Rutger Kuyper
School of Mathematics and Statistics
Victoria University of Wellington, New Zealand
mail@rutgerkuyper.com

Research Stieltjes Prize 2015

An introduction to (algorithmic) randomness

In 2015 Rutger Kuyper has been awarded the Stieltjes Prize, which recognises the best PhD in Mathematics in the Netherlands. The prize was awarded for his thesis entitled *Computability, Probability and Logic*, which he completed at the Radboud University Nijmegen. After receiving his PhD he became a research fellow at the Victoria University of Wellington. In this article he describes his research at Victoria, which focuses on computability theory, with an emphasis on algorithmic randomness.

What does it mean for a sequence to be *random*? You probably have an intuitive idea of what a random sequence is, but how do we formalise this in a mathematically rigorous way? To motivate what comes next, let us look at the (initial digits of) a few infinite sequences of zeroes and ones:

$$x_1 = 00000000000000000000\ldots$$
$$x_2 = 01010101010101010101\ldots$$
$$x_3 = 11001001000011111101\ldots$$
$$x_4 = 00100101110101011101\ldots$$

Take a break and consider for a moment: which of these sequences would you consider to be random sequences?

Would you consider x_1 to be random? Probably not, since it only contains zeroes, no ones, and therefore does not look very random. What about x_2? This time it has both zeroes and ones, but the sequence follows a very clear pattern and therefore again should not be considered to be a random sequence. On the other hand, x_3 looks much more random, since there does not seem to be a clear pattern. However, if we were to write π in binary, we would see that

$$\pi = 11.00100100001111110 1\ldots$$

and if we compare this to x_3, suddenly x_3 does not seem to be random any longer.

So, how do we actually get an example of a random sequence? Something that we can write down easily follows a pattern and is therefore not random by any reasonable definition of randomness. The key to obtaining a random sequence is using a probabilistic method to generate it. To obtain x_4, we flipped a coin, denoting heads by 1 and tails by 0. With a coin flip having an equal probability to land on heads or tails, there should not be any easy way of predicting what x_4 is going to look like, in the sense that even if we know the first 37 elements of the sequence, there is no way of knowing whether the 38th element is going to be a zero or a one. Thus, x_4 should be considered random.

So, we now have three examples of sequences that should not be considered random, and one example of something that should be considered random. Again, we would like to have a mathematically rigorous definition of what exactly it means to be random, and we now have three examples of sequences that should fail this definition, and one example of something that should pass it. We will use this as a starting point and try to build our mathematical definition from the ground up.

The law of large numbers

Recall the *law of large numbers* from probability theory:

If we flip a coin infinitely many times, and we denote by h_n the number of times we see heads amongst the first n flips, then with probability 1,

$$\lim_{n\to\infty} \frac{h_n}{n} = \frac{1}{2}.$$

Our standard example of a random sequence, x_4, was obtained by flipping a coin like this. Thus, our sequence x_4 will (probably) satisfy the following property:

Definition 1. Given an infinite sequence x of zeroes and ones, if we let $\#x\restriction n$ be the number of ones amongst the first n elements of the sequence, we say that x satisfies the *law of large numbers* if

$$\lim_{n\to\infty} \frac{\#x\restriction n}{n} = \frac{1}{2}.$$

More informally, x satisfies the law of large numbers if, when we look far enough in the sequence, roughly half of the elements are zeroes, and roughly half of the elements are ones. Intuitively a random sequence should not be biased towards zeroes or ones, so a random sequence should satisfy the law of large numbers. We could therefore try to use this as a definition of randomness.

Let us go back to our examples at the beginning. The non-random sequence x_5 does not satisfy the law of large numbers, which is a good sign. However, x_2 does satisfy it, while we decided this sequence

should not be called random. Thus, we need a stronger definition.

Normal numbers

Let us have another look at the sequence

$$x_2 = 01010101010101010101\ldots$$

This sequence satisfies the law of large numbers, because the digits 0 and 1 both appear roughly half the time. Let us consider what happens if we consider blocks of two digits in x_2. We see that x_2 contains the blocks:

$$01,10,01,10,01,10,01,10,01,\ldots$$

In other words, we never see the other two blocks of length 2: 00 and 11. Again, thinking about coin flips, if one flips a coin twice and records both the results, then the four possibilities (heads, heads), (tails, tails), (heads, tails) and (tails, heads) all occur with probability $\frac{1}{4}$. Thus, in a random sequence we do not just want each digit to occur roughly half the time, but we also want each block of two digits to occur roughly one quarter of the time. Generalising this to blocks of arbitrary length we get to the following definition.

Definition 2. Let x be an infinite sequence of zeroes and ones. We say that x is *normal* (in base 2) if for every positive natural number k, and every block b of k many zeroes or ones, if we let $N(x,b,n)$ be the number of times the block b occurs amongst the first n digits of x, then

$$\lim_{n\to\infty} \frac{N(x,b,n)}{n} = \frac{1}{2^b}.$$

(The notion of normality — in arbitrary bases — was introduced by Borel, when we talk about normality we mean the special case of normality in base 2.)

Because x_1 does not satisfy the law of large numbers it is certainly not normal, and x_2 is not normal either as argued above. What about x_3? Perhaps surprisingly, it is not known whether the sequence x_3, which we obtained from the binary expansion of π, is normal, so we do not know whether this definition is able to determine the nonrandomness of x_3. On the other hand, the sequence x_4 obtained by flipping coins is (with probability 1) normal.

Does this mean we are done? Let us consider the following sequence:

$$x_5 = 0110111001011101111000100\,1$$
$$1010101011100\ldots$$

Rutger Kuyper in New Zealand

There is a pattern in this sequence that might not be immediately obvious: we write the natural numbers $0,1,2,3,\ldots$ in binary, and form the sequence x_5 by concatenating all these binary numbers together in order. The sequence x_5 we obtain this way is called *Champernowne's sequence*. This sequence is known to be normal, so would be random if we took normality as our definition of randomness. However, it does not seem very reasonable to consider x_5 to be random. After all, it was generated using such an easy pattern and therefore it is very easy to compute the digits of this sequence.

Church stochasticity

Going back to Champernowne's sequence x_5, why exactly is it non-random? We noticed that it is very easy to reconstruct the pattern used to define x_5. In particular, it is very easy to give an infinite set W of positions at which we know for sure the digit is a one. In particular, if we were to restrict x_5 to W, i.e. we throw away all the digits at positions not in W, we get a sequence that only contains ones and therefore no longer even satisfies the law of large numbers.

Again, if we think about a sequence of coin tosses, if we were, for example, to ignore all the even coin tosses and only record the odd ones, the resulting sequence should still satisfy the law of large numbers. Thus, we would like to say that a sequence is random if every infinite subsequence, that is a sequence obtained by removing all but infinitely many digits, satisfies the law of large numbers.

Unfortunately, that is too strong a property: such sequences do not exist. Indeed, every infinite sequence contains either infinitely many zeroes or infinitely many ones, so we can always find an infinite

subsequence which only contains zeroes or ones! How do we resolve this?

Until now, we have been talking about randomness without talking about algorithmic randomness, and this is the point at which the algorithmic part of the title comes into play. In algorithmic randomness, one uses tools from computability theory to define and work with randomness. To do so, one needs a rigorous definition of what it means to be *computable*, which was given by Turing. Instead of giving the rigorous definition here, think about a function $f:\mathbb{N}\to\mathbb{N}$ as being *computable* if there is a procedure such that any capable human could, given enough paper and time, work out $f(n)$ given any number $n\in\mathbb{N}$. Equivalently, such a function f is computable if it can be implemented using your favourite programming language.

Using the notion of a computable function, we can now restrict to certain nice, easy subsequences, which are selected in a computable way using the technical notion of a *computable selection function* (we omit the exact definition). Now, we say that a sequence is *Church stochastic* if every subsequence selected by a computable selection function satisfies the law of large numbers. It is known that every Church stochastic sequence is normal, so this is a strengthening of our previous notion of randomness. Champernowne's sequence is not Church stochastic, but our sequence x_4 obtained using coin flips is (with probability 1).

So, does that mean that Church stochasticity is the definition of randomness we were looking for? Unfortunately not. Ville showed that there is an infinite sequence x_6 which is Church stochastic, but for every positive natural number n, if we look at the first n digits in the sequence, there are always more zeroes than ones (even though the ratio converges to $\frac{1}{2}$). Again, thinking about what happens when you repeatedly flip a coin, if you occasionally take a break when you get tired, sometimes you will have seen more heads than tails up until then, and sometimes you will have seen more tails than heads. Thus, the sequence x_6 should not be considered random.

Martin-Löf randomness

We now take our last step towards our definition of randomness, finally arriving at a notion that is fit to be called random. This notion, called *Martin-Löf randomness*, was originally introduced by Martin-Löf as an abstracting of the computable selection functions defined above. One of the things that makes this definition so robust is that there are many equivalent definitions of it, of which we will here discuss the three most important ones.

Measure-theoretic definition

First, let us introduce Martin-Löf randomness using measure theory, as originally done by Martin-Löf. This definition requires more mathematical prerequisites than the alternatives down below, and the remainder of this article can be understood without understanding this section.

The standard topology on $2^{\mathbb{N}}$ is generated by the basic open sets

$$[\![\sigma]\!]=\{x\in 2^{\mathbb{N}}\mid x \text{ extends } \sigma\},$$

where σ is any finite sequence of zeroes and ones (we denote this set by $2^{<\mathbb{N}}$). In other words, the open sets are the ones that can be written as a union of such basic open sets. Given any open set $U=\bigcup_{\sigma\in X}[\![\sigma]\!]$, we say that X is *prefix-free* if for any two $\sigma,\tau\in X$, σ is not an extension of τ. We may assume without loss of generality that X is prefix-free. Then the measure $\mu(U)$ is $\sum_{\sigma\in X}2^{-\text{length}(\sigma)}$.

Intuitively, a random x should not be in any 'easy' set of measure 0, and the collection of random x should have measure 1. This motivates the following definition.

Definition 3. A G_δ (or $\mathbf{\Pi}^0_2$) set is a set of the form $V=\bigcap_{n\in\mathbb{N}}U_n$, where each U_n is open. We say that V is *effectively* G_δ (or Π^0_2) if there is a computable function $f:\mathbb{N}\times\mathbb{N}\to 2^{<\mathbb{N}}$ such that

$$V=\bigcap_{n\in\mathbb{N}}\bigcup_{m\in\mathbb{N}}[\![f(n,m)]\!].$$

A *Martin-Löf test* is an effective G_δ set V as above such that $\mu(\bigcup_{m\in\mathbb{N}}[\![f(n,m)]\!])\leq 2^{-n}$. Finally, we say that $x\in 2^{\mathbb{N}}$ is *Martin-Löf random* if x is not in any Martin-Löf test.

How does this connect to the previous section? It turns out that every Martin-Löf random x is Church stochastic. Furthermore, Ville's sequence x_6 is not Martin-Löf random.

Thus, we strengthened our notion of randomness, and got rid of our bad example x_6. Does this mean we have reached our goal and defined randomness in a mathematically rigorous way? Of course that is an informal statement that cannot be formalised, let alone proven. However, the fact that we do not know of any counterexamples and have many equivalent definitions of Martin-Löf randomness seem to show that this is the definition we were looking for.

Definition using betting

This time, let us consider a fair casino, in which we are betting on the digits of a fixed infinite sequence of zeroes and ones $x\in 2^{\mathbb{N}}$, which is known by the casino but not by us. Our goal is to make as much money as possible. The rules of the casino are as follows.

- We initially start with one unit of money.
- We then split all our money between betting on 0 or 1.
- The first bit of x is then revealed to us. If this is a 0, the casino pays out twice our bet on 0. If it is 1, the casino pays out twice our bet on 1.
- We again split our (new) money between betting on 0 and 1.
- The second bit of x is now revealed to us. If this is a 0, the casino pays out twice our bet on 0. If it is 1, the casino pays out twice our bet on 1.
- We keep playing for as long as we like.

If the sequence x has a clear pattern on it, we would like to play in this casino, because it would be very easy for us to predict the next digit and make a lot of money this way. On the other hand, if the sequence is truly random, for example if the casino was just flipping a coin to generate the digits of x, we have no way of predicting the outcome and would therefore not expect to be able to make large profits (unless we got really lucky). This brings us to our second definition of Martin-Löf randomness.

Theorem 1 (Schnorr). *A sequence $x\in 2^{\mathbb{N}}$ is Martin-Löf random if and only if there is no semi-effective betting strategy which guarantees arbitrarily large profits when betting on x if we keep playing as long as we want.*

While this is the easiest definition of Martin-Löf randomness to explain, the fact that we need to talk about *semi-effectiveness* is perhaps not very satisfying. We will not go into the technical details here, but

let us mention that if we restricted ourselves to computable betting strategies we would get a notion of randomness called *computable randomness*, which is strictly weaker than Martin-Löf randomness but interesting in its own right.

Compression and randomness
Finally, let us discuss how Martin-Löf randomness relates to compressibility. Consider your favourite compression algorithm. How good is it at compressing these two finite sequences?

$$01010101010101010101$$
$$00100101110101011101$$

The first one is easy to compress: it can be described by saying that 01 should be repeated ten times. The second sequence on the other hand, again obtained by flipping a coin, does not have a clear pattern and hence can only be described by giving the full sequence. In other words, if one looks at any initial segment of a random sequence, one should be unable to compress that segment, or in other words should be unable to give a short description of it. And indeed, this gives us another characterisation of Martin-Löf randomness.

Theorem 2 (Schnorr). *A sequence x is Martin-Löf random if and only if its initial segments cannot be nontrivially compressed by a (partial) computable compression algorithm.*

(For technical reasons, one needs to restrict to *prefix-free* compression algorithms, but we will not go into further details here.)

Applications of randomness
Now that we know what randomness means mathematically, what can we do with this? In the field of algorithmic randomness, one way of studying randomness is to consider how it interacts with the usual tools and concepts of computability theory. A more recent direction of study has been to consider how randomness interacts with usual mathematical theorems about 'almost everywhere' theorems. Consider, for example, the following theorem of Lebesgue:

Theorem 3 (Lebesgue). *Let $f:[0,1]\to\mathbb{R}$ be a function of bounded variation. Then f is differentiable almost everywhere.*

So far, we have been talking about computable functions on the natural numbers, and randomness of infinite binary sequences. There is a natural and robust way to extend and adapt these concepts to real numbers. So, we can talk about computable functions f of bounded variation, and Martin-Löf random elements of the unit interval. If a theorem holds at almost every point, that should be a strong indication that it holds for a random point. Furthermore, if we are lucky the converse might even hold and therefore give us an equivalent definition of Martin-Löf randomness. This is, in fact, true for the theorem just given.

Theorem 4 (Brattka, Miller and Nies [2]). *Let $x\in[0,1]$. The following are equivalent:*
1. *x is Martin-Löf random.*
2. *Every computable $f:[0,1]\to\mathbb{R}$ of bounded variation is differentiable at x.*

Another interesting direction is to look at Brownian motion. It turns out that Martin-Löf randomness is also the right notion to talk about many almost surely theorems about Brownian motion, as studied by Allen, Bienvenu and Slaman [1].

A different approach: Baire category
As we saw above, one way of defining Martin-Löf randomness is by using effective measure 0 sets. Sets which have measure 0 can be seen as sets which are 'small'. There is another way to talk about the smallness or largeness of sets, using category. Recall that a set is *meagre* if it is contained in a countable union of closed nowhere dense sets. We would like to say that a set is 'random' if it is not in any *effective closed nowhere dense set*, which is done in the following definition.

Definition 4. A closed set $V\subseteq 2^{\mathbb{N}}$ is a Π^0_1*-class* if there is a computable function $f:\mathbb{N}\to 2^{<\mathbb{N}}$ such that $V=2^{\mathbb{N}}\setminus\bigcup_{n\in\mathbb{N}}[\![f(n)]\!]$. We say that $x\in 2^{\mathbb{N}}$ is *1-generic* if for every Π^0_1-class V we have that $x\notin V\setminus\mathrm{Int}(V)$.

Thus, 1-genericity is another way of talking about 'random' points, but using category instead of measure. Many theorems about Martin-Löf random reals have analogues in terms of 1-generics, although this is not always the case.

It is easy to adapt this definition to the real numbers, and hence we can talk about generic elements of $[0,1]$ as well. Just as in the previous section, we wonder how this notion interacts with category almost everywhere theorems. In fact, there is a nice analogue of the differentiability characterisation of Martin-Löf randomness given above, obtained by studying a theorem of Bruckner and Leonard.

Theorem 5 (Kuyper and Terwijn [4]). *Let $x\in[0,1]$. The following are equivalent:*
1. *x is 1-generic.*
2. *Every differentiable computable function $f:[0,1]\to\mathbb{R}$ has continuous derivative at x.*

Further reading
The standard reference works on algorithmic randomness are Downey and Hirschfeldt [3] and Nies [5]. There is also the excellent nontechnical (and significantly shorter) account by Terwijn [6].

References

1 K. Allen, L. Bienvenu and T. Slaman, On zeros of Martin-Löf random Brownian motion, *Journal of Logic and Analysis* 6(9) (2014), 1–34.

2 R. Brattka, J.S. Miller and A. Nies, Randomness and differentiability, *Transactions of the American Mathematical Society* 368(1) (2016), 581–605.

3 R.G. Downey and D.R. Hirschfeldt, *Algorithmic Randomness and Complexity*, Springer, 2010.

4 R. Kuyper and S.A. Terwijn, Effective genericity and differentiability, *Journal of Logic and Analysis* 6(4) (2014), 1–14.

5 A. Nies, *Computability and Randomness*, Oxford University Press, 2008.

6 S.A. Terwijn, *The Mathematical Foundations of Randomness, The Challenge of Chance*, N.P. Landsman and E. van Wolde, eds., Springer, 2016, pp. 49–66.