RaNDom!

Verónica Becher
Universidad de Buenos Aires

CLAM 2021, Session on Logic and Computation, September 14, 2021 azar - aléatoire - Zufall - rasgelelik - satunnaisuuden - slumpmässighet - randomness - aleatorietà

Everyone has an intuitive idea about what is randomness, often associated with “gambling” or “luck”.

1. Is there a mathematical definition of randomness?
2. Are there levels of randomness?
3. Examples of randomness?
4. Can a computer produce a sequence that is truly random?

Lady luck is fickle

Think of 0s and 1s.

A sequence is random if it can not be distinguished from independent tosses of a fair coin.

Lady luck is fickle

Would you believe that these have bee obtained by independent toosses?

111111111111111111111111111111111111111111... ✗

Lady luck is fickle

Would you believe that these have bee obtained by independent toosses?

111111111111111111111111111111111111111111... ✗

01001000100001000001000000100000001000000001... ✗

Lady luck is fickle

Would you believe that these have bee obtained by independent toosses?

111111111111111111111111111111111111111111... ✗

01001000100001000001000000100000001000000001... ✗

100101010110001101110100010010101111001001.. ✓

Heads and tails must occur with the same frequency.
Likewise for any combination of heads and tails.
Otherwise we would be able to guess it infinitely many times!

Monkeys and typewritters

Émile Borel. La mécanique statique et l'irréversibilité.
Journal de Physique Théorique et Appliquée, 1913, 3 (1), pp.189-196.

> *[. . .] Concevons quon ait dressé un million de singes à frapper au hasard sur les touches d'une machine à écrire et que, sous la surveillance de contremaîtres illettrés, ces singes dactylographes travaillent avec ardeur dix heures par jour avec un million de machines à écrire de types variés. Les contre-maitres illettrés rassembleraient les feuilles noircies et les relieraient en volumes. Et au bout d'un an, ces volumes se trouveraient renfermer la copie exacte des livres de toute nature et de toutes langues conservés dans les plus riches bibliothéques du monde.*

Randomness is impossibility to guess, to predict, to abbreviate....

By whom?

Randomness is impossibility to guess, to predict, to abbreviate....

By whom?

By a human being? Ugh! we can not formalize it.

Randomness is impossibility to guess, to predict, to abbreviate....

By whom?

By a human being? Ugh! we can not formalize it.

By a universal Turing machine ? It yields the purest notion of randomness.

Randomness is impossibility to guess, to predict, to abbreviate....

By whom?

By a human being? Ugh! we can not formalize it.

By a universal Turing machine ? It yields the purest notion of randomness.

By finite state automata? It yields the most basic notion of randomness: normality.

And there are intermediate notions.

Automata, different abilities

- Finite state atomata
- Stack automata
- Turing machines

Towards a mathematical definition of randomness

A sequence is random (for the class of automata $\mathcal{C}$) when, essentially, the only way to describe the sequence (using an utomata in the class $\mathcal{C}$) is explicitely.

A base is an integer greater than or equal to 2. For a real number x in the unit interval, the expansion of x in base b is a sequence $a_1a_2a_3\dots$ of integers from $\{0,1,\dots,b-1\}$ such that

$$x = \sum_{k\geq 1} \frac{a_k}{b^k} = 0.a_1a_2a_3\dots$$

Normal numbers, the most basic form of randomness

Definition (Borel, 1909)

A real number x is simply normal to base b if, in the expansion of x in base b, each digit occurs with limiting frequency equal to $1/b$.

A real number x is normal to base b if, for every positive integer k, every block of k digits (starting at any position) occurs in the expansion of x in base b with limiting frequency $1/b^k$.

A real number x is absolutely normal if x is normal to every base.

Not normal

$0.01\ 002\ 0003\ 00004\ 000005\ 0000006\ 00000007\ 000000008\ldots$
is not simply normal to base 10.

Not normal

$0.01\ 002\ 0003\ 00004\ 000005\ 0000006\ 00000007\ 000000008\ldots$
is not simply normal to base 10.

$0.0123456789\ \ 0123456789\ \ 0123456789\ \ 0123456789\ \ 0123456789\ldots$
is simply normal to base 10, but not simply normal to base 100.

Not normal

$0.01\ 002\ 0003\ 00004\ 000005\ 0000006\ 00000007\ 000000008\ldots$
is not simply normal to base 10.

$0.0123456789\ \ 0123456789\ \ 0123456789\ \ 0123456789\ \ 0123456789\ldots$
is simply normal to base 10, but not simply normal to base 100.

The numbers is the middle third Cantor set are not simply normal to base 3 (their expansions lack the digit 1).

Not normal

$0.01\ 002\ 0003\ 00004\ 000005\ 0000006\ 00000007\ 000000008\ldots$
is not simply normal to base 10.

$0.0123456789\ \ 0123456789\ \ 0123456789\ \ 0123456789\ \ 0123456789\ldots$
is simply normal to base 10, but not simply normal to base 100.

The numbers is the middle third Cantor set are not simply normal to base 3 (their expansions lack the digit 1).

The rational numbers are not normal to any base.

Not normal

$0.01\ 002\ 0003\ 00004\ 000005\ 0000006\ 00000007\ 000000008\ldots$
is not simply normal to base 10.

$0.0123456789\ \ 0123456789\ \ 0123456789\ \ 0123456789\ \ 0123456789\ldots$
is simply normal to base 10, but not simply normal to base 100.

The numbers is the middle third Cantor set are not simply normal to base 3 (their expansions lack the digit 1).

The rational numbers are not normal to any base.

Liouville's constant $\sum_{n\geq 1} 10^{-n!}$ is not normal to base 10.

Examples of normal numbers?

Theorem (Borel 1909)

Almost all real numbers are absolutely normal.

Problem (Borel 1909)

Give one example of an absolutely normal number.

Examples of normal numbers?

Theorem (Borel 1909)

Almost all real numbers are absolutely normal.

Problem (Borel 1909)

Give one example of an absolutely normal number.

Are the usual mathematical constants, such as π, e, or $\sqrt{2}$, absolutely normal? Or at least simply normal to some base?

Examples of normal numbers?

Theorem (Borel 1909)

Almost all real numbers are absolutely normal.

Problem (Borel 1909)

Give one example of an absolutely normal number.

Are the usual mathematical constants, such as π, e, or $\sqrt{2}$, absolutely normal? Or at least simply normal to some base?

Conjecture (Borel 1950)

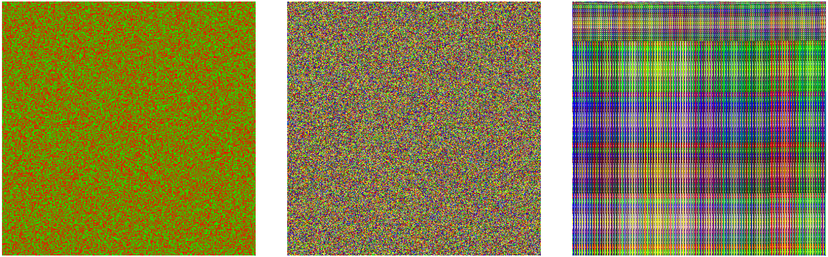
Irrational algebraic numbers are absolutely normal.

Normal to a given base

Theorem (Champernowne, 1933)

$0.123456789101112131415161718192021\ldots$ *is normal to base* 10.

It is unknown if it is normal to bases that are not powers of 10.



base 2 base 6 base 10

Plots of the first 250000 digits of Champernowne's number.

Besicovitch 1935; Copeland and Erdös 1946; Levin 1999;. . . Ugalde 2000; Alvarez, Becher, Ferrari and Yuhjtman 2016.

Absolutely normal

Sierpinski 1917, Lebesgue 1917; Turing 1937; Schmidt 1961; M. Levin 1970; . . . Lutz and Mayordomo 2013,2020; Figueira and Nies 2013, 2020, Becher, Heiber and Slaman 2013.

Theorem

There is an algorithm that computes an absolutely normal number with just above linear time-complexity.

Normal to some bases and not to others

Theorem (Cassels 1959; Schmidt 1961)

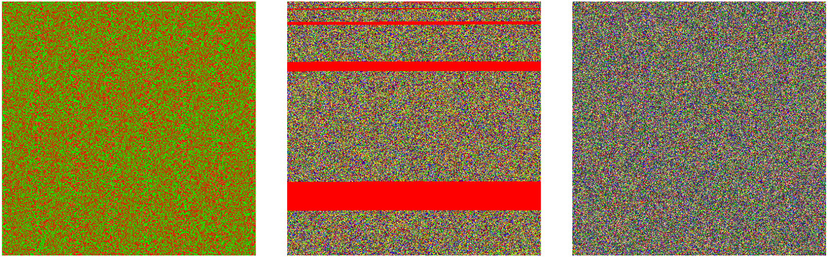
Almost all numbers in the Cantor ternary set are normal to base 2.

Normal to one base, but not to another

Theorem (Stoneham, 1973,Bailey and Borwein 2012)

$$\alpha_{2,3} = \sum_{k \geq 1} \frac{1}{3^k \; 2^{3^k}}$$

is normal to base 2 but not simply normal to base 6.



base 2

base 6

base 10

Plots of the first 250000 digits of Stoneham number $\alpha_{2,3}$.

Normality and finite automata

A deterministic finite transducer T is defined by $\langle Q, A, \delta, q_0 \rangle$ where A is the alphabet, Q is a finite set of states with q_0 the starting state, and $\delta : Q \times A \to Q \times A^*$ is a determinitic transition function.

Every infinite run is accepting (Büchi acceptance condition).

For the result of running T with input $a_1 a_2 a_3 \ldots$ we write $T(a_1 a_2 a_3 \ldots)$.

Normality and finite automata

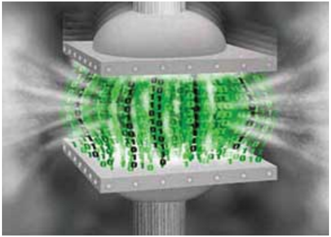
Consider transducer $T = \langle Q, A, \delta, q_0 \rangle$. If $\delta(p, a) = \langle v, q \rangle$ write $p \xrightarrow{a|v} q$.

A sequence $x = a_1 a_2 a_3 \cdots$ is compressible by a finite transducer T if and only if the run in T

$$q_0 \xrightarrow{a_1|v_1} q_1 \xrightarrow{a_2|v_2} q_2 \xrightarrow{a_3|v_3} q_3 \cdots$$

satisfies

$$\liminf_{n \to \infty} \frac{|v_1 v_2 \cdots v_n|}{n} < 1.$$



Normality and finite automata

Theorem (Schnorr, Stimm 1971; Dai, Lathrop, Lutz, Mayordomo 2004)

A sequence is normal if and only if it is incompressible by every one-to-one finite transducer .

Huffman 1959 calls them lossless compressors. A direct proof in Becher and Heiber, 2012.

Theorem (Becher, Carton, Heiber 2013)

Non-deterministic one-to-one finite transducers can not compress normal sequences.

Normality and pushdown automata

Question

Can deterministic pushdown transducers compress normal infinite sequences?

Theorem (Boasson, personal communication 2012)

Non-deterministic puhdown transducers can compress normal sequences.

$0123456789\ 9876543210\ 00\ 01\ 02\ 03\ \ldots 98\ 99\ 99\ 98\ 97 \ldots 03\ 02\ 01\ 00\ 000\ 001\ 002 \ldots$

Normality preservation and finite automata

Let $a_1 a_2 a_3 \cdots$ be an infinite sequence. Consider the infinite sequence obtained by selection of some elements

$$a_1 \; a_2 \; \circled{a_3} \; a_4 \; a_5 \; \circled{a_6} \; \circled{a_7} \; a_8 \; a_9 \; \ldots$$

Theorem (Agafonov 1968)

Prefix selection by a regular set of finite sequences preserves normality.

Normality preservation and finite automata

Let $a_1 a_2 a_3 \cdots$ be an infinite sequence. Consider the infinite sequence obtained by selection of some elements

$$a_1 \ a_2 \ \textcircled{a_3} \ a_4 \ a_5 \ \textcircled{a_6} \ \textcircled{a_7} \ a_8 \ a_9 \ \ldots$$

Theorem (Agafonov 1968)

Prefix selection by a regular set of finite sequences preserves normality.

Theorem (Becher, Carton and Heiber 2013)

Suffix selection by a regular set of infinite sequences preserves normality.

Theorem (Becher, Carton and Heiber 2013)

Two sided selectors do not preserve normality.

Theorem (Merkle and Reimann 2006)

Neither deterministic one-counter sets nor linear sets preserve normality (these are the sets recognized by pushdown finite automata with a unary stack and by one-turn pushdown finite automata, respectively)

Kamae and Weiss (1975) gave a full characterization of the forms of selection that preserve normality.

Problem

What forms of insertion transform normality to base b to normality to base $(b+1)$?

How transform a sequence normal over alphabet A into one normal to alphabet $A \cup \{\sigma\}$, such that the first is a subsequence of the second.

Normality together with other properties

Theorem (Becher, Bugeaud, Slaman 2013)

Let S be any computable set of positive integers, closed by multiplicative dependence and such that, if $b^{km} \in S$ then b^k is in S, and if there are infinitely many k such that $b^k \in S$ then for every $b^m \in S$, for every m. Then, there is a real x which is simply normal to exactly the bases specified by S. Furthermore, the real x is computable from S.

Theorem (after Bugeaud 2002, Becher, Heiber and Slaman 2014)

There is a computable absolutely normal Louville number.

Proof by defining a measure whose Fourier tanform decays quickly.

Theorem (Becher and Madritsch 2021)

There is a computable real x such that x and $1/x$ are absolutely normal and continued fraction normal.

Proof by constructing their continued fraction expansion.

Normal numbers and Descriptive set theory

Asked first by Kechris 1994.

Theorem (Ki and Linton 1994)

The set of real numbers that are normal to any fixed base is Π^0_3*-complete.*

Theorem (Becher, Heiber, Slaman 2014)

The set of real numbers that are absolutely normal is Π^0_3*-complete.*

Descriptive set theory

The set of bases to which a real number can be normal is not tied to any arithmetical properties other than multiplicative dependence.

Theorem (Becher and Slaman 2014)

The set of real numbers that are normal to some base is Σ^0_4-complete in the effective Borel Hierachy on subsets of real numbers.

Achim Ditzen conjectured it in 1994

Theorem (Airey, Jackson and Mance, 2016)

Let N_b be the set of real numbers which are normal to a given base b. The set of real numbers that are normal to base b and preserve normality to base b under addition,

$$\{x : x \in N_b \text{ and } \forall y \in N_b \ (x + y \in N_b)\},$$

is Π^0_3-complete.

Pure randomness

A sequence is purely random if,essentially, its initial segments can only be described explicitely by a Turing machine, thus, requiring each one a different program.

Pure randomness

A sequence is purely random if,essentially, its initial segments can only be described explicitely by a Turing machine, thus, requiring each one a different program.

That is, its initial segments cannot be compressed with a Turing machine. Formally, a sequence is random if its initial segments have almost maximal program-size complexity (Chaitin 1975).

An equivalent definition of randomness

Definition (Martin-Löf 1965, tests of non-randomness)

A sequence is Martin-Löf random if it passes all computably definable tests of non-randomness. Since there is a universal tests, it suffices that to consider just this universal Martin-Löf test.

Technically, a sequence is Martin-Löf random if it belongs to no computably definable null set. Since there is a universal computably definable null set, it suffices to consider this one.

An equivalent definition of randomness

Definition (Martin-Löf 1965, tests of non-randomness)

A sequence is Martin-Löf random if it passes all computably definable tests of non-randomness. Since there is a universal tests, it suffices that to consider just this universal Martin-Löf test.

Technically, a sequence is Martin-Löf random if it belongs to no computably definable null set. Since there is a universal computably definable null set, it suffices to consider this one.

Theorem (Schnorr 1975)

A sequence is random for Chaitin's definition if and only if it does not belong to the universal Martin-Löf null set.

Can a computer produce a purely random sequence?

Can a computer produce a purely random sequence?

No

"Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin."

John von Neumann, 1951

Can a computer produce a purely random sequence?

No

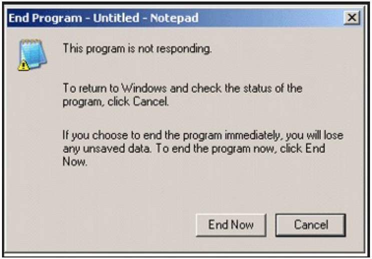
"Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin."

John von Neumann, 1951 (cita Knuth, The Art of Computing Programming)

Every computable sequence is dramatically compressible by a Turing machine. There is a program that outputs each of its symbols, one after the other one. Thus, each initial segment of length n can be obtained with $2\log n+$constant. it is very compressible.

Examples of random sequences

Have you ever experienced that your computer locked up (froze)?



End Program - Untitled - Notepad

This program is not responding.

To return to Windows and check the status of the program, click Cancel.

If you choose to end the program immediately, you will lose any unsaved data. To end the program now, click End Now.

Ω-numbers

Theorem (Chaitin 1975)

The probability that a universal Turing machine with prefix-free domain halts, $\Omega = \sum_{U(p)\,\textit{halts}} 2^{-|p|}$ *is random.*

Similarly, probabilities of other computer behaviours called Ω numbers (Becher,Chaitin 2001,2003; Becher,Grigorieff 2005,2009, Becher,Figueira,Grigorieff,Miller 2006; Barmpalias 2016)

More than 50 years with results program-size complexity (Kolmogorov complexity), computability theory, algorithmic information theory.

Randomness and Birkhoff's ergodic theorem

Theorem (Franklin,Greenberg, Miller,Ng 2012 - Bienvenu,Day,Hoyrup,Mezhirov,Shen 2012)

Let (X, μ) be a computable probability space and let $T : X \to X$ be a computable ergodic map. A point $x \in X$ is random if and only if for every effectively closed subset U of X,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \chi_U(T^n(x)) = \mu(U).$$

Randomness u.d. mod 1

A sequence of reals $(x_n)_{n\geq 1}$ in the unit intevarl is u.d., if for every subinterval $[a,b)$ of the unit interval,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \chi_{[a,b)}(x_n) = b - a$$

A sequence $(x_n)_{n\geq 1}$ of reals in the unit interval is Σ^0_1-u.d. if for every Σ^0_1 set $A \subseteq [0,1]$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \chi_A(x_n) = \mu A,$$

Joint work with Serge Grigorieff based on Koksma Metric Theorem.

Randomness as u.d. $\mod 1$

Theorem (Franklin,Greenberg,Miller,Ng 2012 - Bienvenu,Day,Hoyrup,Mezhirov,Shen 2012)

A real x is random if and only if $(2^n x)_{n \geq 1}$ is Σ^0_1-u.d. mod 1.

Proof using of effective Birkhoff's ergodic theorem

Theorem (Wall 1949)

A real x is normal to base b if and only if $(b^n x)_{n \geq 1}$ is u.d. mod 1.