

The generality of the zero-one laws

Akimichi Takemura¹

takemura@stat.t.u-tokyo.ac.jp, <http://www.e.u-tokyo.ac.jp/~takemura>

Vladimir Vovk²

vovk@cs.rhul.ac.uk, <http://vovk.net>

Glenn Shafer^{3,2}

gshafer@andromeda.rutgers.edu, <http://glennshafer.com>

August 2009

Abstract

We prove game-theoretic generalizations of some well known zero-one laws. Our proofs make the martingales behind the laws explicit, and our results illustrate how martingale arguments can have implications going beyond measure-theoretic probability.

1 Introduction

Using simple martingale arguments, we generalize three zero-one laws:

Kolmogorov’s zero-one law: In an infinite sequence of independent trials, an event whose happening or failing is not affected by a finite number of the trials has either probability zero or probability one. This was first proven by Kolmogorov in an appendix to his *Grundbegriffe*, published in 1933.

Ergodicity of Bernoulli shifts: In an infinite sequence of independent and identically distributed trials, an event that is invariant under shifts has either probability zero or probability one. This remains true for an infinite sequence of states of a homogeneous Markov chain, under certain conditions. See, for example, Cornfeld et al. (1982), §8.1, Theorems 1 and 2.

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

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

³Department of Accounting and Information Systems, Rutgers Business School—Newark and New Brunswick, 180 University Avenue, Newark, NJ 07102, USA.

Hewitt and Savage’s zero-one law: In an infinite sequence of independent and identically distributed trials, a permutable event has either probability zero or probability one. This was first proven by Hewitt and Savage, in 1955 (Hewitt and Savage (1955), Theorem 11.3).

As one would expect from previous literature (cf. the bibliographical notes to Chapter 8 of Cornfeld et al. (1982)), our argument for the ergodicity of Bernoulli and Markov shifts is similar to but simpler than our argument for Kolmogorov’s zero-one law. In the case of Hewitt and Savage’s zero-one law, our game-theoretic generalization concerns only a special case; it is an open question whether it holds for all permutable events.

We formulate our generalizations in the game-theoretic framework introduced by Shafer and Vovk (2001). This means that we use martingale ideas only, without resorting to measure-theoretic assumptions. Our arguments thus make clear that the zero-one laws have a martingale meaning that extends beyond the measure-theoretic setting. Instead of beginning with a probability measure that determines prices (expected values) for all measurable and bounded payoffs, the game-theoretic framework begins directly with prices. As soon as prices are given, we have martingales: a martingale is the capital process resulting from a strategy for gambling at the given prices. As soon as we have martingales, we can prove theorems by constructing strategies. For example, we can prove that an event has probability one by constructing a strategy that multiplies the capital it risks by an infinite factor if the event fails.

If enough prices are given to determine a probability measure, then each event will have a probability, and the capital process for a gambling strategy will be a martingale in the measure-theoretic sense. But we do not assume that the prices determine a probability measure, and so in general events have only upper and lower probabilities. Thus our concept of a martingale is more general than the measure-theoretic concept, and our results are more general than the classical zero-one laws.

As we will see, even a limited number of prices will determine for each subset E of the sample space a lower probability $\underline{\mathbb{P}}(E)$ and an upper probability $\overline{\mathbb{P}}(E)$ satisfying $0 \leq \underline{\mathbb{P}}(E) \leq \overline{\mathbb{P}}(E) \leq 1$ and $\underline{\mathbb{P}}(E) = 1 - \overline{\mathbb{P}}(E^c)$, where E^c is the complement of E . (The latter equality will be our definition of lower probability.) We consider E certain if $\underline{\mathbb{P}}(E) = 1$ or, equivalently, $\overline{\mathbb{P}}(E^c) = 0$. We consider E impossible if E^c is certain. In the special case where there are enough prices to determine a probability measure on a σ -algebra containing E , $\underline{\mathbb{P}}(E)$ and $\overline{\mathbb{P}}(E)$ will both equal the probability the measure gives to E .

In the usual theory, a zero-one law specifies a property of an event E that guarantees that either $\mathbb{P}(E) = 0$ or $\mathbb{P}(E) = 1$. The corresponding game-theoretic zero-one law might say that either $\underline{\mathbb{P}}(E) = 0$ or $\overline{\mathbb{P}}(E) = 1$. This does not assert that anything is certain or impossible, but it reduces to the usual law in the case where the prices determine a probability measure on a σ -algebra containing E , for then upper and lower probabilities are probabilities. Sometimes we will be able to prove a stronger statement, such as the disjunction of $\overline{\mathbb{P}}(E) = 0$, $\underline{\mathbb{P}}(E) = 1$, and $0 = \underline{\mathbb{P}}(E) < \overline{\mathbb{P}}(E) = 1$.

We will first lay out protocols for our games (§2), define upper and lower probabilities in these protocols (§3), and explain what a zero-one law looks like in terms of upper and lower probabilities (§4). Then we give our martingale proofs: first for Kolmogorov’s zero-one law for tail events in independently priced trials (§5), for ergodicity in independently and identically priced trials and Markov trials (§6), and finally for Hewitt and Savage’s zero-one law for permutable events in independently and identically priced trials (§7).

2 Protocols

In the protocols considered in this article, two players, whom we call Skeptic and Reality, play an infinite number of rounds, which we call *trials*. On each trial, Skeptic chooses a gamble and then Reality determines its payoff. Each player sees the other’s moves as they are made. Reality sees Skeptic’s move before determining its payoff, and Skeptic sees Reality’s move before they go on to the next trial.

Formally, Skeptic chooses a real-valued function F on a set Ω , Reality then chooses an element ω of Ω , and Skeptic’s payoff for the trial is $F(\omega)$. We call ω the *outcome* of the trial. We write \mathcal{K}_n for Skeptic’s capital after the n th trial, and we assume that his initial capital is one monetary unit ($\mathcal{K}_0 = 1$).

We assume that Skeptic chooses F from a non-empty set \mathcal{F} of real-valued functions on Ω with these three properties:

1. If $F_1 \in \mathcal{F}$ and $F_2 \in \mathcal{F}$, then $F_1 + F_2 \in \mathcal{F}$.
2. If $c \geq 0$ and $F \in \mathcal{F}$, then $cF \in \mathcal{F}$.
3. There is no $F \in \mathcal{F}$ such that $F(\omega) > 0$ for all $\omega \in \Omega$.

Property 2 guarantees that the function on Ω that is identically equal to 0 is in \mathcal{F} . Properties 1 and 2 are the defining properties of a *cone*; we call property 3 *coherence*.

We illustrate properties 1–3 in the simple case of $\Omega = \{0, 1\}$. Skeptic chooses a pair of real numbers $(F(0), F(1)) \in \mathcal{F} \subseteq \mathbb{R}^2$, where \mathcal{F} is a convex cone in \mathbb{R}^2 (properties 1 and 2). By the coherence property \mathcal{F} cannot meet the strictly positive quadrant $(0, \infty)^2$. Conversely, any convex cone \mathcal{F} not meeting $(0, \infty)^2$ satisfies properties 1–3.

If P is a probability measure on a σ -algebra for Ω , then the set of all measurable real-valued functions on Ω that have expected value zero with respect to P is a coherent cone. Let us call it the *zero cone* for P . The set of all measurable real-valued functions on Ω that have nonpositive expected value with respect to P is also a coherent cone; we may call it the *nonpositive cone* for P . But a coherent cone need not be a zero cone or a nonpositive cone. Coherent cones that are proper subsets of zero cones are studied in Shafer and Vovk (2001). A cone of this type may include a function F such that $-F$ is not in the cone; this means that Skeptic can take only one side of the gamble represented by F . Because

our results apply to any coherent cone, they generalize the measure-theoretic zero-one laws.

We consider three different protocols, which differ only in how \mathcal{F} may change from one trial to the next. In the first protocol, \mathcal{F} is always the same. If \mathcal{F} is the zero cone for a probability measure P , this is a protocol for betting on successive outcomes at odds given by P . If \mathcal{F} is the nonpositive cone for a probability measure P , it is a protocol for betting on successive outcomes at odds no better than those given by P .

PROTOCOL 1. IDENTICALLY PRICED TRIALS

Parameters: set Ω ; coherent cone \mathcal{F} of real-valued functions on Ω

Protocol:

$\mathcal{K}_0 := 1.$

FOR $n = 1, 2, \dots$:

Skeptic announces $F_n \in \mathcal{F}.$

Reality announces $\omega_n \in \Omega.$

$\mathcal{K}_n := \mathcal{K}_{n-1} + F_n(\omega_n).$

END FOR

In the second protocol, the cone from which Skeptic selects his move may change from trial to trial, but the cone \mathcal{F}_n for the n th trial is fixed at the beginning of the game; it does not depend on outcomes of previous trials. In the special case where each \mathcal{F}_n is the zero cone (resp., nonpositive cone) for a probability measure, this is a protocol for betting on successive trials at odds that are fair (resp., not favorable) according to probability measures assigned to the trials at the outset of the game.

PROTOCOL 2. INDEPENDENTLY PRICED TRIALS

Parameters: set Ω ; coherent cones $\mathcal{F}_1, \mathcal{F}_2, \dots$ of real-valued functions on Ω

Protocol:

$\mathcal{K}_0 := 1.$

FOR $n = 1, 2, \dots$:

Skeptic announces $F_n \in \mathcal{F}_n.$

Reality announces $\omega_n \in \Omega.$

$\mathcal{K}_n := \mathcal{K}_{n-1} + F_n(\omega_n).$

END FOR

In the third protocol, the cone for the n th trial may depend on the outcome ω_{n-1} of the preceding trial. To indicate this possible dependence, we designate the cone by $\mathcal{F}(\omega_{n-1})$. In the special case where $\mathcal{F}(\omega)$ is always a zero cone for a probability measure on Ω , this is a protocol for betting on successive outcomes at odds given by a homogeneous Markov chain on Ω . However, since there are no probabilities for the initial state ω_0 , our result for this protocol (part of Theorem 2) is not directly comparable to the standard zero-one law for Markov chains mentioned in §1.

PROTOCOL 3. MARKOV TRIALS

Parameters: set Ω ; for each $\omega \in \Omega$, a coherent cone $\mathcal{F}(\omega)$ of real-valued functions on Ω

Protocol:Reality announces $\omega_0 \in \Omega$. $\mathcal{K}_0 := 1$.FOR $n = 1, 2, \dots$: Skeptic announces $F_n \in \mathcal{F}(\omega_{n-1})$. Reality announces $\omega_n \in \Omega$. $\mathcal{K}_n := \mathcal{K}_{n-1} + F_n(\omega_n)$.

END FOR

3 Events and upper and lower probabilities

Upper and lower probabilities can be defined for any of the protocols used in game-theoretic probability (Shafer and Vovk (2001), Takeuchi (2004)). We now review the definitions assuming, for simplicity, that we are using either Protocol 1 or Protocol 2, where the cone \mathcal{F}_n from which Skeptic chooses on each trial is fixed, independently of how Reality moves earlier in the game.

We call the set Ω^∞ of all infinite sequences of outcomes the *sample space*. We write $\omega_1\omega_2\dots$ for a generic element of Ω^∞ , and we write $\omega_1\dots\omega_n$ for a finite sequence of outcomes.

As we will see in this section, upper and lower probabilities are defined for any subset E of the sample space Ω^∞ . Accordingly, we call any subset of Ω^∞ an *event*. This diverges from standard terminology, in which only elements of a specified σ -algebra are called events. Our definitions of *tail event* (§5), *invariant event* (§6), and *permutable event* (§7) will also make no reference to any σ -algebra.

A strategy for Skeptic specifies his moves F_1, F_2, \dots as functions of the preceding moves by Reality: F_n is a function of $\omega_1\dots\omega_{n-1}$. Once we fix such a strategy, Skeptic's capital process $\mathcal{K}_0, \mathcal{K}_1, \mathcal{K}_2, \dots$ depends on Reality's moves alone: \mathcal{K}_0 remains the constant 1, and \mathcal{K}_n is a function of $\omega_1\dots\omega_n$.

We call a strategy for Skeptic *prudent* if its capital process is everywhere nonnegative—i.e., if

$$\mathcal{K}_n(\omega_1\dots\omega_n) \geq 0 \quad \text{for all } \omega_1\omega_2\dots \in \Omega^\infty \text{ and all } n.$$

In this case, the strategy risks only the initial unit capital. A strategy for Skeptic will satisfy

$$\limsup_{n \rightarrow \infty} \mathcal{K}_n(\omega_1\dots\omega_n) \geq 0 \quad \text{for all } \omega_1\omega_2\dots \in \Omega^\infty \quad (1)$$

if and only if it is prudent. If the strategy is not prudent—i.e., if $\mathcal{K}_n(\omega_1\dots\omega_n) < 0$ for some $\omega_1\omega_2\dots$ and some n , then Reality can violate (1) by making Skeptic's subsequent payoffs nonpositive, which is possible because of the coherence of the cones from which Skeptic selects his moves. The $\limsup_{n \rightarrow \infty}$ in (1) can be replaced by $\liminf_{n \rightarrow \infty}$ or by \inf_n .

In our protocols, for any given event E and any $c > 0$, Skeptic has a prudent strategy that satisfies

$$\liminf_{n \rightarrow \infty} \mathcal{K}_n(\omega_1 \dots \omega_n) \geq c \quad \text{for all } \omega_1 \omega_2 \dots \in E$$

if and only if he has a prudent strategy that satisfies

$$\sup_{n=1,2,\dots} \mathcal{K}_n(\omega_1 \dots \omega_n) \geq c \quad \text{for all } \omega_1 \omega_2 \dots \in E.$$

This is because Skeptic can stop betting (choose $F_n = 0$) once his capital reaches a given level.

For each event E , we set

$$\overline{\mathbb{P}}(E) := \inf \left\{ \epsilon > 0 \mid \text{Skeptic has a prudent strategy for which} \right. \\ \left. \sup_{n=1,2,\dots} \mathcal{K}_n(\omega_1 \dots \omega_n) \geq 1/\epsilon \quad \text{for all } \omega_1 \omega_2 \dots \in E \right\}, \quad (2)$$

and we set

$$\underline{\mathbb{P}}(E) := 1 - \overline{\mathbb{P}}(E^c). \quad (3)$$

We call $\overline{\mathbb{P}}(E)$ the *upper probability* of E , and we call $\underline{\mathbb{P}}(E)$ the *lower probability* of E . Upper and lower probability generalize the standard notion of probability. Let us consider, for simplicity, Protocol 1. If \mathcal{F} is a zero cone for a probability measure P on a σ -algebra for Ω , then $\overline{\mathbb{P}}(E) = \underline{\mathbb{P}}(E) = P^\infty(E)$ for all measurable subsets of Ω (Shafer and Vovk (2001), §8.2), and our theorems below reduce to the standard zero-one laws in measure-theoretic probability.

Lemma 1. $0 \leq \underline{\mathbb{P}}(E) \leq \overline{\mathbb{P}}(E) \leq 1$ for every event E .

Proof. The relation $\overline{\mathbb{P}}(E) \leq 1$ follows from the fact that Skeptic can choose $F_n = 0$ for all n , and $0 \leq \underline{\mathbb{P}}(E)$ then follows by the definition (3).

Suppose $\underline{\mathbb{P}}(E) > \overline{\mathbb{P}}(E)$, i.e., $\overline{\mathbb{P}}(E) + \overline{\mathbb{P}}(E^c) < 1$. Then there exist $\epsilon_1 > 0$, $\epsilon_2 > 0$, and prudent strategies S_1 and S_2 for Skeptic such that $\epsilon_1 + \epsilon_2 < 1$, S_1 guarantees $\sup_n \mathcal{K}_n \geq \mathbb{I}_E / \epsilon_1$, and S_2 guarantees $\sup_n \mathcal{K}_n \geq \mathbb{I}_{E^c} / \epsilon_2$, where \mathbb{I}_E is the indicator function of E , i.e., $\mathbb{I}_E : \Omega^\infty \rightarrow \mathbb{R}$ takes the value 1 on E and the value 0 outside E . Then $(\epsilon_1 S_1 + \epsilon_2 S_2) / (\epsilon_1 + \epsilon_2)$ guarantees

$$\sup_n \mathcal{K}_n \geq \frac{\mathbb{I}_{E \cup E^c}}{\epsilon_1 + \epsilon_2} = \frac{1}{\epsilon_1 + \epsilon_2} > 1.$$

But this is impossible since, by coherence, Reality can choose $\omega_1 \omega_2 \dots$ so that $1 = \mathcal{K}_0 \geq \mathcal{K}_1 \geq \mathcal{K}_2 \geq \dots$. \square

We can also consider capital processes determined by different strategies for Skeptic when his initial capital \mathcal{K}_0 is not necessarily equal to 1. We call any such capital process a *martingale*. The martingales form a cone. We can rephrase our definition of upper probability, (2), by saying that $\overline{\mathbb{P}}(E)$ is the infimum of

all values of ϵ such that there exists a nonnegative martingale starting at ϵ and reaching at least 1 on every sequence $\omega_1\omega_2\dots$ in E .

The preceding definitions are easily adapted to Protocol 3; we simply recognize that Skeptic's strategies and martingales will also depend on ω_0 as well as on $\omega_1\omega_2\dots$. These definitions also apply, with similarly minor modifications, to other protocols used in game-theoretic probability.

The following terminology spells out the intuitive meaning of extreme values for upper and lower probabilities:

- When $\underline{\mathbb{P}}(E) = 0$, we say E is *unsupported*.
- When $\underline{\mathbb{P}}(E) = 1$, we say E is *certain*.
- When $\overline{\mathbb{P}}(E) = 0$, we say E is *impossible*.
- When $\overline{\mathbb{P}}(E) = 1$, we say E is *fully plausible*.
- When E is unsupported and fully plausible, we say it is *fully uncertain*.

When an event is certain, it is also fully plausible. When it is impossible, it is also unsupported. An event being unsupported is equivalent to its complement being fully plausible. An event being certain is equivalent to its complement being impossible. An event is fully uncertain if and only if its complement is fully uncertain.

As we remarked in §1, both $\underline{\mathbb{P}}(E)$ and $\overline{\mathbb{P}}(E)$ will coincide with E 's probability when there are enough prices to determine a probability measure for a σ -algebra containing E ; see Shafer and Vovk (2001), §8.2. In this case, E cannot be fully uncertain.

4 Two types of zero-one law

A measure-theoretic zero-one law says that an event E satisfying specified conditions is either impossible or certain: either $\mathbb{P}(E) = 0$ or $\mathbb{P}(E) = 1$. In the general game-theoretic case, where we have only upper and lower probabilities, we get one of the following weaker statements:

1. E is either fully plausible or unsupported (or both—i.e., fully uncertain).
2. E is certain, impossible, or fully uncertain.

Condition 2 is stronger than Condition 1, because certain implies fully plausible, and impossible implies unsupported.

When the prices determine a probability measure on a σ -algebra containing E , we have $\overline{\mathbb{P}}(E) = \underline{\mathbb{P}}(E) = \mathbb{P}(E)$, and both conditions then imply that $\mathbb{P}(E) = 1$ or $\mathbb{P}(E) = 0$.

Our game-theoretic versions of Kolmogorov's zero-one law and ergodicity will assert Condition 2, but our game-theoretic version of Hewitt and Savage's zero-one law (proven only in a special case) will assert only Condition 1.

5 Kolmogorov's zero-one law

Our game-theoretic version of Kolmogorov's zero-one law is a theorem about Protocol 2. The relevant events are tail events.

An event E 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 —i.e., if $\omega_1\omega_2\dots$ and $\omega'_1\omega'_2\dots$ are either both in E or both not in E whenever $\omega_n = \omega'_n$ except for a finite number of n . It follows immediately from this definition that E is a tail event if and only if its complement E^c is a tail event.

Theorem 1. *Suppose that E is a tail event in Protocol 2 (independently priced trials). Then E is certain, impossible, or fully uncertain.*

Skeptic chooses from \mathcal{F}_n on trial n in Protocol 2. Our proof of Theorem 1 will use the fact that we get another instantiation of Protocol 2 if we start on trial $n + 1$ for some $n \geq 1$ —i.e., if Skeptic chooses from \mathcal{F}_{n+1} on the first trial, from \mathcal{F}_{n+2} on the second trial, and so on. We call this the *shifted protocol*, as opposed to the *original protocol*. The two protocols, the original one and the shifted one, have the same events; in both, any subset of Ω^∞ is an event.

Let $\overline{\mathbb{P}}^{-n}$ denote upper probability in the shifted protocol. (In particular, $\overline{\mathbb{P}}^{-0} = \overline{\mathbb{P}}$.) Given a strategy S for Skeptic in the shifted protocol, write S^{+n} for the strategy in the original protocol that sets Skeptic's first n moves equal to 0 and then plays S .

We write θ for the shift operator, which deletes the first element from a sequence in Ω^∞ :

$$\theta : \omega_1\omega_2\omega_3\dots \mapsto \omega_2\omega_3\dots$$

We write E^{-n} for $\theta^n E$:

$$E^{-n} := \{ \omega_{n+1}\omega_{n+2}\dots \mid \omega_1\omega_2\dots \in E \}.$$

The next two lemmas relate upper probabilities in the original and shifted protocols.

Lemma 2. *Let E be any event in Protocol 2. Then $\overline{\mathbb{P}}^{-n}(E^{-n})$ is non-decreasing in n , i.e., $\overline{\mathbb{P}}(E) \leq \overline{\mathbb{P}}^{-1}(E^{-1}) \leq \overline{\mathbb{P}}^{-2}(E^{-2}) \leq \dots$.*

Proof. Suppose $c > 0$, and suppose S is a prudent strategy in the shifted protocol that achieves

$$\sup_{k=1,2,\dots} \mathcal{K}_k(\omega_1\dots\omega_k) \geq c \quad \text{for all } \omega_1\omega_2\dots \in E^{-n}.$$

Then S^{+n} is evidently also prudent and achieves

$$\sup_{k=1,2,\dots} \mathcal{K}_k(\omega_1\dots\omega_k) \geq c \quad \text{for all } \omega_1\omega_2\dots \in E$$

in the original protocol. Let \mathcal{S}^{+n} denote the set of strategies in the original protocol of the form S^{+n} . It follows that the infimum in (2) over \mathcal{S}^{+n} coincides

with $\overline{\mathbb{P}}^{-n}(E^{-n})$. Since the set \mathcal{S}^{+n} is non-increasing in n , we are taking the infimum over a smaller set of strategies in (2) when n is larger. So $\overline{\mathbb{P}}^{-n}(E^{-n})$ is non-decreasing in n . \square

Lemma 3. *Suppose that E is a tail event in Protocol 2. Then $\overline{\mathbb{P}}(E) = \overline{\mathbb{P}}^{-n}(E^{-n})$ for all n .*

Proof. Suppose $c > 0$, and suppose S is a prudent strategy in the original protocol that achieves

$$\sup_{k=1,2,\dots} \mathcal{K}_k(\omega_1 \dots \omega_k) \geq c \quad \text{for all } \omega_1 \omega_2 \dots \in E. \quad (4)$$

Because E is a tail event, Reality can choose any sequence from Ω^n as her first n moves without affecting whether E happens. By coherence, she can choose these n moves so that S makes no money for Skeptic on the n trials. Condition (4) tells us that the moves specified by S on the $(n+1)$ th and later trials guarantee that Skeptic can make up this loss and still get at least c when $\omega_1 \omega_2 \dots \in E$. In the shifted protocol, Skeptic has capital 1 at the beginning rather than the same or smaller capital resulting from the losses. So these moves still define a prudent strategy and guarantee that

$$\sup_{k=1,2,\dots} \mathcal{K}_k(\omega_1 \dots \omega_k) \geq c \quad \text{for all } \omega_1 \omega_2 \dots \in E^{-n}$$

in the shifted protocol. This implies $\overline{\mathbb{P}}^{-n}(E^{-n}) \leq \overline{\mathbb{P}}(E)$ and together with the previous lemma we obtain $\overline{\mathbb{P}}^{-n}(E^{-n}) = \overline{\mathbb{P}}(E)$. \square

Proof of Theorem 1. We will show that if a tail event is not fully plausible, then it is impossible. This suffices to prove the theorem, because if E is a tail event, then E^c is also a tail event. If E is not fully uncertain, then either E or E^c is not fully plausible, and if one of them is impossible, then E is either impossible or certain.

Suppose, then, that E is not fully plausible: $\overline{\mathbb{P}}(E) < 1$. Assume that Reality chooses a path in E . Choose ϵ such that $\overline{\mathbb{P}}(E) < \epsilon < 1$. Then there is a prudent strategy S for Skeptic that guarantees he will multiply his initial capital of 1 by $1/\epsilon$ in a finite number of trials. Skeptic plays this strategy until his capital reaches at least $1/\epsilon$. Then he starts over, playing $(1/\epsilon)S$ in the shifted game starting at that point. This eventually again multiplies his capital by another factor of $1/\epsilon$ or more. Continuing in this way, he can make his capital arbitrarily large while playing prudently. This demonstrates that $\overline{\mathbb{P}}(E) = 0$ —i.e., that E is impossible. \square

We have just shown that when E is a tail event with $\overline{\mathbb{P}}(E) < 1$,

$$\text{Skeptic has a prudent strategy guaranteeing } \lim_{n \rightarrow \infty} \mathcal{K}_n = \infty \text{ on } E. \quad (5)$$

This implies $\overline{\mathbb{P}}(E) = 0$, but in general it may be stronger. So we have proven a bit more than the theorem asserts. We have proven that a tail event is either

strongly certain, strongly impossible, or fully uncertain, where an event is said to be *strongly impossible* if (5) holds and *strongly certain* if its complement is strongly impossible.

Here is an example of a fully uncertain tail event. Consider Protocol 2 where Ω is a linear space, $\Omega \neq \{0\}$, and $\mathcal{F}_n \equiv \mathcal{F}$ is the set of linear functions on Ω . If Reality chooses the origin $0 \in \Omega$, then $F(0) = 0$ for all $F \in \mathcal{F}$. On the other hand if Skeptic chooses $F \neq 0$, then there exists $\omega \in \Omega$ such that $F(\omega) < 0$. Therefore the protocol is coherent. Let E be the event that Reality chooses $\omega_n = 0$ except for a finite number of n . Clearly E is a tail event. Since $000\dots \in E$, $\overline{\mathbb{P}}(E) = 1$. Now consider E^c , which is the event that $\omega_n \neq 0$ infinitely often. For each choice $F \in \mathcal{F}$, Reality can choose $\omega \neq 0$ such that $F(\omega) \leq 0$. So $\overline{\mathbb{P}}(E^c) = 1$.

Suppose Reality is determined to prevent Skeptic's capital from becoming arbitrarily large. If E is a fully uncertain tail event, Skeptic has no control at all over Reality choosing an element from the event E . On the other hand, there are many examples of tail events in Shafer and Vovk (2001) which are certain (i.e., Skeptic can force these events), such as the strong law of large numbers or the law of the iterated logarithm.

It is interesting to compare our martingale proof with Kolmogorov's measure-theoretic proof, given in an appendix to *Grundbegriffe* and reproduced in many textbooks. Kolmogorov shows that a tail event E is independent of itself, so that $\mathbb{P}(E) = \mathbb{P}(E)^2$ and therefore $\mathbb{P}(E) = 0$ or $\mathbb{P}(E) = 1$. Our martingale proof paints a little larger picture. Having Skeptic start over just once after multiplying his capital by $1/\epsilon$ suffices to show that $\overline{\mathbb{P}}(E)^2 \geq \overline{\mathbb{P}}(E)$, and this implies $\overline{\mathbb{P}}(E) = 0$ or $\overline{\mathbb{P}}(E) = 1$. But by having Skeptic start over again and again, we find that he can become infinitely rich if an event E with $\overline{\mathbb{P}}(E) < 1$ happens—i.e., that such E is strongly impossible.

A proof based on the measure-theoretic martingale convergence theorem is also known (§2.2 of Chow et al. (1971), §14.3 of Williams (1991)), but it is much less constructive and restricted to measurable events. A corresponding constructive game-theoretic proof is desirable. A referee pointed out a result for tail events by Bártfai and Révész (Bártfai and Révész (1967), Theorems 1 and 2) which can be interpreted as an approximate zero-one law that holds under weak long-range dependence. It would be interesting to establish a similar result in the game-theoretic framework by appropriately modifying our protocols.

6 Ergodicity

Now consider Protocol 1, where Skeptic always chooses from the same cone, and Protocol 3, where the cone may depend on Reality's previous move. The relevant events are invariant events.

We call an event E *weakly invariant* if $\theta E = E^{-1} \subseteq E$.

If E is weakly invariant, then by induction E^{-n} is non-increasing in n . In accordance with standard terminology (e.g., Shiryaev (1996), §V.2), we call an event E *invariant* if $E = \theta^{-1}E$.

Lemma 4. *E is invariant if and only if both E and E^c are weakly invariant.*

Proof. 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. \square

Theorem 2. *Suppose that E is a weakly invariant event in Protocol 1 (identically priced trials) or Protocol 3 (Markov trials). Then E is either impossible or fully plausible.*

Proof. It suffices, as in the proof of Theorem 1, to show that $\overline{\mathbb{P}}(E) < 1$ implies $\overline{\mathbb{P}}(E) = 0$. The current proof is, however, significantly simpler than that of Theorem 1: no analogue of Lemma 3 is needed.

Let E be a weakly invariant event in Protocol 1 such that $\overline{\mathbb{P}}(E) < \epsilon < 1$. Assume that Reality chooses a path in E . Skeptic plays a prudent strategy until his capital increases by a factor of $1/\epsilon$ or more at some round n_1 . By the assumption of weak invariance, $E^{-n_1} \subseteq E$. This implies that when he starts over, there will be another time $n_2 > n_1$ such that he multiplies his capital again by $1/\epsilon$ or more. Continuing in this way, he can make his capital arbitrarily large while playing prudently.

In the case of Protocol 3, Skeptic's strategies and martingales depend on ω_0 as well as on $\omega_1\omega_2\dots$, and the definition of $\overline{\mathbb{P}}(E)$, (2), becomes

$$\overline{\mathbb{P}}(E) := \inf \left\{ \epsilon > 0 \mid \text{Skeptic has a prudent strategy for which} \right. \\ \left. \sup_{n=1,2,\dots} \mathcal{K}_n(\omega_0\omega_1\dots\omega_n) \geq 1/\epsilon \text{ for all } \omega_0\omega_1\omega_2\dots \in E \right\}.$$

Similarly, for each $\omega \in \Omega$ we define

$$\overline{\mathbb{P}}_\omega(E) := \inf \left\{ \epsilon > 0 \mid \text{Skeptic has a prudent strategy for which} \right. \\ \left. \sup_{n=1,2,\dots} \mathcal{K}_n(\omega\omega_1\dots\omega_n) \geq 1/\epsilon \text{ for all } \omega_1\omega_2\dots \in E \right\}.$$

We have $\overline{\mathbb{P}}(E) = \sup_{\omega \in \Omega} \overline{\mathbb{P}}_\omega(E)$. Suppose $\overline{\mathbb{P}}(E) < \epsilon < 1$. For each $\omega \in \Omega$, consider the modified protocol in which Reality's first move is ω , and choose a prudent strategy S_ω for Skeptic in this protocol that eventually multiplies the initial unit capital by $1/\epsilon$ or more.

Now we can proceed as before: Skeptic first plays S_{ω_0} until the first n for which $\mathcal{K}_n \geq 1/\epsilon$, then plays a scaled up version of S_{ω_n} until his capital is again multiplied by $1/\epsilon$ or more, etc. \square

In view of Lemma 4 we obtain the following corollary to Theorem 2.

Corollary 1. *Suppose that E is an invariant event in Protocol 1 (identically priced trials) or Protocol 3 (Markov trials). Then E is certain, impossible, or fully uncertain.*

7 Hewitt and Savage's zero-one law

Let us again consider Protocol 1, where Skeptic always chooses from the same cone.

Let us call an event E *permutable* if for any sequence in E and any $n > 1$, any sequence obtained by permuting the first n terms of the sequence is also in E . Let us call E *singly generated* if it is equal to the set of sequences obtained by taking a single sequence and permuting finite initial subsequences in all possible ways.

We may conjecture that any permutable event in Protocol 1 is either fully plausible or unsupported. But we can prove this result only in the case where the permutable event is singly generated.

Proposition 1. *Suppose that E is a singly generated permutable event in Protocol 1 (identically priced trials). Then E is either fully plausible or unsupported.*

Proof of Proposition 1. Suppose E is neither fully plausible nor unsupported. Then we may choose $\epsilon < 1$ such that $\overline{\mathbb{P}}(E) < \epsilon$ and $\overline{\mathbb{P}}(E^c) < \epsilon$. Choose prudent strategies S and S' for Skeptic that multiply his capital by at least $1/\epsilon$ on E and E^c , respectively.

As usual, we derive a contradiction by showing the existence of a prudent strategy for Skeptic that makes his capital tend to infinity if E happens (i.e., showing that E is strongly impossible and so unsupported). Let $\omega_1\omega_2\dots$ be the sequence of outcomes actually chosen by Reality. The strategy begins by playing S until the capital exceeds $1/\epsilon$. This happens on some trial n_1 if $\omega_1\omega_2\dots \in E$. At this point, we ask whether the remaining sequence $\omega_{n_1+1}\omega_{n_1+2}\dots$ is in E or E^c . In the first case, the strategy plays a scaled up version of S ; in the second it plays a scaled up version of S' . Again, the capital will eventually be multiplied by $1/\epsilon$ on some trial n_2 . Etc.

In conclusion, let us check that the initial sequence $\omega_1\dots\omega_n$ (where $n \in \{n_1, n_2, \dots\}$) indeed determines whether the remaining sequence $\omega_{n+1}\omega_{n+2}\dots$ is in E or E^c . Suppose there are two possible continuations

$$\omega'_{n+1}\omega'_{n+2}\dots \in E \quad \text{and} \quad \omega''_{n+1}\omega''_{n+2}\dots \notin E \quad (6)$$

such that both $\omega_1\dots\omega_n\omega'_{n+1}\omega'_{n+2}\dots$ and $\omega_1\dots\omega_n\omega''_{n+1}\omega''_{n+2}\dots$ belong to E . Since E is singly generated, for some $N > n$ it is true that: (a) the sequences $\omega_1\dots\omega_n\omega'_{n+1}\dots\omega'_N$ and $\omega_1\dots\omega_n\omega''_{n+1}\dots\omega''_N$ are permutations of each other; (b) $\omega'_i = \omega''_i$ for $i > N$. Condition (a) means that the corresponding multisets, $\{\omega_1, \dots, \omega_n, \omega'_{n+1}, \dots, \omega'_N\}$ and $\{\omega_1, \dots, \omega_n, \omega''_{n+1}, \dots, \omega''_N\}$, coincide. (We write $\{\dots\}$ rather than $\{\dots\}$ to emphasize that repetitions are allowed.) Therefore, the multisets $\{\omega'_{n+1}, \dots, \omega'_N\}$ and $\{\omega''_{n+1}, \dots, \omega''_N\}$ also coincide. In other words,

the sequences $\omega'_{n+1} \dots \omega'_N$ and $\omega''_{n+1} \dots \omega''_N$ are permutations of each other. In combination with condition (b), this contradicts our assumption (6). \square

Here is an example showing that the conclusion of Proposition 1 cannot be strengthened to say that E is impossible or fully plausible; in particular, to say that E is certain, impossible, or fully uncertain. Let $\Omega := \{-1, 0, 1\}$, let \mathcal{F} consist of functions F taking values $-t, 0, t$ at $-1, 0, 1$, respectively, for some $t \geq 0$, and let E be the set of all sequences in Ω^∞ that do not contain -1 and contain precisely one 1 . Then E is permutable and singly generated; it is generated by the sequence $100\dots$. Reality can keep Skeptic from making any money by choosing the sequence $000\dots$, and since this sequence is in E^c , this implies that $\overline{\mathbb{P}}(E^c) = 1$ and $\mathbb{P}(E) = 0$. This is consistent with the conclusion of the proposition: E is unsupported. But the prudent strategy that does best for Skeptic on E is one that chooses $t = 1$ at the first trial and continues with this choice so long as Reality plays 0 ; this doubles Skeptic's money on E but no more, and so $\overline{\mathbb{P}}(E) = 1/2$. Thus E is neither impossible nor fully plausible.

It would be interesting to extend Proposition 1 to all permutable events or construct a permutable event E for which $0 < \mathbb{P}(E) \leq \overline{\mathbb{P}}(E) < 1$. Note that Proposition 1 can be trivially extended to the case of permutable events generated by two or more sequences that are distinguishable by any initial segment, e.g., generated by a sequence consisting of rational numbers and a sequence consisting of irrational numbers.

References

- [1] Bártfai, P., and Révész, P. (1967). On a zero-one law. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 7, 43–47.
- [2] Chow, Y. S., Robbins, H., Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping*. Boston: Houghton Mifflin.
- [3] Cornfeld, I. P., Fomin, S. V., and Sinai, Y. G. (1982). *Ergodic Theory*. New York: Springer.
- [4] Hewitt, E., Savage, L. J. (1955). Symmetric measures on Cartesian products. *Transactions of the American Mathematical Society*, 80, 470–501.
- [5] Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin: Springer. English translation (1950): *Foundations of the Theory of Probability*. New York: Chelsea.
- [6] Shafer, G., Vovk, V. (2001). *Probability and Finance: It's Only a Game!* New York: Wiley.
- [7] Shiryaev, A. N. (1996). *Probability*. Second edition. New York: Springer.
- [8] Takeuchi, K. (2004). *Mathematics of Betting and Financial Engineering* (in Japanese). Tokyo: Saiensusha.

- [9] Williams, D. (1991). *Probability with Martingales*. Cambridge: Cambridge University Press.