Theory Comput Syst (2012) 50:296–312
DOI 10.1007/s00224-010-9300-9

# On Empirical Meaning of Randomness with Respect to Parametric Families of Probability Distributions

**Vladimir V'yugin**

Published online: 6 November 2010


**Abstract** We study the a priori semimeasure of sets of $P_\theta$-random infinite sequences, where $P_\theta$ is a family of probability distributions depending on a real parameter $\theta$. In the case when for a computable probability distribution $P_\theta$ an effectively strictly consistent estimator exists, we show that Levin's a priory semimeasure of the set of all $P_\theta$-random sequences is positive if and only if the parameter $\theta$ is a computable real number. We show that the a priory semimeasure of the set $\bigcup_\theta I_\theta$, where $I_\theta$ is the set of all $P_\theta$-random sequences and the union is taken over all algorithmically non-random $\theta$, is positive.

**Keywords** Martin-Löf random sequences · A priory semimeasure · Probabilistic machines · Bernoully sequences · Parametric families of probability distributions · Algorithmic information theory · Turing degrees

## 1 Introduction

We use algorithmic randomness theory to analyze "the size" of sets of infinite sequences random with respect to parametric families of probability distributions. We use Levin's [16] a priory lower semicomputable semimeasure as the main tool for this analysis.

Let a parametric family of probability distributions $P_\theta$, where $\theta$ is a real number, be given such that an effectively strictly consistent estimator exists for this family. The Bernoulli family with a real parameter $\theta$ is an example of such family. Theorem 1 shows that Levin's a priory semimeasure of the set of all $P_\theta$-random sequences is positive if and only if the parameter value $\theta$ is a computable real number.

V. V'yugin (✉)
Institute for Information Transmission Problems, Russian Academy of Sciences, Bol'shoi Karetnyi per. 19, GSP-4, Moscow 127994, Russia
e-mail: vyugin@iitp.ru

In this way, the a priory semimeasure of the set of all $P_\theta$-random sequences is zero for non-computable $\theta$. As follows from Theorem 2 the set of $P_\theta$-random sequences have positive a priory semimeasure if these $\theta$ form a set of sequences Martin-Löf random with respect to some computable prior.

We give in Appendix the simple proof of our previous result formulated in Theorem 3 which says that Levin's a priory semimeasure of the set of all infinite binary sequences non-equivalent by Turing to Martin-Löf random sequences is positive. In particular, these sequences are non-random with respect to each computable probability distribution.

We use this result to prove our main result—Theorem 4. This theorem shows that a probabilistic machine can be constructed, which with probability close to one outputs a random $\theta$-Bernoulli sequence such that the parameter $\theta$ is not random with respect to each computable probability distribution. This result can be interpreted such that the Bayesian statistical approach is insufficient to cover all possible "meaningful" cases for $\theta$-random sequences.

This paper is an extended version of the conference paper [15].

## 2 Preliminaries

Let $\Xi$ be the set of all finite binary sequences, $\Lambda$ be the empty sequence, and $\Omega$ be the set of all infinite binary sequences. We write $x \subseteq y$ if a sequence $y$ is an extension of a sequence $x$, $l(x)$ is the length of $x$. For any $\omega \in \Omega$, $\omega^n = \omega_1 \cdots \omega_n$. A real-valued function $P(x)$, where $x \in \Xi$, is called semimeasure if

$$\begin{aligned} P(\Lambda) &\le 1, \\ P(x0) + P(x1) &\le P(x) \end{aligned} \tag{1}$$

for all $x$, and the function $P$ is semicomputable from below; this means that the set $\{(r, x) : r < P(x)\}$, where $r$ is a rational number, is recursively enumerable. A definition of upper semicomputability is analogous.

Solomonoff [11] proposed ideas for defining the a priori probability distribution on the basis of the general theory of algorithms. Levin [3, 16] gave a precise form of Solomonoff's ideas in a concept of a maximal semimeasure semicomputable from below (see also Li and Vitányi [7], Sect. 4.5, Shen et al. [10]). Levin proved that there exists a maximal to within a multiplicative positive constant factor semimeasure $M$ semicomputable from below, i.e. such that for every semimeasure $P$ semicomputable from below a positive constant $c$ exists such that the inequality

$$cM(x) \ge P(x) \tag{2}$$

holds for all $x$. The semimeasure $M$ is called *the a priory* or universal semimeasure.

For any semimeasure $Q$, its *support set* $E_Q$ is a set of all infinite sequences $\omega$ such that $Q(\omega^n) > 0$ for all $n$, i.e., $E_Q = \bigcup_{Q(x)>0} \Gamma_x$.

A function $P$ is a measure if (1) holds, where both inequality signs $\le$ are replaced on $=$. Any function $P$ satisfying (1) (with equalities) can be extended on all Borel subsets of $\Omega$ if we define $P(\Gamma_x) = P(x)$ in $\Omega$, where $x \in \Xi$ and $\Gamma_x = \{\omega \in \Omega :$

$x \subseteq \omega\}$; after that, we use the standard method for extending $P$ to all Borel subsets of $\Omega$. By simple set in $\Omega$ we mean a union of intervals $\Gamma_x$ from a finite set.

A measure $P$ is computable if it is, at one time, lower and upper semicomputable.

For technical reasons, for any semimeasure $P$, we consider the maximal measure $\bar{P}$ such that $\bar{P} \le P$. This measure satisfies

$$\bar{P}(x) = \inf_n \sum_{l(y)=n, x\subseteq y} P(y).$$

In general, the measure $\bar{P}$ is non-computable (and it is not a probability measure). By (2), for each lower semicomputable semimeasure $P$, the inequality $c\bar{M}(A) \ge \bar{P}(A)$ holds for every Borel set $A$, where $c$ is a positive constant.

In the manner of Levin's papers [4–6, 16] (see also [14]), we consider combinations of probabilistic and deterministic processes as the most general class of processes for generating data. With any probabilistic process some computable probability distribution can be assigned. Any deterministic process is realized by means of an algorithm. Algorithmic processes transform sequences generated by probabilistic processes into new sequences. More precise, a probabilistic computer is a pair $(P, F)$, where $P$ is a computable probability distribution, and $F$ is a Turing machine supplied with an additional input tape. In the process of computation this machine reads on this tape a sequence $\omega$ distributed according to $P$ and produces a sequence $\omega' = F(\omega)$ (a correct definition see in [4, 7, 10, 14]). So, we can compute the probability

$$Q(x) = P\{\omega \in \Omega : x \subseteq F(\omega)\}$$

that the result $F(\omega)$ of the computation begins with a finite sequence $x$. It is easy to see that $Q(x)$ is a semimeasure semicomputable from below.

Generally, the semimeasure $Q$ can not be a probability distribution on $\Omega$, since $F(\omega)$ may be finite for some infinite $\omega$.

The converse result is proved in Zvonkin and Levin [16]: for every semimeasure $Q(x)$ semicomputable from below a probabilistic computer $(L, F)$ exists such that

$$Q(x) = L\{\omega | x \subseteq F(\omega)\},$$

for all $x$, where $L(x) = 2^{-l(x)}$ is the uniform probability distribution on the set of all binary sequences.

Analogously, for any Borel set $A \subseteq \Omega$ consisting of infinite sequences, we consider the probability

$$Q(A) = L\{\alpha \in \Omega : F(\alpha) \in A\} \tag{3}$$

of generating a sequence $\omega \in A$ by means of a probabilistic computer $F$. Obviously, we have $c\bar{M}(A) \ge Q(A)$ for all such $A$, where $c$ is a positive constant.

Therefore, by (2) and (3) $M(x)$ and $\bar{M}(A)$ define universal upper bounds of the probability of generating $x$ and $\omega \in A$ by probabilistic computers.

We distinguish between subsets of $\Omega$ of $\bar{M}$-measure 0 and subsets of positive measure $\bar{M}$. If $\bar{M}(A) = 0$ then the probability of generating a sequence $\omega \in A$ by means of any probabilistic computer is equal to 0.

The simplest example of a set of $\bar{M}$-measure 0 is $A = \{\omega\}$, where $\omega$ is a non-computable sequence. Indeed, if $\bar{M}\{\omega\} > 0$ then there exist a rational $r > 0$ such that $M(\omega^n) > r$ for all $n$. Obviously, there are only finite number of uncomparable strings $x$ such that $M(x) > r$. Then there exists an $k$ such that $\omega^k \subseteq x$ and $M(x) > r$ imply $x \subseteq \omega$. We can compute each bit of $\omega$ by enumerating all such $x$.

The sets of $\bar{M}$-measure 0 were described by Levin [4, 5] in terms of quantity of information.

We refer readers to Li and Vitányi [7] and to Shen et al. [10] for the theory of algorithmic randomness. We use the definition of a random sequence in terms of universal probability. Let $P$ be some computable measure in $\Omega$. The deficiency of randomness of a sequence $\omega \in \Omega$ with respect to $P$ is defined as

$$d(\omega|P) = \sup_n \frac{M(\omega^n)}{P(\omega^n)}, \tag{4}$$

where $\omega^n = \omega_1\omega_2\cdots\omega_n$. This definition leads to the same class of random sequences as the original Martin-Löf [8] definition. Let $R_P$ be the set of all infinite binary sequences random with respect to a measure $P$

$$R_P = \{\omega \in \Omega : d(\omega|P) < \infty\}. \tag{5}$$

We also consider *parametric families* of probability distributions $P_\theta(x)$, where $\theta$ is a real number; we suppose that $\theta \in [0, 1]$. An example of such a family is the Bernoulli family $B_\theta(x) = \theta^k(1-\theta)^{n-k}$, where $n$ is the length of $x$ and $k$ is the number of ones in it.

We associate with a binary sequence $\theta_1\theta_2\cdots$ a real number with the binary expansion $0.\theta_1\theta_2\cdots$. When the sequence $\theta_1\theta_2\cdots$ is computable or random with respect to some measure we say that the number $0.\theta_1\theta_2\cdots$ is computable or random with respect to the corresponding measure in [0, 1].

We consider probability distributions $P_\theta$ computable with respect to a parameter $\theta$. Informally, this means that there exists an algorithm enumerating all triples $(x, r_1, r_2)$, where $x \in \Xi$ and $r_1, r_2$ are rational numbers, such that $r_1 < P_\theta(x) < r_2$. This algorithm uses an infinite sequence $\theta$ as an additional input; if some triple $(x, r_1, r_2)$ is enumerated by this algorithm then only a finite initial fragment of $\theta$ was used in the process of computation (for correct definition, see also Shen et al. [10] and Vovk and V'yugin [12]).

Analogously, we consider parametric lower semicomputable semimeasures. It can be proved that there exist a universal parametric lower semicomputable semimeasure $M_\theta$. This means that for each parametric lower semicomputable semimeasure $R_\theta$ there exists a positive constant $C$ such that $CM_\theta(x) \ge R_\theta(x)$ for all $x$ and $\theta$.

The corresponding definition of randomness with respect to a family $P_\theta$ is obtained by relativization of (4) with respect to $\theta$

$$d_\theta(\omega) = \sup_n \frac{M_\theta(\omega^n)}{P_\theta(\omega^n)}$$

(see also [3]). This definition leads to the same class of random sequences as the original Martin-Löf [8] definition relativized with respect to a parameter $\theta$.

For any $\theta$, let

$$I_\theta = \{\omega \in \Omega : d_\theta(\omega) < \infty\}$$

be the set of all infinite binary sequences random with respect to the measure $P_\theta$. In case of Bernoulli family, we call elements of this set *$\theta$-Bernoulli sequences*.

## 3 Randomness with Respect to a Parameter Family

We need some statistical notions (see Cox and Hinkley [2]). Let $P_\theta$ be some computable parametric family of probability distributions. A function $\hat{\theta}(x)$ from $\Xi$ to [0, 1] is called *an estimator*. An estimator $\hat{\theta}$ is called *strictly consistent* if for each parameter value $\theta$ for $P_\theta$-almost all $\omega$,

$$\hat{\theta}(\omega^n) \to \theta$$

as $n \to \infty$.

Let $\epsilon$ and $\delta$ be rational numbers. An estimator $\hat{\theta}$ is called *effectively strictly consistent* if there exists a computable function $N(\epsilon, \delta)$ such that for each $\theta$ for all $\epsilon$ and $\delta$

$$P_\theta\left\{\omega \in \Omega : \sup_{n \geq N(\epsilon,\delta)} |\hat{\theta}(\omega^n) - \theta| > \epsilon\right\} \leq \delta. \tag{6}$$

The strong law of large numbers Borovkov [1] (Chap. 5)

$$B_\theta\left\{\sup_{k\geq n}\left|\frac{1}{k}\sum_{i=1}^{k}\omega_i - \theta\right| \geq \epsilon\right\} < \frac{1}{\epsilon^4 n}$$

shows that the function $\hat{\theta}(\omega^n) = \frac{1}{n}\sum_{i=1}^{n}\omega_i$ is a computable strictly consistent estimator for the Bernoulli family $B_\theta$.

**Proposition 1** *For any effectively strictly consistent estimator $\hat{\theta}$,*

$$\lim_{n\to\infty} \hat{\theta}(\omega^n) = \theta$$

*for each $\omega \in I_\theta$.*

*Proof* Assume an infinite sequence $\omega$ to be Martin-Löf random with respect to $P_\theta$ for some $\theta$.

At first, we prove that $\lim_{n\to\infty}\hat{\theta}(\omega^n)$ exists. Let for $j = 1, 2, \ldots,$

$$W_j = \{\alpha \in \Omega : (\exists n, k \geq N(1/j, 2^{-(j+1)}))|\hat{\theta}(\alpha^n) - \hat{\theta}(\alpha^k)| > 1/j\}.$$

By (6) for any $\theta$, $P_\theta(W_j) < 2^{-j}$ for all $j$. Define $V_i = \bigcup_{j>i} W_j$ for all $i$. By definition for any $\theta$, $P_\theta(V_i) < 2^{-i}$ for all $i$. Also, any set $V_i$ can be represented as a recursively enumerable union of intervals of type $\Gamma_x$. To reduce this definition of

Martin-Löf test to the definition of the test (4) define a sequence of uniform lower semicomputable parametric semimeasures

$$R_{\theta,i}(x) = \begin{cases} 2^i P_\theta(x) & \text{if } \Gamma_x \subseteq V_i, \\ 0 & \text{otherwise} \end{cases}$$

and consider the mixture $R_\theta(x) = \sum_{i=1}^{\infty} \frac{1}{i(i+1)} R_{\theta,i}(x)$.

Suppose that $\lim_{n\to\infty} \hat{\theta}(\omega^n)$ does not exist. Then for each sufficiently large $j$, $|\hat{\theta}(\omega^n) - \hat{\theta}(\omega^k)| > 1/j$ for infinitely many $n$ and $k$. This implies that $\omega \in V_i$ for all $i$, and then for some positive constant $c$,

$$d_\theta(\omega) = \sup_n \frac{M_\theta(\omega^n)}{P_\theta(\omega^n)} \geq \sup_n \frac{R_\theta(\omega^n)}{c P_\theta(\omega^n)} = \infty,$$

i.e., $\omega$ is not Martin-Löf random with respect to $P_\theta$.

Suppose that $\lim_{n\to\infty} \hat{\theta}(\omega^n) \neq \theta$. Then the rational numbers $r_1, r_2$ exist such that $r_1 < \lim_{n\to\infty} \hat{\theta}(\omega^n) < r_2$ and $\theta \notin [r_1, r_2]$. Since the estimator $\hat{\theta}$ is consistent, $P_\theta\{\alpha : r_1 < \lim_{n\to\infty} \hat{\theta}(\alpha^n) < r_2\} = 0$, and we can effectively (using $\theta$) enumerate an infinite sequence of positive integer numbers $n_1 < n_2 < \cdots$ such that for

$$W'_j = \bigcup \{\Gamma_x : l(x) \geq n_j, r_1 < \hat{\theta}(x) < r_2\},$$

we have $P_\theta(W'_j) < 2^{-j}$ for all $j$. Define $V'_i = \bigcup_{j>i} W'_j$ for all $i$. We have $P_\theta(V'_i) \leq 2^{-i}$ and $\omega \in V'_i$ for all $i$. Then $\omega$ can not be Martin-Löf random with respect to $P_\theta$. These two contradictions obtained above prove the proposition. $\square$

The following theorem generalizes the simplest example of a set of $\bar{M}$-measure 0 presented in Sect. 2. It can be interpreted such that $P_\theta$-random sequences with "a pre-specified" non-computable parameter $\theta$ can not be obtained in any combinations of stochastic and deterministic processes.

**Theorem 1** *Assume a computable parametric family $P_\theta$ of probability distributions has an effectively strictly consistent estimator. Then for each $\theta$, $\bar{M}(I_\theta) > 0$ if and only if $\theta$ is computable.*

*Proof* If $\theta$ is computable real number then the probability distribution $P_\theta$ is also computable and by (2) $c\bar{M}(I_\theta) \geq P_\theta(I_\theta) = 1$, where $c$ is a positive constant.

The proof of the converse assertion is more complicated. Let $\hat{\theta}$ be an effectively strictly consistent estimator for a computable parametric family $P_\theta$.

Assume $\bar{M}(I_\theta) > 0$. There exists a simple set $V$ (a union of a finite set of intervals) and a rational number $r$ such that $\frac{1}{2}\bar{M}(V) < r < \bar{M}(I_\theta \cup V)$. For any finite set $X \subseteq \Xi$, let $\bar{X} = \bigcup_{x\in X} \Gamma_x$.

Let $n$ be a positive integer number. We compute a rational approximation $\theta_n$ of $\theta$ up to $\frac{1}{2n}$ as follows. Using the exhaustive search, we find a finite set $X_n$ of pairwise incomparable finite sequences of length $\ge N(1/n, 2^{-n})$ such that

$$\bar{X}_n \subseteq V, \quad \sum_{x \in X_n} M(x) > r,$$
$$|\hat{\theta}(x) - \hat{\theta}(x')| \le \frac{1}{2n} \tag{7}$$

for all $x, x' \in X_n$. If any such set $X_n$ will be found, we put $\theta_n = \hat{\theta}(x)$, where $x$ is the minimal element of $X_n$ with respect to some natural (lexicographic) ordering of all finite binary sequences.

Let us prove that for each $n$ some such set $X_n$ exists. Since $\bar{M}(I_\theta \cap V) > r$, there exists a closed (in the topology defined by intervals $\Gamma_x$) set $E \subseteq I_\theta \cap V$ such that $\bar{M}(E) > r$. Consider the function

$$f_k(\omega) = \inf\left\{n : n \ge k, |\hat{\theta}(\omega^n) - \theta| \le \frac{1}{4n}\right\}.$$

By Proposition 1 this function is defined and continuous on the set $E$ and, since $E$ is compact, it is bounded on $E$. Hence, for each $k$, there exists a finite set $X \subseteq \Xi$ consisting of pairwise incomparable sequences of length $\ge k$ such that $E \subseteq \bar{X}$ and $|\hat{\theta}(x) - \hat{\theta}(x')| \le \frac{1}{2n}$ for all $x, x' \in X$. Since $E \subseteq \bar{X}$, we have $\sum_{x \in X} M(x) > r$. Therefore, the set $X_n$ can be found by exhaustive search.

**Lemma 1** *For any Borel set $V \subseteq \Omega$, $\bar{M}(V) > 0$ and $V \subseteq I_\theta$ imply $P_\theta(V) > 0$.*

*Proof* By definition of $M_\theta$ each computable parametric measure $P_\theta$ is absolutely continuous with respect to the measure $\bar{M}_\theta$, and so, we have representation

$$P_\theta(X) = \int_X \frac{dP_\theta}{d\bar{M}_\theta}(\omega)\, d\bar{M}_\theta(\omega), \tag{8}$$

where $\frac{dP_\theta}{d\bar{M}_\theta}(\omega)$ is the Radon-Nicodim derivative; it exists for $\bar{M}_\theta$-almost all $\omega$.

By definition we have for $\bar{M}_\theta$-almost all $\omega \in I_\theta$

$$\frac{dP_\theta}{d\bar{M}_\theta}(\omega) = \lim_{n\to\infty} \frac{P_\theta}{\bar{M}_\theta}(\omega^n) \ge \liminf_{n\to\infty} \frac{P_\theta}{\bar{M}_\theta}(\omega^n) \ge C_{\theta,\omega} > 0. \tag{9}$$

By definition $c_\theta \bar{M}_\theta(X) \ge \bar{M}(X)$ for all Borel sets $X$, where $c_\theta$ is some positive constant (depending on $\theta$). Then by (8) and (9) the inequality $\bar{M}(X) > 0$ implies $P_\theta(X) > 0$ for each Borel set $X$. □

We rewrite (6) in the form

$$E_n = \left\{\omega \in \Omega : \sup_{N \ge N(1/(2n), 2^{-n})} |\hat{\theta}(\omega^N) - \theta| \ge \frac{1}{2n}\right\}. \tag{10}$$

By definition $P_\theta(E_n) \le 2^{-n}$ for all $n$. We prove that $X_n \not\subseteq E_n$ for almost all $n$. Suppose that the opposite assertion holds. Then there exists an increasing infinite sequence of positive integer numbers $n_1, n_2, \ldots$ such that $X_{n_i} \subseteq E_{n_i}$ for all $i = 1, 2, \ldots$.

This implies $P_\theta(X_{n_i}) \le 2^{-n_i}$ for all $i$. For any $k$, define $U_k = \bigcup_{i \ge k} X_{n_i}$. Clearly, we have for all $k$, $\bar{M}(\bar{U}_k) > r$ and $P_\theta(\bar{U}_k) \le \sum_{i \ge k} 2^{-n_i} \le 2^{-n_k+1}$. Let $U = \bigcap U_k$. Then $P_\theta(U) = 0$ and $\bar{M}(U) \ge r > \frac{1}{2}\bar{M}(V)$. From $U \subseteq V$ and $\bar{M}(I_\theta \cap V) > \frac{1}{2}\bar{M}(V)$ the inequality $\bar{M}(I_\theta \cap U) > 0$ follows. Then the set $I_\theta \cap U$ consists of $P_\theta$-random sequences, $P_\theta(I_\theta \cap U) = 0$ and $\bar{M}(I_\theta \cap U) > 0$. This is a contradiction with Lemma 1.

Assume $X_n \not\subseteq E_n$ for all $n \ge n_0$. Let also, a finite sequence $x_n \in X_n$ is defined such that

$$\Gamma_{x_n} \cap (\Omega \setminus E_n) \neq \emptyset.$$

Then from $l(x_n) \ge N(\frac{1}{2n}, 2^{-n})$ the inequality

$$\left|\hat{\theta}(x_n) - \theta\right| < \frac{1}{2n}$$

follows. By (7) we obtain $|\theta_n - \theta| < \frac{1}{n}$. This means that the real number $\theta$ is computable. Theorem is proved. □

Let $Q$ be a computable probability distribution on $\theta$s (i.e., on the set $\Omega$). Then the Bayesian mixture with respect to the prior $Q$

$$P(x) = \int P_\theta(x)\, dQ(\theta)$$

is also computable probability distribution.

Recall that $R_Q$ is the set of all infinite sequences Martin-Löf random with respect to a computable probability measure $Q$. Obviously, $P(\bigcup_{\theta \in R_Q} I_\theta) = 1$, and then $\bar{M}(\bigcup_{\theta \in R_Q} I_\theta)) > 0$. Moreover, it follows from Corollary 4 of Vovk and V'yugin [12]

**Theorem 2** *For any computable measure $Q$, a sequence $\omega$ is random with respect to the Bayesian mixture $P$ if and only if $\omega$ is random with respect to a measure $P_\theta$ for some $\theta$ random with respect to the measure $Q$; in other words,*

$$R_P = \bigcup_{\theta \in R_Q} I_\theta.$$

Notice that each computable $\theta$ is Martin-Löf random with respect to the computable probability distribution concentrated on this sequence.

## 4 Randomness with Respect to Non-random Parameters

We show in this section that the Bayesian approach is insufficient to cover all possible "meaningful" cases: a probabilistic machine can be constructed, which with probability close to one outputs a random $\theta$-Bernoulli sequence, where the parameter $\theta$ is not random with respect to each computable probability distribution.

Let $\mathcal{P}(\Omega)$ be the set of all computable probability measures on $\Omega$ and let

$$\mathcal{S} = \bigcup_{P \in \mathcal{P}(\Omega)} R_P$$

be the set of all sequences Martin-Löf random with respect to computable probability measures, where $R_P$ is the set of all $P$-random sequences (5). We call these sequences—*stochastic*. Let $\mathcal{S}^c$ be a complement of $\mathcal{S}$—the set of *non-stochastic* sequences.

An infinite binary sequence $\alpha$ is Turing reducible to an infinite binary sequence $\beta$ if $\alpha = F(\beta)$ for some computable operation $F$; we denote this $\alpha \leq_T \beta$. Two infinite sequences $\alpha$ and $\beta$ are Turing equivalent if $\alpha \leq_T \beta$ and $\beta \leq_T \alpha$. Let

$$Cl(\mathcal{S}) = \{\alpha : \exists \beta (\beta \in \mathcal{S} \ \& \ \beta \leq_T \alpha)\}. \tag{11}$$

The complement of the set (11), $Cl(\mathcal{S})^c = \Omega \setminus Cl(\mathcal{S})$, contains all sequences non-random with respect to all computable probability distributions, i.e., $Cl(\mathcal{S})^c \subseteq \mathcal{S}^c$; moreover, it contains all sequences which can not be Turing equivalent to stochastic sequences. Also, no stochastic sequence can be Turing reducible to a sequence from $Cl(\mathcal{S})^c$.

V'yugin [13, 14] proved that $\bar{M}(Cl(\mathcal{S})^c) > 0$.[1]

**Theorem 3** *For any $\epsilon$, $0 < \epsilon < 1$, a lower semicomputable semimeasure $Q$ exists such that $\bar{Q}(E_Q) > 1 - \epsilon$ and $E_Q \subseteq Cl(\mathcal{S})^c$.*

For completeness of presentation we give in Appendix a new simplified proof of this theorem.

We show that result of Theorem 3 can be extended to parameters of the Bernoulli family.

**Theorem 4** *Let $I_\theta$ be the set of all $\theta$-Bernoulli sequences. Then*

$$\bar{M}\Big( \bigcup_{\theta \in Cl(\mathcal{S})^c} I_\theta \Big) > 0.$$

*In terms of probabilistic computers, for any $\epsilon$, $0 < \epsilon < 1$, a probabilistic machine $(L, F)$ can be constructed, which with probability $\geq 1 - \epsilon$ generates an $\theta$-Bernoulli sequence, where $\theta \in Cl(\mathcal{S})^c$ (i.e., $\theta$ is non-stochastic).*

*Proof* For any $\epsilon > 0$, $0 < \epsilon < 1$, we define a lower semicomputable semimeasure $P$ such that

$$\bar{P}\Big( \bigcup_{\theta \in Cl(\mathcal{S})^c} I_\theta \Big) > 1 - \epsilon.$$

The proof of the theorem is based on Theorem 3.

[1] Does $Cl(\mathcal{S}) = \mathcal{S}$ holds is an open question.

Let $Q$ be the semimeasure defined in this theorem. For any $\omega \notin E_Q$ we have $Q(\omega^n) = 0$ for all sufficiently large $n$. For the measure

$$R^-(x) = \int B_\theta(x)\, d\bar{Q}(\theta), \tag{12}$$

where $B_\theta$ is the Bernoulli measure, we have $R^-(\Omega) > 1 - \epsilon$ by Theorem 3, and $R^-(\bigcup_{\theta \in Cl(\mathcal{S})} I_\theta) = 0$.

Unfortunately, we can not conclude that $c\bar{M} \ge R^-$ for some constant $c$, since the measure $R^-$ is not represented in the form $R^- = \bar{P}$ for some lower semicomputable semimeasure $P$. To overcome this problem, we consider some semicomputable approximation of this measure.

For any finite binary sequences $\alpha$ and $x$, let $B^-_\alpha(x) = (\theta^-)^K(1-\theta^+)^{N-K}$, where $N$ is the length of $x$ and $K$ is the number of ones in it, $\theta^-$ is the left side of the subinterval corresponding to the sequence $\alpha$ and $\theta^+$ is its right side. By definition $B^-_\alpha(x) \le B_\theta(x)$ for all $\theta^- \le \theta \le \theta^+$.

Let $\epsilon$ be a rational number such that $0 < \epsilon < 1$. Let $Q^s(x)$ be equal to the maximal rational number $r < Q(x)$ computed in $s$ steps of enumeration of $Q(x)$ from below. Using Theorem 3, we can define for $n = 1, 2, \dots$ and for each $x$ of length $n$ a computable sequence of positive integer numbers $s_x \ge n$ and a sequence of finite binary sequences $\alpha_{x,1}, \alpha_{x,2}, \dots, \alpha_{x,k_x}$ of length $\ge n$ such that the function $P(x)$ defined by

$$P(x) = \sum_{i=1}^{k_x} B^-_{\alpha_{x,i}}(x) Q^{s_x}(\alpha_{x,i}) \tag{13}$$

is a semimeasure, i.e., such that condition (1) holds for all $x$, and such that

$$\sum_{l(x)=n} P(x) > 1 - \epsilon \tag{14}$$

holds for all $n$. These sequences exist, since the limit function $R^-$ defined by (12) is a measure satisfying $R^-(\Omega) > 1 - \epsilon$.

By definition the semimeasure $P(x)$ is lower semicomputable. Then $cM(x) \ge P(x)$ holds for all $x \in \Xi$, where $c$ is a positive constant.

To prove that $\bar{P}(\Omega \setminus \bigcup_\theta I_\theta) = 0$ we consider some probability measure $Q^+ \ge Q$. Since (1) holds, it is possible to define some non-computable measure $Q^+$ satisfying these properties in many different ways. Define the mixture of the Bernoulli measures with respect to $Q^+$

$$R^+(x) = \int B_\theta(x)\, dQ^+(\theta). \tag{15}$$

By definition $R^+(\Omega \setminus \bigcup_\theta I_\theta) = 0$. Using definitions (13) and (15), it can be easily proved that $\bar{P} \le R^+$. Then $\bar{P}(\Omega \setminus \bigcup_\theta I_\theta) = 0$. By Theorem 3 $Cl(\mathcal{S}) \subseteq \Omega \setminus E_Q$, and then $\bar{Q}(Cl(\mathcal{S})) = 0$. By (13) we have $\bar{P}(\bigcup_{\theta \in Cl(\mathcal{S})} I_\theta) = 0$. By (14) we have $\bar{P}(\Omega) > 1 - \epsilon$. Then $\bar{P}(\bigcup_{\theta \in Cl(\mathcal{S})^c} I_\theta) > 1 - \epsilon$. Therefore, $\bar{M}(\bigcup_{\theta \in Cl(\mathcal{S})^c} I_\theta) > 0$. $\square$

Theorem 4 can be easily extended for an arbitrary parametric family $P_\theta$ computable with respect to a real parameter $\theta$. We pass the details of computability with respect to a real parameter $\theta$.

In this case using continuity of $P_\theta$ with respect to the parameter $\theta$, we can define instead of $B^-_\alpha(x)$ a rational approximation $P^-_\alpha(x)$ of $P_\theta(x)$ from below such that (14) holds. The corresponding measure $P(x)$ is defined by (13), where $B^-_\alpha(x)$ is replaced with $P^-_\alpha(x)$.

Moreover, we can strengthen this result replacing $\Omega \setminus Cl(\mathcal{S})^c$ with an arbitrary Borel set $A$ of parameters (where $A \subseteq \Omega$) such that $\bar{M}(A) > 0$. In this case for any $\epsilon > 0$, a lower semicomputable semimeasure $Q$ exists such that $\bar{Q}(A) > 1 - \epsilon/2$. The proof is analogous to the proof of Theorem 4 with an exception that $\bar{Q}(Cl(\mathcal{S})) = 0$ is replaced with $\bar{Q}(\Omega \setminus A) \le \epsilon/2$ and $\bar{P}(\bigcup_{\theta \in Cl(\mathcal{S})} I_\theta) = 0$ is replaced with $\bar{P}(\bigcup_{\theta \in \Omega \setminus A} I_\theta) \le \epsilon/2$.

Therefore, we have proved the following

**Theorem 5** *Assume a parametric family $P_\theta$ of probability distributions computable with respect to a parameter $\theta$ is given. Let also for any $\theta$, $I_\theta$ be the set of all sequences random with respect to a probability measure $P_\theta$. Then for any Borel $A \subseteq \Omega$, if $\bar{M}(A) > 0$ then $\bar{M}(\bigcup_{\theta \in A} I_\theta) > 0$.*

## 5 Conclusion

In this paper we have analyzed parametric probabilistic models in the algorithmic theory framework. We used Levin's a priory lower semicomputable semimeasure as the main tool for this analysis.[2]

We say that a property of infinite sequences has no "empirical meaning" if Levin's a priory semimeasure of the set of all sequences possessing this property is zero. Equivalently, any probabilistic computer can output a sequence possessing this property only with probability 0.

Evidently, the a priory semimeasure of a set of infinite sequences Martin-Löf random with respect to a computable measure is positive. We have proved that the a priory semimeasure of the set of $P$-random sequences is zero if this measure depends on an individual non-computable parameter $\theta$ even if this measure is computable with respect to $\theta$. In this respect, the mathematical model of the biased coin with "a prespecified" probability $\theta$ of head is meaningless when $\theta$ is a non-computable real number. For example, a non-computable real number $\theta$ can be defined by means of a mathematical theory. Non-computable parameters $\theta$ can have empirical meaning only in their totality, i.e., as elements of some uncountable sets. For example, $P_\theta$-random sequences with non-computable $\theta$ can be generated by a Bayesian mixture of these $P_\theta$ using a computable prior. In this case, evidently, the semicomputable

[2] In the algorithmic randomness framework, the problem of generating infinite sequences possessing a given property using probabilistic computers was considered by Zvonkin and Levin in the survey [16] (Sect. 4), where original results of Levin, Petri, Bardzin and Agafonov closely related to this work were presented.

semimeasure of the set of all sequences random with respect to this mixture is positive.

We have presented an example which shows that Bayesian approach is insufficient to cover all possible "meaningful" cases for $\theta$-random sequences. We have shown that a probabilistic machine with probability close to 1 can generate $\theta$-Bernoulli sequences, where $\theta$ is not only non-computable but cannot be Turing equivalent to sequences Martin-Löf random with respect to a computable probability distribution. This result can be considered as a counterexample for Bayesian approach to statistics. This approach is based on assumption that parameters $\theta$ are random with respect to some a priory probability distribution.

**Acknowledgements** This research was partially supported by Russian foundation for fundamental research: 09-07-00180a and 09-01-00709a.

## Appedix: Proof of Theorem 3

Recall that $E_Q$ is the support set of a semimeasure $Q$. In that follows for any $\epsilon > 0$ we define a semicomputable semimeasure $Q$ such that

(1) $\bar{Q}(E_Q) > 1 - \epsilon$;
(2) for each $\omega \in E_Q$ and for each computable operation $F$ such that $F(\omega)$ is infinite, the sequence $F(\omega)$ is not Martin-Löf random with respect to the uniform probability measure $L$ on $\Omega$.

By Theorem 4.2 from [16] for each computable measure $P$ on $\Omega$, there exist two computable operations $F$ and $G$ such that

(3) $F(\omega) \in \Omega$ for each $\omega$ random with respect to $L$, and $G(F(\omega)) = \omega$;
(4) for each sequence $\omega$ random with respect to $P$ (and such that $P\{\omega\} = 0$), the sequence $G(\omega)$ is random with respect to $L$.

By (1)–(4) each sequence $\omega \in E_Q$ can not be Martin-Löf random with respect to any computable probability measure $P$.

We will construct a semicomputable semimeasure $Q$ as some sort of network flow. We define an infinite network on the base of the infinite binary tree. This network has no sink; the top of the tree (empty sequence) is the source.

Each $x \in \Xi$ defines two edges $(x, x0)$ and $(x, x1)$ of length one. In the construction below we will add to the network extra edges $(x, y)$ of length $> 1$, where $x, y \in \Xi$, $x \subseteq y$ and $y \neq x0, x1$. By the length of the edge $(x, y)$ we mean the number $l(y) - l(x)$. For any edge $\sigma = (x, y)$ we denote by $st(\sigma) = x$ its starting vertex and by $ter(\sigma) = y$ its terminal vertex. A computable function $q(\sigma)$ defined on all edges of length one and on all extra edges and taking rational values is called *a network* if for all $x \in \Xi$

$$\sum_{\sigma:\, st(\sigma)=x} q(\sigma) \leq 1.$$

Let $G$ be the set of all extra edges of the network $q$ (it is a part of the domain of $q$). By *q-flow* we mean the minimal semimeasure $P$ such that $P \geq R$, where the function

$R$ is defined by the following recursive equations

$$R(\lambda) = 1;$$
$$R(y) = \sum_{\sigma:\, ter(\sigma)=y} q(\sigma)R(st(\sigma)) \tag{16}$$

for $y \neq \lambda$. It is easy to see that this semimeasure $P$ is lower semicomputable if $q$ is computable.

A network $q$ is called *elementary* if the set of extra edges is finite and $q(\sigma) = 1/2$ for almost all edges of unit length. For any network $q$, we define the *network flow delay* function ($q$-delay function)

$$d(x) = 1 - q(x, x0) - q(x, x1).$$

The construction below works with all programs $i$ computing the operations $F_i(x)$.[3] We define some function $p(n)$ such that for each positive integer number $m$ we have $p(n) = m$ for infinitely many $n$. For example, we can define $p(\langle m, k\rangle) = m$ and $p'(\langle m, k\rangle) = k$ for all $m$ and $k$, where $\langle m, k\rangle$ is some computable one-to-one enumeration of all pairs of nonnegative integer numbers. Then for each step $n$ we compute $\langle i, s\rangle = p(n)$, where $i$ is a program and $s$ is a number (we call $s$ number of a session); so, $i = p(p(n))$ and $s = p'(p(n))$.

Let a program $i$, a number $s$, finite binary sequences $x$ and $y$, an elementary network $q$, and a nonnegative integer number $n$ be given. Define $B(\langle i, s\rangle, x, y, q, n)$ be *true* if the following conditions hold

(i) $l(y) = n$, $x \subseteq y$;
(ii) $d(y^k) < 1$ for all $k$, $1 \le k \le n$, where $d$ is the $q$-delay function and $y^k = y_1 \cdots y_k$;
(iii) $l(F_i(y)) > \langle x, s\rangle$.

Let $B(\langle i, s\rangle, x, y, q, n)$ be *false*, otherwise. Define

$$\beta(x, q, n) = \min\bigl\{y : p(l(y)) = p(l(x)), B(\langle p(p(l(x))), p'(p(l(x)))\rangle, x, y, q, n)\bigr\}$$

Here $p(p(l(x)))$ is a program and $p'(p(l(x)))$ is a number of session; min is considered for lexicographical ordering of strings; we suppose that $\min\emptyset$ is undefined.

**Lemma 2** *For each computable operation $F_i$ and for each finite sequence $x$ such that $F(\omega) \in \Omega$ for some infinite extension $\omega$ of $x$ (i.e., $x \subseteq \omega$), $\beta(x, q, n)$ is defined for all sufficiently large $n$ such that $p(p(n)) = i$.*

*Proof* The needed sequence $y$ exists for all sufficiently large $n$, since $l(F_i(\omega^n)) > \langle x, s\rangle$ holds for all sufficiently large $n$, $p(n) = \langle i, s\rangle$. □

[3] The existence of the effectively computable sequence $\{F_i\}$ such that for each computable operation $F$, $F = F_i$ for some $i$ is proved in [9].

The goal of the construction below is the following. Each extra edge $\sigma$ will be assigned to some task number $I = \langle i, s \rangle$ such that $p(l(st(\sigma))) = p(l(ter(\sigma))) = I$. The goal of the task $I$ is to define a finite set of extra edges $\sigma$ such that for each infinite binary sequence $\omega$ one of the following conditions hold: either $\omega$ contains some extra edge as a subword, or the network flow delay function $d$ equals 1 on some initial fragment of $\omega$. For each extra edge $\sigma$ added to the network $q$, $B(I, st(\sigma), ter(\sigma), q^{n-1}, n)$ is true; it is false, otherwise. Lemma 5 shows that $\bar{Q}(E_Q) > 1 - \epsilon$, where $Q$ is the $q$-flow and $E_Q$ is its support set.

*Construction* Let $\rho(n) = (n + n_0)^2$ for some sufficiently large $n_0$ (the value $n_0$ will be specified below in the proof of Lemma 5).

Using the mathematical induction by $n$, we define a sequence $q^n$ of elementary networks. Put $q^0(\sigma) = 1/2$ for all edges $\sigma$ of length one.

Assume $n > 0$ and a network $q^{n-1}$ is defined. Let $d^{n-1}$ be the $q^{n-1}$-delay function and let $G^{n-1}$ be the set of all extra edges. We suppose also that $l(ter(\sigma)) < n$ for all $\sigma \in G^{n-1}$.

Let us define a network $q^n$. At first, we define a network flow delay function $d^n$ and a set $G^n$.

Let $w(I, q^{n-1})$ be equal to the minimal $m$ such that $p(m) = I$ and $m > l(ter(\sigma))$ for each extra edge $\sigma \in G^{n-1}$ such that $p(l(st(\sigma))) < I$.

The inequality $w(I, q^m) \neq w(I, q^{m-1})$ can be induced by some task $J < I$ that adds an extra edge $\sigma = (x, y)$ such that $l(y) > w(i, q^{m-1})$ and $p(l(x)) = p(l(y)) = J$. Lemma 3 (below) will show that this can happen only at finitely many steps of the construction.

The construction can be split up into two cases.

*Case 1.* $w(p(n), q^{n-1}) = n$ (the goal of this part is to start a new task $I = p(n)$ or to restart the existing task $I = p(n)$ if it was destroyed by some task $J < I$ at some preceding step).

Put $d^n(y) = 1/\rho(n)$ for $l(y) = n$ and define $d^n(y) = d^{n-1}(y)$ for all other $y$. Put also $G^n = G^{n-1}$.

*Case 2.* $w(p(n), q^{n-1}) < n$ (the goal of this part is to process the task $I = p(n)$).

Let $C_n$ be the set of all $x$ such that $w(I, q^{n-1}) \leq l(x) < n$, $0 < d^{n-1}(x) < 1$, the function $\beta(x, q^{n-1}, n)$ is defined[4] and there is no extra edge $\sigma \in G^{n-1}$ such that $st(\sigma) = x$.

In this case for each $x \in C_n$ define $d^n(\beta(x, q^{n-1}, n)) = 0$, and for all other $y$ of length $n$ such that $x \subseteq y$ define

$$d^n(y) = d^{n-1}(x)/(1 - d^{n-1}(x)).$$

Define $d^n(y) = d^{n-1}(y)$ for all other $y$. We add extra edges to $G^{n-1}$, namely, define

$$G^n = G^{n-1} \cup \{(x, \beta(x, q^{n-1}, n)) : x \in C_n\}.$$

[4] In particular, $p(l(x)) = I$ and $l(\beta(x, q^{n-1}, n)) = n$.

We say that the task $I = p(n)$ *adds* the extra edge $(x, \beta(x, q^{n-1}, n))$ to the network and that all existing tasks $J > I$ are destroyed by the task $I$.

After Cases 1 and 2, define for each edge $\sigma$ of unit length

$$q^n(\sigma) = \frac{1}{2}(1 - d^n(st(\sigma)))$$

and $q^n(\sigma) = d^n(st(\sigma))$ for each extra edge $\sigma \in G^n$.

Using this construction, we define the network $q = \lim_{n\to\infty} q^n$, the network flow delay function $d = \lim_{n\to\infty} d^n$, and the set of extra edges $G = \bigcup_n G^n$.

The functions $q$ and $d$ are computable and the set $G$ is recursive by their definitions. Let $Q$ denote the $q$-flow.

The following lemma shows that any task can add new extra edges only at finite number of steps.

Let $G(I)$ be the set of all extra edges added by the task $I$, $w(I, q) = \lim_{n\to\infty} w(I, q^n)$.

**Lemma 3** *The set $G(I)$ is finite and $w(I, q) < \infty$ for all $I$.*

*Proof* Note that if $G(J)$ is finite for all $J < I$ then $w(I, q) < \infty$. Then we must prove that the set $G(I)$ is finite for all $I$. Suppose that the opposite assertion holds. Let $I$ be the minimal number such that $G(I)$ is infinite. By choice of $I$ the sets $G(J)$ for all $J < I$ are finite. Then $w(I, q) < \infty$.

By definition if $d(\omega^m) \neq 0$ then $p_m = 1/d(\omega^m)$ is a positive integer number. Besides, if $(\omega^n, y), (\omega^m, y') \in G(I)$, where $n < m$ and $l(y) = m$, then $p_n > p_m$. Hence, for each $\omega \in \Omega$ a maximal $m$ exists such that $(\omega^m, y) \in G(I)$ for some $y$ or no such extra edge exists. In the latter case put $m = w(I, q)$. Define $u(\omega) = 1/d(\omega^m)$.

By the construction the integer valued function $u(\omega)$ is constant on the interval $\Gamma_{\omega^m}$. Hence, it is continuous in the topology generated by such intervals. Since $\Omega$ is compact in this topology, $u(\omega)$ is bounded. Then for some $m'$, $u(\omega) = u(\omega^{m'})$ for all $\omega$. By the construction if any extra edge of $I$th type was added to $G(I)$ at some step then $d(y) > d(x)$ holds for some new pair $(x, y)$ such that $x \subseteq y$. This is a contradiction if $G(I)$ is infinite. □

An infinite sequence $\alpha \in \Omega$ is called an *$I$-extension* of a finite sequence $x$ if $x \subseteq \alpha$ and $B(I, x, \alpha^n, n)$ is true for almost all $n$.

A sequence $\alpha \in \Omega$ is called *$I$-closed* if $d(\alpha^n) = 1$ for some $n$ such that $p(n) = I$, where $d$ is the $q$-delay function. Note that if $\sigma \in G(I)$ is some extra edge then $B(I, st(\sigma), ter(\sigma), n)$ is true, where $n = l(ter(\sigma))$.

**Lemma 4** *Assume for each initial fragment $\omega^n$ of an infinite sequence $\omega$ some $I$-extension exists. Then either the sequence $\omega$ will be $I$-closed in the process of the construction or $\omega$ contains an extra edge of $I$th type (i.e. such that $ter(\sigma) \subseteq \omega$ for some $\sigma \in G(I)$).*

*Proof* Assume a sequence $\omega$ is not $I$-closed. By Lemma 3 the maximal $m$ exists such that $p(m) = I$ and $d(\omega^m) > 0$. Since the sequence $\omega^m$ has an $I$-extension and

$d(\omega^k) < 1$ for all $k$, by Case 2 of the construction a new extra edge $(\omega^m, y)$ of $I$th type must be added to the binary tree. By the construction $d(y) = 0$ and $d(z) \neq 0$ for all $z$ such that $\omega^m \subseteq z$, $l(z) = l(y)$, and $z \neq y$. By the choice of $m$ we have $y \subseteq \omega$. □

Obviously, $Q(y) = 0$ if and only if $q(\sigma) = 0$ for some edge $\sigma$ of unit length located on $y$ (this edge satisfies $ter(\sigma) \subseteq y$ and $d(st(\sigma)) = 1$). Then the relation $Q(y) = 0$ is recursive and $E_Q = \Omega \setminus \bigcup_{d(x)=1} \Gamma_x$.

**Lemma 5** *It holds* $\bar{Q}(E_Q) > 1 - \epsilon$.

*Proof* We bound $\bar{Q}(\Omega)$ from below. For any $n$, let $q^n$ be the network defined at step $n$, $R^n$ be defined by (16), and $d^n$ be the corresponding $q^n$-delay function. If $w(p(n), q^{n-1}) = n$ (i.e., Case 1 holds at step $n$) then

$$\sum_{l(u)=n} d^n(u)R^n(u) = (n + n_0)^{-2} \sum R^n(u) \leq (n + n_0)^{-2}. \tag{17}$$

Assume Case 2 holds at the step $n$ and $x \in C_n$ such that $(x, y) \in G$ for some $y$, $l(y) = n$. Since by the construction $d^n(y) = 0$,

$$\sum_{l(z)=n, x\subseteq z} d^n(z)R^n(z) \leq \frac{d^{n-1}(x)}{(1 - d^{n-1}(x))} \sum_{l(z)=n, x\subseteq z, z\neq y} R^n(z). \tag{18}$$

We have

$$\sum_{l(z)=n, x\subseteq z, z\neq y} R^{n-1}(z) \leq (1 - d^{n-1}(x))R^{n-1}(x). \tag{19}$$

By the construction $R^n(z) = R^{n-1}(z)$ for $z$ such that $l(z) = n$, $x \subseteq z$, $z \neq y$. Then

$$\sum_{l(z)=n, x\subseteq z} d^n(z)R^n(z) \leq d^{n-1}(x)R^{n-1}(x). \tag{20}$$

By definition $\sum (n + n_0)^{-2} \leq \epsilon$. After that, using (17) and (20) we can prove by the mathematical induction on $n$ that

$$\bar{Q}(\Omega) = \inf_n \sum_{l(u)=n} Q(u) \geq \inf_n \sum_{l(u)=n} R(u) \geq 1 - \epsilon.$$

Lemma is proved. □

**Lemma 6** *For any infinite sequence $\omega \in E_Q$ and for any computable operation $F$ if the sequence $F(\omega)$ is infinite then it is not Martin-Löf random with respect to the uniform probability distribution.*

*Proof* Assume that $\omega$ is an infinite sequence and $F$ is a computable operation such that $F(\omega)$ is infinite. Then $F_i = F$ for some $i$. Define

$$U_s = \bigcup \{\Gamma_{\beta(x, q^{n-1}, n)} : x \in C_n,\ p(n) = \langle i, s\rangle\},$$

where $C_n$ is the set from Case 2 of the construction. By definition

$$L(U_s)=\sum_{x\in C_n}2^{-\langle x,s\rangle}\leq 2^{-cs}$$

for some positive constant $c$, and $F_i(\omega)\in\bigcap_s U_s$. Therefore, the sequence $F(\omega)$ is not Martin-Löf random. Lemma 6 and Theorem 3 are proved. □

## References


1. Borovkov, A.A.: Theory of Probability. Nauka, Moscow (1999). Probability Theory. Gordon and Breach (1998)
2. Cox, D.R., Hinkley, D.V.: Theoretical Statistics. Chapman & Hall, London (1974)
3. Levin, L.A.: On the notion of random sequence. Soviet Math. Dokl. **14**, 1413–1416 (1973)
4. Levin, L.A.: Laws of information conservation (non-growth) and aspects of the foundation of probability theory. Probl. Inf. Transm. **10**, 206–210 (1974)
5. Levin, L.A.: Randomness conservation inequalities; information and independence in mathematical theories. Inf. Control **61**, 15–37 (1984)
6. Levin, L.A., V'jugin, V.V.: Invariant Properties of Informational Bulks. Lect. Notes Comput. Sci. **53**, 359–364 (1977)
7. Li, M., Vitányi, P.: An Introduction to Kolmogorov Complexity and Its Applications, 2nd edn. Springer, New York (1997)
8. Martin-Löf, P.: The definition of random sequences. Inf. Control **9**, 602–619 (1966)
9. Rogers, H.: Theory of Recursive Functions and Effective Computability. McGraw Hill, New York (1967)
10. Shen, A., Uspensky, V.A., Vereshchagin, N.K.: Lecture Notes on Kolmogorov Complexity. http://lpes.math.msu.su/ver/kolm-book (2007)
11. Solomonoff, R.J.: A formal theory of inductive inference. Inf. Control I, II **7**, 1–22 (1964), 224–254
12. Vovk, V.G., V'yugin, V.V.: On the empirical validity of the Bayesian rule. J. R. Stat. Soc. B **55**, 317–351 (1993)
13. V'yugin, V.V.: On Turing invariant sets. Soviet Math. Dokl. **17**, 1090–1094 (1976)
14. V'yugin, V.V.: The algebra of invariant properties of binary sequences. Probl. Inf. Transm. **18**, 147–161 (1982)
15. V'yugin, V.V.: On empirical meaning of randomness with respect to a real parameter. In: Diekert, V., Volkov, M., Voronkov, A. (eds.) CSR 2007. LNCS, vol. 4649, pp. 387–396. Springer, Berlin (2007)
16. Zvonkin, A.K., Levin, L.A.: The complexity of finite objects and the algorithmic concepts of information and randomness. Russ. Math. Surv. **25**, 83–124 (1973)