Randomness and Uniform Distribution Modulo One

Serge Grigorieff
(Joint work with Verónica Becher)
IRIF, CNRS & Université PARIS-CITÉ

Journée en l'honneur de Patrick Cégielski
Sénart, June 28, 2022

Bizarre facts about Probability theory

• **Probabilit theory has for basic intuition that of random objects** (reals, integers, strings,...)

... but it provides no such formal notion

Random variables are formal objects which have nothing to do with random objects:

they are just *measurable functions*

• **An elementary result of probability theory...**

nobody really believes in

if we toss an unbiaised coin 100 times then 100 heads are just as probable as any other outcome!

The axioms of probability theory,
as developped by Kolmogorov in 1933,
do not solve all mysteries
that they are sometimes supposed to.

Peter Gàcs

In particular,

- **what is a random finite string ?**
- **what is a random infinite sequence ?**
- **what is a random real ?**

A quest going back to
Pierre Simon de Laplace (1749–1827)

In this talk, we shall restrain to infinitary objects (infinite sequences, reals)

Martin-Löf formalization of randomness, 1966

Naive approach on which it is based:

$\alpha \in \{0,1\}^{\mathbb{N}}$ *is a random infinite sequence if it avoids every set of measure zero.*

Problem: any singleton set $\{\alpha\}$ has measure zero.

Martin-Löf formalization of randomness

Martin-Löf randomness:
ask α to avoid every "constructively null set"
i.e. sets of the form $\bigcap_{n \in \mathbb{N}} U_n$ where U_n is an open set $U_n = \bigcup_{p \in \mathbb{N}} I_{n,p}$ where the $I_{n,p}$ are open rational intervals and the double indexed sequence $(I_{n,p})_{n,p \in \mathbb{N}}$ is computable and the sequence $(\mathrm{meas}(U_n)_{n \in \mathbb{N}})$ is upper-bounded by a computable function converging towards 0.

Schnorr randomness:
also ask that $(\mathrm{meas}(U_n))_{n \in \mathbb{N}}$ is computable

Fact. *Martin-Löf random sequences* $\alpha \in \{0,1\}^{\mathbb{N}}$

1. *constitute a set of measure* 1
2. *satisfy all usual probability laws*

Same for Schnorr random sequences

Why?

1 As the *complement of the union of countably many null sets*

2 The set of exceptions to the law has measure 0
hence is covered by an open set with measure $\leq \varepsilon$
hence is covered by a union of rational intervals $(I_n)_{n\in\mathbb{N}}$
with total measure $\leq \varepsilon$
The proof of a usual law gives a computable such sequence $(I_n)_{n\in\mathbb{N}}$
hence the set of exceptions is included in a constructively null set

Kolmogorov's approach to randomness, circa 1964

Even after his 1933 axiomatization of probability theory, Kolmogorov (1903–1987) never gave up the project to formalize the notion of random object.

(By the way, he is really the unique probabilist (up to now) to believe that Kolmogorov's axioms for probability theory do not constitute the last word about formalizing randomness...)

Approach via Descriptional complexity

built on the theory of computable functions

$$\text{independently by} \begin{cases} \text{Solomonoff} & (1962/1964) \\ \text{Kolmogorov} & (1963/1965) \\ \text{Chaitin} & (1964/1966) \end{cases}$$

Approach via "Descriptional complexity", also called Kolmogorov complexity

$$\boxed{\begin{array}{l}\textbf{complexity of an objet} \\ \qquad\qquad = \textbf{length of the shortest descriptions}\end{array}}$$

But, *description in which context?*

Care ! Berry's paradox (1908):
"the smallest integer which cannot be described by any sentence with less than twenty words"

Aie, aie, aie...this sentence has 15 words and defines an integer which cannot be defined in less than 20 words!

Kolmogorov complexity

complexity of an objet
$= $ length of the shortest descriptions

Key idea for Kolmogorov complexity:
replace "description" *by* "computation program" *in order to enter the formal framework of computability theory set up by Turing and Church.*

Definition. *Let* $f : \{0,1\}^{<\omega} \to D$ *be a partial function (where* D *is* $\mathbb{N}$ *or* $\{0,1\}^{<\omega}$ *or...)*
Set, for $x \in D$,
$K_f : D \to \mathbb{N} \quad , \quad K_f(x) = \min\{|p| : f(p) = x\}$

Invariance theorem

$$K_f : D \to \mathbb{N} \quad , \quad K_f(x) = \min\{|p| : f(p) = x\}$$

(where D is $\mathbb{N}$ or $\{0,1\}^{<\omega}$ or...)

How to choose f ?

Invariance theorem. *Among the K_f's, where f varies in the family PC^D of partial computable function $\{0,1\}^{<\omega} \to D$, there is a smallest one, up to an additive constant:*

$$\boxed{\exists \varphi \in PC^D \ \forall f \in PC^D \ \exists c \ \forall x \in D \ K_\varphi(x) \le K_f(x) + c}$$

Such a φ is called optimal.

Proof: simple application of the enumeration theorem for partial computable functions.

Kolmogorov complexity

$$\boxed{\text{Kolmogorov complexity} = K_\varphi \text{ for any optimal } \varphi}$$

An integer defined up to a constant. . . !... Fortunately, the constant is uniform in $x \in D$, so asymptotically it is OK.

What Kolmogorov said about the constant:

> *The different "reasonable" [above optimal functions] will lead to "complexity estimates" that will converge on hundreds of bits instead of tens of thousands.*
>
> *Hence, such quantities as the "complexity" of the text of "War and Peace" can be assumed to be defined with what amounts to uniqueness.*

Kolmogorov idea to define randomness

Easy fact. *There exists a constant c such that for every word x we have $K(x) \leq |x| + c$*

Kolmogorov idea: say that an infinite string α is random if $\quad \exists d \quad \forall n \in \mathbb{N} \quad K(\alpha \restriction n) \geq |x| - d$

where $\alpha \restriction n = (\alpha(0)\alpha(1)\dots\alpha(n-1))$

ALAS, THIS IS FALSE FOR ALL α

Theorem (Per Martin-Löf, 1971). *For any α there exists infinitely many n's such that $K(\alpha \restriction n) \leq n - \log(n)$.*

Can replace $\log(n)$ by $f(n)$ where the series $\sum_{n \in \mathbb{N}} 2^{-f(n)}$ is divergent

Nevertheless, this idea – slightly modified – does work

$\boxed{\text{Claus Peter } \textbf{Schnorr's process complexity } S}$

consider partial computable $f : \{0,1\}^{<\omega} \to \{0,1\}^{<\omega}$ which are *monotone*:

$$p \leq_{prefix} q \ \wedge \ p, q \in domain(f) \implies f(p) \leq_{prefix} f(q)$$

The invariance theorem still holds, so we can define a Schnorr complexity S similar to the Kolmogorov complexity

Theorem. (Schnorr, 1973) *α is Martin-Löf random if and only if* $\exists c \; \forall n \; |S(\alpha \upharpoonright n) - n| \leq c$

Nevertheless, this idea – slightly modified – does work

$\boxed{\textbf{Levin-Chaitin complexity } H}$

consider partial computable $f : \{0,1\}^{<\omega} \to \{0,1\}^{<\omega}$ *with prefix-free domain*

The invariance theorem still holds, so we can define a Levin-Chaitin complexity H similar to the Kolmogorov complexity

Theorem. (Levin-Chaitin, 1973) α *is Martin-Löf random if and only if* $\exists c \;\; \forall n \;\; H(\alpha \restriction n) \geq |x| - c$

Nevertheless, this idea – slightly modified in yet another way – does work

Theorem (Per Martin-Löf, 1971). *If the series* $\sum_{n\in\mathbb{N}} 2^{-f(n)}$ *is recursively convergent then*

$$\boxed{\begin{array}{l} \forall\alpha \;\; (\alpha \text{ Martin-Löf random} \\ \qquad \Longrightarrow \quad \exists c\ \forall n \;\; K(\alpha\restriction n) \geq n - f(n) - c \end{array}}$$

Theorem. (Joe Miller & Liang Yu, 2004)
There exists a computable $g : \mathbb{N} \to \mathbb{N}$ *such that the series* $\sum_{n\in\mathbb{N}} 2^{-g(n)}$ *is recursively convergent and* $\forall\alpha$,

$$\boxed{\begin{array}{l} \forall\alpha \;\; (\alpha \text{ Martin-Löf random} \\ \qquad \Longleftrightarrow \quad \exists c\ \forall n \;\; K(\alpha\restriction n) \geq n - g(n) - c) \end{array}}$$

Also, the equivalence holds with H in place of g

Another connected notion: sequence of reals equi-distributed modulo 1

$x = \lfloor x \rfloor + \{x\}$ with $x \in \mathbb{Z}$ and $\{x\} = x \mod 1 \in [0,1)$
(integral and fractional parts of x)

Definition. (Hermann Weyl, 1914)
$(x_n)_{1\le n\le N}$ is uniformly distributed modulo 1 (ud) if
for every rational interval $[a,b) \subseteq [0,1)$
$\lim_{N\to\infty} \dfrac{1}{N} \sharp\{n \mid 1 \le n \le N \text{ and } \{x_n\} \in [a,b)\} = b - a$

Fact. (Bohl, Sierpinski, Weyl, 1909)
$\boxed{x \text{ is irrational} \iff \text{the sequence } (nx)_{n\ge1} \text{ is ud}}$

A connected notion: Borel normality

We now turn towards approaches which have to do with randomness but are really aiming at other characterizations.

- A real x is Borel absolutely normal if for every $b \in \mathbb{N}$, $b \geq 2$, the digits in the base b representation of α are uniformly distributed.

I.e the number of a given digit among the n first digits of x tends to $1/b$

Fact. (Niven,Zuckerman 1951)

$$\boxed{\begin{array}{l} x \text{ is absolutely normal} \iff \\ \qquad\qquad \text{the sequence } (b^n x)_{n\geq 1} \text{ is ud for all } b \geq 2,\ b \in \mathbb{N} \end{array}}$$

From Schnorr randomness to a simple instance of ud

Fact. (Avigad, 2013) • If $(a_n)_{n\geq 1}$ is a computable sequence of pairwise distinct integers then

$$\boxed{\begin{array}{l} x \text{ is Schnorr random} \\ \text{(a fortiori if } x \text{ is Martin-Löf random)} \\ \qquad\qquad \Longrightarrow \text{ the sequence } (a_n x)_{n\geq 1} \text{ is ud} \end{array}}$$

• (Avigad, 2013) It is NOT an equivalence, there are counterexamples

Tool of the theory of uniform distribution: Koksma General Metric Theorem

Definition. (Koksma, 1935) Let $K > 0$. A sequence of functions $u_n : [0,1] \to \mathbb{R}$ is K-Koksma if

- u_n is continuously differentiable for all n
- The difference $u_n' - u_p'$ is monotonous for all n, p
- $|u_n'(x) - u_p'(x)| \geq K$ for all $n \neq p$ and $x \in [0,1]$

Example. If the a_n are pairwise distinct integers then the $(x \mapsto nx)_{n \in \mathbb{N}}$ is 1-Koksma.

Koksma General Metric Theorem. (1935)

$$\boxed{\text{If } (u_n)_{n\geq 1} \text{ is Koksma then for almost all } x \in [0,1] \text{ the sequence } (u_n(x))_{n\in\mathbb{N}} \text{ is ud.}}$$

Extension of Avigad's result to effective Koksma sequences

Definition. A Koksma sequence of functions $u_n : [0,1] \to \mathbb{R}$ is effective Koksma if the sequences $(u_n)_{n\in\mathbb{N}}$ and $(u'_n)_{n\in\mathbb{N}}$ are computable

Theorem. (V.Becher & SG, 2022) If x is Schnorr random then the sequence $(u_n(x))_{n\in\mathbb{N}}$ is ud for every effective Koksma sequence of functions $0,1] \to \mathbb{R}$

This extends Avigad since $(x \mapsto a_n x)_{n\in\mathbb{N}}$ is Koksma if the a_n's are distinct integers and $(a_n)_{n\in\mathbb{N}}$ is computable

What about a reciprocal? Still open

Towards a reciprocal via Σ^0_1-uniform distribution and Lipshitz functions

Σ^0_1 subset of $[0,1] = \bigcup_{p\in\mathbb{N}} I_p$ where $(I_p)_{p\in\mathbb{N}}$ is a computable sequence of rational intervals of $[0,1]$

> A sequence of reals $(x_n)_{n\in\mathbb{N}}$ is Σ^0_1-ud if for every Σ^0_1 set U $\quad \lim_{N\to\infty} \frac{1}{N} \sharp\{n \mid 1 \le n \le N \text{ and } \{x_n\} \in U\} = \mathrm{meas}(U)$

Schnorr-Σ^0_1-ud : ask $\mathrm{meas}(U)$ to be computable

A function $f : [0,1] \to \mathbb{R}$ is ℓ-Lipschitz if $|f(x) - f(y)| \le \ell|x-y|$ for all $x, y \in [0,1]$

A computable sequence $(u_n)_{n\in\mathbb{N}}$ is computably Lipschitz if for some computable sequence $(\ell_n)_{n\in\mathbb{N}}$, for every n the function u_n is ℓ_n-Lipschitz

A reciprocal via Σ^0_1-uniform distribution and Lipshitz functions

Theorem. (V.Becher & SG, 2022)

If $\begin{cases} (u_n)_{n\in\mathbb{N}} \text{ is computably Lipschitz} \\ \text{the sequence } (u_n(x))_{n\in\mathbb{N}} \text{ is } \Sigma^0_1\text{-ud} \end{cases}$

then x is Martin-Löf random

If $\begin{cases} (u_n)_{n\in\mathbb{N}} \text{ is computably Lipschitz} \\ \text{and the sequence } (u_n(x))_{n\in\mathbb{N}} \text{ is Schnorr-}\Sigma^0_1\text{-ud} \end{cases}$

then x is Schnorr random

So,

$$\boxed{\begin{array}{lcl} x \text{ random} & \Longrightarrow & \text{and } (u_n(x))_{n\in\mathbb{N}} \text{ ud for effective Koksma } (u_n)_{n\in\mathbb{N}} \\ x \text{ random} & \Longleftarrow & (u_n(x))_{n\in\mathbb{N}}\ \Sigma^0_1\text{-ud for comput. Lipschitz } (u_n)_{n\in\mathbb{N}} \end{array}}$$

The characterization uses ergodic theory

Let $T : [0,1) \to [0,1)$. A set $A \subseteq [0,1)$ is almost invariant if $T^{-1}(A)$ and A coincide up to a null set

A measure preserving T (i.e. $\mathrm{meas}(T^{-1}(A)) = \mathrm{meas}(A)$) is ergodic if every almost invariant set has measure 0 or 1

Examples: $x \mapsto x + na \mod 1$, $x \mapsto 2^n x \mod 1$

Ergodic Theorem. (Birkhoff, Khinchine, 1931)

> If T is measure preserving and ergodic and A is Lebesgue measurable then for almost all x
> $\lim_{N\to\infty} \frac{1}{N} \sharp\{n \mid 1 \le n \le N \text{ and } T^n(x) \in A\} = \mathrm{meas}(A)$

The characterization uses ergodic theory

Theorem. (V.Becher & SG, 2022)

Equivalent conditions

1 x is Martin-Löf random $\quad 1 \Rightarrow 2$ effective ergodic theorem

2 $(T^n(x))_{n\geq 1}$ is Σ^0_1-ud for every T which is computable, measure preserving and ergodic

$2 \Rightarrow 3, 4$ since 2 applies to $x \mapsto x + a \mod 1$ and $x \mapsto 2x \mod 1$

3 $(x + na)_{n\geq 1}$ is Σ^0_1-ud for some irrational a

4 the sequence $(2^n x)_{n\geq 1}$ is Σ^0_1-ud

$3, 4 \Rightarrow 5$ since $x \mapsto x + a \mod 1$ and $x \mapsto 2x \mod 1$ are Lipschitz

5 for some computably Lipschitz sequence $(u_n)_{n\in\mathbb{N}}$ the sequence $(u_n(x))_{n\in\mathbb{N}}$ is Σ^0_1-ud

$5 \Rightarrow 1$ our previous theorem

Thank you for your attention