Algorithmic Randomness

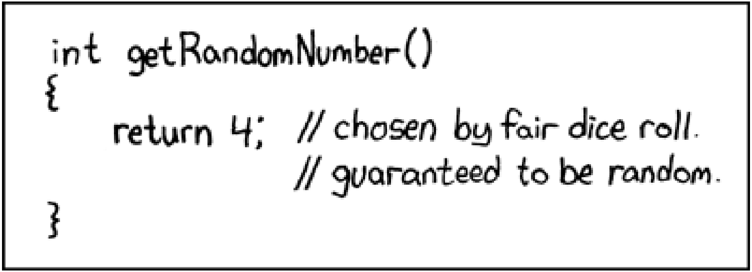
Daniel Turetsky

Victoria University of Wellington

January 2021

Randomness?

What does it mean to say a number is random?



(From Randall Munroe, xkcd.com)

Real numbers

$$\cdots \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \quad \cdots$$

We'll look at real numbers.

Real numbers

We'll look at real numbers.

We'll focus on the unit interval.

Motivation

Many theorems hold "almost surely".

Theorem (Lebesgue)

Every nondecreasing function $f : [0,1] \to \mathbb{R}$ *is differentiable almost everywhere.*

Motivation

Many theorems hold "almost surely".

Theorem (Poincaré)

Let $(\mathcal{X}, \mu)$ be a probability space, $E \subseteq \mathcal{X}$ be measurable, and $T : \mathcal{X} \to \mathcal{X}$ be measure preserving. Then for almost every $y \in E$, there are infinitely many n with $T^n(y) \in E$.

Motivation

Many theorems hold "almost surely".

Theorem (Lebesgue)

If $E \subseteq \mathbb{R}$ is measurable, then for almost every $y \in E$,

$$\lim_{\delta \to 0} \frac{\mu(E \cap [y-\epsilon, y+\epsilon])}{2\epsilon} = 1.$$

Motivation

Many theorems hold "almost surely".

Theorem (Birkhoff)

Let $(\mathcal{X}, \mu)$ be a probability space, $E \subseteq \mathcal{X}$ be measurable, and $T : \mathcal{X} \to \mathcal{X}$ be ergodic. Then for almost every $y \in \mathcal{X}$,

$$\lim_{n\to\infty} \frac{\#\{i : i < n \text{ and } T^i(y) \in E\}}{n} = \mu(E).$$

Motivation

Many theorems hold "almost surely".

Theorem (Birkhoff)

Let $(\mathcal{X}, \mu)$ be a probability space, $E \subseteq \mathcal{X}$ be measurable, and $T : \mathcal{X} \to \mathcal{X}$ be ergodic. Then for almost every $y \in \mathcal{X}$,

$$\lim_{n\to\infty} \frac{\#\{i : i < n \text{ and } T^i(y) \in E\}}{n} = \mu(E).$$

So if we choose a point "at random", it will satisfy the theorem.

How random does it need to be? Can we compare the amount of randomness required?

Cantor space

Cantor space: $\{0,1\}^{\mathbb{N}}$, i.e. the space of infinite binary sequences

Cantor space can be identified with the unit interval via binary expansion: $X \in \{0,1\}^{\mathbb{N}}$ corresponds to $0.X \in \mathbb{R}$.

Finite binary strings: $\{0,1\}^{<\mathbb{N}}$. $\langle\rangle$ is the empty string.

If σ is a finite binary string, $[\sigma]$ is the set of all infinite binary sequences beginning with σ. Give $[\sigma]$ the fair coin measure:

$$\mu([\sigma]) = 2^{-|\sigma|}$$

Typicality

- First attempt: a random sequence should not have any rare (measure 0) properties.
  - Problem: every sequence has such a property: being itself.
- Second attempt: a random sequence should not have any rare (measure 0) properties *that can be described via computability theory*.

A key fact

If $E \subseteq [0,1]$ is null, then there is a sequence of open sets $A_0, A_1, \dots$ with:

- $|A_n| \leq 2^{-n}$;
- $E \subseteq \bigcap_n A_n$.

We will describe a measure 0 set by describing a sequence of open sets of this sort.

Computability theory

A *partial computable function* is a partial function given by an algorithm, i.e. a human could follow the instructions and calculate it with enough pencils, paper and time.

Important: there are only countably many computable functions!

A *computably enumerable (c.e.) set* is the range of a partial computable function.

A c.e. set is a black box that every so often claims elements.

Martin-Löf randomness

A *Martin-Löf test* is a c.e. set $D \subseteq \{0,1\}^{<\mathbb{N}} \times \mathbb{N}$ such that for

$$V_n = \bigcup_{(\sigma,n)\in D} [\sigma],$$

$\mu(V_n) \leq 2^{-n}$.

$X \in \{0,1\}^{\mathbb{N}}$ *passes* the Martin-Löf test if $X \notin \bigcap_n V_n$.

X is *Martin-Löf random* if it passes every Martin-Löf test.

An example

$$\begin{array}{ll} V_0 : & \langle\rangle \\ V_1 : & 0 \\ V_2 : & \overbrace{000 \qquad\qquad 010} \\ V_3 : & \overbrace{00000 \quad 00010} \qquad \overbrace{01000 \quad 01010} \\ \vdots & \end{array}$$

Having a 0 in every other position is atypical.

Normality

Similarly, every Martin-Löf random obeys the law of large numbers:

$$\lim_{n\to\infty} \frac{\#\{i < n : X(i) = 1\}}{n} = \frac{1}{2}.$$

More generally, every Martin-Löf random is normal in every base.

Schnorr randomness

A *Schnorr test* is a Martin-Löf test where $\mu(V_n) = 2^{-n}$.

X is *Schnorr random* if it passes every Schnorr test.

Unpredictability

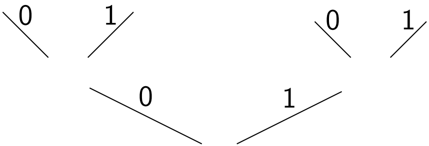
- Idea: a random sequence should be impossible to predict.
- There should be no (computable) betting system by which a gambler can make money betting on the next value.

Martingales

A *martingale* is a function $m : \{0,1\}^{<\mathbb{N}} \to [0,\infty)$ such that

$$m(\sigma) = \frac{m(\sigma * 0) + m(\sigma * 1)}{2}.$$

Example:

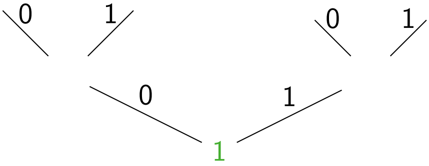


Martingales

A *martingale* is a function $m : \{0,1\}^{<\mathbb{N}} \to [0,\infty)$ such that

$$m(\sigma) = \frac{m(\sigma * 0) + m(\sigma * 1)}{2}.$$

Example:



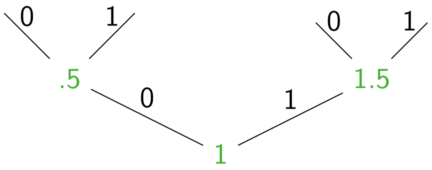
A gambler starts with 1 dollar. $m(\langle\rangle) = 1$

Martingales

A *martingale* is a function $m : \{0,1\}^{<\mathbb{N}} \to [0,\infty)$ such that

$$m(\sigma) = \frac{m(\sigma * 0) + m(\sigma * 1)}{2}.$$

Example:



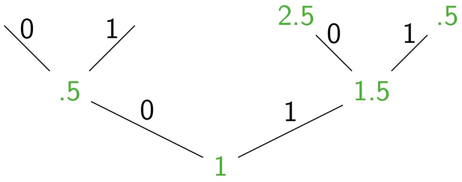
They bet .5 that the first bit is 1. $m(0) = .5$; $m(1) = 1.5$

Martingales

A *martingale* is a function $m : \{0,1\}^{<\mathbb{N}} \to [0,\infty)$ such that

$$m(\sigma) = \frac{m(\sigma * 0) + m(\sigma * 1)}{2}.$$

Example:



If it was 1, they bet 1 that the next bit is 0.
$m(00) = 2.5$; $m(01) = .5$

Martingale success

Exercise

For any martingale m and $c > 0$,

$$\mu\{X \in \{0,1\}^{\mathbb{N}} : \exists n\, m(X\upharpoonright_n) \geq cm(\langle\rangle)\} \leq \frac{1}{c}.$$

A martingale *succeeds* on X if $\liminf_n m(X\upharpoonright_n) = \infty$.

By the exercise, a martingale only succeeds on a null set.

Randomness from martingales

X is *computably random* if no computable martingale succeeds on it.

A martingale is *left c.e.* if $\{(q, \sigma) \in \mathbb{Q} \times \{0,1\}^{<\mathbb{N}} : q < m(\sigma)\}$ is a c.e. set.

Theorem (Schnorr)

X is Martin-Löf random iff no left c.e. martingale succeeds on it.

Proof of $\Rightarrow$.

If m is a left c.e. martingale, define

$$V_n = \bigcup_{m(\sigma) > 2^n m(\langle\rangle)} [\sigma].$$

If $\liminf_n m(X\restriction_n) = \infty$ (or $\limsup$), then $X \in V_n$ for all n. $\square$

Comparing and using

Martin-Löf randoms $\subset$ computable randoms $\subset$ Schnorr randoms

Theorem	Randomness
Nondecr. fns differentiable	Computable randomness[1]
Poincaré Rec.	Martin-Löf randomness[2]
Birkhoff's Theorem	Schnorr randomness[3]
Lebesgue Density (with $\limsup$)	Complicated (Martin-Löf suffices)
Lebesgue Density (with $\lim$)	Complicated (Martin-Löf does not suffice)

[1] Brattka, J. Miller, Nies
[2] Hoyrup
[3] Gács, Hoyrup, Rojas

Incompressibility

- Lossless compression algorithms work by recognizing patterns in the data.
- Randoms should have no patterns.
- Thus randoms should be incompressible.

Kolmogorov complexity

A partial function $f : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ is *prefix-free* if no element of its domain extends another.

Think of f as a decompression function.

The *f-Kolmogorov complexity* of a string σ is

$$K_f(\sigma) = \min\{|\rho| : f(\rho) = \sigma\}.$$

Randomness from Kolmogorov complexity

Theorem (Schnorr)

X is Martin-Löf random iff for every partial computable, prefix-free f, $\sup_n[n - K_f(X \upharpoonright_n)] < \infty$.

Proof of $\Rightarrow$.

For $\sigma \in \{0,1\}^{<\mathbb{N}}$, define

$$m_\sigma(\tau) = \begin{cases} 0 & \text{if } \sigma \perp \tau, \\ 2^{\min\{|\sigma|,|\tau|\}} & \text{otherwise.} \end{cases}$$

Define

$$M = \sum_{f(\rho)=\sigma} 2^{-|\rho|} m_\sigma.$$

If $n - K_f(X \upharpoonright_n) > b$, then for all $\ell > n$,

$$M(X \upharpoonright_\ell) \geq 2^{-|K_f(X \upharpoonright_n)|} M_{X \upharpoonright_n}(X \upharpoonright_\ell) > 2^b.$$

$\square$

An example

There are algorithms to calculate π to any precision.

So define $f(0^i1) = \pi \upharpoonright_{2i}$. This is computable and prefix free.

For $n = 2i$,

$$n - K_f(X \upharpoonright_n) = 2i - i = i.$$

This tends to infinity as n does.

So π is not Martin-Löf random. Nor is $e, \sqrt{2}, \varphi$, or any other computable real.

What kind of functions are K_f?

Given σ, we can search for a τ such that $f(\tau) = \sigma$. If we find one, we know $K_f(\sigma) \leq |\tau|$. But there might be a shorter ρ with $f(\rho) = \sigma$.

In general, K_f is not computable but *computable from above*: from σ, we can compute a decreasing sequence of (extended) integers which stops at $K_f(\sigma)$, but we'll never know when we reach the end of the sequence.

Equivalently, $\{(\sigma, n) : K_f(\sigma) \leq n\}$ is c.e.

Also,

$$\sum_{\sigma \in \{0,1\}^{<\mathbb{N}}} 2^{-K_f(\sigma)} \leq \sum_{\tau \in \mathrm{dom}(f)} 2^{-|\tau|} = \sum_{\tau \in \mathrm{dom}(f)} \mu([\tau]) \leq \mu(\{0,1\}^{\mathbb{N}}) = 1.$$

An optimal K

Fact (Kleene?)

There is a partial computable, prefix-free function U such that for every g which is computable from above and has $\sum_\sigma 2^{-g(\sigma)} < \infty$,

$$\sup_{\sigma \in \{0,1\}^{<\mathbb{N}}} [K_U(\sigma) - g(\sigma)] < \infty.$$

We denote K_U by simply K.

So X is Martin-Löf random iff $\sup_n [n - K(X \upharpoonright_n)]$.

By the earlier proofs, there is a best left c.e. martingale and a best Martin-Löf test (a *universal* Martin-Löf test).

Halting probability

Definition

$$\Omega = \sum_{\tau \in \operatorname{dom} U} 2^{-|\tau|}.$$

Intuitively: pick an OS. Generate a file 1 bit at a time by flipping a fair coin. What is the probability that you eventually generate a well-formed program that runs and halts?

Ω is computable from below: we can build a computable increasing sequence of rationals which converges to Ω.

Ω is a fixed real... but we can influence it?

Fix $q_0, q_1, \dots$ computable, increasing, converging to Ω.

We build a g based on this sequence which is computable from above and summable. So there is b with $K_U(\sigma) \leq g(\sigma) + b$.

At some stage s, we pick σ not yet in the range of U and define $g(\sigma) = n$. U must eventually reveal a new string of length at most $n + b$ in its domain.

So there is $t \geq s$ with $q_{t+1} - q_t \geq 2^{-n-b}$.

Finally, a random

Theorem (Chaitin)

Ω is Martin-Löf random.

Proof.

Fix $q_0, q_1, \ldots$ computable, increasing, converging to Ω. Fix $V_0, V_1, \ldots$ the universal Martin-Löf test.

Let b be as in the previous discussion.

When we see some $[\tau] \subseteq V_b$ containing the current q_s, we trigger an increasing of at least $2^{-|\tau|-b-1}$. This moves some q_t beyond $[\tau]$.

By topological considerations, $\Omega \notin V_b$. $\square$

It's powerful

Theorem (Calude and Nies)

Ω *computes every c.e. set.*

Proof.

Fix a c.e. set A.

If we see n enter A at stage s, trigger an increase of at least $\epsilon 2^{-n}$.

With oracle Ω, to decide if $n \in A$, wait until $\Omega - q_s < \epsilon 2^{-n}$. If n hasn't entered A by stage s, it never will. $\square$