
DOI: https://doi.org/10.4213/tvp806

Use of the all-Russian mathematical portal Math-Net.Ru implies that you have read and agreed to these terms of use
http://www.mathnet.ru/eng/agreement

Download details:
IP: 54.70.40.11
December 2, 2018, 04:55:35
PURELY GAME-THEORETIC RANDOM SEQUENCES: I. STRONG LAW OF LARGE NUMBERS AND LAW OF THE ITERATED LOGARITHM

1. Introduction

1.1. Random sequences. Since the pioneering works of Kolmogorov, Chaitin and Solomonoff in the sixties, random sequences (see, for a review, [10]) are usually defined with respect to a probability distribution \( P \) (a \( \sigma \)-additive set function, normed to one, defined over a \( \sigma \)-algebra) assuming the traditional axioms for probability theory of Kolmogorov [3]. In this paper, alternatively to these traditional investigations, we propose a definition of random (typical) sequences, without assuming Kolmogorov’s axiomatics, by taking as primitive the notion of a martingale. With respect to a probability distribution \( P \), using Kolmogorov’s classical probability framework, martingale processes have been widely used (see, e.g., [6]) as
test functions for defining random (typical) sequences on the line of the typicalness approach first proposed, in a rigorous mathematical theory, by Martin-Löf [4]. Here, on the other hand, instead of defining a typical sequence with respect to a probability distribution $P$, or to any other partially or fully specified alternative measure, we will define a typical sequence with respect to a sequence of measurable functions, which we declare to be a martingale, without introducing any probability distribution. The intuitive concept of probability, or better of «fairness», will be directly embodied by the limiting behaviour of appropriate transformations of these sequences, by applying directly to them the «principle of the excluded gambling strategy».

This purely game-theoretic approach to the definition of typical sequences has been mostly inspired by the investigations of Vovk [14], [15] on martingale foundations for probability theory. Similar foundations, either based on systems of conditional probability distributions, or on systems of «martingales», have been investigated in a series of papers by Shafer (see, e.g., [7]). An algorithmic probability framework based on the martingale calculus of Vovk [14] is presented in [16].

1.2. Notation and organization. The set of positive integers $1, 2, \ldots$ is denoted by $\mathbb{N}$, and the sets of rational and real numbers are denoted respectively by $\mathbb{Q}$ and $\mathbb{R}$. For any set $\Omega$, $\Omega^*$ is the set of finite sequences $\omega_1 \omega_2 \cdots \omega_n$ of elements of $\Omega$; $\Omega^*$ includes the empty sequence $\square$. The concatenation of a sequence $x \in \Omega^*$ with an element $\omega \in \Omega$ is denoted by $x * \omega$. The set of infinite sequences $\omega_1 \omega_2 \cdots$ of elements of $\Omega$ is denoted by $\Omega^\infty$. If $\xi = \omega_1 \omega_2 \cdots$ is an infinite sequence of elements of $\Omega$ and $n \in \mathbb{N}$, $\xi^n$ is the initial segment $\omega_1 \omega_2 \cdots \omega_n$ of $\xi$ of length $n$. For every $x \in \Omega^*$, if the length of $x$ is denoted by $l(x)$, the cylinder set $\Gamma_x \subseteq \Omega^\infty$ is defined by $\Gamma_x = \{ \xi : \xi \in \Omega^\infty, \xi(l(x)) = x \}$. For any real number $z$, its absolute value is denoted by $|z|$, and the largest integer not greater than $z$ is denoted by $\lfloor z \rfloor$. We indicate with $\ln x$ the natural (base $e$) logarithm of $x$.

The paper is organized as follows. In Section 2 we recall Vovk’s purely game-theoretic probability framework conveniently adapted to our purposes, and in Section 3 we give our definition of typical sequences. In Section 4 we give a direct algorithmic version for our typical sequences of Kolmogorov’s strong law of large numbers (SLLN) whereas in Section 5 we give, in the case of a binary martingale, a direct algorithmic version for our typical sequences of the upper half of Kolmogorov’s law of the iterated logarithm (LIL).

2. The game-theoretic probability framework

Due to the mathematical requirements of constructiveness of any algorithmic approach to the definition of random sequences, in what follows we will restrict ourselves to consider only filtered spaces $(\Omega^\infty, (\mathcal{F}_0, \mathcal{F}_1, \ldots), \mathcal{F})$, where $\mathcal{F}_0 = \{ \emptyset, \Omega^\infty \}$, in which each $\sigma$-algebra $\mathcal{F}_n$
is generated by a countable partition. In particular, we will always consider the set $\Omega$ to be a subset of $Q$ (a classical example being $\Omega = \{0, 1\}$), and, for a fixed set $\Omega$, we will always consider the family of $\sigma$-algebras $(\mathcal{F}_0, \mathcal{F}_1, \ldots)$ on $\Omega^\infty$ to be the filtration generated by the cylinder sets $\Gamma_x$, $x \in \Omega^*$.

To a fixed filtered space $(\Omega^\infty, (\mathcal{F}_0, \mathcal{F}_1, \ldots), \mathbb{P})$, where $\mathcal{F}_0 = \{\emptyset, \Omega^\infty\}$, we add an $\mathbb{R}^k$-valued, $k \in \mathbb{N}$, stochastic sequence $M$, that is, a sequence of $\mathbb{R}^k$-valued random elements $M_0, M_1, \ldots$, defined on $\Omega^\infty$, such that $M_0 = 0$, and each $M_n$ is $\mathcal{F}_n$-measurable. The stochastic sequence $M$ is called the basic martingale, and the stochastic sequence $M_1 - M_0, M_2 - M_1, \ldots$ is called the basic martingale difference sequence. If $V = (V_1, V_2, \ldots)$ is an $\mathbb{R}^k$-valued predictable sequence (that is, a sequence of $\mathbb{R}^k$-valued random elements such that each $V_n$ is $\mathcal{F}_{n-1}$-measurable), then the stochastic sequence defined by $(V \cdot M)_n = \sum_{i=1}^{n} V_i \cdot (M_i - M_{i-1})$, where $\cdot$ in the right-hand side stands for the inner product of vectors in $\mathbb{R}^k$, is called the martingale transform $V \cdot M$. Stochastic sequences of the form $c + V \cdot M$, with $c \in \mathbb{R}$, are called $M$-martingales.

To give an interpretation to the framework, we attach to these definitions the following classical gambling picture. Each element of the, in general multivariate, basic martingale $M$ represents the evolution of our capital in an infinite sequence of «fair games» against an infinitely rich bookmaker in which at each trial we bet a unit of money. Each scalar element of the multivariate value $M_n$ is interpreted as our capital after $n$ games, in the corresponding infinite sequence of fair games. A strategy for simultaneously varying the sizes of the bets in the finite collection of $k$ infinite sequences of fair games represented by $M$ is then represented by a predictable sequence $V$, and the evolution of our capital corresponding to the application of this strategy is represented by the martingale transform $V \cdot M$. In general, an $M$-martingale of the form $c + V \cdot M$ represents the evolution of our capital corresponding to the application of the strategy $V$ starting from an initial capital $c$.

**Definition 1 [15].** A set $E \subseteq \Omega^\infty$ is $M$-null if there is a non-negative $M$-martingale $S$ such that $S_0 = 1$ and $S_n(\xi) \to \infty$, as $n \to \infty$, for all $\xi \in E$.

Note that a null set is not necessarily a measurable subset of $\Omega^\infty$. In accordance with this definition, we also say that a set $E \subseteq \Omega^\infty$ is $M$-almost sure if the set $\Omega^\infty \setminus E$ is $M$-null. This definition, which provides the basis of the present game-theoretic probability framework, is interpreted in terms of the (infinitary) principle of the excluded gambling strategy [14, Section 3]. If $S$ is a pre-specified non-negative $M$-martingale such that $S_0 = 1$ and $\xi$ is the realized outcome, then, as long as we believe in our basic martingale $M$, we can be practically sure that $S_n(\xi)$ does not tend to infinity, as $n \to \infty$. So, if a set $E$ is $M$-null, we can be practically sure that it will not happen.

This definition of an $M$-null set does not make use of any probability
distribution, and so does not make use of Kolmogorov's axioms of probability. Nevertheless, consider extending our filtered space to a probability model by adding a probability distribution \( P \) in \( \mathcal{F} \).

**Theorem 1** [15]. If \( M \) is a local martingale (see, e.g., [8]) with respect to a probability distribution \( P \) in \( (\Omega^\infty, \mathcal{F}) \) and an event \( E \in \mathcal{F} \) is \( M \)-null, then \( P(E) = 0 \).

The converse statement that, for any \( E \in \mathcal{F} \), if \( P(E) = 0 \), then \( E \) is \( M \)-null, is not true in general. For instance, if the basic martingale \( M \) is such that \( \min(M_n | \mathcal{F}_{n-1}) \leq M_{n-1} \leq \max(M_n | \mathcal{F}_{n-1}) \), for all \( n \in \mathbb{N} \), then no \( \mathcal{F}_n \)-measurable event, for \( n \) finite, can be \( M \)-null.

### 3. \( M \)-typical sequences

To have a meaningful definition of a computable stochastic sequences, we need to identify every stochastic sequence \( S \), defined over the set of infinite sequences \( \Omega^\infty \), with a single function from the set of finite initial segments \( \Omega^* \) into \( \mathbb{R}^k \) \((k \geq 1)\). In the sequel, a stochastic sequence \( S \) will be a function \( S: \Omega^* \rightarrow \mathbb{R}^k \), and a predictable sequence \( V \) will be a function \( V: \Omega^* \rightarrow \mathbb{R}^k \) such that, for every fixed \( n \in \mathbb{N} \), \( V(\xi^n \ast \omega') = V(\xi^n \ast \omega''), \) for all \( \omega', \omega'' \in \Omega \). From here on, quantities like \( S(\xi^i) \), \( V(\xi^i) \), etc., will often be abbreviated to \( S_i \), \( V_i \), etc., and, in particular, \( S(\emptyset) \) will always be abbreviated to \( S_0 \). So, when we will be asking to consider a stochastic sequence \( M_n = \sum_{i=1}^{n} X_i \), we will actually mean a stochastic sequence \( M \), and a stochastic sequence \( X \), for which \( M(\xi^n) = \sum_{i=1}^{n} X(\xi^i) \). We say that a stochastic sequence \( S: \Omega^* \rightarrow \mathbb{R} \) is **computable** if there is an algorithm \( \mathcal{U} \) which transforms any input and positive integer \( n \) into a rational number \( r \) satisfying \( |S(x) - r| < 2^{-n} \). That is, the stochastic sequence \( S \) is computable if its values can be computed arbitrarily accurately by some fixed algorithm. (An in depth treatment of computable functions can be found, for instance, in [5].)

Under the usual Kolmogorov axiomatics for probability theory, every definition of typical sequences, with respect to a probability distribution \( P \), corresponds to a definition of an effectively null set (see, e.g., [10, Section 2.1]). In our purely game-theoretic framework, we define an effectively null set as follows (cf. Definition 1).

**Definition 2.** A set \( E \subseteq \Omega^\infty \) is **effectively \( M \)-null** if there is a computable non-negative \( M \)-martingale \( S \) such that \( S_0 = 1 \) and \( S(\xi^n) \rightarrow \infty \), as \( n \rightarrow \infty \), for all \( \xi \in E \).

As before, we also say that a set \( E \subseteq \Omega^\infty \) is **effectively \( M \)-almost sure** if the set \( \Omega^\infty \setminus E \) is effectively \( M \)-null. Let us note that, to define an effectively \( M \)-null set, we did not assume that the basic martingale \( M \) is computable. We now give our purely game-theoretic definition of a random sequence.

**Definition 3.** An infinite sequence \( \xi \in \Omega^\infty \) is **\( M \)-typical** if, for any computable non-negative \( M \)-martingale \( S \), such that \( S_0 = 1 \), \( S(\xi^n) \) does not tend to infinity.
An $M$-typical sequence may be interpreted as follows. Let us suppose that in our gambling picture, starting with an initial amount of money equal to one, at each trial we bet a fraction of our capital, with the constraint that we can never incur a debt. Of course, for any possible sequence of outcomes there would always be a winning strategy among all possible strategies. But if we limit ourselves to only those strategies which are not too complicated, viz. those which can be effectively calculated by means of some algorithm, then we would expect that, whatever betting strategy we might decide to employ, we will never become richer and richer as the game goes on. If it really happens that we become richer and richer, this is because the actual sequence of outcomes is not random at all.

Let us denote the set of all $M$-typical sequences by $T_M$. This set is the intersection of all effectively $M$-almost sure sets or, equivalently, the set $\Omega^\infty \setminus T_M$ is the union of all effectively $M$-null sets.

**Lemma 1.** The set $T_M$ of all $M$-typical sequences is $M$-almost sure.

**Proof.** We have to prove that the set $\Omega^\infty \setminus T_M$ is $M$-null. That is, we have to find a non-negative $M$-martingale $S$ such that $S_0 = 1$ and $S(\xi^n) \to \infty$, as $n \to \infty$, for all $\xi \in \Omega^\infty \setminus T_M$. Now, since the set of all computable $M$-martingales is countable, we can consider an enumeration of all computable non-negative $M$-martingales starting at one, $S^{(1)}, S^{(2)}, \ldots$ say, and consider the non-negative $M$-martingale $S_n = \sum_{i=1}^\infty 2^{-i} S_n^{(i)}$. The non-negative $M$-martingale $S$ is such that $S_0 = 1$ and $S(\xi^n) \to \infty$, as $n \to \infty$, whenever there is an $i$ such that $S^{(i)}(\xi^n) \to \infty$, as $n \to \infty$. Lemma 1 is proved.

**Remark 1.** Let us note that Lemma 1 just states that the set $T_M$ is $M$-almost sure, and not that $T_M$ is effectively $M$-almost sure, that is, that the union of all effectively $M$-null sets is also an effectively $M$-null set. The existence of such a «maximal» effectively null set is guaranteed, for instance, under Kolmogorov’s axiomatics, by the classical definition of typical sequences proposed by Martin-Löf [4, Section III]. In our framework, to attain an effectively $M$-almost sure set of typical sequences (smaller than the set of $M$-typical sequences $T_M$), we would have to consider lower semicomputable $M$-supermartingales instead of computable $M$-martingales in Definitions 2 and 3 of an effectively $M$-null set and of an $M$-typical sequence respectively. This would also allow to prove the existence of a «maximal» lower semicomputable non-negative $M$-supermartingale.

Let $M_n = \sum_{i=1}^n X_i$ be an $\mathbb{R}^k$-valued ($k \geq 1$) basic martingale, and consider the $M$-martingales $N_n^{(j)} = c_j + \sum_{i=1}^n V^{(j)}_i \cdot X_i$, $j = 1, 2, \ldots, J$, where $c_j \in \mathbb{R}$ and $V^{(j)}$ are $\mathbb{R}^k$-valued ($k \geq 1$) predictable sequences, and define the stochastic sequence $N_n = [N^{(1)}_n, N^{(2)}_n, \ldots, N^{(J)}_n]^T$. Then every sequence $\xi \in \Omega^\infty$ which is $M$-typical is also $N$-typical. (For an arbitrary stochastic sequence $N$, an infinite sequence $\xi \in \Omega^\infty$ is $N$-typical if Definition 3 holds.
for $\xi$ when $N$ is taken to be the basic martingale.) Note that, in general, we cannot say that a sequence $\xi \in \Omega^\infty$ is $[M^{(1)}, M^{(2)}]$-typical just because it is simultaneously $M^{(1)}$-typical and $M^{(2)}$-typical, where $M^{(1)}$ and $M^{(2)}$ are two stochastic sequences.

4. The strong law of large numbers

4.1. The convergence lemma. The following lemma, which will be used in the proof of the SLLN and in the proof of the upper half of the LIL, is an analogue of Doob's martingale convergence theorem [1, Ch. VII, Section 4], restricted to non-negative martingales.

**Lemma 2.** If $S$ is a computable non-negative $M$-martingale with $S_0 > 0$ and $\xi$ is an $M$-typical sequence, then the limit $\lim_{n \to \infty} S(\xi^n)$ (of the sequence of values of $S$ for the $M$-typical sequence $\xi$) exists and is finite.

**Proof.** (Reductio ad absurdum.) Suppose the limit $\lim_{n \to \infty} S(\xi^n)$ does not exist. Then there are rational numbers $a_1, a_2, b_1, b_2$ such that $\liminf_{n \to \infty} S(\xi^n) < a_1 < a_2 < b_1 < b_2 < \limsup_{n \to \infty} S(\xi^n)$. Fix an algorithm computing $S$, $\mathcal{U}_S$ say, and consider the algorithm, which takes as input $(\xi^n, m), m = 1, 2, \ldots$, and yields as output zero or one for $n = 1, 2, \ldots$ and never stops for $n = 0$, represented by the flow chart of Fig. 1.

![Flow chart of the algorithm $\mathcal{U}_V$ computing the binary predictable sequence $V$ in the proof of Lemma 2.](image-url)
Purely game-theoretic random sequences

(Note that, in the flow chart, the value of \( m \) is not actually used, and \( r \) is the rational provided by the algorithm \( \mathcal{U}_S \) when the input is \((\xi^i, k)\), \( k \in \mathbb{N} \).) This algorithm, \( \mathcal{U}_V \) say, is completely determined by the rational approximations to \( S \) given by \( \mathcal{U}_S \). It defines a binary predictable sequence, \( V \) say, which is computable, by definition, by \( \mathcal{U}_V \) itself (here whatever the value of \( m, V \) is always computed exactly). The sequence \( V \), starting at one, takes value one until \( S \) crosses \( b_1 \) or \( b_2 \), then takes value zero until \( S \) crosses \( a_1 \) or \( a_2 \), then again it takes value one until \( S \) crosses \( b_1 \) or \( b_2 \), and so forth.

We define

\[
S^*(\xi^n) = 1 + \frac{1}{S_0} \sum_{i=1}^{n} V(\xi^i) [S(\xi^i) - S(\xi^{i-1})].
\]

The function \( S^* \) is a computable non-negative \( M \)-martingale, with \( S_0^* = 1 \), such that \( S^*(\xi^n) \to \infty \), as \( n \to \infty \), for the \( M \)-typical sequence \( \xi \). So, by contradiction, the limit \( \lim_{n \to \infty} S(\xi^n) \) exists. To prove that this limit is also finite, just consider the computable non-negative \( M \)-martingale \( S/S_0 \). Lemma 2 is proved.

4.2. The strong law of large numbers. Let us consider, under a classical probability distribution \( \mathbb{P} \), the following generalization of Kolmogorov's SLLN. If \( X_1, X_2, \ldots \) is a scalar martingale difference sequence with respect to a filtration \((\mathcal{F}_0, \mathcal{F}_1, \ldots)\), then

\[
\sum_{i=1}^{\infty} \mathbb{E}(X_i^2 | \mathcal{F}_{i-1}) < \infty \implies \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X_i = 0,
\]

almost surely. Here we will consider the «pointwise» algorithmic version of this result based on the definition of \( M \)-typical sequences. In proving this version, we will parallel the proof given by Vovk [15] which resembles, in turn, the proof of the SLLN given in [8, Theorem VII.5.4].

Let us consider the following additional definitions. An \( M \)-submartingale is defined as a stochastic sequence of the form \( T = S + A \), where \( S: \Omega^* \to \mathbb{R} \) is an \( M \)-martingale and \( A: \Omega^* \to \mathbb{R} \) is a non-decreasing predictable sequence. If \( S \) and \( A \) are both computable, then \( T \) is computable as well. Any such sequence \( A \) is called a compensator of \( T \).

**Lemma 3.** If \( T \) is a computable non-negative \( M \)-submartingale, \( A \) is one of its computable compensators, and \( \xi \) is an \( M \)-typical sequence, then

\[
A(\xi^n) < \infty \implies T(\xi^n) \text{ converges.}
\]

**Proof.** Suppose that the \( M \)-typical sequence \( \xi \) is such that \( A(\xi^n) < \infty \). Then there is a \( C \in \mathbb{N} \) such that \( A(\xi^n) < C - 1 \), for all \( n \).

Fix an algorithm computing \( A \), \( \mathcal{U} \) say, and consider the algorithm \( \mathcal{U}_A \) that takes as input \((\xi^n, m)\), \( m, n = 1, 2, \ldots \), feeds \( \mathcal{U} \) with \((\xi^{n-1} * \omega, m)\), where
ω is a fixed element of Ω, and yields as output the rational supplied by $\mathcal{U}$. This algorithm computes $A$ yielding as output rational approximations that are predictable. We build out of $\mathcal{U}_A$ the algorithm $\mathcal{U}_V$, which takes as input $(\xi^n, m)$, $m, n = 1, 2, \ldots$, and yields as output zero or one, given by the flow chart of Fig. 2.

\begin{figure}[h]
\centering
\begin{tikzpicture}[node distance=2cm, auto]
  \node (input) {$(\xi^n, m)$};
  \node (init) [below of=input] {$i := 1$};
  \node (cond) [below of=init] {$i > 1 \quad V_{i-1} = 0$};
  \node (assign) [below of=cond] {$i := i + 1$};
  \node (decision) [right of=cond] {$r := \mathcal{U}_A(\xi^i, 2)$};
  \node (zero) [below of=assign] {$V_i := 0$};
  \node (final) [right of=decision] {$V_i := \begin{cases} 1 \text{ if } r + \frac{1}{k} < C \\ 0 \text{ otherwise} \end{cases}$};
  \node (end) [below of=final] {$V_n := V_i$};

  \draw[->] (input) -- (init);
  \draw[->] (init) -- (cond);
  \draw[->] (cond) node[anchor=west] {yes} -- (assign) node[anchor=east] {no};
  \draw[->] (assign) -- (decision);
  \draw[->] (input) -- (zero) node[below] {$i := i + 1$};
  \draw[->] (zero) -- (decision) node[anchor=west] {yes} node[anchor=east] {no};
  \draw[->] (decision) -- (final);
  \draw[->] (final) -- (end);
\end{tikzpicture}
\caption{Flow chart of the algorithm $\mathcal{U}_V$ computing the binary predictable sequence $V$ in the proof of Lemma 3.}
\end{figure}

This algorithm specifies a computable predictable sequence $V$ with values zero and one. The realizations of the sequence $V$ are step functions, with only one jump, such that, for $n = 1, 2, \ldots$, $V_n = 1$, for $A_n \leq C - 1$, $V_n = 0$, for $A_n \geq C$, and $V_n$ equals 1 or 0, for $C - 1 < A_n < C$.

Call $S$ the computable $M$-martingale $T - A$ and consider the stochastic sequence

$$S_n^{(C)} = 1 + \sum_{i=1}^{n} V_i^{(C)}(S_i - S_{i-1}),$$

where $V^{(C)}$ is the computable predictable sequence defined by $V_n^{(C)} = V_n/(S_0 + C)$. It is easy to see that $S^{(C)}$ is a computable non-negative $M$-martingale with $S_0^{(C)} = 1$. By Lemma 2, the limit $\lim_{n \to \infty} S^{(C)}(\xi^n)$ exists and is finite and since, for $n = 1, 2, \ldots$, $S(\xi^n) = S^{(C)}(\xi^n)[S_0 + C] - C$ (for the $M$-typical sequence $\xi$), also the limit $\lim_{n \to \infty} S(\xi^n)$ exists and is finite. So,
Purely game-theoretic random sequences

\[ T(\xi^n) = S(\xi^n) + A(\xi^n) \] converges, \( A(\xi^n) \) being a bounded non-decreasing sequence by hypothesis. Lemma 3 is proved.

**Lemma 4.** Let \( S \) be a computable \( M \)-martingale, let \( S^2 \) be a computable \( M \)-submartingale, and let \( A \) be one of its computable compensators. Then, for every \( M \)-typical sequence \( \xi \),

\[ A(\xi^\infty) < \infty \implies S(\xi^n) \] converges.

**Proof.** If \( A \) is a computable compensator of the computable \( M \)-submartingale \( S^2 \), then \( (S_n + 1)^2 = (N_n + 2S_n + 1) + A_n \), where \( N \) is a computable \( M \)-martingale, and so \( A \) is also a computable compensator of the computable \( M \)-submartingale \((S + 1)^2\). By Lemma 3, \( S^2(\xi^n) \) and \( (S(\xi^n) + 1)^2 \) converge when \( A(\xi^\infty) < \infty \). Then, since \( S(\xi^n) = \frac{1}{2} [ (S(\xi^n) + 1)^2 - S^2(\xi^n) - 1] \), \( S(\xi^n) \) converges as well. Lemma 4 is proved.

**Theorem 2.** Consider the bivariate computable basic martingale

\[ M_n = \left[ \sum_{i=1}^{n} X_i, \sum_{i=1}^{n} (X_i^2 - d_i) \right], \]

where \( X \) is a computable stochastic sequence and \( d \) is a computable non-negative predictable sequence. For every \( M \)-typical sequence \( \xi \),

\[ \sum_{i=1}^{\infty} \frac{d(\xi^i)}{i^2} < \infty \implies \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X(\xi^i) = 0. \]

**Proof.** Let \( S_0 = 0, S_i - S_{i-1} = X_i/i; S \) is a computable \( M \)-martingale. Then, since

\[ \left( \sum_{i=1}^{n} \frac{X_i}{i} \right)^2 = \left[ \sum_{i=1}^{n} \frac{(X_i^2 - d_i)}{i^2} + 2 \sum_{i \geq j} \frac{X_i X_j}{i j} \right] + \sum_{i=1}^{n} \frac{d_i}{i^2}, \]

and the term in square brackets is a computable \( M \)-martingale, \( S^2 \) is a computable \( M \)-submartingale and \( d/i^2 \) is one of its computable compensator difference sequences. By Lemma 4, applied to the computable \( M \)-martingale \( S \), we have that if \( \sum_{i=1}^{\infty} d(\xi^i)/i^2 < \infty \) then \( \sum_{i=1}^{\infty} X(\xi^i)/i \) converges. Finally, by Kronecker’s lemma [8, Lemma IV.3.2] applied to this last series of real numbers we have the result. Theorem 2 is proved.

**4.3. Variants of the strong law of large numbers.** Following more or less directly from this result, we may have many variants of the SLLN. Consider the bivariate computable basic martingale \( M \) of Theorem 2. Then for a computable predictable sequence \( V \), it is easy to see that for every \( M \)-typical sequence \( \xi \),

\[ \sum_{i=1}^{\infty} V^2(\xi^i) d(\xi^i)/i^2 < \infty \implies \lim_{n \to \infty} (1/n) \sum_{i=1}^{n} V(\xi^i) X(\xi^i) = 0. \]
Also, considering \( \{n_k\} \), where \( n_0 = 0 \), to be a computable predictable subsequence of \( \{n\} \), we may prove that, for every \( M \)-typical sequence \( \xi \),

\[
\sum_{j=1}^{\infty} \frac{(V_{n_j+1} d_{n_j+1} + \cdots + V_{n_j} d_{n_j})}{j^2} < \infty
\]

\[
\implies \lim_{k \to \infty} \frac{1}{k} \sum_{j=1}^{k} (V_{n_j+1} X_{n_j+1} + \cdots + V_{n_j} X_{n_j}) = 0.
\]

5. The upper half of the law of the iterated logarithm

In the classical probability setting, let \( X_1, X_2, \ldots \) be a sequence of independent, not necessarily identically distributed, random variables with \( \mathbb{E}(X_i) = 0 \) and finite variance, and let \( S_n = \sum_{i=1}^{n} X_i \). Kolmogorov [2] proved that, if, as \( n \to \infty \), \( V_n = D(S_n) \to \infty \), and \( |X_n| \leq \varepsilon_n \sqrt{V_n / \ln \ln V_n} \), almost surely, for some constants \( \varepsilon_n \to 0 \), then

\[
\lim_{n \to \infty} \sup \frac{S_n}{\sqrt{2V_n \ln \ln V_n}} = 1,
\]

almost surely. This remarkable result provided the best possible refinement of the SLLN. Under Kolmogorov’s probability axiomatics, a martingale version of Kolmogorov’s LIL was later obtained by Stout [9].

In this section, we consider a direct algorithmic version for \( M \)-typical sequences of the upper half of Kolmogorov’s LIL in the case of a univariate binary basic martingale. In proving our result, we adapt the proofs of the martingale extensions of the upper half of Kolmogorov’s LIL given by Vovk [12, Theorem 4]; [13, Theorem 6], which are based on an idea originally due to Ville [11, ch. V, 1re Section, § 3].

5.1. The upper half of the law of the iterated logarithm.

Let us note that, when each observation is binary, specifying a univariate basic martingale is essentially equivalent to specifying a full probability distribution \( \mathbb{P} \) over \( \Omega^\infty \). For a univariate binary computable basic martingale \( M_n = \sum_{i=1}^{n} X_i, X_i \in \{a_i, b_i\} \), where \( a \) and \( b \) are a negative and a positive computable predictable sequence respectively, the higher powers of \( X_i \) can all be given by

\[
X_i^2 = -a_i b_i + (a_i + b_i) X_i, \\
X_i^3 = -a_i b_i a_i b_i + (a_i^2 + a_i b_i + b_i^2) X_i, \ldots
\]

and any martingale-like property involving these higher powers can be written as a martingale transform. For instance, \( \sum_{i=1}^{n} (X_i^3 + a_i b_i) = \sum_{i=1}^{n} V_i X_i \), where the right-hand side is a computable martingale transform involving the computable predictable sequence \( V_i = a_i + b_i \). 

Theorem 3. Let us consider the univariate binary computable basic martingale \( M_n = \sum_{i=1}^{n} X_i, X_i \in \{a_i, b_i\} \), where \( a \) and \( b \) are a negative and
a positive computable predictable sequence respectively. Let also
\[ V_n = -\sum_{i=1}^{n} a_i b_i, \quad W_n = \sum_{i=1}^{n} \frac{b_i |a_i|^3 - a_i |b_i|^3}{b_i - a_i}, \]
and consider an M-typical sequence \( \xi \) such that, as \( n \to \infty \),
\[ V(\xi^n) \to \infty, \quad \Delta(\xi^n) = \max_{\omega} |X(\xi^{n-1} * \omega)| = O\left( \sqrt[\ln \ln V(\xi^n)]{V(\xi^n)} \right), \tag{1} \]
\[ W(\xi^n) = o\left( \sqrt[\ln \ln V(\xi^n)]{V^3(\xi^n)} \right). \tag{2} \]

Then
\[ \limsup_{n \to \infty} \frac{|M(\xi^n)|}{\sqrt{2V(\xi^n)\ln \ln V(\xi^n)}} \leq 1. \]

Proof. Let \( \xi \in \Omega^\infty \) be an M-typical sequence for which condition (1) and (2) are satisfied. We shall first prove that, for an arbitrarily small \( \varepsilon > 0 \), from some \( n \) on \( M(\xi^n) \leq (1 + \varepsilon)\sqrt{2V(\xi^n)\ln \ln V(\xi^n)} \).

For a rational \( \alpha > 0 \), define the non-negative stochastic sequence \( S^{(\alpha)} \) given by \( S^{(\alpha)}_0 = 1 \) and
\[ S^{(\alpha)}_n = \prod_{i=1}^{n} \frac{e^{\alpha X_i}}{(b_i - a_i)^{-1}(b_i e^{\alpha a_i} - a_i e^{\alpha b_i})} = \prod_{i=1}^{n} \left( 1 + \frac{e^{\alpha b_i} - e^{\alpha a_i}}{b_i e^{\alpha a_i} - a_i e^{\alpha b_i}} X_i \right), \tag{3} \]
for \( n = 1, 2, \ldots \). It is easy to check that
\[ S^{(\alpha)}_n = 1 + \sum_{i=1}^{n} U^{(\alpha)}_i X_i, \quad \text{where } U^{(\alpha)}_i = S^{(\alpha)}_{i-1} \frac{e^{\alpha b_i} - e^{\alpha a_i}}{b_i e^{\alpha a_i} - a_i e^{\alpha b_i}}, \]
and so, that \( S^{(\alpha)} \) is an M-martingale. Moreover, since \( a \) and \( b \) are two computable sequences, the predictable sequence \( U^{(\alpha)} \) and the M-martingale \( S^{(\alpha)} \) are also computable.

Now, let \( \varepsilon > 0 \) be arbitrarily small, let \( \delta > 0 \) be a fixed rational small compared to \( \varepsilon \), and consider the non-negative stochastic sequence
\[ S_n = c_1 \sum_{k=1}^{\infty} \frac{1}{k^{1+\delta}} S^{(\alpha(k))}_n, \]
where \( \alpha(k) \) are rational approximations, to a given precision, to the computable function \( \sqrt{2(1 + \delta)^{-k} \ln k} \), and \( c_1 > 0 \) is a rational chosen so that to ensure that \( 0 < S_0 \leq 1 \). It is easy to check that, since for every \( k = 1, 2, \ldots \), \( S^{(\alpha(k))} \) is an M-martingale, also \( S \) is an M-martingale. As far as the computability of \( S \) is concerned, note that both quantities \( \sqrt{2(1 + \delta)^{-k} \ln k} \) and \( k^{-(1+\delta)} \) are computable functions in \( k \).

Then, given an algorithm computing the function \( \sqrt{2(1 + \delta)^{-k} \ln k} \), whatever the precision of its rational approximations \( \alpha(k) \), the sequence of stochastic
sequences defined by \( k^{-(1+\delta)}S'(\alpha(k)), k = 1, 2, \ldots \), is a computable sequence of computable functions, and since each one of them is non-negative and for every given \( n \) the series in \( k \) converges, we have that \( S \) is computable.

For the \( M \)-typical sequence \( \xi \), by Lemma 2, the limit \( \lim_{n \to \infty} S(\xi^n) \) exists and is finite, that is, there exists a constant \( c_2 \) such that \( S(\xi^n) \leq c_2 \), for every \( n = 1, 2, \ldots \). So, since \( S \) is a sum of non-negative stochastic sequences, for every \( k = 1, 2, \ldots \),

\[
\frac{1}{K^{1+\delta}} S(\alpha(k))(\xi^n) \leq c_2, \quad n = 1, 2, \ldots,
\]
or \( S(\alpha(k))(\xi^n) \leq c_3 k^{1+\delta}, n = 1, 2, \ldots \). (The constants \( c_1, c_2, c_3 \) above and \( c_4 \) below depend only on \( \delta \).)

Consider now a sufficiently large \( n \) and choose the non-negative \( M \)-martingale \( S(\alpha(k)) \) where \( \alpha(k) \) is sufficiently close to \( 2(1 + \delta)^{-k} \ln k \) and \( k = [\log_{1+\delta} V_n] \). Then \( \ln S_n(\alpha(k)) \leq (1 + \delta) \ln k + \ln c_3 = (1 + \delta) \ln \ln V_n + c_4 \), and, by the definition of \( S(\alpha) \), \( \ln S_n(\alpha) = \delta M_n - \sum_{i=1}^{n} \ln((b_i e^{\alpha a_i} - a_i e^{\alpha b_i})/(b_i - a_i)) \), and so the last inequality is equivalent to

\[
\alpha M_n \leq \sum_{i=1}^{n} \ln \left( \frac{b_i e^{\alpha a_i} - a_i e^{\alpha b_i}}{b_i - a_i} \right) + (1 + \delta) \ln \ln V_n + c_4,
\]

where \( \alpha = \alpha(k) \). Using the inequalities \( e^t \leq 1 + t + t^2/2 + |t|^3 e^{|t|}/6, \ln(1 + t) \leq t \), we obtain

\[
\alpha M_n \leq \frac{\alpha^2}{2} V_n + \frac{\alpha^3}{6} \sum_{i=1}^{n} e^{\alpha \Delta_i} \frac{b_i |a_i|^3 - a_i |b_i|^3}{b_i - a_i} + (1 + \delta) \ln \ln V_n + c_4,
\]

and, defining \( \Delta_i^* = \max_{i \leq n} \Delta_i \),

\[
\alpha M_n \leq \frac{\alpha^2}{2} V_n + \frac{\alpha^3}{6} e^{\alpha \Delta_i^*} W_n + (1 + \delta) \ln \ln V_n + c_4.
\]

Now, since \( V_n \to \infty \), as \( n \to \infty \), if \( n \) had been chosen sufficiently large, and the rational \( \alpha(k) \) computed with sufficient accuracy,

\[
\frac{1}{1 + \delta} \sqrt{\frac{2 \ln \ln V_n}{V_n}} \leq \alpha(k) \leq (1 + \delta) \sqrt{\frac{2 \ln \ln V_n}{V_n}}.
\]

Also, since \( \Delta_n = O((V_n/\ln \ln V_n)^{1/2}) \) and \( W_n = o((V_n^3/\ln \ln V_n)^{1/2}) \), for sufficiently large \( n \),

\[
\frac{\Delta_n}{\sqrt{V_n/\ln \ln V_n}} < B, \quad \text{and} \quad W_n \leq \delta \sqrt{\frac{V_n^3}{\ln \ln V_n}},
\]

for some \( B \). Thus, we get

\[
M_n \leq \sqrt{2V_n \ln \ln V_n} \left\{ \frac{1}{2} (1 + \delta) + \frac{\sqrt{2}}{6} (1 + \delta)^2 \delta e^{(1+\delta) B \sqrt{2}} + \frac{1}{2} (1 + \delta)^2 + \frac{c_4(1 + \delta)}{2 \ln \ln V_n} \right\},
\]
which for sufficiently small $\delta$ implies $M_n \leq \sqrt{2V_n \ln \ln V_n (1 + \varepsilon)}$, that is, 
\[
\limsup_{n \to \infty} M(\xi^n)/\sqrt{2V(\xi^n) \ln \ln V(\xi^n)} \leq 1.
\]

To complete the proof, consider the stochastic sequences $X' = -X$ and $M' = -M$. Any computable non-negative $M'$-martingale is also a computable non-negative $M$-martingale, and so, the $M$-typical sequence $\xi$ is also $M'$-typical. Since $X_i' \in \{a_i', b_i'\}$, where $a_i' = -a_i > 0$, $b_i' = -b_i < 0$, it is easy to check that we can apply the obtained result to $M'$ to have

\[
\limsup_{n \to \infty} \frac{M'(\xi^n)}{\sqrt{2V'(\xi^n) \ln \ln V'(\xi^n)}} = \limsup_{n \to \infty} \frac{-M(\xi^n)}{\sqrt{2V(\xi^n) \ln \ln V(\xi^n)}} \leq 1,
\]

and the conclusion then follows. Theorem 3 is proved.

### 5.2. Variants of the upper half of the law of the iterated logarithm.

In the above result we considered a univariate binary computable basic martingale $M_n = \sum_{i=1}^{n} X_i$, $X_i \in \{a_i, b_i\}$, where $a$ and $b$ are a negative and a positive computable predictable sequence respectively. Of course, we could have just assumed that $a_i \neq 0$, $b_i \neq 0$, and that, for every $i$, $a_i$ and $b_i$ are opposite in sign, without wondering about which one is negative and which one is positive. Easily, given an $M$-typical sequence $\xi$ with respect to a basic martingale $M$ for which the latter assumption is true, we have that $\xi$ is also typical with respect to a computable stochastic sequence for which the former assumption is true.

The situation, however, is completely different if we allow $a_i$ and, or $b_i$ to be zero. In this case, the $M$-martingales $S^{(a)}$ defined in (3) would not be computable any more, since the denominators could be equal to zero. To guarantee the existence of an algorithm giving rational approximations to $S^{(a)}$ to any desired precision we would need to treat separately the case in which $a_i$ and $b_i$ are equal to zero.

**Definition 4.** A stochastic sequence $S: \Omega^* \to \mathbb{R}$ is strongly computable if there exists an algorithm which when fed with $(\xi^n, m)$ yields the symbol $\odot$ if $S(\xi^n) = 0$ and an approximation to $2^{-m}$ if $S(\xi^n) \neq 0$.

With this definition we can now give the following result.

**Theorem 4.** Let us consider the univariate binary computable basic martingale $M_n = \sum_{i=1}^{n} X_i$, where $X_i \in \{a_i, b_i\}$, $a_i \leq 0$, $b_i \geq 0$, is a strongly computable stochastic sequence. Let also $V_n$ and $W_n$ be as in Theorem 3. Then for every $M$-typical sequence $\xi$ such that, as $n \to \infty$, conditions (1) and (2) are true, we have that

\[
\limsup_{n \to \infty} \frac{|M(\xi^n)|}{\sqrt{2V(\xi^n) \ln \ln V(\xi^n)}} \leq 1.
\]

**Proof.** The proof is the same as in Theorem 3, but for the following note. We set $(b_i - a_i)^{-1}(b_ie^{a_i} - a_ie^{b_i}) = 1$, when either $a_i$ or $b_i$, or both,
are equal to zero, so that \( S(\alpha) \) is now defined as \( S_0(\alpha) = 1 \),

\[
S_n(\alpha) = \begin{cases} 
S_{n-1}(\alpha) & a_n < 0, \ b_n > 0, \\
\frac{e^{\alpha X_n}}{(b_n - a_n)^{-1}}(b_n e^{\alpha a_n} - a_n e^{\alpha b_n}) & a_n, \ b_n, \text{ or both, are zero,}
\end{cases}
\]

\( n = 1, 2, \ldots \). In this way, due to the strong computability of \( X \), the \( \mathcal{M} \)-martingales \( S(\alpha) \) are still computable and the previous proof remains perfectly valid. Theorem 4 is proved.

Similarly to the case of the SLLN, it is easy to derive variants of the above theorems for generic martingale transforms.

Acknowledgements. I would like to acknowledge that this research was carried out as part of the author Ph. D. at University College London and I express my gratitude to Professor A. Ph. Dawid and to Dr. V. G. Vovk. I am also grateful to the «Centro Nazionale delle Ricerche» for partial financial support.

REFERENCES

1. Дуб. Дж.Л. Вероятностные процессы. М.: ИЛ, 1956, 605 с.
8. Ширяев А. Н. Вероятность. М.: Наука, 1988, 640 с.