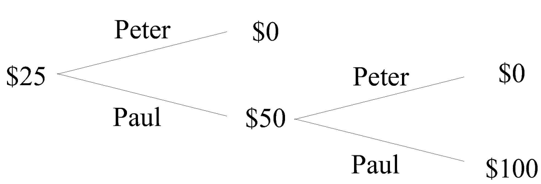
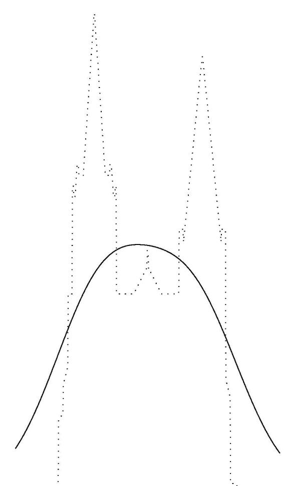
Lévy's zero-one law in game-theoretic probability

Glenn Shafer, Vladimir Vovk, and Akimichi Takemura





The Game-Theoretic Probability and Finance Project

Working Paper #29

First posted May 3, 2009. Last revised April 15, 2018.

Project web site:
http://www.probabilityandfinance.com

Abstract

We prove a non-stochastic version of Lévy's zero-one law, and deduce several corollaries from it, including non-stochastic versions of 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.

Contents

1 Introduction

In this article, we prove a game-theoretic version of Lévy's zero-one law. It applies in situations where standard statements of Lévy's zero-one law ([8], Section 41) do not apply, because we do not postulate a probability measure on outcomes. This is typical for game-theoretic probability: see, e.g., [9]. Upper and lower probabilities do emerge naturally in prediction protocols considered in game-theoretic probability, but in many cases lower probabilities are strictly less than the corresponding upper probabilities, and so they fall short of defining a probability measure.

The investigation of zero-one laws of game-theoretic probability was started in [10]. From our game-theoretic version of Lévy's zero-one law we deduce game-theoretic versions of Kolmogorov's zero-one law ([7], Appendix), the ergodicity of Bernoulli shifts (see, e.g., [4], Section 8.1, Theorem 1), and Bártfai and Révész's [2] zero-one law. The first two results have been established in [10], 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 basic prediction protocol and defining the game-theoretic notions of expectation and probability (upper and lower); these definitions are explored in Section 3. In Section 4 we prove Lévy's zero-one law for this protocol. In the second part of Section 4 we consider the special case of an event whose upper probability coincides with its lower probability; our version of Lévy's zero-one law for such events looks much more similar to the standard statement. Section 5 describes a useful application of Lévy's zero-one law to the foundations of game-theoretic probability theory. In Section 6 we derive two other non-stochastic zero-one laws (Kolmogorov's zero-one law and the ergodicity of Bernoulli shifts) as corollaries. In Section 7 we explain that our prediction protocol covers as special cases seemingly more general protocols considered in literature. This helps us to derive a non-stochastic version of one more zero-one law (Bártfai and Révész's) in Section 8.

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 $\overline{\mathbb{R}} := [-\infty, \infty]$, obtained from $\mathbb{R}$ by adding $-\infty$ or ∞ (or both). We set $0 \times \infty := 0$ and $\infty + (-\infty) := \infty$ (the operation $+$ extended to $\overline{\mathbb{R}}$ by $\infty + (-\infty) := \infty$ is denoted by $\dot{+}$ in [6]; in this article we will omit the dot in $\dot{+}$ since we do not need any other extensions of $+$). 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 $<$.

2 Game-theoretic expectation and probability

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

Let X be a set, and let $\overline{\mathbb{R}}^X$ stand for the set of all functions $f : X \to \overline{\mathbb{R}}$. A function $\mathcal{E} : \overline{\mathbb{R}}^X \to \overline{\mathbb{R}}$ is called an *outer probability content* if it satisfies the following four axioms:

1. If $f, g \in \overline{\mathbb{R}}^X$ satisfy $f \le g$, then $\mathcal{E}(f) \le \mathcal{E}(g)$.
2. If $f \in \overline{\mathbb{R}}^X$ and $c \in (0, \infty)$, then $\mathcal{E}(cf) = c\mathcal{E}(f)$.
3. If $f, g \in \overline{\mathbb{R}}^X$, then $\mathcal{E}(f + g) \le \mathcal{E}(f) + \mathcal{E}(g)$.
4. For each $c \in \mathbb{R}$, $\mathcal{E}(c) = c$, where the c in parentheses is the function in $\overline{\mathbb{R}}^X$ that is identically equal to c.

The function $\mathcal{E}$ is called a *superexpectation functional* (or *superexpectation* for brevity) if, in addition, it satisfies the following axiom (sometimes referred to as *σ-subadditivity on* $[0, \infty]^X$).

5. For any sequence of positive functions $f_1, f_2, \dots$ in $\overline{\mathbb{R}}^X$,

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

Replacing the $=$ in Axiom 2 with $\le$ leads to an equivalent statement because, for $c \in (0, \infty)$, $\mathcal{E}(cf) \le c\mathcal{E}(f) = c\mathcal{E}((1/c)cf) \le \mathcal{E}(cf)$. In presence of Axiom 1, we can allow $c \in \overline{\mathbb{R}}$ in Axiom 4 without changing the content of the latter. Axiom 4 implies $\mathcal{E}(0) = 0$ (so that we can allow $c = 0$ in Axiom 2). This, in combination with Axiom 1, implies

$$f \ge 0 \Longrightarrow \mathcal{E}(f) \ge 0. \tag{2}$$

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, for any $c \in \mathbb{R}$,

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

Axioms 1–5 are relaxations of the standard properties of the expectation functional: cf., e.g., Axioms 1–5 in [12] (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 [9]. Axiom 5 is convenient and often makes proofs easier. However, most of the results in this article hold without it, as pointed out by a referee. For our principal results, we first prove them assuming Axiom 5 but then give an additional argument to get rid of the reliance on it.

The most noticeable difference between what we call superexpectation functionals and the standard expectation functionals is that the former are defined for all functions $f : \mathcal{X} \to \overline{\mathbb{R}}$ whereas the latter are defined only for functions that are measurable w.r.t. to a given σ-algebra. The notion of superexpectation functional is more general since every expectation functional can be extended to the whole of $\overline{\mathbb{R}}^X$ 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. Superexpectation functionals have been studied in the past in many different contexts under different names; in our terminology we mainly follow [6], except that we abbreviate "outer probability content σ-subadditive on $[0,\infty]^X$" to "superexpectation functional". Upper previsions studied in the theory of imprecise probabilities are closely related to superexpectation functionals, one difference being that upper previsions are only defined on the bounded functions in $\overline{\mathbb{R}}^X$. There is a burgeoning literature, started by [1], on coherent measures of risk, which are close to being mappings $f \mapsto \mathcal{E}(-f)$, where $\mathcal{E}$ is an outer probability content; coherent measures of risk, however, are usually defined only for functions f that are measurable and do not take values $\pm\infty$.

Remark. It is sometimes useful to have the stronger form

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

of (2). Even the strong form (5) follows from Axioms 1–5. Indeed, if $f > 0$ but $\mathcal{E}(f) = 0$, Axioms 1–5 imply

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

(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 main prediction protocol that we consider in this article is as follows.

Protocol 1. Basic prediction protocol

Parameters: non-empty set $\mathcal{X}$
and outer probability contents $\mathcal{E}_1, \mathcal{E}_2, \dots$ on $\mathcal{X}$
Protocol:
Skeptic announces $\mathcal{K}_0 \in \overline{\mathbb{R}}$.
FOR $n = 1, 2, \dots$:
Skeptic announces $f_n \in \overline{\mathbb{R}}^{\mathcal{X}}$ such that $\mathcal{E}_n(f_n) \le \mathcal{K}_{n-1}$.
World announces $x_n \in \mathcal{X}$.
$\mathcal{K}_n := f_n(x_n)$.
END FOR

The set $\mathcal{X}$ will be called the *outcome space*; this set and the other parameters of the game (namely, the outer probability contents $\mathcal{E}_1, \mathcal{E}_2, \dots$ on $\mathcal{X}$) will be fixed until Section 7. An important special case is where $\mathcal{E}_1, \mathcal{E}_2, \dots$ are superexpectations on $\mathcal{X}$, but as we said, our principal results will not require this assumption.

At the beginning of each trial n Skeptic chooses a gamble represented as a function f_n on $\mathcal{X}$. After that World chooses the outcome x_n of this trial, which determines the payoff $f_n(x_n)$ of Skeptic's gamble. The gambles available to Skeptic at trial n are determined by $\mathcal{E}_n$. 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. Our definitions imply that Skeptic is allowed not to play at trial n (and thus keep his money intact) by choosing $f_n \equiv K_{n-1}$ (cf. Axiom 4). Property (4) reflects what is sometimes called the "coherence" of the protocol; in its absence the protocol becomes a money machine for Skeptic. (See also Lemma 1 below.) Protocol 1 covers the apparently more general case where the superexpectations $\mathcal{E}_n$ are not fixed in advance but chosen by a third player: see Section 7 below.

Remark. In [10] we considered a different but essentially equivalent prediction protocol ([10], Protocol 2). For connections with Peter Walley's theory of imprecise probabilities, see the recent article [5].

Remark. An apparently more general version of Protocol 1 is where, at each trial n, World chooses the outcome x_n from a set $\mathcal{X}_n$ which may depend on n and $\mathcal{E}_n$ is an outer probability content on $\mathcal{X}_n$. However, this version immediately reduces to our Protocol 1 by setting $\mathcal{X} := \bigcup_{n=1}^{\infty} \mathcal{X}_n$ and extending each $\mathcal{E}_n$ to $\overline{\mathbb{R}}^{\mathcal{X}}$ as $\mathcal{E}'_n(f) := \mathcal{E}_n(f|_{\mathcal{X}_n})$, $f|_{\mathcal{X}_n}$ being the restriction of f to $\mathcal{X}_n$.

We call the set $\Omega := \mathcal{X}^\infty$ of all infinite sequences of World's moves the *sample space*. The elements of the set $\mathcal{X}^* := \bigcup_{n=0}^{\infty} \mathcal{X}^n$ of all finite sequences of World's moves are called *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. If s is a situation and $x \in \mathcal{X}$, sx is the situation obtained from s by adding x on the right; therefore, $sx = x_1 \dots x_n x$ when $s = x_1 \dots x_n$. If s and t are two situations, we write $s \subseteq t$ when s is a prefix of t, and we write $s \subset t$ when $s \subseteq t$ and $s \ne t$. We will also be using derived notation such as $s \subseteq u \subseteq t$, $s \not\subseteq t$, and $s \supseteq t$.

The *length* $|s|$ of a situation $s \in \mathcal{X}^n$ is n (i.e., $|s|$ is the length of s as a finite sequence); in particular, $|\Box| = 0$. If $\omega \in \Omega$ and $n \in \{0, 1, \ldots\}$, ω^n is defined to be the unique situation of length n that is a prefix of ω. For $\omega = x_1 x_2 \ldots \in \Omega$ and $n \in \mathbb{N}$, we let $\omega_n \in \mathcal{X}$ stand for x_n.

A strategy Σ for Skeptic is a pair (Σ_0, Σ_1), where $\Sigma_0 \in \overline{\mathbb{R}}$ (informally, this is the initial capital chosen by Skeptic) and $\Sigma_1 : \mathcal{X}^* \to \overline{\mathbb{R}}^{\mathcal{X}}$ is a function satisfying $\mathcal{E}_1(\Sigma_1(\Box)) \le \Sigma_0$ and $\mathcal{E}_n(\Sigma_1(sx)) \le \Sigma_1(s)(x)$ for all $n \ge 2$, $s \in \mathcal{X}^{n-2}$, and $x \in \mathcal{X}$ (informally, for each situation s, $\Sigma_1(s)$ is the function chosen by Skeptic in situation s). If we fix a strategy $\Sigma = (\Sigma_0, \Sigma_1)$ for Skeptic, his capital $\mathcal{K}_n$ becomes a function of the current situation $s = x_1 \ldots x_n$ of length n. We write $\mathcal{K}^\Sigma(s)$ for $\mathcal{K}_n$ resulting from Skeptic following Σ and from World playing s. Formally, the function $\mathcal{K}^\Sigma : \mathcal{X}^* \to \overline{\mathbb{R}}$ is defined by $\mathcal{K}^\Sigma(\Box) := \Sigma_0$ and $\mathcal{K}^\Sigma(sx) := \Sigma_1(s)(x)$ for all $s \in \mathcal{X}^*$ and all $x \in \mathcal{X}$. This function will be called the *capital process* of Σ. 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.

Notice that a function $\mathcal{S}$ is a supermartingale if and only if $\mathcal{S} : \mathcal{X}^* \to \overline{\mathbb{R}}$ and, for all $n \in \mathbb{N}$ and all situations $s \in \mathcal{X}^{n-1}$ of length $n-1$,

$$\mathcal{E}_n\left(\mathcal{S}(s\,\cdot)\right) \le \mathcal{S}(s),$$

where, as usual, $\mathcal{S}(s\,\cdot) : \mathcal{X} \to \overline{\mathbb{R}}$ is the function mapping each $x \in \mathcal{X}$ to $\mathcal{S}(sx)$. A supermartingale $\mathcal{S}$ is a *martingale* if $\mathcal{E}_n(\mathcal{S}(s\,\cdot)) = \mathcal{S}(s)$ for all $n \in \mathbb{N}$ and $s \in \mathcal{X}^{n-1}$.

Remark. Martingales are less useful for us than supermartingales 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. Under Axiom 5, even a countable sum of positive supermartingales is a supermartingale.

The following useful property of supermartingales follows easily from (4).

Lemma 1. *For each supermartingale $\mathcal{S}$ and each situation s,*

$$\mathcal{S}(s) \ge \inf_{\omega \in \Gamma(s)} \overline{\mathcal{S}}(\omega),$$

where $\overline{\mathcal{S}}(\omega)$ is defined to be $\limsup_{n\to\infty} \mathcal{S}(\omega^n)$ for all $\omega \in \Omega$.

Proof. Let $s = x_1 \ldots x_k \in \mathcal{X}^k$ and let $r > \mathcal{S}(s)$ be given. Since $\mathcal{E}_{k+1}(\mathcal{S}(s\,\cdot)) \le \mathcal{S}(s) < r$, by (4) there exists $x_{k+1} \in \mathcal{X}$ such that $\mathcal{S}(x_1 \ldots x_{k+1}) < r$. Repeating the argument we can find $x_{k+1} x_{k+2} \ldots$ such that $\mathcal{S}(x_1 \ldots x_n) < r$ for all $n \ge k$. Setting $\omega := x_1 x_2 \ldots$, we have $\omega \in \Gamma(s)$ and $\mathcal{S}(\omega^n) < r$ for all $n \ge k$. This completes the proof. $\square$

For each function $\xi : \Omega \to \overline{\mathbb{R}}$ and each situation s, we define the (conditional) *upper expectation* of ξ given s by

$$\overline{\mathbb{E}}(\xi \mid s) := \inf\Big\{a \;\Big|\; \exists \mathcal{S} : \mathcal{S}(s) = a \text{ and}$$

$$\left.\liminf_{n\to\infty} \mathcal{S}(\omega^n) \ge \xi(\omega) \text{ for all } \omega \in \Gamma(s)\right\} \quad (6)$$

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

$$\underline{\mathbb{E}}(\xi \mid s) := -\overline{\mathbb{E}}\,(-\xi \mid s)\,. \quad (7)$$

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

$$\overline{\mathbb{P}}(E \mid s) := \overline{\mathbb{E}}(\mathbb{I}_E \mid s), \quad \underline{\mathbb{P}}(E \mid s) := \underline{\mathbb{E}}(\mathbb{I}_E \mid s), \quad (8)$$

respectively. In what follows we sometimes refer to sets $E \subseteq \Omega$ as *events*.

Lemma 2. *For each situation s, $\overline{\mathbb{E}}(\cdot \mid s) : \overline{\mathbb{R}}^\Omega \to \overline{\mathbb{R}}$ is an outer probability content.*

Proof. It is evident that Axiom 1 is satisfied for $\overline{\mathbb{E}}(\cdot \mid s)$. Axiom 3 for $\overline{\mathbb{E}}(\cdot \mid s)$ follows from the fact that, by Axiom 3 applied to $\mathcal{E}_1, \mathcal{E}_2, \ldots$, the sum $\mathcal{U} := \mathcal{S} + \mathcal{T}$ of two bounded below supermartingales $\mathcal{S}$ and $\mathcal{T}$ is again a bounded below supermartingale and that it satisfies

$$\liminf_{n\to\infty} \mathcal{S}(\omega^n) + \liminf_{n\to\infty} \mathcal{T}(\omega^n) \le \liminf_{n\to\infty} \mathcal{U}(\omega^n)$$

for all $\omega \in \Gamma(s)$. In the same manner, we can use Axiom 2 applied to $\mathcal{E}_1, \mathcal{E}_2, \ldots$ to deduce that $\overline{\mathbb{E}}(\cdot \mid s)$ satisfies Axiom 2. Let $c \in \mathbb{R}$. Since the function on $\overline{\mathbb{R}}^{\mathcal{X}}$ that is identically equal to c is a supermartingale, $\overline{\mathbb{E}}(c \mid s) \le c$. Let $\mathcal{S}$ be a bounded below supermartingale satisfying $\liminf_{n\to\infty} \mathcal{S}(\omega^n) \ge c$ for all $\omega \in \Gamma(s)$. By Lemma 1, we then have $\mathcal{S}(s) \ge c$. Therefore, $\overline{\mathbb{E}}(c \mid s) \ge c$, which completes the proof of Axiom 4 for $\overline{\mathbb{E}}(\cdot \mid s)$. □

Lemma 2 immediately implies the following statement (cf. [6], (5.4)).

Corollary 1. *For all situations s and all functions $\xi : \Omega \to \overline{\mathbb{R}}$, $\underline{\mathbb{E}}(\xi \mid s) \le \overline{\mathbb{E}}(\xi \mid s)$. In particular, $\underline{\mathbb{P}}(E \mid s) \le \overline{\mathbb{P}}(E \mid s)$ for all events $E \subseteq \Omega$.*

Proof. Suppose $\underline{\mathbb{E}}(\xi \mid s) > \overline{\mathbb{E}}(\xi \mid s)$, i.e., $\overline{\mathbb{E}}(\xi \mid s) + \overline{\mathbb{E}}(-\xi \mid s) < 0$. By Axiom 3 applied to $\overline{\mathbb{E}}(\cdot \mid s)$ this implies $\overline{\mathbb{E}}(0 \mid s) < 0$, which contradicts Axiom 4 for $\overline{\mathbb{E}}(\cdot \mid s)$. □

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 [9] 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 3. *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) and Lemma 2, 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 will be used in the proof of Lemma 5 stating that upper expectation is a superexpectation functional in the case where $\mathcal{E}_1, \mathcal{E}_2, \ldots$ are superexpectation functionals.

Lemma 4. *The right-hand side of (6) will not change when $\Gamma(s)$ is replaced by Ω.*

Proof. It suffices to prove that if a bounded below supermartingale $\mathcal{S}$ satisfies $\mathcal{S}(s) < r$ and $\liminf_{n\to\infty} \mathcal{S}(\omega^n) \geq \xi(\omega)$ for all $\omega \in \Gamma(s)$, then there exists another bounded below supermartingale $\mathcal{S}'$ that satisfies $\mathcal{S}'(s) < r$ and $\liminf_{n\to\infty} \mathcal{S}'(\omega^n) \geq \xi(\omega)$ for all $\omega \in \Omega$. Such an $\mathcal{S}'$ can be defined by

$$\mathcal{S}'(t) := \begin{cases} \mathcal{S}(t) & \text{if } s \subseteq t \\ \infty & \text{otherwise.} \end{cases} \qquad \square$$

Lemma 5. *Suppose $\mathcal{E}_1, \mathcal{E}_2, \ldots$ are superexpectations, and let s be a situation. Then $\overline{\mathbb{E}}(\cdot \mid s)$ is also a superexpectation. In particular, for any sequence of events $E_1, E_2, \ldots$, it is true that*

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

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

Proof. In view of Lemma 2, our goal is to prove

$$\overline{\mathbb{E}}\left(\sum_{k=1}^{\infty} \xi_k \;\middle|\; s\right) \leq \sum_{k=1}^{\infty} \overline{\mathbb{E}}\left(\xi_k \mid s\right),$$

where $\xi_1, \xi_2, \ldots$ are positive functions in $\overline{\mathbb{R}}^{\Omega}$. Let $\epsilon > 0$ be arbitrarily small. For each $k \in \mathbb{N}$ choose a supermartingale $\mathcal{S}_k$ (automatically positive, by Lemma 1) such that $\liminf_n \mathcal{S}_k(\omega^n) \geq \xi_k(\omega)$ for all $\omega \in \Omega$ (cf. Lemma 4) and $\mathcal{S}_k(s) \leq \overline{\mathbb{E}}(\xi_k \mid s) + \epsilon/2^k$. Since all $\mathcal{E}_n$ are superexpectations, the sum $\mathcal{S} := \sum_{k=1}^{\infty} \mathcal{S}_k$ will be a supermartingale (cf. (1)); this supermartingale will satisfy $\mathcal{S}(s) \leq \sum_{k=1}^{\infty} \overline{\mathbb{E}}(\xi_k \mid s) + \epsilon$ and, by Fatou's lemma,

$$\liminf_n \mathcal{S}(\omega^n) = \liminf_n \sum_{k=1}^{\infty} \mathcal{S}_k(\omega^n) \geq \sum_{k=1}^{\infty} \liminf_n \mathcal{S}_k(\omega^n) \geq \sum_{k=1}^{\infty} \xi_k(\omega)$$

for all $\omega \in \Omega$. Therefore,

$$\overline{\mathbb{E}}\left(\sum_{k=1}^{\infty} \xi_k \;\middle|\; s\right) \le \mathcal{S}(s) \le \sum_{k=1}^{\infty} \overline{\mathbb{E}}\left(\xi_k \mid s\right) + \epsilon.$$

It remains to remember that ϵ can be taken arbitrarily small. $\square$

3 Equivalent definitions of game-theoretic expectation and probability

The following proposition, which is our main statement of equivalence, gives two equivalent definitions of upper game-theoretic expectation.

Theorem 1. *For all $\xi : \Omega \to \overline{\mathbb{R}}$ and all situations s,*

$$\overline{\mathbb{E}}(\xi \mid s) = \inf\left\{\mathcal{S}(s) \;\middle|\; \limsup_{n\to\infty} \mathcal{S}(\omega^n) \ge \xi(\omega) \text{ for all } \omega \in \Gamma(s)\right\}$$

(i.e., on the right-hand side of (6), we can replace $\liminf$ *by* $\limsup$*), where $\mathcal{S}$ ranges over the supermartingales that are bounded below, and*

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

where $\mathcal{S}$ ranges over the class $\mathbf{L}$ *of all bounded below supermartingales for which* $\lim_{n\to\infty} \mathcal{S}(\omega^n)$ *exists in* $(-\infty, \infty]$ *for all* $\omega \in \Omega$.

Proof. Let a bounded below supermartingale $\mathcal{S}$ satisfy the inequality

$$\forall\omega \in \Gamma(s_0) : \limsup_{n\to\infty} \mathcal{S}(\omega^n) \ge \xi(\omega)$$

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

$$\mathcal{S}^*(s_0) \le \mathcal{S}(s_0) + \epsilon \quad \text{and} \quad \forall\omega \in \Gamma(s_0) : \lim_{n\to\infty} \mathcal{S}^*(\omega^n) \ge \xi(\omega). \tag{9}$$

Without loss of generality we assume $\mathcal{S}(s_0) < \infty$ (if $\mathcal{S}(s_0) = \infty$, set $\mathcal{S}^*(s) := \infty$ for all s). Setting $\mathcal{S}' := (\mathcal{S} - C)/(\mathcal{S}(s_0) - C)$, where C is any constant satisfying $C < \inf \mathcal{S}$, we obtain a positive supermartingale satisfying $\mathcal{S}'(s_0) = 1$.

The idea is now to use the standard proof of Doob's convergence theorem (see, e.g., [9], Lemma 4.5). But first we need to give more definitions, which will also be used in the proof of Theorem 2.

If s and t are two situations such that $s \subseteq t$, we define the "intervals"

$$[s,t] := \{u \mid s \subseteq u \subseteq t\}, \qquad [s,t) := \{u \mid s \subseteq u \subset t\},$$
$$(s,t] := \{u \mid s \subset u \subseteq t\}, \qquad (s,t) := \{u \mid s \subset u \subset t\}.$$

Two situations s and t are said to be *comparable* if $s \subseteq t$ or $t \subseteq s$; otherwise they are *incomparable*. A *cut* is a set of situations that are pairwise incomparable (cuts are analogous to stopping times in measure-theoretic probability). If σ and τ are two cuts, we write $\sigma \le \tau$ to mean $\forall t \in \tau\, \exists s \in \sigma : s \subseteq t$, and we write $\sigma < \tau$ to mean $\forall t \in \tau\, \exists s \in \sigma : s \subset t$. In the case $\sigma \le \tau$, we define the "time intervals"

$$\begin{aligned}
[\sigma,\tau] &:= \{u \mid [\Box,u] \cap \sigma \ne \emptyset, [\Box,u) \cap \tau = \emptyset\},\\
[\sigma,\tau) &:= \{u \mid [\Box,u] \cap \sigma \ne \emptyset, [\Box,u] \cap \tau = \emptyset\},\\
(\sigma,\tau] &:= \{u \mid [\Box,u) \cap \sigma \ne \emptyset, [\Box,u) \cap \tau = \emptyset\},\\
(\sigma,\tau) &:= \{u \mid [\Box,u) \cap \sigma \ne \emptyset, [\Box,u] \cap \tau = \emptyset\}.
\end{aligned}$$

Notice that for all stopping times σ, τ, ρ such that $\sigma \le \tau \le \rho$,

$$[\sigma,\tau) \cap [\tau,\rho) = \emptyset \text{ and } [\sigma,\tau) \cup [\tau,\rho) = [\sigma,\rho).$$

If s is a situation and τ is a cut, s^τ stands for the unique (when it exists) situation $t \in \tau$ such that $t \subseteq s$. Similarly, if $\omega \in \Omega$ and τ is a cut, ω^τ stands for the unique (when it exists) situation $t \in \tau$ that is a prefix of ω. (The case where ω^τ does not exist is analogous to the case where a stopping time takes value ∞ in measure-theoretic probability.) Notice that our notation ω^n for $n = 0, 1, \dots$ can be regarded as a special case of the new notation: we can interpret the upper index n as the cut consisting of all situations of length n. We will also be using the notation s^n, where $n = 0, 1, \dots$ and s is a situation, in the same sense.

Let $[a_i, b_i]$, $i = 1, 2, \dots$, be an enumeration of all intervals with $0 \le a_i < b_i < \infty$ and both end-points rational. For each i one can define a positive supermartingale $\mathcal{S}^i$ with $\mathcal{S}^i(s_0) = 1$ such that $\mathcal{S}^i(\omega^n)$ converges to ∞ as $n \to \infty$ when $\liminf_n \mathcal{S}'(\omega^n) < a_i$ and $\limsup_n \mathcal{S}'(\omega^n) > b_i$. The construction of $\mathcal{S}^i$ is standard. First we define two sequences of cuts $\tau_0^i, \tau_1^i, \dots$ and $\sigma_1^i, \sigma_2^i, \dots$ by setting $\tau_0^i := \{s_0\}$ and, for $k = 1, 2, \dots$,

$$\begin{aligned}
\sigma_k^i &:= \{s \mid \mathcal{S}'(s) > b_i, \exists t \subset s : t \in \tau_{k-1}^i, \forall u \in (t,s) : \mathcal{S}'(u) \le b_i\},\\
\tau_k^i &:= \{s \mid \mathcal{S}'(s) < a_i, \exists t \subset s : t \in \sigma_k^i, \forall u \in (t,s) : \mathcal{S}'(u) \ge a_i\}.
\end{aligned}$$

Now we define $\mathcal{S}^i$ by the requirement that, for all situations $s \supseteq s_0$ and all $x \in \mathcal{X}$,

$$\mathcal{S}^i(sx) := \begin{cases} \mathcal{S}^i(s) + \mathcal{S}'(sx) - \mathcal{S}'(s) & \text{if } \mathcal{S}^i(s) < \infty \text{ and } \exists k : s \in [\tau_{k-1}^i, \sigma_k^i) \\ \mathcal{S}^i(s) & \text{otherwise;} \end{cases} \tag{10}$$

in conjunction with $\mathcal{S}^i(s_0) = 1$ this determines $\mathcal{S}^i$ uniquely on the situations $s \supseteq s_0$. If $s \not\supseteq s_0$, set $\mathcal{S}^i(s) := \infty$. We have $\mathcal{E}_n(\mathcal{S}^i(s\,\cdot)) \le \mathcal{S}^i(s)$, where $n := |s| + 1$, in both cases considered in (10); e.g., since $\mathcal{S}'$ is a supermartingale,

$$\mathcal{E}_n(\mathcal{S}^i(s\,\cdot)) = \mathcal{E}_n(\mathcal{S}^i(s) + \mathcal{S}'(s\,\cdot) - \mathcal{S}'(s)) \le \mathcal{S}^i(s)$$

when $\mathcal{S}^i(s) < \infty$ and $\exists k : s \in [\tau_{k-1}^i, \sigma_k^i)$.

Let us check that each supermartingale $\mathcal{S}^i$ is positive. There are three (overlapping) cases:

- If $s \in [\tau_0^i, \sigma_1^i]$,
$$\mathcal{S}^i(s) \ge \mathcal{S}'(s) \ge 0$$
(we write $\mathcal{S}^i(s) \ge \mathcal{S}'(s)$ rather than $\mathcal{S}^i(s) = \mathcal{S}'(s)$ because of the possibility that $\mathcal{S}'(s) < \infty$ but $\mathcal{S}'(t) = \infty$ for some $t \subset s$).

- If $s \in [\sigma_k^i, \tau_k^i]$ for some $k = 1, 2, \ldots$,
$$\mathcal{S}^i(s) \ge 1 + \Big(\mathcal{S}'(s^{\sigma_1^i}) - \mathcal{S}'(s_0)\Big) + \Big(\mathcal{S}'(s^{\sigma_2^i}) - \mathcal{S}'(s^{\tau_1^i})\Big) + \cdots + \Big(\mathcal{S}'(s^{\sigma_k^i}) - \mathcal{S}'(s^{\tau_{k-1}^i})\Big) \ge b_i + (k-1)(b_i - a_i) \ge 0. \quad (11)$$

- If $s \in [\tau_k^i, \sigma_{k+1}^i]$ for some $k = 1, 2, \ldots$,
$$\mathcal{S}^i(s) \ge 1 + \Big(\mathcal{S}'(s^{\sigma_1^i}) - \mathcal{S}'(s_0)\Big) + \Big(\mathcal{S}'(s^{\sigma_2^i}) - \mathcal{S}'(s^{\tau_1^i})\Big) + \cdots + \Big(\mathcal{S}'(s^{\sigma_k^i}) - \mathcal{S}'(s^{\tau_{k-1}^i})\Big) + \Big(\mathcal{S}'(s) - \mathcal{S}'(s^{\tau_k^i})\Big) \ge b_i + (k-1)(b_i - a_i) + \mathcal{S}'(s) - a_i \ge k(b_i - a_i) \ge 0. \quad (12)$$

Equations (11) and (12) also show that $\mathcal{S}^i(\omega^n)$ indeed converges to ∞ as $n \to \infty$ whenever $\liminf_n \mathcal{S}'(\omega^n) < a_i$ and $\limsup_n \mathcal{S}'(\omega^n) > b_i$ for $\omega \in \Gamma(s_0)$.

Now we can set
$$\mathcal{T} := \sum_{i=1}^{\infty} 2^{-i} \mathcal{S}^i \quad (13)$$
and $\mathcal{S}^* := \mathcal{S} + \epsilon\mathcal{T}$. Assume, for a moment, that $\mathcal{E}_1, \mathcal{E}_2, \ldots$ are superexpectations. In this case $\mathcal{T}$, being a countable sum of positive supermartingales, is a positive supermartingale itself. Being a sum of two supermartingales, $\mathcal{S}^*$ is a supermartingale itself.

Let us check that $\mathcal{S}^* \in \mathbf{L}$; this will imply the second inequality in (9) (the first inequality holds by the definition of $\mathcal{S}^*$). Since $\mathcal{S}^*$ is bounded below, we are only required to check that $\mathcal{S}^*(\omega^n)$ converges in $(-\infty, \infty]$ as $n \to \infty$ for all $\omega \in \Omega$. Fix $\omega \in \Omega$.

If $\mathcal{S}(\omega^n) = \infty$ for some n, there exists i such that $\mathcal{S}^i(\omega^n) = \infty$ from some n on (take any i such that $a_i = 0$ and $b_i > \max_{k<n} \mathcal{S}'(\omega^k)$, where n is the smallest number such that $\mathcal{S}(\omega^n) = \infty$), and so we have $\mathcal{T}(\omega^n) = \infty$ and $\mathcal{S}^*(\omega^n) = \infty$ from some n on. Therefore, we will assume that $\mathcal{S}(\omega^n) < \infty$ for all n.

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

Suppose $\mathcal{S}(\omega^n)$ and, therefore, $\mathcal{S}'(\omega^n)$ converge in $\mathbb{R}$ but $\mathcal{S}^*(\omega^n)$ does not converge in $(-\infty, \infty]$. Choose a non-empty interval $(a, b) \subseteq \mathbb{R}$ such that $\liminf_n \mathcal{S}^*(\omega^n) < a < b < \limsup_n \mathcal{S}^*(\omega^n)$ and set $c := b - a$. Take any $N \in \mathbb{N}$

such that $\mathcal{S}^*(\omega^N) > b$ and $|\mathcal{S}(\omega^n) - \mathcal{S}(\omega^m)| < c/2$, $|\mathcal{S}'(\omega^n) - \mathcal{S}'(\omega^m)| < c/4$ for all $n, m \geq N$. Since $\mathcal{S}'(\omega^n) - \mathcal{S}'(\omega^m) > -c/4$ for all $n, m \geq N$, we will have

$$\mathcal{S}^i(\omega^n) - \mathcal{S}^i(\omega^N) > -c/2 \tag{14}$$

for all i and all $n \geq N$. Indeed, there are five cases (overlapping):

- If $\omega^N \in [\sigma_l^i, \tau_l^i]$ for some $l = 1, 2, \ldots$ and $\omega^n \in [\sigma_k^i, \tau_k^i]$ for some $k = l, l+1, \ldots$:

$$\begin{aligned}\mathcal{S}^i(\omega^n) - \mathcal{S}^i(\omega^N) &= \left(\mathcal{S}'(\omega^{\sigma_{l+1}^i}) - \mathcal{S}'(\omega^{\tau_l^i})\right) + \left(\mathcal{S}'(\omega^{\sigma_{l+2}^i}) - \mathcal{S}'(\omega^{\tau_{l+1}^i})\right) \\ &\quad + \cdots + \left(\mathcal{S}'(\omega^{\sigma_k^i}) - \mathcal{S}'(\omega^{\tau_{k-1}^i})\right) \geq (k-l)(b_i - a_i) \geq 0.\end{aligned}$$

- If $\omega^N \in [\sigma_l^i, \tau_l^i]$ for some $l = 1, 2, \ldots$ and $\omega^n \in [\tau_k^i, \sigma_{k+1}^i]$ for some $k = l, l+1, \ldots$:

$$\begin{aligned}\mathcal{S}^i(\omega^n) - \mathcal{S}^i(\omega^N) &= \left(\mathcal{S}'(\omega^{\sigma_{l+1}^i}) - \mathcal{S}'(\omega^{\tau_l^i})\right) + \left(\mathcal{S}'(\omega^{\sigma_{l+2}^i}) - \mathcal{S}'(\omega^{\tau_{l+1}^i})\right) \\ &\quad + \cdots + \left(\mathcal{S}'(\omega^{\sigma_k^i}) - \mathcal{S}'(\omega^{\tau_{k-1}^i})\right) + \left(\mathcal{S}'(\omega^n) - \mathcal{S}'(\omega^{\tau_k^i})\right) \\ &\quad > (k-l)(b_i - a_i) - c/4 \geq -c/4.\end{aligned}$$

- If $\omega^N \in [\tau_{l-1}^i, \sigma_l^i]$ for some $l = 1, 2, \ldots$ and $\omega^n \in [\sigma_k^i, \tau_k^i]$ for some $k = l, l+1, \ldots$:

$$\begin{aligned}\mathcal{S}^i(\omega^n) - \mathcal{S}^i(\omega^N) &= \left(\mathcal{S}'(\omega^{\sigma_l^i}) - \mathcal{S}'(\omega^N)\right) + \left(\mathcal{S}'(\omega^{\sigma_{l+1}^i}) - \mathcal{S}'(\omega^{\tau_l^i})\right) + \\ &\quad \left(\mathcal{S}'(\omega^{\sigma_{l+2}^i}) - \mathcal{S}'(\omega^{\tau_{l+1}^i})\right) + \cdots + \left(\mathcal{S}'(\omega^{\sigma_k^i}) - \mathcal{S}'(\omega^{\tau_{k-1}^i})\right) \\ &\quad > -c/4 + (k-l)(b_i - a_i) \geq -c/4.\end{aligned}$$

- If $\omega^N \in [\tau_{l-1}^i, \sigma_l^i]$ for some $l = 1, 2, \ldots$ and $\omega^n \in [\tau_k^i, \sigma_{k+1}^i]$ for some $k = l, l+1, \ldots$:

$$\begin{aligned}\mathcal{S}^i(\omega^n) - \mathcal{S}^i(\omega^N) &= \left(\mathcal{S}'(\omega^{\sigma_l^i}) - \mathcal{S}'(\omega^N)\right) + \left(\mathcal{S}'(\omega^{\sigma_{l+1}^i}) - \mathcal{S}'(\omega^{\tau_l^i})\right) \\ &\quad + \left(\mathcal{S}'(\omega^{\sigma_{l+2}^i}) - \mathcal{S}'(\omega^{\tau_{l+1}^i})\right) + \cdots + \left(\mathcal{S}'(\omega^{\sigma_k^i}) - \mathcal{S}'(\omega^{\tau_{k-1}^i})\right) \\ &\quad + \left(\mathcal{S}'(\omega^n) - \mathcal{S}'(\omega^{\tau_k^i})\right) > -c/4 + (k-l)(b_i - a_i) - c/4 \geq -c/2.\end{aligned}$$

- If $\omega^N, \omega^n \in [\tau_{l-1}^i, \sigma_l^i]$ for some $l = 1, 2, \ldots$:

$$\mathcal{S}^i(\omega^n) - \mathcal{S}^i(\omega^N) = \mathcal{S}'(\omega^n) - \mathcal{S}'(\omega^N) > -c/4.$$

In all five cases we have (14). This implies $\mathcal{T}(\omega^n) - \mathcal{T}(\omega^N) > -c/2$, and so $\mathcal{S}^*(\omega^n) - \mathcal{S}^*(\omega^N) > -c$ for all $n \geq N$ (remember that $\epsilon < 1$). The latter contradicts the fact that $\mathcal{S}^*(\omega^n) - \mathcal{S}^*(\omega^N) < -c$ for some $n \geq N$ (namely, for any $n \geq N$ satisfying $\mathcal{S}^*(\omega^n) < a$).

This completes the proof in the case where $\mathcal{E}_1, \mathcal{E}_2, \ldots$ are superexpectations. However, we do not really need Axiom 5: despite the appearance of an infinite sum in (13), for each situation s and each $x \in \mathcal{X}$ the increment $\mathcal{T}(sx) - \mathcal{T}(s)$ of $\mathcal{T}$ can be represented (assuming $\mathcal{T}(s) < \infty$) as

$$\mathcal{T}(sx) - \mathcal{T}(s) = \sum_{i=1}^{\infty} 2^{-i}(\mathcal{S}^i(sx) - \mathcal{S}^i(s)) = \left(\sum_{i=1}^{\infty} w_i\right)(\mathcal{S}'(sx) - \mathcal{S}'(s)),$$

where $w_i \in \{0, 2^{-i}\}$ (this makes the series $\sum_{i=1}^{\infty} w_i$ convergent in $\mathbb{R}$) are defined by

$$w_i := \begin{cases} 2^{-i} & \text{if } \exists k : s \in [\tau^i_{k-1}, \sigma^i_k) \\ 0 & \text{otherwise.} \end{cases}$$

Since $\mathcal{S}'$ is a supermartingale, $\mathcal{E}_{|s|+1}(\mathcal{T}(s\,\cdot) - \mathcal{T}(s)) \leq 0$. This argument for $\mathcal{T}$ being a supermartingale does not depend on Axiom 5. $\square$

Replacing the $\liminf_{n\to\infty}$ in (6) by $\inf_n$ or $\sup_n$ does change the definition. If we replace the $\liminf_{n\to\infty}$ by $\inf_n$, we will have $\overline{\mathbb{E}}(\xi \mid s) = \sup_{\omega\in\Gamma(s)} \xi(\omega)$. In the following example we consider replacing $\liminf_{n\to\infty}$ by $\sup_n$.

Example 1. Set

$$\overline{\mathbb{E}}_1(\xi) := \inf \left\{ \mathcal{S}(\Box) \;\middle|\; \forall \omega \in \Omega : \sup_n \mathcal{S}(\omega^n) \geq \xi(\omega) \right\},$$

$\mathcal{S}$ ranging over the bounded below supermartingales. It is always true that $\overline{\mathbb{E}}_1(\xi) \leq \overline{\mathbb{E}}(\xi)$. Consider the coin-tossing protocol ([9], Section 8.2), which is the special case of Protocol 1 with $\mathcal{X} = \{0,1\}$ and $\mathcal{E}_n(f) = (f(0) + f(1))/2$ for all $n \in \mathbb{N}$ and $f \in \overline{\mathbb{R}}^{\mathcal{X}}$. For each $\epsilon \in (0,1]$ there exists a bounded positive function ξ on Ω such that $\overline{\mathbb{E}}(\xi) = 1$ and $\overline{\mathbb{E}}_1(\xi) = \epsilon$.

Proof. Let us demonstrate the following equivalent statement: for any $C \geq 1$ there exists a bounded positive function ξ such that $\overline{\mathbb{E}}_1(\xi) = 1$ and $\overline{\mathbb{E}}(\xi) = C$. Fix such a C. Define $\Xi : \Omega \to [0, \infty]$ by the requirement $\Xi(\omega) := 2^n$ 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 $\overline{\mathbb{E}}_1(\Xi) = 1$ and $\overline{\mathbb{E}}(\Xi) = \infty$. However, Ξ is unbounded. We can always find $A \geq 1$ such that $\overline{\mathbb{E}}(\min(\Xi, A)) = C$ (as the function $a \mapsto \overline{\mathbb{E}}(\min(\Xi, a))$ is continuous). Since $\overline{\mathbb{E}}_1(\min(\Xi, A)) = 1$, we can set $\xi := \min(\Xi, A)$. $\square$

Game-theoretic probability is a special case of game-theoretic expectation, and in this special case it is possible to replace $\liminf_{n\to\infty}$ not only by $\limsup_{n\to\infty}$ but also by $\sup_n$, provided we restrict our attention to positive supermartingales (simple examples show that this qualification is necessary). By

Lemma 4, the definition of conditional upper probability $\overline{\mathbb{P}}$ can be rewritten as

$$\overline{\mathbb{P}}(E \mid s) := \inf\left\{\mathcal{S}(s) \;\middle|\; \liminf_{n\to\infty} \mathcal{S}(\omega^n) \ge 1 \text{ for all } \omega \in E \cap \Gamma(s)\right\}, \qquad (15)$$

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

Lemma 6. *The definition of upper probability will not change if we replace the* $\liminf_{n\to\infty}$ *in (15) 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 of Lemma 6. It suffices to prove that the definition will not change if we replace the $\liminf_{n\to\infty}$ in (15) by $\sup_n$. Consider a strategy for Skeptic resulting in a positive capital process. If this strategy ensures $\sup_n \mathcal{K}_n > 1$ when $x_1 x_2 \ldots \in E \cap \Gamma(s)$ (it is obvious that it does not matter whether we have $\ge$ or $>$ in (15)), Skeptic can also ensure $\liminf_{n\to\infty} \mathcal{K}_n > 1$ when $x_1 x_2 \ldots \in E \cap \Gamma(s)$ by stopping (i.e., always choosing f_n identically equal to his current capital) after his capital $\mathcal{K}_n$ exceeds 1. $\square$

Remark. The basic notion of this article is that of a bounded below supermartingale; in particular, upper and lower expectation and probability are defined in terms of bounded below supermartingales. To define the latter it would be sufficient to start, instead of outer probability contents, from functionals $\mathcal{F}$ that satisfy Axioms 1–4 and whose domain consists of the functions in $\overline{\mathbb{R}}^X$ that are bounded from below. To see that no generality is lost when one starts from outer probability contents, it is sufficient to check that any such $\mathcal{F}$ can be extended to an outer probability content. One possible extension is $\mathcal{E}(f) := \lim_{a\to-\infty} \mathcal{F}(\max(f, a))$. Axioms 1–4 are easy to check for $\mathcal{E}$; e.g., Axiom 3 follows from the inequality $\max(f + g, 2a) \le \max(f, a) + \max(g, a)$ and Axiom 3 for $\mathcal{F}$:

$$\begin{aligned} \mathcal{E}(f+g) &= \lim_{a\to-\infty} \mathcal{F}(\max(f+g, a)) = \lim_{a\to-\infty} \mathcal{F}(\max(f+g, 2a)) \\ &\le \lim_{a\to-\infty} \mathcal{F}(\max(f, a) + \max(g, a)) \le \lim_{a\to-\infty} (\mathcal{F}(\max(f, a)) + \mathcal{F}(\max(g, a))) \\ &= \lim_{a\to-\infty} \mathcal{F}(\max(f, a)) + \lim_{a\to-\infty} \mathcal{F}(\max(g, a)) = \mathcal{E}(f) + \mathcal{E}(g). \end{aligned}$$

It is also easy to check that $\mathcal{E}$ will be a superexpectation functional whenever $\mathcal{F}$ is a superexpectation functional.

4 Lévy's zero-one law

The following simple theorem is our main result.

Theorem 2. *Let $\xi : \Omega \to (-\infty, \infty]$ be bounded from below. For almost all $\omega \in \Omega$,*

$$\liminf_{n\to\infty} \overline{\mathbb{E}}(\xi \mid \omega^n) \geq \xi(\omega). \tag{16}$$

This theorem is a game-theoretic version of Lévy's zero-one law. Its name derives from its well-known connections with various zero-one phenomena, some of which will be explored in the next section and Section 8.

Proof of Theorem 2. This proof will be similar to the proof of Theorem 1 in that it will be based on the idea used in the standard proof of Doob's martingale convergence theorem. However, this idea will be applied in a less familiar mode ("multiplicative" rather than "additive"), and so before giving a detailed proof we explain the intuition behind it making the simplifying assumption that $\mathcal{E}_1, \mathcal{E}_2, \ldots$ are superexpectations.

By Lemma 6, it suffices to construct a positive supermartingale starting from 1 that is unbounded on $\omega \in \Omega$ for which (16) is not true. (We say that a supermartingale is unbounded on a sequence ω if it is unbounded on its prefixes ω^n as $n \to \infty$.) Without loss of generality we will assume ξ to be positive (we can always redefine $\xi := \xi - \inf \xi$). According to Lemma 5 (last statement), we can, without loss of generality, replace "for which (16) is not true" by

$$\liminf_{n\to\infty} \overline{\mathbb{E}}(\xi \mid \omega^n) < a < b < \xi(\omega) \tag{17}$$

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 for Skeptic. Let $\omega \in \Omega$ be the sequence of moves chosen by World. 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 $f_n \equiv \mathcal{K}_{n-1}$). As soon as this happens, choose a positive supermartingale $\mathcal{S}_1$ starting from a, $\mathcal{S}_1(\omega^n) = a$, whose upper limit $\psi \in \Omega \mapsto \limsup_{m\to\infty} \mathcal{S}_1(\psi^m)$ 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 (17) Skeptic's capital will be unbounded.

We start the formal proof by setting $\xi' := \xi - C$, where C is any constant satisfying $C < \inf \xi$. Let $[a_i, b_i]$, $i = 1, 2, \ldots$, be an enumeration of all intervals with $0 \leq a_i < b_i < \infty$ and both end-points rational. For each i we will define a positive supermartingale $\mathcal{S}^i$ with $\mathcal{S}^i(\Box) = 1$ such that $\mathcal{S}^i(\omega^n)$ converges to ∞ as $n \to \infty$ when

$$\liminf_{n\to\infty} \overline{\mathbb{E}}(\xi' \mid \omega^n) < a_i < b_i < \xi'(\omega) \tag{18}$$

(cf. (17)). First we define two sequences of cuts $\sigma_0^i, \sigma_1^i, \sigma_2^i, \dots$ and $\tau_1^i, \tau_2^i, \dots$ and a family of supermartingales $\mathcal{S}_s$, $s \in \cup_{k=1}^{\infty} \tau_k^i$ (the dependence of $\mathcal{S}_s$ on i is not indicated explicitly but should always be borne in mind). Set $\sigma_0^i := \{\Box\}$. For $k = 1, 2, \dots$, set

$$\tau_k^i := \{s \mid \overline{\mathbb{E}}(\xi' \mid s) < a_i, \exists t \subset s : t \in \sigma_{k-1}^i, \forall u \in (t, s) : \overline{\mathbb{E}}(\xi' \mid u) \geq a_i\},$$

choose for each $s \in \tau_k^i$ a positive supermartingale $\mathcal{S}_s$ satisfying $\mathcal{S}_s(s) < a_i$ and, for all $\omega \in \Gamma(s)$, $\liminf_n \mathcal{S}_s(\omega^n) \geq \xi'(\omega)$, and set

$$\sigma_k^i := \{s \mid \exists t \subset s : t \in \tau_k^i, \mathcal{S}_t(s) > b_i, \forall u \in (t, s) : \mathcal{S}_t(u) \leq b_i\}.$$

This definition is inductive: the two cuts and the family of supermartingales are defined in the indicated order: first σ_0^i, then τ_1^i, then $\mathcal{S}_s$ for $s \in \tau_1^i$, then σ_1^i, then τ_2^i, then $\mathcal{S}_s$ for $s \in \tau_2^i$, then σ_2^i, etc. Now we define $\mathcal{S}^i$ inductively. Set $\mathcal{S}^i(\Box) := 1$. For all situations s and all $x \in \mathcal{X}$, define $\mathcal{S}^i(sx)$ via $\mathcal{S}^i(s)$ as follows:

- First suppose that $\mathcal{S}^i(s) < \infty$ and, for some $k \in \mathbb{N}$, $s \in [\tau_k^i, \sigma_k^i)$. Let k be the unique value satisfying $s \in [\tau_k^i, \sigma_k^i)$. Set $t := s^{\tau_k^i}$ and $\mathcal{S}^i(sx) := \mathcal{S}^i(s)\mathcal{S}_t(sx)/\mathcal{S}_t(s)$ (notice that, by induction, $\mathcal{S}^i(s) < \infty$ implies $\mathcal{S}_t(s) < \infty$).
- Otherwise, set $\mathcal{S}^i(sx) := \mathcal{S}^i(s)$.

Since each $\mathcal{S}_s$ is a supermartingale (strictly positive, by Lemma 1), Axiom 2 shows that $\mathcal{S}^i$ is also a supermartingale. It is positive by construction.

Let us check that each supermartingale $\mathcal{S}^i$ satisfies $\limsup_n \mathcal{S}^i(\omega^n) = \infty$ for $\omega \in \Omega$ satisfying (18). For all $k \in \mathbb{N}$ and all ω satisfying (18), $\omega^{\sigma_k^i}$ exists and satisfies

$$\mathcal{S}^i(\omega^{\sigma_k^i}) = \frac{\mathcal{S}_{\omega^{\tau_1^i}}(\omega^{\sigma_1^i})}{\mathcal{S}_{\omega^{\tau_1^i}}(\omega^{\tau_1^i})} \frac{\mathcal{S}_{\omega^{\tau_2^i}}(\omega^{\sigma_2^i})}{\mathcal{S}_{\omega^{\tau_2^i}}(\omega^{\tau_2^i})} \cdots \frac{\mathcal{S}_{\omega^{\tau_k^i}}(\omega^{\sigma_k^i})}{\mathcal{S}_{\omega^{\tau_k^i}}(\omega^{\tau_k^i})} \geq (b_i/a_i)^k \to \infty \qquad (19)$$

as $k \to \infty$. Setting

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

and assuming that $\mathcal{E}_1, \mathcal{E}_2, \dots$ are superexpectations, we obtain a positive supermartingale $\mathcal{T}$ with $\mathcal{T}(\Box) = 1$ that is unbounded on the complement of (16). Application of Lemma 6 completes the proof under the assumption that $\mathcal{E}_1, \mathcal{E}_2, \dots$ are superexpectations.

It remains to get rid of the assumption. To this end we modify the definition of the supermartingales $\mathcal{S}_s$: now in each situation $s \in \mathcal{X}^*$ such that $\overline{\mathbb{E}}(\xi' \mid s) < \infty$ we fix a strictly positive supermartingale $\mathcal{S}_s$ such that $\mathcal{S}_s(s) < \overline{\mathbb{E}}(\xi' \mid s) + 2^{-|s|}$ and, for all $\omega \in \Gamma(s)$, $\liminf_n \mathcal{S}_s(\omega^n) \geq \xi'(\omega)$. This definition does not depend on i anymore. Using the new definition of $\mathcal{S}_s$, define stopping times σ_k^i, τ_k^i, supermartingales $\mathcal{S}^i$, and a function $\mathcal{T}$ as before; remember that in the definition of $\mathcal{S}^i$, $\mathcal{S}^i(sx) := \infty$ whenever $\mathcal{S}^i(s) = \infty$. For each $i \in \mathbb{N}$ set

$$A_i := \left\{ s \in \mathcal{X}^* \mid \mathcal{S}^i(s) < \infty, \exists k \in \mathbb{N} : s \in [\tau_k^i, \sigma_k^i) \right\}$$

(this is the set of situations in which $\mathcal{S}^i$ is "active"), and for each $s \in A_i$ set $T(s,i) := s^{\tau_k^i}$, where k satisfies $s \in [\tau_k^i, \sigma_k^i)$ (there is only one such k).

Fix an arbitrary situation $s \in \mathcal{X}^{n-1}$, for some $n \in \mathbb{N}$; our next goal is to prove $\mathcal{E}_n(\mathcal{T}(s\,\cdot)) \leq \mathcal{T}(s)$. Without loss of generality, assume $\mathcal{T}(s) < \infty$. Setting, for $t \subseteq s$,

$$w_t := \sum_{i \in \mathbb{N}: s \in A_i, T(s,i) = t} 2^{-i} \mathcal{S}^i(s) \in [0, \infty],$$

we have

$$\begin{aligned}
\mathcal{T}(sx) = \sum_{i=1}^{\infty} 2^{-i} \mathcal{S}^i(sx) = \sum_{i: s \in A_i} 2^{-i} \mathcal{S}^i(s) \frac{\mathcal{S}_{T(s,i)}(sx)}{\mathcal{S}_{T(s,i)}(s)} + \sum_{i: s \notin A_i} 2^{-i} \mathcal{S}^i(s) \\
= \sum_{t \subseteq s} w_t \frac{\mathcal{S}_t(sx)}{\mathcal{S}_t(s)} + \sum_{i: s \notin A_i} 2^{-i} \mathcal{S}^i(s)
\end{aligned}$$

(the denominators $\mathcal{S}_{T(s,i)}(s)$ and $\mathcal{S}_t(s)$ are finite because of our assumption $\mathcal{T}(s) < \infty$), which implies, since the sum $\sum_{t \subseteq s}$ is finite (namely, contains $|s|+1$ addends),

$$\begin{aligned}
\mathcal{E}_n(\mathcal{T}(s\,\cdot)) = \sum_{t \subseteq s} w_t \frac{\mathcal{E}_n(\mathcal{S}_t(s\,\cdot))}{\mathcal{S}_t(s)} + \sum_{i: s \notin A_i} 2^{-i} \mathcal{S}^i(s) \\
\leq \sum_{t \subseteq s} w_t \frac{\mathcal{S}_t(s)}{\mathcal{S}_t(s)} + \sum_{i: s \notin A_i} 2^{-i} \mathcal{S}^i(s) = \mathcal{T}(s).
\end{aligned}$$

Therefore, $\mathcal{T}$ is a supermartingale.

Notice that (19) will still hold for ω satisfying (18) if we replace $(b_i/a_i)^k$ by

$$\prod_{j=1}^{k} \frac{b_i}{a_i + 2^{-|\omega^{\tau_j^i}|}}; \tag{20}$$

since $|\omega^{\tau_j^i}| \to \infty$ as $j \to \infty$, the product (20) still tends to ∞ as $k \to \infty$. Therefore, $\mathcal{T}$ is unbounded on the complement of (16). □

In the rest of the article we will derive a series of corollaries from Theorem 2. First of all, specializing Theorem 2 to the indicators of events, we obtain:

Corollary 2. *Let E be any event. For almost all $\omega \in E$,*

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

as $n \to \infty$.

It is easy to check that we cannot replace the $\geq$ in (16) by $=$, even when ξ is the indicator of an event. For example, suppose that $\mathcal{X} = \{0, 1\}$ and each $\mathcal{E}_n$ is the sup functional: $\mathcal{E}_n(f) := \sup_{x \in \mathcal{X}} f(x)$ for all $n \in \mathbb{N}$ and $f \in \overline{\mathbb{R}}^{\mathcal{X}}$. 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,

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

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 [8] (pp. 128–130), Lévy states his zero-one law in terms of a property E that a sequence $X_1, X_2, \dots$ of random 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 $X_1, \dots, X_n$ 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 $X_1, X_2, \dots$ 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 functions ξ satisfying $\overline{\mathbb{E}}(\xi) = \underline{\mathbb{E}}(\xi)$ and events E satisfying $\overline{\mathbb{P}}(E) = \underline{\mathbb{P}}(E)$.

Lemma 7. *Suppose $\mathcal{E}_1, \mathcal{E}_2, \dots$ are superexpectations. Let a function $\xi : \Omega \to \mathbb{R}$ satisfy $\overline{\mathbb{E}}(\xi) = \underline{\mathbb{E}}(\xi) \in \mathbb{R}$. 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 strictly positive ϵ, there exist bounded below supermartingales $\mathcal{S}_1$ and $\mathcal{S}_2$ such that

$$\mathcal{S}_1(\Box) < \overline{\mathbb{E}}(\xi) + \epsilon/2, \quad \mathcal{S}_2(\Box) < \overline{\mathbb{E}}(-\xi) + \epsilon/2$$

and, for all $\omega \in \Omega$,

$$\liminf_{n \to \infty} \mathcal{S}_1(\omega^n) \geq \xi(\omega), \quad \liminf_{n \to \infty} \mathcal{S}_2(\omega^n) \geq -\xi(\omega).$$

Set $\mathcal{S} := \mathcal{S}_1 + \mathcal{S}_2$. The assumption $\overline{\mathbb{E}}(\xi) = \underline{\mathbb{E}}(\xi) \in \mathbb{R}$ can also be written $\overline{\mathbb{E}}(\xi) + \overline{\mathbb{E}}(-\xi) = 0$. So the supermartingale $\mathcal{S}$ satisfies $\mathcal{S}(\Box) < \epsilon$ and $\liminf_{n \to \infty} \mathcal{S}(\omega^n) \geq 0$ for all $\omega \in \Omega$; by Lemma 1, the supermartingale $\mathcal{S}$ is positive.

Fix n and $\delta > 0$, and let E 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,

$$\mathcal{S}_1(\omega^n) \geq \overline{\mathbb{E}}(\xi \mid \omega^n) \quad \text{and} \quad \mathcal{S}_2(\omega^n) \geq \overline{\mathbb{E}}(-\xi \mid \omega^n).$$

So $\mathcal{S}(\omega^n) > \delta$ for all $\omega \in E$. So, by Lemma 6, the upper probability of E is less than ϵ/δ. Since ϵ may be as small as we like for fixed δ, this shows that E has upper probability zero. Letting δ range over the strictly positive rational numbers and n over $\{0, 1, 2, \dots\}$ and applying the last part of Lemma 5, we obtain the statement of the lemma. ☐

Corollary 3. *Suppose $\mathcal{E}_1, \mathcal{E}_2, \dots$ are superexpectations. Let $\xi : \Omega \to \mathbb{R}$ be a bounded function 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 2,

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

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 (8) make it easy to obtain the following corollary for events.

Corollary 4. *Suppose $\mathcal{E}_1, \mathcal{E}_2, \dots$ are superexpectations. 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 $n \to \infty$.*

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

Let $\xi : \Omega \to \mathbb{R}$ be a bounded function, 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 $\mathcal{S}(s) := \overline{\mathbb{E}}(\xi \mid s)$, $s \in \mathcal{X}^*$. (This fact is the key technical tool used in [11].) In view of Theorem 2, to prove this statement it suffices to check that $\mathcal{S}$ is indeed a supermartingale (it will be bounded below by Lemma 1). We will prove a slightly stronger statement.

Lemma 8. *Let $\xi : \Omega \to \overline{\mathbb{R}}$. Then $\mathcal{S} := \overline{\mathbb{E}}(\xi \mid \cdot)$ is a supermartingale.*

Proof. Let $n \in \mathbb{N}$, $s \in \mathcal{X}^{n-1}$, and $r > \mathcal{S}(s)$. By Lemma 4, there exists a bounded below supermartingale $\mathcal{T}$ such that $\mathcal{T}(s) < r$ and $\liminf_n \mathcal{T}(\omega^n) \geq \xi(\omega)$ for all $\omega \in \Omega$. Then we have $\mathcal{S}(sx) \leq \mathcal{T}(sx)$ for all $x \in \mathcal{X}$, and so we have

$$\mathcal{E}_n(\mathcal{S}(s\,\cdot)) \leq \mathcal{E}_n(\mathcal{T}(s\,\cdot)) \leq \mathcal{T}(s) < r.$$

Letting $r \to \mathcal{S}(s)$ (if $\mathcal{S}(s) < \infty$) shows that $\mathcal{E}_n(\mathcal{S}(s\,\cdot)) \leq \mathcal{S}(s)$, which proves the supermartingale property. ☐

The following simple example shows that replacing "for all $\omega \in \Gamma(s)$" in (6) by "for almost all $\omega \in \Omega$" is essential if we want the infimum to be attained.

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

Proof. By Lemma 1, such an $\mathcal{S}$ would be positive. Let L be the uniform probability measure on $\{0,1\}^\infty$ equipped with the Borel σ-algebra. Since $\overline{\mathbb{P}}(E) = L(E)$ for all Borel sets in $\{0,1\}^\infty$ ([9], Proposition 8.5), we would have $\mathcal{S}(\Box) = \overline{\mathbb{E}}(\xi) = 0$. A positive supermartingale with initial value 0 in the coin-tossing protocol must be a constant. $\Box$

6 More explicit zero-one laws

In this section we will deduce two corollaries from our game-theoretic version of Lévy's zero-one law: Kolmogorov's zero-one law and the ergodicity of Bernoulli shifts. Both corollaries were proved in [10] directly. These two results are more general than the corresponding measure-theoretic results; see [9], Section 8.1, for relations between measure-theoretic results and their game-theoretic counterparts.

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 $n \geq N$). By a *tail event* we mean an element of $\cap_N \mathcal{F}_N$. In other words, an event $E \subseteq \Omega$ is a tail event if any sequence in Ω that agrees from some point onwards with a sequence in E is also in E.

Kolmogorov's zero-one law

Lévy's zero-one law immediately implies the following game-theoretic version of Kolmogorov's zero-one law.

Corollary 5 ([10])**.** *For all tail events $E \subseteq \Omega$, $\overline{\mathbb{P}}(E) \in \{0,1\}$.*

Proof. First we will check that, for each $N \in \mathbb{N}$ and each $E \in \mathcal{F}_N$, $\overline{\mathbb{P}}(E \mid s)$ does not depend on $s \in \mathcal{X}^{N-1}$. Indeed, let $s, t \in \mathcal{X}^{N-1}$ and $\overline{\mathbb{P}}(E \mid s) < r$. Choose a bounded below supermartingale $\mathcal{S}$ such that $\mathcal{S}(s) < r$ and $\liminf_n \mathcal{S}(\omega^n) \geq \mathbb{I}_E(\omega)$ for all $\omega \in \Gamma(s)$. We will write ab for the concatenation of two situations $a \in \mathcal{X}^*$ and $b \in \mathcal{X}^*$, and $a\omega$ for the concatenation of $a \in \mathcal{X}^*$ and $\omega \in \Omega$. The supermartingale

$$\mathcal{S}'(u) := \begin{cases} \mathcal{S}(sv) & \text{if } u = tv \text{ for some (uniquely determined) } v \in \mathcal{X}^* \\ \infty & \text{otherwise} \end{cases}$$

witnesses that $\overline{\mathbb{P}}(E \mid t) < r$, in the sense that $\mathcal{S}'(t) = \mathcal{S}(s) < r$ and, for all $\omega \in \Omega$,

$$\liminf_n \mathcal{S}'(t\omega^n) = \liminf_n \mathcal{S}(s\omega^n) \geq \mathbb{I}_E(s\omega) = \mathbb{I}_E(t\omega).$$

Since this is true for all $s, t \in \mathcal{X}^{N-1}$, $\overline{\mathbb{P}}(E \mid s)$ cannot depend on $s \in \mathcal{X}^{N-1}$.

By Lemma 8 and Axiom 4, this implies $\overline{\mathbb{P}}(E \mid \omega^{N-1}) = \overline{\mathbb{P}}(E)$ for $N \in \mathbb{N}$, $E \in \mathcal{F}_N$, and $\omega \in \Omega$. By Corollary 2, for $E \in \mathcal{F}$ we have $\overline{\mathbb{P}}(E) = 1$ for almost all $\omega \in E$, which is equivalent to $\overline{\mathbb{P}}(E) \in \{0, 1\}$. □

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 5.

Corollary 6 ([10]). *If $E \subseteq \Omega$ is a tail event, then E is almost certain, almost impossible, or fully unprobabilized.*

Ergodicity of Bernoulli shifts

In this subsection we consider a special case of Protocol 1 where $\mathcal{E}_1 = \mathcal{E}_2 = \cdots$. 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 $E \subseteq \Omega$ *weakly invariant* if $\theta E \subseteq E$. In accordance with standard terminology, an event E is *invariant* if $E = \theta^{-1} E$.

Lemma 9. *An event E is invariant if and only if both E and E^c are weakly invariant.*

Proof. We will give the simple argument from [10]. 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 7 ([10]). *Suppose $\mathcal{E}_1 = \mathcal{E}_2 = \cdots$. For all weakly invariant events E, $\overline{\mathbb{P}}(E) \in \{0, 1\}$.*

Proof. For any weakly invariant event E and any situation s, $\overline{\mathbb{P}}(E \mid s) \le \overline{\mathbb{P}}(E)$. Indeed, let $\overline{\mathbb{P}}(E) < r$. Choose a bounded below supermartingale $\mathcal{S}$ such that $\mathcal{S}(\Box) < r$ and $\liminf_n \mathcal{S}(\omega^n) \ge \mathbb{I}_E(\omega)$ for all $\omega \in \Omega$. Define a new bounded below supermartingale $\mathcal{S}'$ by $\mathcal{S}'(st) := \mathcal{S}(t)$ for all $t \in \mathcal{X}^*$ and $\mathcal{S}'(t) := \infty$ for all $t \in \mathcal{X}^*$ such that $s \not\sqsubseteq t$. This supermartingale witnesses that $\overline{\mathbb{P}}(E \mid s) < r$, in the sense that $\mathcal{S}'(s) = \mathcal{S}(\Box) < r$ and

$$\liminf_n \mathcal{S}'(s\omega^n) = \liminf_n \mathcal{S}(\omega^n) \ge \mathbb{I}_E(\omega) \ge \mathbb{I}_E(s\omega), \quad \forall \omega \in \Omega,$$

the last inequality following from $s\omega \in E \Rightarrow \omega \in E$.

Therefore, we have $\overline{\mathbb{P}}(E \mid \omega^n) \le \overline{\mathbb{P}}(E)$ when E is weakly invariant. By Corollary 2, 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 9 we obtain the following corollary to Corollary 7.

Corollary 8 ([10]). *Suppose* $\mathcal{E}_1 = \mathcal{E}_2 = \cdots$. *If* E *is an invariant event, then* E *is almost certain, almost impossible, or fully unprobabilized.*

Since each invariant event is a tail event, Corollary 8 also follows from Corollary 6.

7 The generality of the basic prediction protocol

Let $\mathbf{E}(X)$ be the set of all outer probability contents on a set X. Protocol 1 is a special case of the following apparently more general protocol.

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

As compared with Protocol 1, Protocol 2 involves another player, Forecaster; World is now called Reality. (We will see later that another interpretation is that World is split into two players: Reality and Forecaster; cf. [9], p. 90.) At the beginning of each trial Forecaster gives his prediction p_n for Reality's move 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 $\mathcal{E} : \mathcal{P} \to \mathbf{E}(\mathcal{X})$).

Protocol 1 is a special case of Protocol 2 obtained by taking distinct one-element sets $\mathcal{P}_1, \mathcal{P}_2, \ldots$. In some sense Protocol 1 describes independent trials (since $\mathcal{E}_1, \mathcal{E}_2, \ldots$ are given in advance) whereas Protocol 2 describes dependent trials (Forecaster has a say in choosing $\mathcal{E}_{p_1}, \mathcal{E}_{p_2}, \ldots$).

Remark. Another version of the prediction protocol with Forecaster is where Forecaster chooses the superexpectation functional directly. This is a special case of our protocol with $\mathcal{P}_n = \mathbf{E}(\mathcal{X})$ for all n and with $\mathcal{E} : \mathcal{P} \to \mathbf{E}(\mathcal{X})$ 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.

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 *clearing situations*, and the elements of the set $\bigcup_{n=0}^{\infty}\left(\prod_{i=1}^{n}(\mathcal{P}_i \times \mathcal{X}) \times \mathcal{P}_{n+1}\right)$ are called *betting situations*. We will be mostly interested in clearing situations. For each clearing situation s we let $\Gamma(s) \subseteq \Omega$ stand for the set of all infinite extensions in Ω of s and let $\Box$ be the empty clearing situation.

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

A function $\mathcal{S}$ defined on the clearing situations and taking values in $\overline{\mathbb{R}}$ is called a *supermartingale* if, for each $n \in \mathbb{N}$, each clearing situation s at level $n-1$, and each $p \in \mathcal{P}_n$,

$$\mathcal{E}_p(\mathcal{S}(sp\,\cdot)) \le \mathcal{S}(s).$$

For each function $\xi : \Omega \to \overline{\mathbb{R}}$ and each clearing situation s, we define the (conditional) *upper expectation* of ξ given s by the same formula (6), where $\mathcal{S}$ ranges over the supermartingales that are bounded below, and we define the *lower expectation* of ξ given s by (7). As before, upper and lower probabilities of sets are defined by (8).

Lemma 10. *Theorem 2 and, therefore, Corollary 2 continue to hold under the definitions of this section.*

Proof. Protocol 2 can be embedded in Protocol 1 as follows. Let the parameters of Protocol 2 be $\mathcal{X}$, $\mathcal{P}_1, \mathcal{P}_2, \dots$, and $\mathcal{E}$; as before, $\mathcal{P} := \cup_n \mathcal{P}_n$. For each $n \in \mathbb{N}$, define an outer probability content $\mathcal{E}_n$ on $\mathcal{X}' := \mathcal{P} \times \mathcal{X}$ by

$$\mathcal{E}_n(f) := \sup_{p \in \mathcal{P}_n} \mathcal{E}_p(f(p, \cdot)), \quad f : \mathcal{X}' \to \overline{\mathbb{R}}.$$

(Axioms 1–4 are easy to check for $\mathcal{E}_n$; e.g., Axiom 3 follows from $\sup(f+g) \le \sup f + \sup g$.) The parameters of Protocol 1 will be $\mathcal{X}'$ and $\mathcal{E}_1, \mathcal{E}_2, \dots$.

Our goal is to prove (16) in Protocol 2. Let $\xi : \Omega \to (-\infty, \infty]$ be bounded below. To each ω in the sample space Ω of Protocol 2 corresponds the same sequence in the sample space $\Omega' := (\mathcal{X}')^{\infty}$ of Protocol 1; therefore, $\Omega \subseteq \Omega'$ (perhaps $\Omega \subset \Omega'$). Let $\xi' : \Omega' \to (-\infty, \infty]$ be the extension of ξ defined by, say, $\xi'(\omega) := 0$ for $\omega \notin \Omega$. Since the analogue

$$\liminf_{n \to \infty} \overline{\mathbb{E}}(\xi' \mid \omega^n) \ge \xi'(\omega), \text{ for almost all } \omega \in \Omega',$$

of (16) holds in Protocol 1, we are only required to prove two statements:

1. $\overline{\mathbb{E}}(\xi \mid \omega^n) \ge \overline{\mathbb{E}}(\xi' \mid \omega^n)$ for all $\omega \in \Omega$ and $n = 0, 1, \dots$, where the $\overline{\mathbb{E}}$ on the left-hand side refers to Protocol 2 and the $\overline{\mathbb{E}}$ on the right-hand side refers to Protocol 1.

2. If an event $E \subseteq \Omega'$ is null in Protocol 1, $E \cap \Omega$ will be null in Protocol 2.

First we prove Statement 1. Fix $\omega \in \Omega$ and $n \in \{0, 1, \ldots\}$. For each $\epsilon > 0$ there is a bounded below supermartingale $\mathcal{S}$ in Protocol 2 such that $\mathcal{S}(\omega^n) \leq \overline{\mathbb{E}}(\xi \mid \omega^n) + \epsilon$ and $\liminf_{n\to\infty} \mathcal{S}(\omega^n\psi) \geq \xi(\omega^n\psi)$ for all $\psi \in \Omega$. Let $\mathcal{S}'$ be the extension of $\mathcal{S}$ to $(\mathcal{X}')^*$ defined as ∞ on the situations in Protocol 1 that are not clearing situations in Protocol 2. By the definition of $\mathcal{E}_n$, $\mathcal{S}'$ will be a supermartingale in Protocol 1: if s is a clearing situation in Protocol 2 (the case where it is not is trivial), we have

$$\mathcal{E}_n(\mathcal{S}(s\,\cdot)) = \sup_{p\in\mathcal{P}_n} \mathcal{E}_p(\mathcal{S}(sp\,\cdot)) \leq \sup_{p\in\mathcal{P}_n} \mathcal{S}(s) = \mathcal{S}(s),$$

where $n = |s| + 1$. The supermartingale $\mathcal{S}'$ witnesses that $\overline{\mathbb{E}}(\xi' \mid \omega^n) \leq \overline{\mathbb{E}}(\xi \mid \omega^n) + \epsilon$: indeed, we have $\liminf_{m\to\infty} \mathcal{S}'(\omega^n\psi^m) = \liminf_{m\to\infty} \mathcal{S}(\omega^n\psi^m) \geq \xi(\omega^n\psi) = \xi'(\omega^n\psi)$ for $\psi \in \Omega$ since $\mathcal{S}'$ is an extension of $\mathcal{S}$ and ξ' is an extension of ξ, and we have $\liminf_{m\to\infty} \mathcal{S}'(\omega^n\psi^m) = \infty \geq 0 = \xi'(\omega^n\psi)$ for $\psi \in \Omega' \setminus \Omega$ since in this case $\omega^n\psi^m$ is not a clearing situation in Protocol 2 from some m on. Setting $\epsilon \to 0$ completes the proof of Statement 1.

To prove Statement 2, it suffices to check that for any supermartingale $\mathcal{S}'$ in Protocol 1 its restriction to the clearing situations in Protocol 2 will be a supermartingale in Protocol 2. This follows immediately from the definition of $\mathcal{E}_1, \mathcal{E}_2, \ldots$: if $n \in \mathbb{N}$, s is a clearing situation at level $n-1$, and $p \in \mathcal{P}_n$,

$$\mathcal{E}_p\left(\mathcal{S}'(sp\,\cdot)\right) \leq \sup_{p\in\mathcal{P}_n} \mathcal{E}_p\left(\mathcal{S}'(sp\,\cdot)\right) = \mathcal{E}_n\left(\mathcal{S}'(s\,\cdot)\right) \leq \mathcal{S}'(s)$$

(the first two $\cdot$ stand for an element of $\mathcal{X}$ and the last $\cdot$ stands for an element of $\mathcal{X}'$). ☐

Notice that Theorem 2 is a special case of Lemma 10, corresponding to distinct one-element sets $\mathcal{P}_1, \mathcal{P}_2, \ldots$ in Protocol 2. The argument in the proof of Lemma 10 (which is due to a referee) demonstrates that Protocols 1 and 2 are essentially equivalent.

8 Bártfai and Révész's zero-one law

In this section we will illustrate Lévy's zero-one law by deducing a simple game-theoretic analogue of a zero-one law [2] for dependent random variables. Intuitively, the role of the sample space will now be played by the set $\mathcal{X}^\infty$ of all moves by Reality, and the role of situations will be played by elements of $\mathcal{X}^*$. If $\chi = x_1x_2\ldots \in \mathcal{X}^\infty$, we let χ_n stand for $x_n \in \mathcal{X}$ for $n \in \mathbb{N}$, and let χ^n stand for $x_1 \ldots x_n \in \mathcal{X}^n$ for $n \in \{0, 1, \ldots\}$.

A *forecasting system* Φ is any function $\Phi : \mathcal{X}^* \to \mathcal{P}$ such that $\Phi(\chi^{n-1}) \in \mathcal{P}_n$ for all $\chi \in \mathcal{X}^\infty$ and $n \in \mathbb{N}$. A forecasting system can serve as a strategy for Forecaster in Protocol 2, giving Forecaster's move as function of Reality's moves. For each $\chi \in \mathcal{X}^\infty$ define

$$\chi_\Phi := \Phi(\Box)\chi_1\Phi(\chi^1)\chi_2\Phi(\chi^2)\chi_3 \ldots \in \Omega.$$

For each $E \subseteq \mathcal{X}^\infty$ define $E_\Phi := \{\chi_\Phi \mid \chi \in E\}$. For $E \subseteq \mathcal{X}^\infty$, $\chi \in \mathcal{X}^\infty$, and $n \in \{0, 1, \ldots\}$, set

$$\overline{\mathbb{P}}_\Phi(E \mid \chi^n) := \overline{\mathbb{P}}(E_\Phi \mid (\chi_\Phi)^n), \quad \underline{\mathbb{P}}_\Phi(E \mid \chi^n) := 1 - \overline{\mathbb{P}}_\Phi(E^c \mid \chi^n).$$

As before, "$\mid \square$" may be omitted, so that $\overline{\mathbb{P}}_\Phi(E) = \overline{\mathbb{P}}(E_\Phi)$ and $\underline{\mathbb{P}}_\Phi(E) = \underline{\mathbb{P}}(E_\Phi)$. An $E \subseteq \mathcal{X}^\infty$ holds Φ-*almost surely* (Φ-*a.s.*) if $\overline{\mathbb{P}}_\Phi(E^c) = 0$.

For each $N \in \mathbb{N}$, let $\mathcal{F}_N$ be the set of all $E \subseteq \mathcal{X}^\infty$ such that, for all $\chi, \chi' \in \Omega$,

$$(\chi \in E, \forall n \geq N : \chi'_n = \chi_n) \Longrightarrow \chi' \in E.$$

Let us say that a forecasting system Φ is δ-*mixing*, for $\delta \in [0, 1)$, if there exists a function $a : \mathbb{N} \to \mathbb{N}$ such that

$$\overline{\mathbb{P}}_\Phi(E \mid \chi^n) - \overline{\mathbb{P}}_\Phi(E) \leq \delta \quad \Phi\text{-a.s.} \tag{21}$$

for each $n \in \mathbb{N}$ and each $E \in \mathcal{F}_{n+a(n)}$. By a *tail set* we mean an element of $\cap_N \mathcal{F}_N$. Now we can state an approximate zero-one law, which is a game-theoretic analogue of the main result of [2].

Corollary 9. *Suppose $\mathcal{E}_p$ is a superexpectation for all $p \in \mathcal{P}$. Let $\delta \in [0, 1)$ and let Φ be a δ-mixing forecasting system. If $E \subseteq \mathcal{X}^\infty$ is a tail set, then $\overline{\mathbb{P}}_\Phi(E) = 0$ or $\overline{\mathbb{P}}_\Phi(E) \geq 1 - \delta$.*

Proof. Fix a tail set $E \subseteq \mathcal{X}^\infty$; (21) then holds for all n. By the last part of Lemma 5 (which is also valid in Protocol 2), there is $A \subseteq \mathcal{X}^\infty$ such that $\overline{\mathbb{P}}_\Phi(A) = 0$ and

$$\overline{\mathbb{P}}_\Phi(E \mid \chi^n) - \overline{\mathbb{P}}_\Phi(E) \leq \delta \tag{22}$$

holds for all n and all $\chi \notin A$. By definition, (22) means

$$\overline{\mathbb{P}}(E_\Phi \mid (\chi_\Phi)^n) - \overline{\mathbb{P}}(E_\Phi) \leq \delta. \tag{23}$$

By Corollary 2 and Lemma 10, there is a set $B \subseteq \Omega$ such that $\overline{\mathbb{P}}(B) = 0$ and

$$\overline{\mathbb{P}}(E_\Phi \mid \omega^n) \to 1 \quad (n \to \infty) \tag{24}$$

for all $\omega \in E_\Phi \setminus B$. Letting $n \to \infty$ in (23) and using (24), we can see that $\overline{\mathbb{P}}(E_\Phi) \geq 1 - \delta$ (i.e., $\overline{\mathbb{P}}_\Phi(E) \geq 1 - \delta$) for all χ such that $\chi \notin A$ (i.e., $\chi_\Phi \notin A_\Phi$) and $\chi_\Phi \in E_\Phi \setminus B$.

Suppose $\overline{\mathbb{P}}_\Phi(E) \geq 1 - \delta$ is violated. Then there are no χ satisfying $\chi_\Phi \notin A_\Phi$ and $\chi_\Phi \in E_\Phi \setminus B$. In other words, $E_\Phi \setminus B \subseteq A_\Phi$, which implies $E_\Phi \subseteq A_\Phi \cup B$, which in turn implies $\overline{\mathbb{P}}(E_\Phi) = 0$, i.e., $\overline{\mathbb{P}}_\Phi(E) = 0$. □

Let us say that a set $E \subseteq \mathcal{X}^\infty$ is Φ-*unprobabilized* if $\underline{\mathbb{P}}_\Phi(E) < \overline{\mathbb{P}}_\Phi(E)$. An important special case of Corollary 9 is the following zero-one law for "weakly dependent" trials (cf. Corollary 1 in [2]).

Corollary 10. *Suppose $\mathcal{E}_p$ is a superexpectation for all $p \in \mathcal{P}$. Let $\delta \in [0, 1/2)$ and let Φ be a δ-mixing forecasting system. Every tail set $E \subseteq \mathcal{X}^\infty$ satisfies $\underline{\mathbb{P}}_\Phi(E) = 1$, satisfies $\overline{\mathbb{P}}_\Phi(E) = 0$, or is Φ-unprobabilized.*

Proof. It suffices to apply Corollary 9 to the tail sets E and E^c. □

It is easy to strengthen Corollary 9 by modifying the notion of a δ-mixing forecasting system. Let us say that the forecasting system is *asymptotically δ-mixing*, for $\delta \in [0,1)$, if (21) holds for each $n \in \mathbb{N}$ and each tail set E. Bártfai and Révész [2] 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 [2].

Corollary 11. *Suppose $\mathcal{E}_p$ is a superexpectation for all $p \in \mathcal{P}$. Let $\delta \in [0,1)$. The following two conditions are equivalent:*

1. *A forecasting system Φ is asymptotically δ-mixing.*
2. *Every tail set $E \subseteq \mathcal{X}^\infty$ satisfies $\overline{\mathbb{P}}_\Phi(E) = 0$ or $\overline{\mathbb{P}}_\Phi(E) \ge 1 - \delta$.*

Proof. The argument of Corollary 9 shows that the first condition implies the second. Let us now assume the second condition and deduce the first. Let $n \in \mathbb{N}$ and E be a tail set. If $\overline{\mathbb{P}}_\Phi(E) = 0$, then $\overline{\mathbb{P}}_\Phi(E \mid \chi^n) = 0$ Φ-a.s. can be proved similarly to the proof of Lemma 7, and so (21) holds. If $\overline{\mathbb{P}}_\Phi(E) \ge 1 - \delta$, (21) is vacuous. □

Acknowledgements

Our thinking about Lévy's zero-one law was influenced by a preliminary draft of [3]. We are grateful to Gert de Cooman for his questions that inspired some of the results in Section 3 of this article, and to an anonymous referee of [10] who pointed out to us the zero-one law in [2].This article has benefitted very much from a close reading by its anonymous referee, whose penetrating comments have led to a greatly improved presentation (in particular, Lemmas 1, 2, 4, 5 and the final statements of Theorem 2 and Lemma 8 are due to him or her) and helped us correct a vacuous statement in a previous version. Andrzej Ruszczyński has brought to our attention the literature on coherent measures of risk. Our work has been supported in part by EPSRC grant EP/F002998/1.

References

[1] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

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

[3] 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. Available on-line at `http://www.jehps.net/`.

[4] Isaac P. Cornfeld, Sergei V. Fomin, and Yakov G. Sinai. *Ergodic Theory.* Springer, New York, 1982.

[5] Gert de Cooman and Filip Hermans. Imprecise probability trees: bridging two theories of imprecise probability. *Artificial Intelligence*, 172:1400–1427, 2008.

[6] Jørgen Hoffmann-Jørgensen. The general marginal problem. In Svetozar Kurepa, Hrvoje Kraljević, and Davor Butković, editors, *Functional Analysis II*, volume 1242 of *Lecture Notes in Mathematics*, pages 77–367. Springer, Berlin, 1987.

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

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

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

[10] Akimichi Takemura, Vladimir Vovk, and Glenn Shafer. The generality of the zero-one laws. *Annals of the Institute of Statistical Mathematics*, 63:873–885, 2011.

[11] Vladimir Vovk and Alexander Shen. Prequential randomness and probability. *Theoretical Computer Science*, 411:2632–2646, 2010.

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