

Departamento de Computación, Facultad de Ciencias Exactas y Naturales, UBA

Azar y Automatas

Clase 8: Aleatoriedad pura (Test de Martin-Löf)
Relación con distribución uniforme

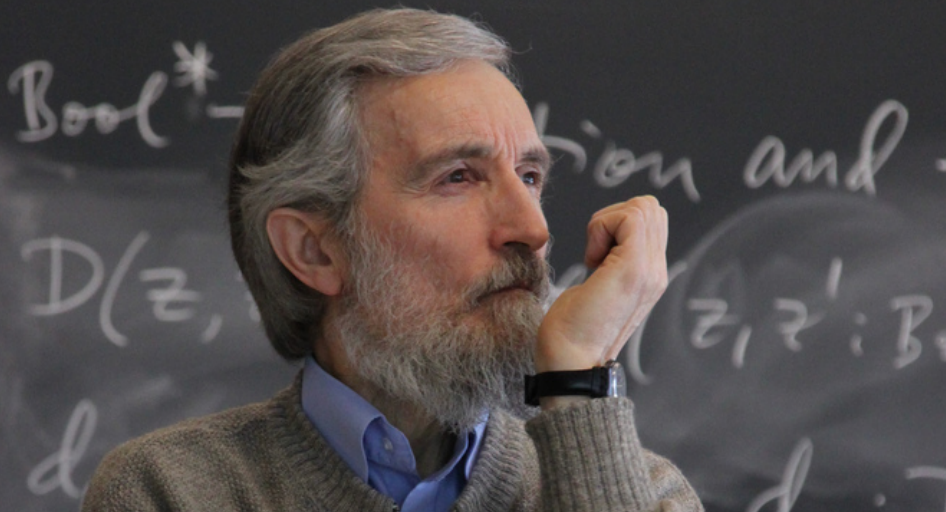
The Definition of Random Sequences

PER MARTIN-LÖF

Institute of Mathematical Statistics, University of Stockholm, Stockholm, Sweden

Kolmogorov has defined the conditional complexity of an object y when the object x is already given to us as the minimal length of a binary program which by means of x computes y on a certain asymptotically optimal machine. On the basis of this definition he has proposed to consider those elements of a given large finite population to be random whose complexity is maximal. Almost all elements of the population have a complexity which is close to the maximal value.

In this paper it is shown that the random elements as defined by Kolmogorov possess all conceivable statistical properties of randomness. They can equivalently be considered as the elements which withstand a certain universal stochasticity test. The definition is extended to infinite binary sequences and it is shown that the non random sequences form a maximal constructive null set. Finally, the Kollektivs introduced by von Mises obtain a definition which seems to satisfy all intuitive requirements.



Intuition for randomness

A real number is random if it belongs to not set of probability 0.

Intuition for randomness

A real number is random if it belongs to not set of probability 0.

A literal reading is not good: no real number would be random.

Martin-Löf randomness

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.

Equivalently,

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

Martin-Löf random reals

Definition (Martin-Löf 1966)

A real x is *random* if for every computable sequence $(V_n)_{n \geq 1}$ of computably enumerable open sets of reals such that $\mu(V_n) < 2^{-n}$,

$$x \notin \bigcap_{n \geq 1} V_n.$$

Almost all (for Lebesgue measure) reals are Martin-Löf random.

How do we know that the definition is right?

The definition of **randomness** was accepted when two different formulations were shown to be equivalent.

(This is similar to what happened with the notion of **algorithm** in 1930s with Church-Turing thesis.)

How do we know that the definition is right?

Theorem (Schnorr 1975)

Martin-Löf and Chaitin definitions coincide.

How is randomness related to theory of uniform distribution?

Über die Gleichverteilung von Zahlen mod. Eins.*)

Von

HERMANN WEYL in Zürich.

§ 1.

Grundlagen. Der lineare Fall.

Es seien auf der Geraden der reellen Zahlen unendlich viele Punkte

$$\alpha_1, \alpha_2, \alpha_3, \dots$$

markiert; wir rollen die Gerade auf einen Kreis vom Umfange 1 auf und fragen, ob dabei die an den Stellen α_n befindlichen Marken schließlich den Umfang des Kreises überall gleich dicht bedecken. Dies würde dann

Uniform distribution modulo one

For a real x , $\{x\} = x - \lfloor x \rfloor$.

Definition

A sequence of reals $(x_n)_{n \geq 1}$ is uniformly distributed modulo one, abbreviated *u.d. mod 1*, if for all $a, b \in [0, 1]$,

$$\lim_{N \rightarrow \infty} \frac{\#\{n : 1 \leq n \leq N, \{x_n\} \in [a, b)\}}{N} = b - a$$

Weyl's criterion

A sequence $(x_n)_{n \geq 1}$ of real numbers is u.d. mod 1 if for every Riemann integrable function f ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(x_n) = \int_0^1 f(x) dx$$

Theorem (Weyl 1916)

A sequence $(x_n)_{n \geq 1}$ of real numbers is u.d. mod 1 if and only if for every non-zero integer h ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i h x_n} = 0$$

Hermann Weyl on a seesaw at a Gasthaus in Nikolausberg, Germany in 1932



Examples

Theorem (Bohl; Sierpiński; Weyl 1909-1910)

A real x is irrational if and only if $(nx)_{n \geq 1}$ is u.d. mod 1.

Theorem (Wall 1949)

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

Koksma's General Metric Theorem

Given a real x in $[0, 1]$ and $(u_n : [0, 1] \rightarrow \mathbb{R})_{n \geq 1}$ consider $(u_n(x))_{n \geq 1}$.

Definition (Koksma 1935)

Let \mathcal{K}^{all} be the class of sequences $(u_n : [0, 1] \rightarrow \mathbb{R})_{n \geq 1}$ such that

1. $u_n(x)$ is continuously differentiable for every n ,
2. $u'_m(x) - u'_n(x)$ is monotone on x for all $m \neq n$,
3. there exists $K > 0$ such that for all $x \in [0, 1]$ and all $m \neq n$,
 $|u'_m(x) - u'_n(x)| \geq K$.

Examples:

$$(nx)_{n \geq 1}$$

$$(2^n x)_{n \geq 1}$$

$(a_n x)_{n \geq 1}$ where $(a_n)_{n \geq 1}$ is a sequence of distinct integers.

Koksma's General Metric Theorem

Theorem (Koksma General Metric Theorem 1935)

Let $(u_n : [0, 1] \rightarrow \mathbb{R})_{n \geq 1}$ in \mathcal{K}^{all} . Then, for almost all (Lebesgue measure) reals x in $[0, 1]$, $(u_n(x))_{n \geq 1}$ is u.d. mod 1.

Avigad's Theorem

Theorem (Avigad 2013)

*If a real x is random then for every **computable** sequence $(a_n)_{n \geq 1}$ of distinct integers, $(a_n x)_{n \geq 1}$ is u.d. mod 1.*

Avigad's Theorem

Theorem (Avigad 2013)

*If a real x is random then for every **computable** sequence $(a_n)_{n \geq 1}$ of distinct integers, $(a_n x)_{n \geq 1}$ is u.d. mod 1.*

Actually Avigad's theorem holds for Schnorr randomness which is weaker than Martin-Löf randomness.

Effective Koksma class \mathcal{K}

Definition

Let \mathcal{K} be the class of *computable* sequences $(u_n : [0, 1] \rightarrow \mathbb{R})_{n \geq 1}$ in \mathcal{K}^{all} such that the sequence of derivatives $(u'_n : [0, 1] \rightarrow \mathbb{R})_{n \geq 1}$ is also *computable*.

Strict inclusion

Theorem 1

Let x be a real in $[0, 1]$. If x is random then for every $(u_n : [0, 1] \rightarrow \mathbb{R})_{n \geq 1}$ in \mathcal{K} the sequence $(u_n(x))_{n \geq 1}$ is u.d. mod 1.

Strict inclusion

Theorem 1

Let x be a real in $[0, 1]$. If x is random then for every $(u_n : [0, 1] \rightarrow \mathbb{R})_{n \geq 1}$ in \mathcal{K} the sequence $(u_n(x))_{n \geq 1}$ is u.d. mod 1.

The reverse of Theorem 1 does **not** hold (proved by Avigad for a smaller class).

Σ_1^0 -u.d. mod 1

Definition

A sequence $(x_n)_{n \geq 1}$ of reals is Σ_1^0 -u.d. mod 1 if for every computably enumerable open set $A \subseteq [0, 1]$,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \# \left\{ n : 1 \leq n \leq N, \{x_n\} \in A \right\} = \mu(A).$$

Σ_1^0 -u.d. mod 1 is different from u.d. mod 1

Proposition

If x is computable and irrational then $(nx)_{n \geq 1}$ is u.d. mod 1 but not Σ_1^0 -u.d. mod 1.

Σ_1^0 -u.d. mod 1 is different from u.d. mod 1

Proposition

If x is computable and irrational then $(nx)_{n \geq 1}$ is u.d. mod 1 but not Σ_1^0 -u.d mod 1.

Proof. Let x be computable and irrational, for example π .

$$A = \bigcup_{n \geq 1} \left(\{nx\} - 2^{-n-3}, \{nx\} + 2^{-n-3} \right)$$

Then,

$$\mu(A) \leq \sum_{n \geq 1} 2 \cdot 2^{-n-3} = 1/2 \quad \text{and} \quad \frac{1}{N} \# \left\{ n : 1 \leq n \leq N, \{x_n\} \in A \right\} = 1.$$

Hence, $(nx)_{n \geq 1}$ is not Σ_1^0 -u.d. mod 1.

Almost all sequences are Σ_1^0 -u.d. mod 1

Consider Lebesgue measure μ on $[0, 1]$ and the product measure μ_∞ on $[0, 1]^\mathbb{N}$.

Proposition (easy extension of Hlawka, 1956)

μ_∞ -almost all elements in $[0, 1]^\mathbb{N}$ are Σ_1^0 -u.d. in the unit interval.

Inclusion

Theorem 2

Let x be a real number in $[0, 1]$. If there is $(u_n : [0, 1] \rightarrow \mathbb{R})_{n \geq 1}$ in \mathcal{K} such that $(u_n(x))_{n \geq 1}$ is Σ_1^0 -u.d. mod 1 then x is random.

Characterization

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

A real x is random if and only if $(2^n x)$ is Σ_1^0 -u.d. mod 1.

Randomness and uniform distribution

exists $(u_n)_{n \geq 1}$ in \mathcal{K} , $(u_n(x))_{n \geq 1}$ is Σ_1^0 -u.d. mod 1

\Downarrow $\Uparrow?$

$(2^n x)_{n \geq 1}$ is Σ_1^0 -u.d. mod 1

\Downarrow \Uparrow

x is random

\Downarrow \nexists

for all $(u_n)_{n \geq 1}$ in \mathcal{K} is $(u_n(x))_{n \geq 1}$ is u.d. mod 1

Discrepancy associated to random reals

Problem

Is there a random real x such that $(2^n x)_{n \geq 1}$ has discrepancy $O((\log N)/N)$?

Discrepancy associated random reals

Definition

$$D_N((x_n)_{n \geq 1}) = \sup_{0 \leq u < v \leq 1} \left| \frac{\#\{n : 1 \leq n \leq N, u \leq \{x_n\} < v\}}{N} - (v - u) \right|$$

Discrepancy associated random reals

Definition

$$D_N((x_n)_{n \geq 1}) = \sup_{0 \leq u < v \leq 1} \left| \frac{\#\{n : 1 \leq n \leq N, u \leq \{x_n\} < v\}}{N} - (v - u) \right|$$

Thus, $(x_n)_{n \geq 1}$ is u.d. mod 1 if $\lim_{N \rightarrow \infty} D_N((x_n)_{n \geq 1}) = 0$.

Discrepancy associated random reals

Definition

$$D_N((x_n)_{n \geq 1}) = \sup_{0 \leq u < v \leq 1} \left| \frac{\#\{n : 1 \leq n \leq N, u \leq \{x_n\} < v\}}{N} - (v - u) \right|$$

Thus, $(x_n)_{n \geq 1}$ is u.d. mod 1 if $\lim_{N \rightarrow \infty} D_N((x_n)_{n \geq 1}) = 0$.

Schmidt, 1972, proved that there is a constant C such that for every $(x_n)_{n \geq 1}$ there are infinitely many N s with

$$D_N((x_n)_{n \geq 1}) \geq C \frac{\log N}{N}.$$

There are Van der Corput sequences such that there is C such that for cofinitely many N s,

$$D_N((x_n)_{n \geq 1}) \leq C \frac{\log N}{N}.$$

Questions and answers about random sequences

Are random sequences normal?

Questions and answers about random sequences

Are random sequences normal?

Yes. Incompressibility by a Turing machine implies incompressibility by a finite automaton.

Questions and answers about random sequences

Are random sequences normal?

Yes. Incompressibility by a Turing machine implies incompressibility by a finite automaton.

Yes. Another proof: Construct a Martin-Löf test that covers all non-normal sequences.

Questions and answers about random sequences

Are almost all sequences random?



Questions and answers about random sequences

Are almost all sequences random?

Yes. By definition, the set of random sequences is the whole set minus the effectively defined universal null set. Then, with probability 1 an arbitrary sequence belongs to the set of random sequences.

Questions and answers about random sequences

Is the spell of good luck (or bad luck) necessarily short?

Questions and answers about random sequences

Is the spell of good luck (or bad luck) necessarily short?

Yes (“Nothing lasts forever...”).

Proof: Think of 0s and 1s. Suppose a random sequence starts $a_1a_2\dots a_n$. If there is a run of 0's longer than $\log n$, then $a_1a_2\dots a_n$ is compressible. Randomness ensures that this will happen only finitely many times.

Questions and answers about random sequences

Can a computer output a random sequence?

Questions and answers about random sequences

Can a computer output a random sequence?

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

John Von Neumann (1951). Various techniques used in connection with random digits.

Questions and answers about random sequences







Can a computer output a random sequence?

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

John Von Neumann (1951). Various techniques used in connection with random digits.

Proof: Every computable sequence is dramatically compressible by a Turing machine! An initial segment of length n can be compressed to $2 \log n + \text{constant}$. Hence, computable sequences are not random.

References

-  J. Avigad. Uniform distribution and algorithmic randomness. *Journal of Symbolic Logic*, 78(1):334–344, 2013.
-  Y. Bugeaud. *Distribution modulo one and Diophantine approximation*, volume 193 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2012.
-  M. Drmota and R. Tichy. *Sequences, discrepancies and applications*. Lecture Notes in Mathematics. 1651. Springer, Berlin, 1997.
-  J. F. Koksma. Ein mengentheoretischer satz über die gleichverteilung modulo eins. *Compositio Math*, 2:250–258, 1935.
-  L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. Dover, 2006.
-  W. Schmidt. Irregularities of distribution VII. *Acta Arithmetica*, 21:45–50, 1972.