How much randomness is needed for statistics?

Bjørn Kjos-Hanssen[a], Antoine Taveneaux[b], Neil Thapen[c]

[a] *University of Hawai'i at Mānoa, Honolulu, HI 96822, USA,*
[b] *LIAFA, Université Paris Diderot - Paris 7, 75205 Paris Cedex 13, France,*
[c] *Academy of Sciences of the Czech Republic, 115 67 Praha 1, Czech Republic,*

Abstract

In algorithmic randomness, when one wants to define a randomness notion with respect to some non-computable measure λ, a choice needs to be made. One approach is to allow randomness tests to access the measure λ as an oracle (which we call the "classical approach"). The other approach is the opposite one, where the randomness tests are completely effective and do not have access to the information contained in λ (we call this approach "Hippocratic"). While the Hippocratic approach is in general much more restrictive, there are cases where the two coincide. The first author showed in 2010 that in the particular case where the notion of randomness considered is Martin-Löf randomness and the measure λ is a Bernoulli measure, classical randomness and Hippocratic randomness coincide. In this paper, we prove that this result no longer holds for other notions of randomness, namely computable randomness and stochasticity.

Keywords: Hippocratic randomness, martingales, Bernoulli measures

1. Introduction

In algorithmic randomness theory we are interested in which almost sure properties of an infinite sequence of bits are effective or computable in some sense. Martin-Löf defined randomness with respect to the uniform fair-coin measure μ on 2^ω as follows.

> A sequence $X \in 2^\omega$ is *Martin-Löf random* if we have $X \notin \bigcap_{n\in\mathbb{N}} \mathcal{U}_n$ for every sequence of uniformly Σ^0_1 (or effectively open) subsets of 2^ω such that $\mu(\mathcal{U}_n) \leq 2^{-n}$.

Now if we wish to consider Martin-Löf randomness for a Bernoulli measure μ_p (that is, a measure such that the i^{th} bit is the result of a Bernoulli trial with parameter p), we have have two possible ways to extend the previous definition.

Email addresses: bjoernkh@hawaii.edu (Bjørn Kjos-Hanssen), taveneaux@calculabilite.fr (Antoine Taveneaux), thapen@math.cas.cz (Neil Thapen)

The first option is to consider p as an oracle (with an oracle p we can compute μ_p) and relativize everything to this oracle. Then X is μ_p-Martin-Löf random if for every sequence $(\mathcal{U}_n)_{n\in\mathbb{N}}$ of uniformly $\Sigma^0_1[p]$ sets such that $\mu_p(\mathcal{U}_n) \leq 2^{-n}$ we have $X \notin \bigcap_{n\in\mathbb{N}} \mathcal{U}_n$. We will call this approach the *classical* notion of Martin-Löf randomness relative to μ_p.

Another option is to keep the measure μ_p "hidden" from the process which describes the sequence $(\mathcal{U}_n)$. One can merely replace μ by μ_p in Martin-Löf's definition but still require $(\mathcal{U}_n)$ to be uniformly Σ^0_1 in the unrelativized sense. This notion of randomness was introduced by Kjos-Hanssen [K10] who called it *Hippocratic randomness*; Bienvenu, Doty and Stephan [BDS09] used the term *blind randomness*.

Kjos-Hanssen showed that for Bernoulli measures, Hippocratic and classical randomness coincide in the case of Martin-Löf randomness. Bienvenu, Gács, Hoyrup, Rojas and Shen [BGHRS11] extended Kjos-Hanssen's result to other classes of measures. Here we go in a different direction and consider weaker randomness notions, such as computable randomness and stochasticity. We discover the contours of a dividing line for the type of betting strategy that is needed in order to render the probability distribution superfluous as a computational resource.

We view *statistics* as the discipline concerned with determining the underlying probability distribution μ_p by looking at the bits of a random sequence. In the case of Martin-Löf randomness it is possible to determine p ([K10]), and therefore Hippocratic randomness and classical randomness coincide. In this sense, Martin-Löf randomness is sufficient for statistics to be possible, and it is natural to ask whether smaller amounts of randomness, such as computable randomness, are also sufficient.

An earlier version of this paper appeared in the proceedings of the Computability in Europe 2012 conference [KTT12].

Notation.. Our notation generally follows Nies' monograph [Nie09]. We write 2^n for $\{0,1\}^n$, and for sequences $\sigma \in 2^{\leq\omega}$ we will also use σ to denote the real with binary expansion $0.\sigma$, that is, the real $\sum_{i=1}^{\infty} \sigma(i)2^{-i}$. We use ε to denote the empty word, $\sigma(n)$ for the n^{th} element of a sequence and $\sigma\upharpoonright n$ for the sequence formed by the first n elements. For sequences ρ, σ we write $\sigma \prec \rho$ if σ is a proper prefix of ρ and denote the concatenation of σ and ρ by $\sigma.\rho$ or simply $\sigma\rho$. Throughout the paper we set $n' = n(n-1)/2$.

1.1. Hippocratic martingales

Formally a martingale is a function $\mathcal{M} : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ satisfying

$$\mathcal{M}(\sigma) = \frac{\mathcal{M}(\sigma 0) + \mathcal{M}(\sigma 1)}{2}.$$

Intuitively, such a function arises from a betting strategy for a fair game played with an unbiased coin (a sequence of Bernoulli trials with parameter $1/2$). In each round of the game we can choose our stake, that is, how much of our capital

we will bet, and whether we bet on heads (1) or tails (0). A coin is tossed, and if we bet correctly we win back twice our stake.

Suppose that our betting strategy is given by some fixed function S of the history σ of the game up to that point. Then it is easy to see that the function $\mathcal{M}(\sigma)$ giving our capital after a play σ satisfies the above equation. On the other hand, from any $\mathcal{M}$ satisfying the equation we can recover a corresponding strategy S.

More generally, consider a biased coin which comes up heads with probability $p \in (0,1)$. In a fair game played with this coin, we would expect to win back $1/p$ times our stake if we bet correctly on heads, and $1/(1-p)$ times our stake if we bet correctly on tails. Hence we define a p-martingale to be a function satisfying

$$\mathcal{M}(\sigma) = p\mathcal{M}(\sigma 1) + (1-p)\mathcal{M}(\sigma 0).$$

We can generalize this further, and for any probability measure μ on 2^ω define a μ-martingale to be a function satisfying

$$\mu(\sigma)\mathcal{M}(\sigma) = \mu(\sigma 1)\mathcal{M}(\sigma 1) + \mu(\sigma 0)\mathcal{M}(\sigma 0).$$

For the Bernoulli measure with parameter p, we say that a sequence $X \in 2^\omega$ is p-computably random if for every total, p-computable p-martingale $\mathcal{M}$, the sequence $(\mathcal{M}(X\upharpoonright n))_n$ is bounded.

This is the classical approach to p-computable randomness. Under the Hippocratic approach, the bits of the parameter p should not be available as a computational resource. The obvious change to the definition would be to restrict to p-martingales $\mathcal{M}$ that are computable without an oracle for p. However this does not give a useful definition, as p can easily be recovered from any non-trivial p-martingale. Instead we will define p-Hippocratic computable martingales in terms of their stake function (or strategy) S.

We formalize S as a function $2^{<\omega} \to [-1,1] \cap \mathbb{Q}$. The absolute value $|S(\sigma)|$ gives the fraction of our capital we put up as our stake, and we bet on 1 if $S(\sigma) \geq 0$ and on 0 if $S(\sigma) < 0$. Given $\alpha \in (0,1)$, the α-martingale $\mathcal{M}^\alpha$ arising from S is then defined inductively by

$$\begin{aligned}
\mathcal{M}^\alpha(\varepsilon) &= 1 \\
\mathcal{M}^\alpha(\sigma 1) &= \mathcal{M}^\alpha(\sigma)\Big(1 - |S(\sigma)| + \frac{|S(\sigma)|}{\alpha} 1_{\{S(\sigma)\geq 0\}}\Big) \\
\mathcal{M}^\alpha(\sigma 0) &= \mathcal{M}^\alpha(\sigma)\Big(1 - |S(\sigma)| + \frac{|S(\sigma)|}{1-\alpha} 1_{\{S(\sigma)< 0\}}\Big)
\end{aligned}$$

where, for a formula T, we use the notation $1_{\{T\}}$ to mean the function which takes the value 1 if T is true and 0 if T is false.

We define a p-Hippocratic computable martingale to be a p-martingale $\mathcal{M}^p$ arising from some total computable (without access to p) stake function S. We say that a sequence $X \in 2^\omega$ is p-Hippocratic computably random if for every p-Hippocratic computable martingale $\mathcal{M}$, the sequence $(\mathcal{M}(X\upharpoonright n))_n$ is bounded.

In Section 2 below we show that for all $p \in \mathrm{MLR}$ the set of p-Hippocratic computably random sequences is strictly bigger than the set of p-computable

random sequences. More precisely, we show that we can compute a sequence $Q \in 2^\omega$ from p such that Q is p-Hippocratic computably random. In a nutshell, the proof works as follows. We use the number p in two ways. To compute the i^{th} bit of Q, the first i bits of p are treated as a parameter $r = 0.p_0 \dots p_i$, and we pick the i^{th} bit of Q to look like it has been chosen at random in a Bernoulli trial with bias r. To do this, we use some fresh bits of p (which have not been used so far in the construction of Q) and compare them to r, to simulate the trial. Since these bits of p were never used before, if we know only the first $i-1$ bits of Q they appear random, and thus the i^{th} bit of Q indeed appears to be chosen at random with bias r. Since $r = 0.p_1p_2 \dots p_i$ converges quickly to p (this convergence is faster than the deviations created by statistical noise in a real sequence of Bernoulli trials with parameter p), we are able to show that Q overall looks p-random as long as we do not have access to p, in other words, that Q is p-Hippocratic computably random.

1.2. Hippocratic stochasticity and KL randomness

In Section 3 we consider another approach to algorithmic randomness, known as stochasticity. It is reasonable to require that a random sequence satisfies the law of large numbers, that is, that the proportion of 1s in the sequence converges to the bias p. But, for an unbiased coin, the string

$$010101010\dots$$

satisfies this law but is clearly not random. Following this idea, we say that a sequence X is p-Kolmogorov-Loveland (or p-KL) stochastic if there is no p-computable way to select infinitely many bits from X, where we are not allowed to know the value of a bit before we select it, without the selected sequence satisfying the law of large numbers (see Definition 1 for a formal approach).

For this paradigm the Hippocratic approach is clear: we consider only selection functions which are computable without an oracle for p. We show in Theorem 7 that for $p \in \Delta^0_2 \cap \mathrm{MLR}$ there exists a sequence Q which is p-Hippocratic KL stochastic but not p-KL stochastic. Again we use p as a random bit generator and create a sequence Q that appears random for a sequence of Bernoulli trials, where the bias of the i^{th} trial is q_i for a certain sequence $(q_i)_i$ converging to p. Intuitively, the convergence is so slow that it is impossible to do (computable) statistics with Q to recover p, and we are able to show that without access to p the sequence Q is p-KL stochastic.

At the end of Section 3 we consider another notion, Kolmogorov-Loveland randomness. We give a simple argument to show that if we can compute p from every p-Hippocratic KL random sequence, then the p-Hippocratic KL random sequences and the p-KL random sequences are the same (and vice versa).

2. Computable randomness

In this section we show that for any Martin-Löf random bias p, p-computable randomness is a stronger notion than p-Hippocratic computable randomness.

Theorem 1. *Let $\alpha \in MLR$. There exists a sequence $Q \in 2^\omega$, computable in polynomial time from α, such that Q is α-Hippocratic computably random.*

Before giving the proof, we remark that a stronger version of the theorem is true: the sequence Q is in fact α-Hippocratic partial computably random (meaning that we allow the martingale to be a partial computable function, see Definition 7.4.3 in [DH10]).

Also, a sceptic could (reasonably) complain that it is not really natural for us to make bets without any idea about our current capital. However if we add an oracle to give the integer part of our capital at each step (or even an approximation with accuracy 2^{-n} when we bet on the n^{th} bit), Theorem 1 remains true and the proof is the same. In the same spirit we could object that it is more natural to have a stake function giving the amount of our bet (to be placed only if we have a capital large enough) and not the proportion of our capital. For this definition of a Hippocratic computable martingale, similarly the theorem remains true and the proof is the same.

Proof. Let $\alpha \in$ MLR. Then α is not rational and cannot be represented by a finite binary sequence and we can suppose that $0 < \alpha < 1/2$. Recall that $n' = n(n-1)/2$ and that we freely identify a sequence X (finite or infinite) with the real number with the binary expansion $0.X$.

The proof has the following structure. First, we describe an algorithm to compute a sequence Q from α. To compute each bit Q_n of Q we will use a finite initial segment of α as an approximation of α, and we will compare this with some other fresh bits of α which we treat as though they are produced by a random bit generator. In this way Q_n will approximate the outcome of a Bernoulli trial with bias α.

Second, we will suppose for a contradiction that there is an α-Hippocratic computable martingale (that is, a martingale that arises from a stake function computable without α) such that the capital of this martingale is not bounded on Q. We will show that we can use this stake function to construct a Martin Löf test $(U_n)_n$ such that α does not pass this test.

So let $Q = Q_1 Q_2 \dots$ be defined by the condition that:

$$Q_n = \begin{cases} 0 & \text{if } 0.\alpha_{n'+1} \dots \alpha_{n'+n} \geq 0.\alpha_1 \dots \alpha_n, \\ 1 & \text{otherwise.} \end{cases}$$

We can compute Q in polynomial time from α, as we can compute each bit Q_n in time $O(n^2)$.

Now let $S : 2^{<\omega} \to \mathbb{Q} \cap [-1, 1]$ be a computable stake function. We will write $\mathcal{M}^X$ for the X-martingale arising from S. Suppose for a contradiction that

$$\limsup_{n \to \infty} \mathcal{M}^\alpha(Q \upharpoonright n) = \infty.$$

Our goal is to use this to define a Martin-Löf test which α fails. The classical argument (see Theorem 6.3.4 in [DH10]) would be to consider the sequence of sets

$$V_j = \{X \in 2^\omega | \exists n \;\; \mathcal{M}^\alpha(X \upharpoonright n) > 2^j\},$$

but without oracle access to α this is not Σ^0_1, and does not define a Martin-Löf test. However it turns out that we can use a similar sequence of sets, based on the idea that, although we cannot compute $\mathcal{M}^\alpha$ precisely, we can approximate it using the approximation $\alpha_1 \dots \alpha_{n'}$ of α. For this we will use the following lemma. The proof is rather technical and we postpone it to later in this section.

Lemma 2. *For $\alpha \in$ MLR, there exists $m \in \mathbb{N}$ such that if $\sigma \succcurlyeq (\alpha \upharpoonright m')$ and $\tau \succcurlyeq (\alpha \upharpoonright m')$ then for all $\eta \in 2^{<\omega}$ and all $n \geq m$ we have:*

$$\text{if } \ 0 < \tau - \sigma < 2^{-n'} \ \text{ and } \ |\eta| \leq n+1 \ \text{ then } \ |\mathcal{M}^\sigma(\eta) - \mathcal{M}^\tau(\eta)| \leq 2^{-n}.$$

Let m be given by Lemma 2 and let ρ be $\alpha \upharpoonright m'$. Without loss of generality we may assume $2^{-m} < \rho$. Let $\Gamma : 2^{\leq\omega} \to 2^{\leq\omega}$ be the operator which converts $\alpha_1 \dots \alpha_{n'}$ into $Q_1 \dots Q_n$. That is, $\Gamma(\alpha_1 \dots \alpha_k) = Q_1 \dots Q_n$ where n is the biggest integer such that $n' \leq k$. This notation naturally extends to infinite sequences so we may write $\Gamma(\alpha) = Q$. We consider the uniform sequence of Σ^0_1 sets

$$U'_j = \{X_1 \dots X_{k'} | \rho \preccurlyeq X_1 \dots X_{k'} \text{ and } \mathcal{M}^{X_1 \dots X_{k'}}(\Gamma(X_1 \dots X_{k'})) > 2^j\}.$$

We let U_j denote the set of infinite sequences with a prefix in U'_j. By Lemma 2,

$$|\mathcal{M}^\alpha(\Gamma(\alpha_1 \dots \alpha_{k'})) - \mathcal{M}^{\alpha_1 \dots \alpha_{k'}}(\Gamma(\alpha_1 \dots \alpha_{k'}))| < 2^{-k} \leq 1$$

for all sufficiently large k. Since $\mathcal{M}^\alpha$ increases unboundedly on $Q = \Gamma(\alpha)$ it follows that $\alpha \in U_j$ for all j.

To show that (U_j) is a Martin-Löf test, it remains to show that the measure of U_j is small. Since $\sigma \mapsto \mathcal{M}^\sigma(\sigma)$ is almost a α-martingale, where σ runs over the prefixes of α, we will use a lemma similar to the Kolmogorov inequality (see Theorem 6.3.3 in [DH10]). Again we postpone the proof to later in this section.

Lemma 3. *For any number $n \geq m$, any extension $\sigma \succcurlyeq \rho$ of length n' and any prefix-free set $Z \subseteq \bigcup_{k \in \mathbb{N}} \{0,1\}^{k'}$ of extensions of σ, we have*

$$\sum_{\tau \in Z} 2^{-|\tau|} \mathcal{M}^\tau(\Gamma(\tau)) \leq 2^{-|\sigma|} e^2 \left[1 + \mathcal{M}^\sigma(\Gamma(\sigma))\right].$$

Now fix j and let W_j be a prefix-free subset of U'_j with the property that the set of infinite sequences with a prefix in W_j is exactly U_j. Then by the definition of U'_j, if $\tau \in W_j$ then $\mathcal{M}^\tau(\Gamma(\tau)) \geq 2^j$. Hence by Lemma 3 we have:

$$\mu(U_j) = \sum_{\tau \in W_j} 2^{-|\tau|} \leq \sum_{\tau \in W_j} \frac{\mathcal{M}^\tau(\Gamma(\tau))}{2^j} 2^{-|\tau|} \leq \frac{2^{-|\rho|} e^2 \left(1 + \mathcal{M}^\rho(\Gamma(\rho))\right)}{2^j}.$$

Since $2^{-|\rho|}\left(1 + \mathcal{M}^\rho(\Gamma(\rho))\right)$ is constant, this shows that (U_j) is a Martin-Löf test. As $\alpha \in \bigcap_j U_j$ it follows that $\alpha \notin$ MLR. This is a contradiction. $\square$

Notice that this proof makes use of the fact that in our betting strategy we have to proceed monotonically from left to right through the string, making a decision for each bit in turn as we come to it. This is why our construction is able to use α as a random bit generator, because at each step it can use bits that were not used to compute the previous bits of Q. Following this idea the question naturally arises: if we are allowed to use a non-monotone strategy, then are the classical and Hippocratic random sequences the same? We explore this question in Section 3.

We now return to the postponed proofs. We will need a couple of technical lemmas, the first one giving, roughly speaking, a modulus of continuity for the map $(\alpha, X) \mapsto \mathcal{M}^\alpha(X)$.

Lemma 4. *Let $\epsilon > 0$. Then there exists $r \in \mathbb{N}$ such that for all sufficiently large k, for all $\alpha, \beta \in 2^{\leq\omega}$ with $\epsilon < \alpha < \beta < 1 - \epsilon$ and for all non-empty $\sigma \in 2^{<\omega}$,*

$$0 < \beta - \alpha < 2^{-k} \Rightarrow \ |\mathcal{M}^\alpha(\sigma) - \mathcal{M}^\beta(\sigma)| < 2^{-k+r|\sigma|}.$$

Proof. Since $0 < \epsilon < \alpha < \beta < \alpha + 2^{-k}$ we have

$$\frac{1}{\alpha + 2^{-k}} < \frac{1}{\beta} < \frac{1}{\alpha} < \frac{1}{\epsilon}$$

and hence

$$0 < \frac{1}{\alpha} - \frac{1}{\beta} < \frac{1}{\alpha} - \frac{1}{\alpha + 2^{-k}} = \frac{2^{-k}}{\alpha(\alpha + 2^{-k})} < \frac{2^{-k}}{\alpha^2} < 2^{-k}\epsilon^{-2}.$$

It follows, since $|S(X)|$ is less than or equal to 1, that

$$0 \leq \Big(1 - |S(\sigma)| + \frac{|S(\sigma)|}{\alpha} 1_{\{S(\sigma)\geq 0\}}\Big) - \Big(1 - |S(\sigma)| + \frac{|S(\sigma)|}{\beta} 1_{\{S(\sigma)\geq 0\}}\Big) \leq 2^{-k}\epsilon^{-2}$$

and symmetrically, since $0 < \epsilon < 1 - \beta < 1 - \alpha < (1 - \beta) + 2^{-k}$, also that

$$0 \leq \Big(1 - |S(\sigma)| + \frac{|S(\sigma)|}{1 - \beta} 1_{\{S(\sigma)<0\}}\Big) - \Big(1 - |S(\sigma)| + \frac{|S(\sigma)|}{1 - \alpha} 1_{\{S(\sigma)<0\}}\Big) \leq 2^{-k}\epsilon^{-2}.$$

Hence if we write R_i^X for $\frac{\mathcal{M}^X(\sigma|i)}{\mathcal{M}^X(\sigma|(i-1))}$ (with the convention $0/0 = 0$) we have for all $i \leq |\sigma|$ that $|R_i^\alpha - R_i^\beta| < 2^{-k}\epsilon^{-2}$. Furthermore, take s to be a positive integer such that $2^s > 1 + 1/\epsilon$. Then we know that R_i^α and R_i^β are both always smaller than 2^s.

We can now bound $|\mathcal{M}^\alpha(\sigma) - \mathcal{M}^\beta(\sigma)|$. Consider the case when

$\mathcal{M}^\alpha(\sigma) \geq \mathcal{M}^\beta(\sigma)$ (the other case is symmetrical). Then, writing n for $|\sigma|$,

$$\begin{aligned}
\mathcal{M}^\alpha(\sigma) - \mathcal{M}^\beta(\sigma) &= \prod_{i=1}^{n} R_i^\alpha - \prod_{i=1}^{n} R_i^\beta \\
&\leq \prod_{i=1}^{n} (R_i^\beta + 2^{-k}\epsilon^{-2}) - \prod_{i=1}^{n} R_i^\beta \\
&= \left[\prod_{i=1}^{n} R_i^\beta + \sum_{\substack{Z \subseteq [1,n] \\ |Z|<n}} (2^{-k}\epsilon^{-2})^{n-|Z|} \prod_{i \in Z} R_i^\beta \right] - \prod_{i=1}^{n} R_i^\beta \\
&\leq 2^n (2^{-k}\epsilon^{-2})(2^s)^n,
\end{aligned}$$

where for the last inequality we are assuming that k is large enough that $2^{-k}\epsilon^{-2} < 1$. The result follows. □

Lemma 5. *For $s \in \mathbb{R}$, $s > 0$ we have $\prod_{n=1}^{\infty}(1 + s2^{-n}) < e^s$.*

Proof. It is enough to show that

$$\sum_{n=1}^{\infty} \ln\left(\frac{2^n + s}{2^n}\right) = \sum_{n=1}^{\infty} [\ln(2^n + s) - \ln(2^n)] < s.$$

The derivative of ln is the decreasing function $x \mapsto 1/x$ so by the mean value theorem we have that $\ln(2^n + s) - \ln(2^n) < s/2^n$, which gives the inequality. □

We are now able to prove the lemmas used in the proof of Theorem 1.

Restatement of Lemma 2. *For $\alpha \in MLR$, there exists $m \in \mathbb{N}$ such that if $\sigma \succcurlyeq (\alpha \restriction m')$ and $\tau \succcurlyeq (\alpha \restriction m')$ then for all $\eta \in 2^{<\omega}$ and all $n \geq m$ we have:*

$$\text{if } \ 0 < \tau - \sigma < 2^{-n'} \ \text{ and } \ |\eta| \leq n+1 \ \text{ then } \ |\mathcal{M}^\sigma(\eta) - \mathcal{M}^\tau(\eta)| \leq 2^{-n}.$$

Proof. Since $\alpha \in$ MLR we can find $m_0 \in \mathbb{N}$ and $\epsilon > 0$ such that $\epsilon < \alpha \restriction m_0' < (\alpha \restriction m_0') + 2^{-m_0'} < 1 - \epsilon$. Let τ, σ, η and n satisfy the assumptions of the lemma, using m_0 in place of m. We must have $\epsilon < \sigma < \tau < 1 - \epsilon$, hence by Lemma 4 there exists r depending only on ϵ such that

$$|\mathcal{M}^\sigma(\eta) - \mathcal{M}^\tau(\eta)| \leq 2^{-n' + r|\phi|} \leq 2^{-n' + r(n+1)}.$$

It is clear that we can find $m_1 \in \mathbb{N}$ such that, for $n \geq m_1$,

$$r(n+1) - n' = r(n+1) - \frac{n(n-1)}{2} < -n.$$

For the lemma take $m = \max(m_0, m_1)$. □

Now we suppose $\alpha \in \mathrm{MLR}$, $\alpha < 1/2$ and let m be as given by Lemma 2. We write ρ for $\alpha \upharpoonright m'$. Without loss of generality we may assume $2^{-m} < \rho$.

Restatement of Lemma 3. *For any number $n \geq m$, any extension $\sigma \succcurlyeq \rho$ of length n' and any prefix-free set $Z \subseteq \bigcup_{k\in\mathbb{N}} \{0,1\}^{k'}$ of extensions of σ, we have*

$$\sum_{\tau\in Z} 2^{-|\tau|}\mathcal{M}^\tau(\Gamma(\tau)) \leq 2^{-|\sigma|}e^2\left[1+\mathcal{M}^\sigma(\Gamma(\sigma))\right].$$

Proof. It is enough to show this for every finite Z. We will use induction on the size p of Z, with our inductive hypothesis that for all $n \geq m$, all extensions $\sigma \succcurlyeq \rho$ of length n' and all suitable sets Z of size p,

$$\sum_{\tau\in Z} 2^{-|\tau|}\mathcal{M}^\tau(\Gamma(\tau)) \leq 2^{-|\sigma|}\left[\sum_{i=n}^{\infty} 2e^2 2^{-i} + \mathcal{M}^\sigma(\Gamma(\sigma))\prod_{i=n}^{\infty}\left(1+2{\cdot}2^{-i}\right)\right].$$

Note that by Lemma 5 the right hand side is bounded by $2^{-|\sigma|}e^2[1+\mathcal{M}^\sigma(\Gamma(\sigma))]$ (as long as $n \geq 2$).

The base case $|Z| = 0$ is trivial. Now suppose that the hypothesis is true for all sets of size less than or equal to p and suppose that $|Z| = p+1$. Let ν be the longest extension of σ which has length of the form k' for some $k \in \mathbb{N}$ and which is such that all strings in Z are extensions of ν. Then for each string θ of length k there are fewer than $p+1$ strings in Z beginning with $\nu\theta$. Recall that $|\nu\theta| = k' + k = (k+1)'$ and that $\Gamma(\nu)$ and $\Gamma(\nu\theta)$ are strings of length respectively k and $k+1$. Applying the inductive hypothesis, we have

$$\begin{aligned}
\sum_{\tau\in Z} 2^{-|\tau|}\mathcal{M}^\tau(\Gamma(\tau)) &\leq \sum_{\theta\in\{0,1\}^k}\sum_{\substack{\tau\in Z\\ \tau\succcurlyeq\nu\theta}} 2^{-|\tau|}\mathcal{M}^\tau(\Gamma(\tau)) \\
&\leq \sum_{\theta\in\{0,1\}^k} 2^{-|\nu\theta|}\left[\sum_{i=k+1}^{\infty} 2e^2 2^{-i} + \mathcal{M}^{\nu\theta}(\Gamma(\nu\theta))\prod_{i=k+1}^{\infty}\left(1+2{\cdot}2^{-i}\right)\right] \\
&\leq \sum_{\theta\in\{0,1\}^k} 2^{-|\nu\theta|}\left[\sum_{i=k+1}^{\infty} 2e^2 2^{-i} + e^2 2^{-k} + \mathcal{M}^{\nu}(\Gamma(\nu\theta))\prod_{i=k+1}^{\infty}\left(1+2{\cdot}2^{-i}\right)\right]
\end{aligned}$$

where for the last inequality we are using that, by Lemma 2, $\mathcal{M}^{\nu\theta}(\Gamma(\nu\theta)) \leq \mathcal{M}^\nu(\Gamma(\nu\theta)) + 2^{-k}$. Rearranging the last line, we get

$$2^{-|\nu|}\left[\sum_{i=k+1}^{\infty} 2e^2 2^{-i} + e^2 2^{-k} + \left(\sum_{\theta\in\{0,1\}^k} 2^{-k}\mathcal{M}^\nu(\Gamma(\nu\theta))\right)\prod_{i=k+1}^{\infty}\left(1+2{\cdot}2^{-i}\right)\right].$$

Now we will find an upper bound for the term in round brackets.

We will write $\widehat{\nu}$ for $\nu \upharpoonright k$ and S for $S(\Gamma(\nu))$. By the definition of Γ, if $\theta \leq \widehat{\nu}$ (as real numbers) then $\Gamma(\nu\theta) = \Gamma(\nu).1$. Hence, by the definition of $\mathcal{M}$ and the

fact that $\nu \geq \widehat{\nu}$, summing only over $\theta \in \{0,1\}^k$ we have

$$\sum_{\theta \leq \widehat{\nu}} 2^{-k} \mathcal{M}^\nu(\Gamma(\nu\theta)) = \widehat{\nu}\, \mathcal{M}^\nu(\Gamma(\nu))\big(1 - |S| + \frac{1}{\nu}\,|S| \cdot 1_{\{S \geq 0\}}\big)$$
$$\leq \mathcal{M}^\nu(\Gamma(\nu))\big(\widehat{\nu}\,\big(1 - |S|\big) + |S| \cdot 1_{\{S \geq 0\}}\big).$$

Observing that $\nu \leq \widehat{\nu} + 2^{-k}$ and $1 - \nu \geq 1/2$ and that hence

$$\frac{1 - \widehat{\nu}}{1 - \nu} \leq \frac{1 - \nu + 2^{-k}}{1 - \nu} \leq 1 + 2 \cdot 2^{-k},$$

we similarly get that

$$\sum_{\theta > \widehat{\nu}} 2^{-k} \mathcal{M}^\nu(\Gamma(\nu\theta)) = (1 - \widehat{\nu})\mathcal{M}^\nu(\Gamma(\nu))\big(1 - |S| + \frac{1}{1 - \nu}\,|S| \cdot 1_{\{S < 0\}}\big)$$
$$\leq \mathcal{M}^\nu(\Gamma(\nu))\big((1 - \widehat{\nu})\big(1 - |S|\big) + |S| \cdot 1_{\{S<0\}} + 2 \cdot 2^{-k}\big).$$

Summing these gives

$$\sum_{\theta \in \{0,1\}^k} 2^{-k} \mathcal{M}^\nu(\Gamma(\nu\theta)) \leq \mathcal{M}^\nu(\Gamma(\nu))\big(1 - |S| + |S| \cdot 1_{\{S \geq 0\}} + |S| \cdot 1_{\{S<0\}} + 2 \cdot 2^{-k}\big)$$
$$= \mathcal{M}^\nu(\Gamma(\nu))(1 + 2 \cdot 2^{-k}).$$

Combining this with our earlier bound, we now have

$$\sum_{\tau \in Z} 2^{-|\tau|} \mathcal{M}^\tau(\Gamma(\tau)) \leq 2^{-|\nu|} \left[\sum_{i=k+1}^{\infty} 2e^2 2^{-i} + e^2 2^{-k} + \mathcal{M}^\nu(\Gamma(\nu)) \prod_{i=k}^{\infty} \big(1 + 2 \cdot 2^{-i}\big) \right].$$

Finally, let $r = k - n$ so that $|\nu| = k' = (n + r)' \geq n' + nr = |\sigma| + nr$. Recall that $1 - \nu > \nu \succcurlyeq \rho > 2^{-n}$, which means that a ν-martingale can multiply its capital by at most 2^n in one round. Hence, also using Lemma 2,

$$2^{-|\nu|} \mathcal{M}^\nu(\Gamma(\nu)) \leq 2^{-|\sigma| - nr} (2^n)^r \mathcal{M}^\nu(\Gamma(\sigma)) \leq 2^{-|\sigma|}\big(2^{-n} + \mathcal{M}^\sigma(\Gamma(\sigma))\big).$$

This gives us the bound

$$2^{-|\sigma|} \left[\sum_{i=k+1}^{\infty} 2e^2 2^{-i} + e^2 2^{-k} + e^2 2^{-n} + \mathcal{M}^\sigma(\Gamma(\sigma)) \prod_{i=k}^{\infty} \big(1 + 2 \cdot 2^{-i}\big) \right],$$

which is less than or equal to

$$2^{-|\sigma|} \left[\sum_{i=n}^{\infty} 2e^2 2^{-i} + \mathcal{M}^\sigma(\Gamma(\sigma)) \prod_{i=n}^{\infty} \big(1 + 2 \cdot 2^{-i}\big) \right].$$

This completes the induction. □

3. Kolmogorov-Loveland stochasticity and randomness

We define Kolmogorov-Loveland stochasticity and show that, in this setting, the Hippocratic and classical approaches give different sets. We also consider whether this is true for Kolmogorov-Loveland randomness, and relate this to a statistical question.

3.1. Definitions

For a finite string $\sigma \in \{0,1\}^n$, we write $\#0(\sigma)$ for $|\{k < n | \sigma(k) = 0\}|$ and $\#1(\sigma)$ for $|\{k < n | \sigma(k) = 1\}|$. We write $\Phi(\sigma)$ for $\#1(\sigma)/n$, the frequency of 1s in σ.

Definition 1 (Selection function). A KL selection function is a partial function

$$f : 2^{<\omega} \to \{\text{scan}, \text{select}\} \times \mathbb{N}.$$

We write $f(\sigma)$ as a pair $(s(\sigma), n(\sigma))$ and in this paper we insist that for all σ and $\rho \succ \sigma$ we have $n(\rho) \neq n(\sigma)$, so that each bit is read at most once.

Given input X, we write (V_f^X) for the sequence of strings seen (with bits either scanned or selected) by f, so that

$$V_f^X(0) = X(n(\varepsilon))$$
$$V_f^X(k+1) = V_f^X(k).X(n(V_f^X(k))).$$

We write U_f^X for the subsequence of bits selected by f. Formally U_f^X is the limit of the monotone sequence of strings (T_f^X) where

$$T_f^X(0) = \varepsilon$$
$$T_f^X(k+1) = \begin{cases} T_f^X(k) & \text{if } s(V_f^X(k)) = \text{scan} \\ T_f^X(k).n(V_f^X(k)) & \text{if } s(V_f^X(k)) = \text{select}. \end{cases}$$

Informally, the function is used to select bits from X in a non-monotone way. If V is the string of bits we have read so far, $n(V)$ gives the location of the next bit of X to be read. Then "$s(V) = \text{scan}$" means that we will just read this bit, whereas "$s(V) = \text{select}$" means that we will add it to our string T of selected bits.

Definition 2 (p-KL stochastic sequence). A sequence X is p-KL stochastic if for all p-computable KL selection functions f (notice that f can be a partial function) such that the limit U_f^X of (T_f^X) is infinite, we have

$$\lim_{k\to\infty} \Phi(T_f^X(k)) = p.$$

A sequence X is p-Hippocratic KL stochastic if for all KL selection functions f, computable without an oracle p, such that U_f^X is infinite, we have

$$\lim_{k\to\infty} \Phi(T_f^X(k)) = p.$$

Definition 3 (Generalized Bernoulli measure). A generalized Bernoulli measure λ on 2^ω is determined by a sequence (b_i^λ) of real numbers in $(0,1)$. For each i, the event $\{X_1X_2\ldots|X_i=1\}$ has probability b_i^λ, and these events are all independent. In other words, for all finite strings w the set $[w]$ of strings beginning with w has measure

$$\lambda([w]) = \prod_{\substack{i<|w|\\ w_i=1}} b_i^\lambda \prod_{\substack{i<|w|\\ w_i=0}} (1-b_i^\lambda).$$

We say the measure is computable if the sequence (b_i^λ) is uniformly computable.

In some sense a generalized Bernoulli measure treats sequences as though they arise from a sequences of independent Bernoulli trials with parameter b_i^λ for the i^{th} bit.

Recall that for a measure λ, a λ-martingale is a function $\mathcal{M} : 2^\omega \to \mathbb{R}$ satisfying

$$\lambda(\sigma)\mathcal{M}(\sigma) = \lambda(\sigma 1)\mathcal{M}(\sigma 1) + \lambda(\sigma 0)\mathcal{M}(\sigma 0).$$

We now define the notion of a KL martingale, which will be able to select which bit it will bet on next, in a generalized Bernoulli measure. We use the notation from Definition 1.

Definition 4 (λ-KL randomness). Let λ be a generalized Bernoulli measure. A λ-KL martingale is a pair $(f,\mathcal{M})$ where f is a selection function (δ, n) and $\mathcal{M}$ is a function $2^{<\omega} \to \mathbb{R}$ such that, for every sequence $X \in 2^\omega$ for which f select infinitely many bits of X,

$$\mathcal{M}\left(T_f^X(k)\right) = b_i^\lambda \mathcal{M}\left(T_f^X(k).1\right) + (1-b_i^\lambda)\mathcal{M}\left(T_f^X(k).0\right)$$

for all $k \in \mathbb{N}$, where $i = n(V_f^X(k))$.

We say that X is λ-KL random if, for every λ-KL martingale computable with oracle (b_i^λ), the sequence $(\mathcal{M}(T_f^X(k)))$ is bounded. For a sequence Y, we say that X is λ-KL^Y random if this is true even when the λ-KL martingale is also given oracle access to Y.

3.2. Hippocratic stochasticity is not stochasticity

We will show that, despite the fact that we now allow non-monotone strategies, once again there exist sequences computable from α which are α-Hippocratic KL stochastic, for $\alpha \in \text{MLR} \cap \Delta_2^0$ (recall that Chaitin's constant Ω is the prototypical example of such an α).

We remark that our proof shows also that for $\alpha \in \text{MLR} \cap \Delta_2^0$ the Hippocratic and classical versions of MWC-stochasticity are different (see Definition 7.4.1 in [DH10] for a formal definition).

We first need a lemma:

Lemma 6 ([MSU98] and in [DH10] p.311). *If X is Martin-Löf random for a computable generalized Bernoulli measure λ, then X is λ-KL random.*

PROOF (SEE [DH10]). Consider the set of sequences in which the player achieves capital greater than j when he started with capital 1. For obvious reasons, this is an effective open set of measure less than $1/j$. □

Theorem 7. *Let $\alpha \in MLR \cap \Delta^0_2$. There exists a sequence $Q \in 2^\omega$, computable from α, such that Q is α-Hippocratic KL stochastic.*

PROOF. We will first define the sequence Q, and then show that Q is λ-KL random for a certain generalized Bernoulli measure λ for which the parameters (b^λ_i) converge to α. Finally we will show that it follows that Q is actually α-Hippocratic KL stochastic.

Since $\alpha \in \Delta^2_0$, by Shoenfield's limit lemma α is the limit of a computable sequence of real numbers (although the convergence must be extremely slow, since α is not computable). In particular there exists a computable sequence of finite strings (β^k) such that $\beta^k \in \{0,1\}^k$ and

$$\lim_{k\to\infty} \beta^k = \alpha.$$

We define Q_k by

$$Q_k = \begin{cases} 1 & \text{if } 0.\beta^k \geq 0.\alpha_{k'+1} \ldots \alpha_{k'+k} \\ 0 & \text{otherwise.} \end{cases}$$

We set $Q = Q_1 Q_2 \ldots$. Intuitively, as in the proof of Theorem 1, we are using α as a random bit generator to simulate a sequence of Bernoulli trials with parameter β^k.

Notice that the transformation mapping α to Q is a total computable function. We know that in general if g is total computable, and X is (Martin-Löf) random for the uniform measure μ, then $g(X)$ is random for the measure $\mu \circ g^{-1}$ (see [S86] for a proof of this fact). Since $\alpha \in$ MLR, in our case this tell us that Q is random for exactly the generalized Bernoulli measure λ given by $b^\lambda_i = \beta^i$.

It follows from Lemma 6 that Q is λ-KL random. Finally by Lemma 8 below we can conclude that Q is α-Hippocratic KL stochastic, completing the argument. □

Lemma 8. *Let λ be a computable generalized Bernoulli measure and suppose*

$$\lim_{i\to\infty} b^\lambda_i = p.$$

Then every λ-KL random sequence is p-Hippocratic KL stochastic.

PROOF. We prove the contrapositive. Without loss of generality we assume that $0 < p < 1/2$. Suppose that the sequence X is not p-Hippocratic KL stochastic. Then there is a selection function f, computed without an oracle for p, for which the selected sequence U^X_f is infinite and $\Phi(T^X_f(k))$ does not tend to the limit p. We will define a λ-KL martingale which wins on X.

Without loss of generality, there is a rational number $\tau > 0$ such that $p + 2\tau < 1$ and

$$\limsup_{k\to\infty} \Phi(T_f^A(k)) \geq p + 2\tau.$$

Since (b_k^λ) converges to p, by changing the selection function if necessary, we may assume without loss of generality that $b_k^\lambda < p + \tau$ for all locations k read by the selection function. We let

$$\gamma = \frac{p + 2\tau}{p + \tau} - 1 > 0.$$

We let δ be a rational in $(0, 1)$ satisfying both

$$\delta \ \frac{(1 - p - \tau)^2}{(p + \tau)^2} \leq \tau \qquad \text{and} \qquad \log(1 - \delta) > -\frac{\delta}{\ln 2}(1 + \gamma/2)$$

(where log is to base 2). Such a δ exists because $\log(1 - \delta)/\delta$ converges to $-1/\ln 2$ as $\delta > 0$ tends to 0. Note that $\delta(1 - p - \tau)/(p + \tau) < 1$.

Let $\mathcal{M}$ be the λ-KL martingale which begins with capital 1 and then, using selection function f, bets every turn a fraction δ of its current capital on the next bit being 1. Formally, writing T_k for $T_f^X(k)$ and i for $n(V_f^X(k))$, we put $\mathcal{M}(\varepsilon) = 1$ and for each k

$$\begin{aligned}
\mathcal{M}(T_k.0) &= \mathcal{M}(T_k)\Big(1 - \delta\Big), \\
\mathcal{M}(T_k.1) &= \mathcal{M}(T_k)\left(1 - \delta + \frac{\delta}{b_i^\lambda}\right) \geq \mathcal{M}(T_k)\left(1 + \delta\left(\frac{1}{p + \tau} - 1\right)\right).
\end{aligned}$$

We do not care how $\mathcal{M}$ is defined elsewhere.

By induction,

$$\mathcal{M}(T_k) \geq \left(1 + \delta\ \frac{1 - p - \tau}{p + \tau}\right)^{\#1(T_k)} (1 - \delta)^{\#0(T_k)}$$

and thus

$$\frac{\log(\mathcal{M}(T_k))}{k} \geq \frac{\#1(T_k)}{k} \log\left(1 + \delta\ \frac{1 - p - \tau}{p + \tau}\right) + \frac{\#0(T_k)}{k} \log\left(1 - \delta\right).$$

In the following we use standard properties of the logarithm together with our definitions of τ, δ and γ. In particular, note that if we let $x = \delta(1 - p -$

$\tau)/(p+\tau)$ then $0 < x < 1$ so we have $\ln(1+x) > x - x^2/2 > x - x^2$. We have

$$\begin{aligned}
\limsup_{k\to\infty} \frac{\log \mathcal{M}(T_k)}{k} &\geq (p+2\tau)\log\left(1+\delta\,\frac{1-p-\tau}{p+\tau}\right) + (1-p-2\tau)\log(1-\delta) \\
&\geq \frac{p+2\tau}{\ln 2}\left(\delta\,\frac{1-p-\tau}{p+\tau} - \delta^2\frac{(1-p-\tau)^2}{(p+\tau)^2}\right) - \frac{\delta(1+\gamma/2)}{\ln 2}(1-p-2\tau) \\
&\geq \frac{\delta}{\ln 2}\left((1-p-\tau)(1+\gamma) - \delta\,\frac{(1-p-\tau)^2}{(p+\tau)^2} - (1+\gamma/2)(1-p-2\tau)\right) \\
&\geq \frac{\delta}{\ln 2}\left(\tau + \frac{\gamma}{2}(1-p-\tau) - \delta\,\frac{(1-p-\tau)^2}{(p+\tau)^2}\right) \\
&\geq \frac{\delta\gamma\,(1-p-\tau)}{2\ln 2} \\
&> 0.
\end{aligned}$$

Hence $\log(\mathcal{M}(T_k)) \geq ck$ infinitely often, for some strictly positive constant c. Therefore our martingale is unbounded on U_f^X. □

3.3. Kolmogorov-Loveland randomness

We have shown that, for computable randomness and non-monotone stochasticity, whether a string is random can depend on whether or not we have access to the actual bias of the coin. It is natural to ask if this remains true for Kolmogorov-Loveland randomness.

Lemma 9 ([MMNRS06]). *For sequences $X, Y \in 2^\omega$, $X \oplus Y$ is p-KL random if and only if both X is p-KL^Y-random and Y is p-KL^X-random, and this remains true in the Hippocratic setting (that is, where the KL martingales do not have oracle access to p).* □

The proof is a straightforward adaptation of the proof given in [MMNRS06] (Proposition 11). Using Lemma 9 we can show an equivalence between our question and a statistical question.

Theorem 10. *The following two sentences are equivalent.*

1. *The p-Hippocratic KL random and p-KL random sequences are the same.*
2. *From every p-Hippocratic KL random sequence X, we can compute p.*

Proof. For (1) ⇒ (2), we know that if X is p-KL random then it must satisfy the law of the iterated logarithm (see [W00] for this result). Hence we know how quickly $\Phi(X\upharpoonright k)$ converges to p and using this we can (non-uniformly) compute p from X.

For (2) ⇒ (1), suppose that X is p-Hippocratic KL random but not p-KL random. Let $X = Y \oplus Z$. Then by Lemma 9, Y (say) is not p-KL^Z random, meaning that there is a KL martingale $(\mathcal{M}, f)$ which, given oracle access to p and Z, wins on Y. On the other hand, both Y and Z remain p-Hippocratic KL random, so in particular by (2) if we have oracle access to Z then we can compute

p. But this means that we can easily convert $(\mathcal{M}, f)$ into a p-Hippocratic KL martingale which wins on X, since to answer oracle queries to either Z or p it is enough to scan Z and do some computation. □

Acknowledgements

The authors would like to thank Samuel Buss for his invitation to University of California, San Diego. Without his help and the university's support this paper would never exist. Taveneaux's research has been helped by a travel grant of the "Fondation Sciences Mathématiques de Paris". Kjos-Hanssen's research was partially supported by NSF (USA) grant no. DMS-0901020. Thapen's research was partially supported by grant IAA100190902 of GA AV ČR, by Center of Excellence CE-ITI under grant P202/12/G061 of GA CR and RVO: 67985840 and by a visiting fellowship at the Isaac Newton Institute for the Mathematical Sciences in the programme "Semantics and Syntax".

References

[BDS09] Laurent Bienvenu, David Doty, and Frank Stephan, Constructive dimension and Turing degrees. *Theory of Computing Systems*, Vol. 45(4), 2009, pp. 740–755.

[BGHRS11] Laurent Bienvenu, Peter Gács, Mathieu Hoyrup, Cristóbal Rojas and Alexander Shen, Algorithmic tests and randomness with respect to a class of measures. *Proceedings of the Steklov Institute of Mathematics*, Vol. 271, 2011, pp. 41–102.

[DH10] Rod G. Downey and Denis Hirschfeldt. *Algorithmic Randomness and Complexity*. Springer, 2010.

[K10] Bjørn Kjos-Hanssen, The probability distribution as a computational resource for randomness testing. *J. Log. Anal.*, Vol.2, 2010, issn 1759-9008.

[KTT12] Bjørn Kjos-Hanssen, Antoine Taveneaux and Neil Thapen, How much randomness is needed for statistics? *Computability in Europe 2012*, LNCS Vol. 7318, 2012, pp. 395-404.

[Nie09] André Nies. *Computability and Randomness*. Oxford University Press, 2009.

[MMNRS06] Wolfgang Merkle, Joseph S. Miller, André Nies, Jan Reimann and Frank Stephan. Kolmogorov-Loveland randomness and stochasticity. *Ann. Pure Appl. Logic*, Vol. 138, 2006, pp. 183–210, issn 0168-0072.

[MSU98] Andrei A. Muchnik, Alexei L. Semenov and Vladimir A. Uspensky. *Mathematical metaphysics of randomness*. In *Theoret. Comput. Sci.*, Vol. 207, 1998, pp. 263–317, issn 0304-3975.

[S86] Alexander Shen. *One more definition of random sequence with respect to computable measure.* In *First World Congress of the Bernoulli Society on Math. Statictics and Probability theory*, Tashkent, 1986.

[W00] Yongge Wang. *Resource bounded randomness and computational complexity.* In *Theoretical Computer Science*, Vol. 237, 2000, pp. 33–55, issn 0304-3975.