [9] (Axioms 2 and 3 are weaker than the corresponding standard axioms, Axioms 1 and 4 are stronger than the corresponding standard axioms but follow from standard Axioms 1–4, and Axiom 5 follows from standard Axiom 5 in the presence of our Axiom 3). The most controversial axiom is Axiom 5. It is satisfied in many interesting cases, such as in the case of finite $\mathcal{X}$ and for many protocols in [6].

The most noticeable difference between our superexpectation functionals and the standard expectation functionals is that the former are defined for all functions $f : \mathcal{X} \to (-\infty, \infty]$ whereas the latter are defined only for functions that are measurable w.r. to a given σ-algebra. The notion of superexpectation functional is more general since every expectation functional can be extended to the whole of $(-\infty, \infty]^{\mathcal{X}}$ as the corresponding upper integral. Namely, $\mathcal{E}(f)$ can be defined as the infimum of the expectation of g (taken to be ∞ whenever the expectation of $\max(g, 0)$ is ∞) over all measurable functions $g \geq f$. The extension may no longer be an expectation functional but is still a superexpectation functional.

Remark. It is sometimes useful to have the stronger form

$$f > 0 \Longrightarrow \mathcal{E}(f) > 0 \tag{5}$$

of (2). (In this article we do not really need (5): it would merely slightly simplify the proof of Lemma 1 below.) Even the strong form (5) follows from Axioms 1–5. Indeed, if $f > 0$ but $\mathcal{E}(f) = 0$, Axioms 1–5 imply

$$1 \overset{4}{=} \mathcal{E}\,\mathbb{I}_{\{f>0\}} = \mathcal{E}\,\mathbb{I}_{\cup_{n=1}^{\infty}\{nf\geq 1\}} \overset{1}{\leq} \mathcal{E} \sum_{n=1}^{\infty} \mathbb{I}_{\{nf\geq 1\}} \overset{5}{\leq} \sum_{n=1}^{\infty} \mathcal{E}\,\mathbb{I}_{\{nf\geq 1\}} \overset{1}{\leq} \sum_{n=1}^{\infty} \mathcal{E}(nf) \overset{2}{=} 0$$

(over each relation symbol we write the ordinal number of the axiom that justifies it; we could avoid using Axiom 2 by using (2) and Axiom 3 instead).

The most general protocol that we consider in this article is as follows.

Protocol 1. General prediction protocol
Parameters: non-empty sets $\mathcal{P}_1, \mathcal{P}_2, \ldots$ and function $\mathcal{E} : p \in \cup_n \mathcal{P}_n \mapsto \mathcal{E}_p \in \mathbf{E}$
Protocol:
Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
  Forecaster announces $p_n \in \mathcal{P}_n$.
  Skeptic announces f_n such that $\mathcal{E}_{p_n}(f_n) \leq \mathcal{K}_{n-1}$.
  Reality announces $x_n \in \mathcal{X}$.
  $\mathcal{K}_n := f_n(x_n)$.
END FOR

At the end of each trial n Reality chooses the outcome x_n of this trial. At the beginning of this trial Forecaster gives his prediction p_n for x_n; the prediction is chosen from a set $\mathcal{P}_n$, the *prediction space* for trial n. We will use the notation $\mathcal{P}$ for $\cup_n \mathcal{P}_n$. After Forecaster's move Skeptic chooses a gamble, which we represent as a function f_n on $\mathcal{X}$: $f_n(x)$ is the payoff of the gamble if Reality chooses x as the trial's outcome. The gambles available to Skeptic are determined by Forecaster's prediction (via the function $\mathcal{E} : \mathcal{P} \to \mathbf{E}$). Skeptic's capital after the nth trial is denoted $\mathcal{K}_n$. He is allowed to choose his initial capital $\mathcal{K}_0$ and, implicitly, also allowed to throw away part of his capital at each trial.

Remark. Another version of the general prediction protocol is where Forecaster chooses the superexpectation functional directly. This is a special case of our protocol with $\mathcal{P}_n = \mathbf{E}$ for all n and with $\mathcal{E} : \mathcal{P} \to \mathbf{E}$ the identity function. The reader will also notice that allowing $\mathcal{E}$ to depend not only on Forecaster's last move but also on his and Reality's previous moves is straightforward but does not lead to stronger results: the seemingly more general results easily follow from our results.

Remark. In [7] we considered a different but essentially equivalent prediction protocol.

We call the set $\Omega := \prod_{n=1}^{\infty}(\mathcal{P}_n \times \mathcal{X})$ of all infinite sequences of Forecaster's and Reality's moves the *sample space*. The elements of the set $\bigcup_{n=0}^{\infty} \prod_{i=1}^{n}(\mathcal{P}_i \times \mathcal{X})$ of all finite sequences of Forecaster's and Reality's moves are called *post-situations*, and the elements of the set $\bigcup_{n=0}^{\infty} (\prod_{i=1}^{n}(\mathcal{P}_i \times \mathcal{X}) \times \mathcal{P}_{n+1})$ are called *pre-situations*. The term *situation* will be applied to both pre-situations and post-situations. For each situation s we let $\Gamma(s) \subseteq \Omega$ stand for the set of all infinite extensions in Ω of s (i.e., $\Gamma(s)$ is the set of all $\omega \in \Omega$ such that s is a prefix of ω). Let $\Box$ be the empty situation.

The *level* of a situation s is the number of predictions in s. In other words, n is the level of pre-situations of the form $p_1 x_1 \ldots p_{n-1} x_{n-1} p_n$ and of post-situations of the form $p_1 x_1 \ldots p_n x_n$. The level of $\Box$ is 0. If $\omega \in \Omega$ and $n \in \{0, 1, \ldots\}$, ω^n is defined to be the unique post-situation of level n that is a prefix of ω.

If we fix a strategy Σ for Skeptic, Skeptic's capital $\mathcal{K}_n$ becomes a function of the current post-situation s of level n. We write $\mathcal{K}^{\Sigma}(s)$ for $\mathcal{K}_n$ resulting from Skeptic following Σ and from Forecaster and Reality playing s. The function $\mathcal{K}^{\Sigma}$, defined on the set of all post-situations and taking values in $(-\infty, \infty]$, will be called the *capital process* of Σ; we will omit Σ when it is clear from context. A function $\mathcal{S}$ is called a (game-theoretic) *supermartingale* if it is the capital process, $\mathcal{S} = \mathcal{K}^{\Sigma}$, of some strategy Σ for Skeptic. Sometimes we will extend the domain of game-theoretic supermartingales $\mathcal{S}$ to include pre-situations: if s is a pre-situation, $\mathcal{S}(s)$ is interpreted as $\mathcal{S}(s^-)$, where s^- is s with the last prediction removed. We will often write $\mathcal{S}_n(\omega)$ to mean $\mathcal{S}(\omega^n)$.

Remark. The definition of a martingale is obtained from that of a supermartingale by replacing the condition $\mathcal{E}_{p_n}(f_n) \le \mathcal{K}_{n-1}$ in Protocol 1 by $\mathcal{E}_{p_n}(f_n) = \mathcal{K}_{n-1}$. Martingales are less useful for us since the sum of two martingales may fail to be a martingale (the inequality in Axiom 3 may be strict), whereas the sum of two supermartingales is always a supermartingale.

For each function $\xi : \Omega \to [-\infty, \infty]$ and each (pre- or post-) situation s, we define the (conditional) *upper expectation* of ξ given s by

$$\overline{\mathbb{E}}(\xi \mid s) := \inf\Bigl\{a \mid \exists \mathcal{S} : \mathcal{S}(s) = a \text{ and } \liminf_{n\to\infty} \mathcal{S}_n(\omega) \ge \xi(\omega) \text{ for all } \omega \in \Gamma(s)\Bigr\} \tag{6}$$

where $\mathcal{S}$ ranges over the supermartingales that are bounded below, and we define the *lower expectation* of ξ given s by

$$\underline{\mathbb{E}}(\xi \mid s) := -\overline{\mathbb{E}}(-\xi \mid s).$$

If E is any subset of Ω, its *upper* and *lower probability* given a situation s are defined by

$$\overline{\mathbb{P}}(E \mid s) := \overline{\mathbb{E}}(\mathbb{I}_E \mid s), \quad \underline{\mathbb{P}}(E \mid s) := \underline{\mathbb{E}}(\mathbb{I}_E \mid s), \tag{7}$$

respectively. In what follows we sometimes refer to functions $\xi : \Omega \to \mathbb{R}$ as *variables*, to functions $\xi : \Omega \to [-\infty, \infty]$ as *extended variables*, and to sets $E \subseteq \Omega$ as *events*.

Lemma 1. *For all situations s and all variables $\xi : \Omega \to (-\infty, \infty)$, $\underline{\mathbb{E}}(\xi \mid s) \leq \overline{\mathbb{E}}(\xi \mid s)$. In particular, $\underline{\mathbb{P}}(E \mid s) \leq \overline{\mathbb{P}}(E \mid s)$ for all events $E \subseteq \Omega$.*

Proof. Suppose $\underline{\mathbb{E}}(\xi \mid s) > \overline{\mathbb{E}}(\xi \mid s)$, i.e., $\overline{\mathbb{E}}(\xi \mid s) + \overline{\mathbb{E}}(-\xi \mid s) < 0$. Then there exist supermartingales $\mathcal{S}^1$ and $\mathcal{S}^2$ such that $\mathcal{S}^1(s) + \mathcal{S}^2(s) < 0$ and, on $\Gamma(s)$, $\liminf_n \mathcal{S}^1_n \geq \xi$ and $\liminf_n \mathcal{S}^2_n \geq -\xi$. Then the supermartingale $\mathcal{S} := \mathcal{S}^1 + \mathcal{S}^2$ satisfies $\mathcal{S}(s) < 0$ and $\liminf \mathcal{S}_n(\omega) \geq 0$ for all $\omega \in \Gamma(s)$. Let us show that this is impossible. Set $\epsilon := -\mathcal{S}(s)$. By coherence (see (4)), Reality can choose the outcomes after the situation s so that

$$\mathcal{S}(s) = \mathcal{S}(\omega^k) \geq \mathcal{S}(\omega^{k+1}) - \epsilon/4 \geq \mathcal{S}(\omega^{k+2}) - \epsilon/4 - \epsilon/8 \geq \cdots,$$

where k is the level of s and $\omega \in \Gamma(s)$ is the realized path; this path ω will satisfy $\liminf_n \mathcal{S}_n(\omega) \leq \limsup_n \mathcal{S}_n(\omega) \leq \mathcal{S}_k(\omega) + \epsilon/2 < 0$. $\square$

Important special cases are where $s = \square$ (unconditional upper and lower expectations and probabilities). We set $\overline{\mathbb{E}}(\xi) := \overline{\mathbb{E}}(\xi \mid \square)$, $\underline{\mathbb{E}}(\xi) := \underline{\mathbb{E}}(\xi \mid \square)$, $\overline{\mathbb{P}}(E) := \overline{\mathbb{P}}(E \mid \square)$, and $\underline{\mathbb{P}}(E) := \underline{\mathbb{P}}(E \mid \square)$. We say that an event E is *almost certain*, or happens *almost surely* (a.s.), if $\underline{\mathbb{P}}(E) = 1$; in this case we will also say that E, considered as a property of $\omega \in \Omega$, holds for *almost all* ω. More generally, we say that E holds *almost surely on* B (or *for almost all* $\omega \in B$), for another event B, if the event $(B \Rightarrow E) := (B^c \cup E)$ is almost certain. An event E is *almost impossible*, or *null*, if $\overline{\mathbb{P}}(E) = 0$.

In [6] we defined the lower probability of an event E as $1 - \overline{\mathbb{P}}(E^c)$. The following lemma says that this definition is equivalent to our current definition.

Lemma 2. *For each event $E \subseteq \Omega$ and each situation s,*

$$\underline{\mathbb{P}}(E \mid s) = 1 - \overline{\mathbb{P}}(E^c \mid s).$$

Proof. By (3), we have $\overline{\mathbb{E}}(\xi + c \mid s) = \overline{\mathbb{E}}(\xi \mid s) + c$ for all $\xi : \Omega \to \mathbb{R}$ and $c \in \mathbb{R}$. Therefore,

$$\begin{aligned}\underline{\mathbb{P}}(E \mid s) &= \underline{\mathbb{E}}(\mathbb{I}_E \mid s) = -\overline{\mathbb{E}}(-\mathbb{I}_E \mid s) \\ &= 1 - \overline{\mathbb{E}}(1 - \mathbb{I}_E \mid s) = 1 - \overline{\mathbb{E}}(\mathbb{I}_{E^c} \mid s) = 1 - \overline{\mathbb{P}}(E^c \mid s). \quad \square\end{aligned}$$

The following lemma establishes the fundamental property of subadditivity of game-theoretic probability.

Lemma 3. *For any sequence of events $E_1, E_2, \dots$ and any situation s, it is true that*

$$\overline{\mathbb{P}}\left(\bigcup_{k=1}^{\infty} E_k \mid s\right) \le \sum_{k=1}^{\infty} \overline{\mathbb{P}}\left(E_k \mid s\right).$$

In particular, the union of a sequence of null events is null.

Proof. Assume, without loss of generality, that $s = \Box$. Let $\epsilon > 0$ be arbitrarily small. For each k choose a supermartingale $\mathcal{S}^k$ (automatically positive, by coherence) such that $\liminf_n \mathcal{S}_n^k \ge \mathbb{I}_{E_k}$ and $\mathcal{S}_0^k \le \overline{\mathbb{P}}(E_k) + \epsilon/2^k$. It is easy to check that the sum $\sum_{k=1}^{\infty} \mathcal{S}^k$ will be a supermartingale (cf. (1)) that satisfies $\mathcal{S}_0 \le \sum_{k=1}^{\infty} \overline{\mathbb{P}}(E_k) + \epsilon$ and $\liminf_n \mathcal{S}_n \ge \mathbb{I}_{E_k}$ for all k. $\Box$

Equivalent definitions of game-theoretic expectation and probability

The following proposition is our main statement of equivalence.

Lemma 4. *The definition of upper expectation will not change if we replace the* $\liminf$ *in (6) by* $\limsup$. *This definition is also equivalent to*

$$\overline{\mathbb{E}}(\xi \mid s) := \inf\left\{\mathcal{S}_0 \mid \forall \omega \in \Gamma(s) : \lim_{n\to\infty} \mathcal{S}_n(\omega) \ge \xi(\omega)\right\}$$

where $\mathcal{S}$ ranges over the class $\mathbf{L}$ of all bounded below supermartingales for which $\lim_{n\to\infty} \mathcal{S}_n(\omega)$ exists (with $\lim_{n\to\infty} \mathcal{S}_n(\omega) = \infty$ allowed) for all $\omega \in \Omega$.

Proof. We will only consider the case where s is a post-situation; the case of a pre-situation is completely analogous. Without loss of generality assume $s = \Box$. Let a bounded below supermartingale $\mathcal{S}$ satisfy the inequality

$$\forall \omega \in \Omega : \limsup_{n\to\infty} \mathcal{S}_n(\omega) \ge \xi(\omega)$$

(cf. (6)) and let $\epsilon \in (0,1)$. It suffices to show that there exists $\mathcal{S}^* \in \mathbf{L}$ such that $\mathcal{S}_0^* \le \mathcal{S}_0 + \epsilon$ and

$$\forall \omega \in \Omega : \lim_{n\to\infty} \mathcal{S}_n^*(\omega) \ge \xi(\omega).$$

Setting $\mathcal{S}' := (\mathcal{S} - C)/(\mathcal{S}_0 - C)$, where C is any constant satisfying $C < \inf \mathcal{S}$, we obtain a positive supermartingale satisfying $\mathcal{S}_0' = 1$.

The idea is now to use the standard proof of Doob's convergence theorem (see, e.g., [6], Lemma 4.5). Let $[a_i, b_i]$, $i = 1, 2, \dots$, be an enumeration of all intervals with $0 < a_i < b_i$ and both end-points rational. For each i one can define a positive supermartingale $\mathcal{S}^i$ with $\mathcal{S}_0^i = 1$ diverging to ∞ when $\liminf_n \mathcal{S}_n' < a_i$ and $\limsup_n \mathcal{S}_n' > b_i$. The construction of $\mathcal{S}^i$ is standard: set $\tau_0^i := 0$ and, for $k = 1, 2, \dots$,

$$\sigma_k^i := \min\{n > \tau_{k-1}^i \mid \mathcal{S}_n' > b_i\}, \quad \tau_k^i := \min\{n > \sigma_k^i \mid \mathcal{S}_n' < a_i\}; \tag{8}$$

define $\mathcal{S}^i$ by the requirement that

$$\mathcal{S}^i_n := \begin{cases} \mathcal{S}^i_{n-1} + \mathcal{S}'_n - \mathcal{S}'_{n-1} & \text{if } \mathcal{S}^i_{n-1} < \infty \text{ and } \exists k : \tau^i_{k-1} < n \le \sigma^i_k \\ \mathcal{S}^i_{n-1} & \text{otherwise} \end{cases} \tag{9}$$

for all $n \in \mathbb{N}$. Now we can set

$$\mathcal{T}_n := \sum_{i=1}^{\infty} 2^{-i} \mathcal{S}^i_n \tag{10}$$

and $\mathcal{S}^* := \mathcal{S} + \epsilon\mathcal{T}$.

Let us check that $\mathcal{S}^* \in \mathbf{L}$. Since $\mathcal{S}^*$ is bounded below, we are only required to check that $\mathcal{S}^*_n(\omega)$ converges (perhaps to ∞) as $n \to \infty$ for all $\omega \in \Omega$. Fix $\omega \in \Omega$.

If $\mathcal{S}_n(\omega) = \infty$ for some n, there exists i such that $\mathcal{S}^i_n(\omega) = \infty$ from some n on (take any i such that $\mathcal{S}_n(\omega) < a_i$ strictly before $\mathcal{S}_n(\omega)$ becomes infinite for the first time), and so we have $\mathcal{T}_n(\omega) = \infty$ and $\mathcal{S}^*_n(\omega) = \infty$ from some n on. Therefore, we will assume that $\mathcal{S}_n(\omega) < \infty$ for all n.

If $\mathcal{S}_n(\omega)$ converges to ∞, $\mathcal{S}^*_n(\omega)$ also converges to ∞. If $\mathcal{S}_n(\omega)$ does not converge, there exists i such that $\mathcal{S}^i_n(\omega) \to \infty$ (take any i satisfying $\liminf_n \mathcal{S}_n(\omega) < a_i < b_i < \limsup_n \mathcal{S}_n(\omega)$), and so we have $\mathcal{T}_n(\omega) \to \infty$ and $\mathcal{S}^*_n(\omega) \to \infty$. It remains to consider the case $\mathcal{S}_n(\omega) \to p$ for some $p \in \mathbb{R}$.

Suppose $\mathcal{S}_n(\omega) \to p$, $p \in \mathbb{R}$, but $\mathcal{S}^*_n(\omega)$ does not converge (not even to ∞). Choose a non-empty interval (a, b) such that $\liminf_n \mathcal{S}^*_n(\omega) < a < b < \limsup_n \mathcal{S}^*_n(\omega)$ and set $c := b - a$. Take any $N \in \mathbb{N}$ such that $\mathcal{S}^*_N(\omega) > b$ and $|\mathcal{S}_n(\omega) - p| < c/4$ for all $n \ge N$. It is clear that $\mathcal{S}_n(\omega) - \mathcal{S}_N(\omega) > -c/2$ and $\mathcal{S}^i_n(\omega) - \mathcal{S}^i_N(\omega) > -c/2$ for all i and all $n \ge N$. This implies $\mathcal{S}^*_n(\omega) - \mathcal{S}^*_N(\omega) > -c$ for all $n \ge N$, and so contradicts the fact that $\mathcal{S}^*_n(\omega) - \mathcal{S}^*_N(\omega) < -c$ for some $n \ge N$ (namely, for any $n \ge N$ satisfying $\mathcal{S}^*_n(\omega) < a$). $\square$

Notice that in the proof of Lemma 4 we do not really need Axiom 5: countable combinations of gambles are irrelevant. Despite the appearance of an infinite sum in (10), for each n the increment of $\mathcal{T}_n$ can be represented (assuming $\mathcal{T}_{n-1} < \infty$) as

$$\mathcal{T}_n - \mathcal{T}_{n-1} = \sum_{i=1}^{\infty} 2^{-i} (\mathcal{S}^i_n - \mathcal{S}^i_{n-1}) = \left(\sum_{i=1}^{\infty} w_i \right) (\mathcal{S}'_n - \mathcal{S}'_{n-1}),$$

where $w_i \in \{0, 2^{-i}\}$ (which makes the series $\sum_{i=1}^{\infty} w_i$ convergent). Since $\mathcal{S}'$ is a supermartingale, $\mathcal{T}_n - \mathcal{T}_{n-1}$, as function of the nth outcome x_n, has negative $\mathcal{E}_{p_n}$-superexpectation. This argument for $\mathcal{T}$ being a supermartingale does not depend on Axiom 5.

Replacing the $\liminf$ in (6) by inf or sup does change the definition. If we replace the $\liminf$ by inf, we will have $\overline{\mathbb{E}}(\xi \mid s) = \sup_{\omega \in \Gamma(s)} \xi(\omega)$.

Example 1. Set

$$\overline{\mathbb{E}}_1(\xi) := \inf\left\{\mathcal{S}_0 \mid \forall\omega \in \Omega : \sup_n \mathcal{S}_n(\omega) \geq \xi(\omega)\right\}.$$

It is always true that $\overline{\mathbb{E}}_1(\xi) \leq \overline{\mathbb{E}}(\xi)$. Consider the coin-tossing protocol ([6], Section 8.2), which is the special case of Protocol 1 with $\mathcal{X} = \{0, 1\}$, $\mathcal{P}$ a one-element set, and $\mathcal{E}_p(f) = (f(0) + f(1))/2$. The sample space Ω can now be identified with $\{0, 1\}^\infty$: as the predictions are not informative, we can omit them. For each $\epsilon \in (0, 1)$ there exists a bounded positive variable ξ such that $\overline{\mathbb{E}}(\xi) = 1$ and $\overline{\mathbb{E}}_1(\xi) = \epsilon$.

Proof. Let us demonstrate the following equivalent statement: for any $C > 1$ there exists a bounded positive variable ξ such that $\overline{\mathbb{E}}_1(\xi) = 1$ and $\overline{\mathbb{E}}(\xi) = C$. Fix such a C. Define $\psi : \Omega \to [0, \infty]$ by the requirement $\psi(\omega) := 2^n$ where n is the number of 1s at the beginning of ω: $n := \max\{i \mid \omega_1 = \cdots = \omega_i = 1\}$. It is obvious that $\overline{\mathbb{E}}_1(\psi) = 1$ and $\overline{\mathbb{E}}(\psi) = \infty$. However, ψ is unbounded. We can always find $A > 1$ such that $\overline{\mathbb{E}}(\min(\psi, A)) = C$ (as the function $a \mapsto \overline{\mathbb{E}}(\min(\psi, a))$ is continuous). Since $\overline{\mathbb{E}}_1(\min(\psi, A)) = 1$, we can set $\xi := \min(\psi, A)$. $\square$

Game-theoretic probability is a special case of game-theoretic expectation, and in this special case it is possible to replace $\liminf$ not only by $\limsup$ but also by $\sup$, provided we restrict our attention to positive supermartingales (simple examples show that this qualification is necessary). By coherence, the definition of conditional upper probability $\overline{\mathbb{P}}$ can be rewritten as

$$\overline{\mathbb{P}}(E \mid s) := \inf\left\{\mathcal{S}_0 \mid \liminf_{n\to\infty} \mathcal{S}_n(\omega) \geq 1 \text{ for all } \omega \in E \cap \Gamma(s)\right\}, \tag{11}$$

$\mathcal{S}$ ranging over the positive supermartingales.

Lemma 5. *The definition of upper probability will not change if we replace the* $\liminf_{n\to\infty}$ *in (11) by* $\limsup_{n\to\infty}$ *or by* $\sup_n$.

It is obvious that the definition will change if we replace the $\liminf_{n\to\infty}$ by $\inf_n$: in this case we will have

$$\overline{\mathbb{P}}(E \mid s) = \begin{cases} 0 & \text{if } E \cap \Gamma(s) = \emptyset \\ 1 & \text{otherwise.} \end{cases}$$

Proof. It suffices to prove that the definition will not change if we replace the $\liminf_{n\to\infty}$ in (11) by $\sup_n$. This is obvious: if a prudent (i.e., resulting in a positive capital) strategy for Skeptic ensures $\forall\omega \in E \cap \Gamma(s) : \sup_n \mathcal{K}_n(\omega) > 1$ (it is obvious that it does not matter whether we have $\geq$ or $>$ in (11)), Skeptic can also ensure $\forall\omega \in E \cap \Gamma(s) : \liminf_{n\to\infty} \mathcal{K}_n(\omega) > 1$ by stopping whenever his capital $\mathcal{K}_n$ exceeds 1. $\square$

3 Lévy's zero-one law

The following simple theorem is our main result.

Theorem 1. *Let $\xi : \Omega \to (-\infty, \infty]$ be bounded from below. For almost all $\omega \in \Omega$,*

$$\liminf_{n\to\infty} \overline{\mathbb{E}}(\xi \mid \omega^n) \geq \xi(\omega). \tag{12}$$

This theorem is a game-theoretic version of Lévy's zero-one law. The name of this result might look puzzling now; connections with zero-one phenomena will be explored in the next section.

Proof of Theorem 1. It suffices to construct a positive supermartingale starting from 1 that tends to ∞ on the paths $\omega \in \Omega$ for which (12) is not true. Without loss of generality we will assume ξ to be positive (we can always redefine $\xi := \xi - \inf \xi$). According to Lemma 3 (second statement), we can, without loss of generality, replace "for which (12) is not true" by

$$\liminf_{n\to\infty} \overline{\mathbb{E}}(\xi \mid \omega^n) < a < b < \xi(\omega) \tag{13}$$

where a and b are given positive rational numbers such that $a < b$. The supermartingale is defined as the capital process of the following strategy. Let $\omega \in \Omega$ be the sequence of moves chosen by Forecaster and Reality. Start with 1 monetary unit. Wait until $\overline{\mathbb{E}}(\xi \mid \omega^n) < a$ (if this never happens, do nothing, i.e., always choose constant $f_n = \mathcal{K}_{n-1}$). As soon as this happens, choose a positive supermartingale $\mathcal{S}_1$ starting from a, $\mathcal{S}_1(\omega^n) = a$, whose upper limit exceeds ξ on $\Gamma(\omega^n)$. Maintain capital $\mathcal{S}_1/a$ until $\mathcal{S}_1$ reaches a value $m_1 > b$ (at which point Skeptic's capital is $m_1/a > b/a$). After that do nothing until $\overline{\mathbb{E}}(\xi \mid \omega^n) < a$. As soon as this happens, choose a positive supermartingale $\mathcal{S}_2$ starting from a, $\mathcal{S}_2(\omega^n) = a$, whose upper limit exceeds ξ on $\Gamma(\omega^n)$. Maintain capital $(m_1/a^2)\mathcal{S}_2$ until $\mathcal{S}_2$ reaches a value $m_2 > b$ (at which point Skeptic's capital is $m_1 m_2/a^2 > (b/a)^2$). After that do nothing until $\overline{\mathbb{E}}(\xi \mid \omega^n) < a$. As soon as this happens, choose a positive supermartingale $\mathcal{S}_3$ starting from a whose upper limit exceeds ξ on $\Gamma(\omega^n)$. Maintain capital $(m_1 m_2/a^3)\mathcal{S}_3$ until $\mathcal{S}_3$ reaches a value $m_3 > b$ (at which point Skeptic's capital is $m_1 m_2 m_3/a^3 > (b/a)^3$). And so on. On the event (13) Skeptic's capital will tend to infinity. $\square$

Specializing Theorem 1 to the indicators of events, we obtain:

Corollary 1. *Let E be any event. For almost all $\omega \in E$,*

$$\overline{\mathbb{P}}(E \mid \omega^n) \to 1 \tag{14}$$

as $n \to \infty$.

It is easy to check that we cannot replace the $\geq$ in (12) by $=$, even when ξ is the indicator of an event. For example, suppose that $\mathcal{X} = \{0, 1\}$, $\mathcal{P}$ is a one-element set, and $\mathcal{E}_p$ is the sup functional: $\mathcal{E}_p(f) := \sup_{x\in\mathcal{X}} f(x)$ for all f.

Since the predictions are not informative, Ω can be identified with $\{0,1\}^\infty$. If E consists of binary sequences containing only finitely many 1s, $\overline{\mathbb{P}}(E \mid \omega^n) = 1$ for all ω and n; therefore,

$$\liminf_{n\to\infty} \overline{\mathbb{P}}(E \mid \omega^n) = \mathbb{I}_E(\omega)$$

is violated for all $\omega \in E^c$, and $\overline{\mathbb{P}}(E^c) = 1$.

The case of a determinate expectation or probability

In Section 41 of [5] (pp. 128–130), Lévy states his zero-one law in terms of a property E that a sequence $X_1, X_2, \ldots$ of variables might or might not have. He writes $\mathrm{Pr.}\{E\}$ for the initial probability of E, and $\mathrm{Pr}_n\{E\}$ for its probability after $X_1, \ldots, X_n$ is known. He remarks that if $\mathrm{Pr.}\{E\}$ is well defined (i.e., if E is measurable), then the conditional probabilities $\mathrm{Pr}_n\{E\}$ are also well defined. Then he states the law as follows (our translation from the French):

> Except in cases that have probability zero, if $\mathrm{Pr.}\{E\}$ is determined, then $\mathrm{Pr}_n\{E\}$ tends, as n tends to infinity, to one if the sequence $X_1, X_2, \ldots$ verifies the property E, and to zero in the contrary case.

In this subsection we will derive a game-theoretic result that resembles Lévy's statement of his result. We will be concerned with variables ξ satisfying $\overline{\mathbb{E}}(\xi) = \underline{\mathbb{E}}(\xi)$ and events E satisfying $\overline{\mathbb{P}}(E) = \underline{\mathbb{P}}(E)$.

Lemma 6. *Suppose a variable ξ satisfies $\overline{\mathbb{E}}(\xi) = \underline{\mathbb{E}}(\xi)$. Then it is almost certain that it also satisfies $\overline{\mathbb{E}}(\xi \mid \omega^n) = \underline{\mathbb{E}}(\xi \mid \omega^n)$ for all n.*

Proof. For any positive ϵ, there exist supermartingales $\mathcal{S}_1$ and $\mathcal{S}_2$ that start at $\overline{\mathbb{E}}(\xi) + \epsilon/2$ and $\overline{\mathbb{E}}(-\xi) + \epsilon/2$, respectively, and tend to ξ or more and to $-\xi$ or more, respectively. Set $\mathcal{S} := \mathcal{S}_1 + \mathcal{S}_2$. The assumption $\overline{\mathbb{E}}(\xi) = \underline{\mathbb{E}}(\xi)$ can also be written $\overline{\mathbb{E}}(\xi) + \overline{\mathbb{E}}(-\xi) = 0$. So the positive (by coherence) supermartingale $\mathcal{S}$ begins at ϵ and tends to 0 or more on all $\omega \in \Omega$.

Fix n and $\delta > 0$, and let A be the event that

$$\overline{\mathbb{E}}(\xi \mid \omega^n) + \overline{\mathbb{E}}(-\xi \mid \omega^n) > \delta.$$

By the definition of conditional upper expectation,

$$\mathcal{S}_1(\omega^n) \geq \overline{\mathbb{E}}(\xi \mid \omega^n)$$

and

$$\mathcal{S}_2(\omega^n) \geq \overline{\mathbb{E}}(-\xi \mid \omega^n).$$

So $\mathcal{S}$ exceeds δ on A. So the upper probability of A is at most ϵ/δ. Since ϵ may be as small as we like for fixed δ, this shows that A has upper probability zero. Letting δ range over the positive rational numbers and n over $\{0,1,2,\ldots\}$ and applying the second part of Lemma 3, we obtain the statement of the lemma. $\square$

Corollary 2. *Let ξ be a bounded variable for which $\overline{\mathbb{E}}(\xi) = \underline{\mathbb{E}}(\xi)$. Then, almost surely, $\overline{\mathbb{E}}(\xi \mid \omega^n) = \underline{\mathbb{E}}(\xi \mid \omega^n) \to \xi(\omega)$ as $n \to \infty$.*

Proof. By Theorem 1,

$$\liminf_{n\to\infty} \overline{\mathbb{E}}(\xi \mid \omega^n) \geq \xi(\omega)$$

for almost all $\omega \in \Omega$ and (applying the theorem to $-\xi$)

$$\limsup_{n\to\infty} \underline{\mathbb{E}}(\xi \mid \omega^n) \leq \xi(\omega)$$

for almost all $\omega \in \Omega$. □

Our definitions (7) make it easy to obtain the following corollaries for events.

Corollary 3. *Suppose an event E satisfies $\overline{\mathbb{P}}(E) = \underline{\mathbb{P}}(E)$. Then, almost surely, it also satisfies $\overline{\mathbb{P}}(E \mid \omega^n) = \underline{\mathbb{P}}(E \mid \omega^n)$ for each n.*

Corollary 4. *Let E be an event for which $\overline{\mathbb{P}}(E) = \underline{\mathbb{P}}(E)$. Then, almost surely, $\overline{\mathbb{P}}(E \mid \omega^n) = \underline{\mathbb{P}}(E \mid \omega^n) \to \mathbb{I}_E$ as $n \to \infty$.*

4 More explicit zero-one laws

In this section we will prove Bártfai and Révész's [1] zero-one law for our general prediction protocol (Protocol 1), and then deduce two corollaries for the case of independent trials (to be defined later): Kolmogorov's zero-one law and the ergodicity of Bernoulli shifts. The first corollary is deduced from Bártfai and Révész's zero-one law, and the second one from Lévy's zero-one law. Both corollaries were proved in [7] directly. All these results are more general than the corresponding measure-theoretic results; see [6], Section 8.1, for relations between measure-theoretic results and their game-theoretic counterparts.

For $\omega = p_1 x_1 p_2 x_2 \ldots \in \Omega$ and $n \in \mathbb{N}$, we let $\omega_n \in \mathcal{P}_n \times \mathcal{X}$ stand for the pair (p_n, x_n).

Bártfai and Révész's zero-one law

For each $N \in \mathbb{N}$, let $\mathcal{F}_N$ be the set of all events E that are properties of $(\omega_N, \omega_{N+1}, \ldots)$ only (i.e., E such that, for all $\omega, \omega' \in \Omega$, $\omega' \in E$ whenever $\omega \in E$ and $\omega'_n = \omega_n$ for all $n \geq N$). Let us say that a general prediction protocol (determined by the sequence $\mathcal{P}_1, \mathcal{P}_2, \ldots$ and the function $\mathcal{E}$) is *δ-mixing*, for $\delta \in [0, 1)$, if there exists a function $a : \mathbb{N} \to \mathbb{N}$ such that

$$\overline{\mathbb{P}}(E \mid \omega^n) - \overline{\mathbb{P}}(E) \leq \delta \quad \text{a.s.} \tag{15}$$

for each $n \in \mathbb{N}$ and each $E \in \mathcal{F}_{n+a(n)}$.

First we state an approximate zero-one law (a game-theoretic analogue of the main result of [1]). By a *tail event* we mean an element of $\cap_N \mathcal{F}_N$.

Theorem 2. *Let $\delta \in [0,1)$ and let a sequence $\mathcal{P}_1, \mathcal{P}_2, \dots$ and a function $\mathcal{E}$ be such that the general prediction protocol is δ-mixing. If E is a tail event, then $\overline{\mathbb{P}}(E) = 0$ or $\overline{\mathbb{P}}(E) \geq 1 - \delta$.*

Proof. Since E is a tail event, (15) holds for all n. By Corollary 1, $1 - \overline{\mathbb{P}}(E) \leq \delta$ holds on E almost surely. In other words, $1 - \overline{\mathbb{P}}(E) \leq \delta$ unless E is null. □

Let us say that an event E is *unprobabilized* if $\underline{\mathbb{P}}(E) < \overline{\mathbb{P}}(E)$. An important special case of Theorem 2 is the following zero-one law for "weakly dependent" trials (cf. Corollary 1 in [1]).

Corollary 5. *Let $\delta \in [0,1/2)$ and let a sequence $\mathcal{P}_1, \mathcal{P}_2, \dots$ and a function $\mathcal{E}$ be such that Protocol 1 is δ-mixing. Every tail event is almost certain, almost impossible, or unprobabilized.*

Proof. It suffices to apply Theorem 2 to the tail events E and E^c. □

It is easy to strengthen Theorem 2 by modifying the notion of a δ-mixing protocol. Let us say that the protocol is *asymptotically δ-mixing*, for $\delta \in [0,1)$, if (15) holds for each $n \in \mathbb{N}$ and each tail event E. Bártfai and Révész [1] do not introduce this notion (more precisely, its measure-theoretic version) explicitly, but they do introduce two notions intermediate between δ-mixing and asymptotic δ-mixing, which they call stochastic δ-mixing and δ-mixing in mean. The following proposition is similar to (but much simpler than) Theorems 2 and 3 in [1].

Proposition 1. *Let $\delta \in [0,1)$. The following two conditions on the general prediction protocol are equivalent:*

1. *The protocol is asymptotically δ-mixing.*
2. *Every tail event E satisfies $\overline{\mathbb{P}}(E) = 0$ or $\overline{\mathbb{P}}(E) \geq 1 - \delta$.*

Proof. The argument of Theorem 2 shows that the first condition implies the second. Let us now assume the second condition and demonstrate the first. Let $n \in \mathbb{N}$ and E be a tail event. If $\overline{\mathbb{P}}(E) = 0$, then $\overline{\mathbb{P}}(E \mid \omega^n) = 0$ a.s. can be proved similarly to the proof of Lemma 6, and so (15) holds. If $\overline{\mathbb{P}}(E) \geq 1 - \delta$, (15) is vacuous. □

Kolmogorov's zero-one law

In this section we will apply Theorem 2 to deduce a game-theoretic version of Kolmogorov's zero-one law. In our next protocol, Forecaster's moves p_n, $n = 1, 2, \dots$, are fixed at the beginning of the game; in other words, $\mathcal{P}_n$ are one-element sets. Since the predictions are no longer informative, we can remove Forecaster from the protocol, and include his superexpectation functionals as parameters of the protocol.

PROTOCOL 2. INDEPENDENT TRIALS
Parameters: superexpectation functionals $\mathcal{E}_1, \mathcal{E}_2, \dots$

Protocol:
Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
  Skeptic announces $f_n \in (-\infty, \infty]^{\mathcal{X}}$ such that $\mathcal{E}_n(f_n) \leq \mathcal{K}_{n-1}$.
  Reality announces $x_n \in \mathcal{X}$.
  $\mathcal{K}_n := f_n(x_n)$.
END FOR

Now we can redefine $\Omega := \mathcal{X}^\infty$. As before, an event $E \subseteq \Omega$ is called a *tail event* if any sequence in Ω that agrees from some point onwards with a sequence in E is also in E.

Corollary 6 ([7])**.** *For all tail events E in the protocol of independent trials, $\overline{\mathbb{P}}(E) \in \{0, 1\}$.*

Proof. For each $n \in \mathbb{N}$ and each $E \in \mathcal{F}_{n+1}$, $\overline{\mathbb{P}}(E \mid \omega^n)$ does not depend on ω. This implies $\overline{\mathbb{P}}(E \mid \omega^n) = \overline{\mathbb{P}}(E)$. Therefore, the protocol is 0-mixing, and it remains to apply Theorem 2. □

We say that an event E is *fully unprobabilized* if $\underline{\mathbb{P}}(E) = 0$ and $\overline{\mathbb{P}}(E) = 1$. Since complements of tail events are also tail events, we obtain the following corollary to Corollary 6.

Corollary 7 ([7])**.** *If E is a tail event in Protocol 2 (independent trials), then E is almost certain, almost impossible, or fully unprobabilized.*

Ergodicity of Bernoulli shifts

The protocol of this subsection is even more specialized than Protocol 2: Forecaster always chooses the same prediction.

PROTOCOL 3. IDENTICALLY PRICED TRIALS
Parameter: superexpectation functional $\mathcal{E}$
Protocol:
Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
  Skeptic announces $f_n \in (-\infty, \infty]^{\mathcal{X}}$ satisfying $\mathcal{E}(f_n) \leq \mathcal{K}_{n-1}$.
  Reality announces $x_n \in \mathcal{X}$.
  $\mathcal{K}_n := f_n(x_n)$.
END FOR

We write θ for the shift operator, which deletes the first element from a sequence in $\mathcal{X}^\infty$:

$$\theta : x_1 x_2 x_3 \ldots \mapsto x_2 x_3 \ldots.$$

We call an event E in Protocol 3 *weakly invariant* if $\theta E \subseteq E$. In accordance with standard terminology, we call an event E *invariant* if $E = \theta^{-1} E$.

Lemma 7. *E is invariant if and only if both E and E^c are weakly invariant.*

Proof. We will give the simple argument from [7]. If E is invariant, then E^c is also invariant, because the inverse map commutes with complementation. Hence in this case both E and E^c are weakly invariant.

Conversely suppose that $\theta E \subseteq E$ and $\theta E^c \subseteq E^c$. The first inclusion is equivalent to $E \subseteq \theta^{-1}E$ and the second is equivalent to $E^c \subseteq \theta^{-1}E^c$. Since the right-hand sides of the last two inclusions are disjoint, these inclusions are in fact equalities. ☐

The following corollary asserts the ergodicity of Bernoulli shifts.

Corollary 8 ([7])**.** *For all weakly invariant events E in the protocol of identically priced trials, $\overline{\mathbb{P}}(E) \in \{0, 1\}$.*

Proof. For weakly invariant events, $\overline{\mathbb{P}}(E \mid \omega^n) \leq \overline{\mathbb{P}}(E)$. Let E be weakly invariant. By (14), for almost all $\omega \in E$ it is true that $\overline{\mathbb{P}}(E) = 1$. Therefore, $\overline{\mathbb{P}}(E)$ is either 0 or 1. ☐

In view of Lemma 7 we obtain the following corollary to Corollary 8.

Corollary 9 ([7])**.** *If E is an invariant event in Protocol 3 (identically priced trials), then E is almost certain, almost impossible, or fully unprobabilized.*

5 An implication for the foundations of game-theoretic probability theory

In this section we return to Protocol 1. Let ξ be a bounded variable, and let $s := \Box$. We will obtain an equivalent definition of the upper expectation $\overline{\mathbb{E}}(\xi \mid s) = \overline{\mathbb{E}}(\xi)$ if we replace the phrase "for all $\omega \in \Gamma(s)$" in (6) by "for almost all $\omega \in \Omega$". It turns out that if we do so, the infimum in (6) becomes attained; namely, it is attained by the supermartingale $\mathcal{S}_n(\omega) := \overline{\mathbb{E}}(\xi \mid \omega^n)$. (This fact is the key technical tool used in [8].) In view of Theorem 1, to prove this statement it suffices to check that $\mathcal{S}_n(\omega) := \overline{\mathbb{E}}(\xi \mid \omega^n)$ is indeed a supermartingale.

Theorem 3. *Let $\xi : \Omega \to (-\infty, \infty]$ be bounded from below. Then $\mathcal{S}_n(\omega) := \overline{\mathbb{E}}(\xi \mid \omega^n)$ is a supermartingale.*

Proof. As a first step, let us check that, for any $\epsilon > 0$, $\mathcal{S}^\epsilon_n(\omega) := \overline{\mathbb{E}}(\xi \mid \omega^n) + \epsilon 2^{-n}$ is a supermartingale, i.e., that

$$\mathcal{E}_p \, \overline{\mathbb{E}}(\xi \mid \omega^{n-1}px) \leq \overline{\mathbb{E}}(\xi \mid \omega^{n-1}) + \epsilon 2^{-n}$$

for all $p \in \mathcal{P}_n$, $\omega \in \Omega$, and $n \in \mathbb{N}$. The last inequality follows from the existence of a supermartingale $\mathcal{T}$ that starts from $\overline{\mathbb{E}}(\xi \mid \omega^{n-1}) + \epsilon 2^{-n}$ in the situation ω^{n-1} and ensures $\liminf_n \mathcal{T}_n \geq \xi$ on $\Gamma(\omega^{n-1})$: it is clear that such $\mathcal{T}$ will satisfy $\mathcal{T}(\omega^{n-1}px) \geq \overline{\mathbb{E}}(\xi \mid \omega^{n-1}px)$.

It remains to notice that the infimum of any set of supermartingales is again a supermartingale and that $\mathcal{S} = \inf_{\epsilon > 0} \mathcal{S}^\epsilon$. ☐

Example 2. Consider the coin-tossing protocol, as in Example 1. Let E be the set of all $\omega \in \Omega$ containing only finitely many 1s and let $\xi := \mathbb{I}_E$. The infimum in (6) is not attained: there exist no supermartingale $\mathcal{S}$ satisfying $\mathcal{S}_0 = \overline{\mathbb{E}}(\xi) = 0$ and $\liminf_{n\to\infty} \mathcal{S}_n(\omega) \geq \xi(\omega)$ for all $\omega \in \Omega$.

Proof. By coherence, such an $\mathcal{S}$ would be positive. Since its initial value is 0, $\mathcal{S}$ would be constant. □

Acknowledgements

Our thinking about Lévy's zero-one law was influenced by a preliminary draft of [2]. We are grateful to Gert de Cooman for his questions that inspired some of the results in Section 2 of this article, and to an anonymous referee of the journal version of [7] who pointed out to us the zero-one law in [1]. This work has been supported in part by EPSRC grant EP/F002998/1.

References

[1] Pál Bártfai and Pál Révész. On a zero-one law. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 7:43–47, 1967.

[2] Bernard Bru and Salah Eid. Jessen's theorem and Lévy's lemma: A correspondence. *Electronic Journal for History of Probability and Statistics*, 5(1), June 2009. To appear.

[3] Isaac P. Cornfeld, Sergei V. Fomin, and Yakov G. Sinai. *Ergodic Theory*, volume 245 of *Grundlehren der Mathematischen Wissenschaften*. Springer, New York, 1982.

[4] Andrei N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin, 1933. English translation: *Foundations of the Theory of Probability*. Chelsea, New York, 1950.

[5] Paul Lévy. *Théorie de l'addition des variables aléatoires*. Gauthier-Villars, Paris, 1937. Second edition: 1954.

[6] Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's Only a Game!* Wiley, New York, 2001.

[7] Akimichi Takemura, Vladimir Vovk, and Glenn Shafer. The generality of the zero-one laws. Submitted for publication in *Annals of the Institute of Statistical Mathematics*, April 2009.

[8] Vladimir Vovk. Prequential probabability: game-theoretic = measure-theoretic. The Game-Theoretic Probability and Finance project, `http://probabilityandfinance.com`, Working Paper 27, January 2009.

[9] Peter Whittle. *Probability via Expectation*. Springer, New York, fourth edition, 2000.