Randomness interacts with computable and polynomial time analysis

André Nies

The University of Auckland

January 25, 2012

Main thesis

- Martingales are related to nondecreasing functions on the unit interval.
- Concepts about martingales have corresponding analytical notions.
- The martingale concepts are the restrictions of these analytical notions to dyadic rationals.

“Almost everywhere” theorems

Several theorems in real analysis make a statement that holds at almost every real. Often they state that a function of a certain type is well-behaved at almost every input.

Theorem (Lebesgue, 1904)

Let $f : [0,1] \to \mathbb{R}$ be non-decreasing.
Then the derivative $f'(z)$ exists for almost every real z,
that is, with (uniform) probability 1.

Denjoy alternative

For a function f, define the Dini derivatives by

$$\begin{aligned} \overline{D}f(x) &= \limsup_{h\to 0} (f(x+h)-f(x))/h \\ \underline{D}f(x) &= \liminf_{h\to 0} (f(x+h)-f(x))/h \end{aligned}$$

The function f is differentiable at x iff $\underline{D}f(x) = \overline{D}f(x)$ and this value is finite.

Let λ denote Lebesgue measure.

Theorem (Denjoy-Young-Saks; strengthens Lebesgue)

Let f be an arbitrary function $[0,1] \to \mathbb{R}$. Then λ-almost surely,

either $f'(x)$ exists,

or $\overline{D}f(x) = \infty$ and $\underline{D}f(x) = -\infty$.

Rademacher's theorem

Theorem (Rademacher)

Let $f : [0,1]^n \to \mathbb{R}$ be Lipschitz.
Then the derivative $Df(z)$ (an element of $\mathbb{R}^n$) exists for almost every vector $z \in [0,1]^n$.

Ergodic theory

Definition

A measurable operator T on a probability space $(M, \mathcal{A}, \mu)$ is called ergodic if for each $X \in \mathcal{A}$,

- $\mu(X) = \mu(T^{-1}(X))$, and
- $T^{-1}(X) = X$ implies $\mu(X) = 0$ or $\mu(X) = 1$.

Examples: Cantor space $\{0,1\}^{\mathbb{N}}$, T is shift map.
Unit interval, $T(x) =$ fractional part of $x + \alpha$, where $\alpha > 0$ is irrational.

Theorem (Ergodic Theorem; Birkhoff, 1932)

Let $f \in L^1(\mu)$. Then for almost every $x \in M$, the "time average" $\frac{1}{N}\sum_{i=0}^{N-1} f \circ T^i(x)$ converges to the "space average" $\int f d\mu$.

Connection to computable analysis, and algorithmic randomness

For each theorem of this kind:

- Find a framework in which the given objects are computable.
- Now the "almost everywhere" property may correspond to an algorithmic randomness notion. Try to figure out which one.

- For Lebesgue's theorem, the notion is "computable randomness", which is based on effective betting strategies.
- The same holds for Denjoy-Young-Saks.
- For Rademacher, I conjecture it is computable randomness in $[0, 1]^n$.
- For the Ergodic Theorem, the notion is Schnorr randomness.

Background on computable analysis and algorithmic randomness

Computable reals

In the definition of computable functions, $\mathbb{N}$ can be replaced by domains that are effectively encoded by natural numbers, such as the rationals $\mathbb{Q}$.

- A real $r \in \mathbb{R}$ is computable if there is a computable sequence $(q_n)_{n\in\mathbb{N}}$ of rational numbers such that $|r - q_n| < 2^{-n-1}$ for each n.
- Computable reals are $\sqrt{2}, \pi, e, \ldots$
- To define a non-computable real, one needs computability theory. Examples of such reals are
  - $\sum_{n\in\mathcal{H}} 2^{-n}$, where $\mathcal{H}$ is the halting problem
  - Chaitin's Ω.

Computable functions on the unit interval

Definition

We say that a function $f\colon [0,1] \to \mathbb{R}$ is computable if

(a) For each rational $q \in [0,1]$, the real $f(q)$ is computable uniformly in q.

(b) f is effectively uniformly continuous: for input a rational $\epsilon > 0$ we can compute a rational $\delta > 0$ such that

$$|x - y| < \delta \text{ implies } |f(x) - f(y)| < \epsilon.$$

In general, the condition (a) by itself is too weak. However,

- if a nondecreasing function f satisfies (a) and is continuous, then it is already computable.
- For a Lipschitz function f, (a) is also sufficient.

For instance, the functions e^x, and $\sqrt{x}$, and $\sin x$ are computable.

Randomness via betting strategies

Computable betting strategies are certain computable functions M from binary strings to the non-negative reals.

- Let Z be a sequence of bits. When the player has seen the string σ of the first n bits of Z, she can make a bet q, where $0 \leq q \leq M(\sigma)$, on what the next bit $Z(n)$ is.
- If she is right, she gets q. Otherwise she loses q. Thus, we have
  $$M(\sigma 0) + M(\sigma 1) = 2M(\sigma)$$
  for each string σ.
- She wins on Z if M is unbounded along Z. We call a set Z computably random if no computable betting strategy wins on Z.

Martin-Löf random $\Rightarrow$ computably random, but not conversely.

Effective versions of almost everywhere theorems

Effective form of 's theorem

Theorem (Brattka, Miller, N; submitted)

Let $f : [0,1] \to \mathbb{R}$ be non-decreasing and computable.
Then $f'(z)$ exists for every computably random real z.

If we merely assume that f has bounded variation, then $f'(z)$ exists for each Martin-Löf random real z (Demuth, 1975).

In BMN (submitted) we obtained a new proof of this, combining our result above with Jordan (f of bounded variation is the difference of two nondecreasing functions).

I don't know how Demuth proved it. He may have used "Bauer's thesis".

Andrej Bauer's thesis

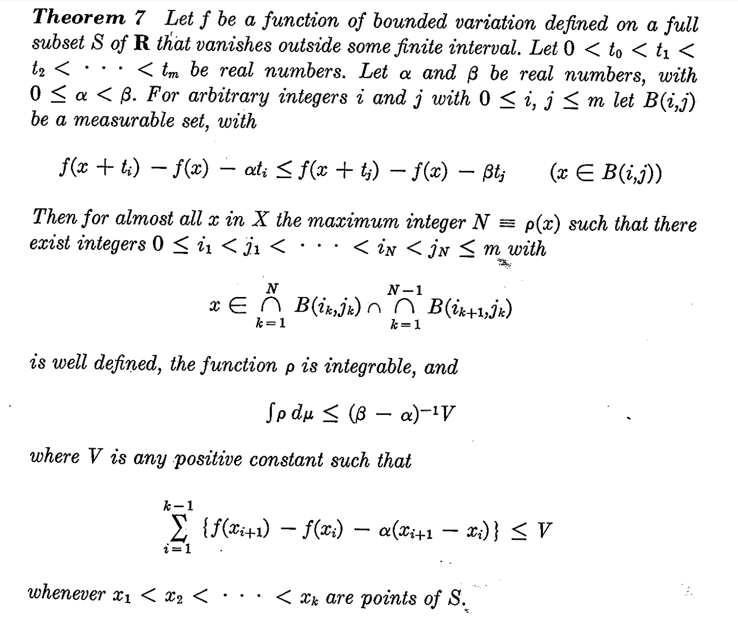
Computable mathematics is the “realizability interpretation” (Kleene) of constructive mathematics.

According to this, you can take the proof of any theorem from Bishop's 1967 book (Foundations of constructive Analysis) and use it to prove a theorem in computable analysis.

V'yugin (1997/1998) used Bishop's proof of the Birkhoff ergodic theorem to show that it holds at each ML-random. (This was pointed out to me by Jason Rute.)

Maybe Demuth did the same to prove his theorem?

The theorem from Bishop's book



Theorem 7 *Let f be a function of bounded variation defined on a full subset S of $\mathbf{R}$ that vanishes outside some finite interval. Let $0 < t_0 < t_1 < t_2 < \cdots < t_m$ be real numbers. Let α and β be real numbers, with $0 \leq \alpha < \beta$. For arbitrary integers i and j with $0 \leq i, j \leq m$ let $B(i,j)$ be a measurable set, with*

$$f(x + t_i) - f(x) - \alpha t_i \leq f(x + t_j) - f(x) - \beta t_j \qquad (x \in B(i,j))$$

Then for almost all x in X the maximum integer $N \equiv \rho(x)$ such that there exist integers $0 \leq i_1 < j_1 < \cdots < i_N < j_N \leq m$ with

$$x \in \bigcap_{k=1}^{N} B(i_k,j_k) \cap \bigcap_{k=1}^{N-1} B(i_{k+1},j_k)$$

is well defined, the function ρ is integrable, and

$$\int \rho \, d\mu \leq (\beta - \alpha)^{-1} V$$

where V is any positive constant such that

$$\sum_{i=1}^{k-1} \{f(x_{i+1}) - f(x_i) - \alpha(x_{i+1} - x_i)\} \leq V$$

whenever $x_1 < x_2 < \cdots < x_k$ are points of S.

Proving the effective form of 's theorem

Theorem (Brattka, Miller, N; submitted)

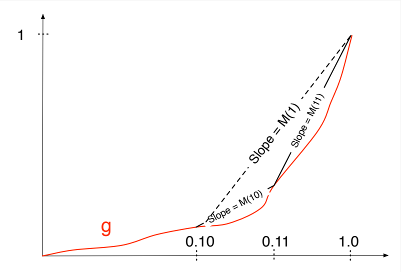
Let $f : [0,1] \to \mathbb{R}$ be non-decreasing and computable.
Then $f'(z)$ exists for every computably random real z.

Turning nondecreasing functions into martingales

For the simplest case suppose that $\overline{D}g(z) = \infty$ for g computable nondecreasing. Then martingale M succeeds on z, where for a string σ, we let

$$M(\sigma) = \frac{g(0.\sigma + 2^{-|\sigma|}) - g(0.\sigma)}{2^{-|\sigma|}}.$$

Thus $M(\sigma)$ is the slope of g between the points $0.\sigma$ and $0.\sigma + 2^{-|\sigma|}$. It is clear that this is a martingale. For instance, the following shows $2M(1) = M(10) + M(11)$.



To prove the full result is much harder

It may happen that $f'(z)$ fails to exist, but the dyadic slope M_f doesn't "notice".

Solution: use finitely many computable martingales. How many depends on $\overline{D}f(z)/\underline{D}f(z)$.

- Shifts: $M_{f,\alpha}(\sigma) =$ slope at interval $\alpha + [0.\sigma, 0.\sigma + 2^{-|\sigma|}]$, $\alpha \in \mathbb{Q}$ fixed
- we also need scaling of intervals by a fixed rational $\delta \in [1/2, 1]$.

We think the scaling is necessary in general. However, sometimes shifting definitely suffices:

Theorem (N and Solecki)

*Suppose f is a nondecreasing function which is not differentiable at $z \in [0,1]$. Suppose that $\underline{D}f(z) = 0$, or $\overline{D}f(z)/\underline{D}f(z) > 72$.
Then one of the martingales M_f, $M_{f,1/3}$ does not converge on the binary expansion of z.*

Converse of the effective form of Lebesgue's theorem

.

Theorem (BMN, submitted)

If z is not computably random, then there is a computable nondecreasing function g such that $g'(z)$ does not exist.

Sketch the reason why.

Turning martingales into nondecreasing functions

- Let M be a computable betting strategy (also called a computable martingale).
- Let μ be the measure induced by M. It is determined by its values on the basic clopen sets: $\mu([\sigma]) = M(\sigma)2^{-|\sigma|}$.
- Let
  $$g_M(x) = \mu[0, x).$$
  Then the function g_M is nondecreasing, and $g_M(q)$ is uniformly computable in a dyadic rational q.
- We have $M(\sigma) = (g_M(0.\sigma 1) - g_M(0.\sigma))/2^{-|\sigma|}$. So if M succeeds on a real z and μ_M is atomless, we have $\overline{D}g_M(z) = \infty$.

Now suppose z is not computably random. Then some computable martingale M with μ_M atomless succeeds on z.
In this case g_M is continuous, and hence computable. And $g'_M(z)$ does not exist because $\overline{D}g_M(z) = \infty$.

Computable randomness and differentiability: Lipschitz functions

Different classes of functions can describe the same randomness notion: instead of monotone, we can take Lipschitz functions.

Theorem (Freer, Kjos-Hanssen, N: in prep)

A real z is computably random $\iff$ each computable Lipschitz function $f\colon [0,1] \to \mathbb{R}$ is differentiable at z.

$\Rightarrow$ follows from the BMN result because $f(x) + cx$ is nondecreasing computable, where $c \in \mathbb{N}$ is a Lipschitz constant.

For $\Leftarrow$ we have to work harder than in the previous case: if computable martingale M succeeds on the binary expansion z, first change it into a bounded computable martingale N that oscillates between 1 and 2 along z. Then g_N is Lipschitz and $g'_N(z)$ doesn't exist.

Computable randomness and the Denjoy alternative

Recall that f satisfies the Denjoy alternative at x if

$$\text{either } f'(x) \text{ exists,}$$

$$\text{or } \overline{D}f(x) = \infty \text{ and } \underline{D}f(x) = -\infty.$$

Theorem (Bienvenu, Hoelzl, Miller, Nies, STACS 2012)

A real z is computably random $\Longleftrightarrow$ *each computable function $f\colon [0,1] \to \mathbb{R}$ satisfies the Denjoy alternative at z.*

$\Leftarrow$ follows from BMN result.

$\Rightarrow$ uses a result of Demuth I don't know how to prove at present.

Summary of correspondence between martingales and nondecreasing functions (1)

Given martingale M

$g_M(x)$ is $\mu_M[0,x)$ where μ_M is the measure on $[0,1]$ corresponding to M.

M_g is the slope of g at dyadic rationals:

$$M_g(\sigma) = \frac{g(0.\sigma + 2^{-|\sigma|}) - g(0.\sigma)}{2^{-|\sigma|}}.$$

Given nondecreasing function g

Correspondence between martingales and nondecreasing functions (2)

M succeeds on binary expansion of *z*	$\overline{D}g(z) = \infty$
M oscillates (i.e., fails to converge) on binary expansion of *z*, but doesn't succeed	$\underline{D}g(z) < \overline{D}g(z) < \infty$.

BUT, remember that functions are the more complex objects because the martingales only look at dyadic rationals. We need finitely many martingales (with shift + scaling) to detect non-differentiablity of the function.

Other models of computation

- Polynomial time: from string σ can compute the rational $M(\sigma)$ in poly. time
- computable: from string σ can compute the real $M(\sigma)$
- hyperarithmetical, or Δ^1_1: from string σ can compute an index for $M(\sigma)$ as a Δ^1_1 real.

Definition

A real z is BLA random (where BLA $\in \{$ polytime, computable, Δ^1_1 $\}$) if no BLA martingale succeeds on its binary expansion.

What's known on poly time random reals?

- Can be computable; in fact exists in all time classes properly containing P
- Incomparable with Schnorr randomness!
- Has nice statistical properties, such as absolutely normal.

Effectiveness notions for functions $g\colon [0,1] \to \mathbb{R}$

We can define effectiveness notions similar to the ones for martingales, for functions g.

- Represent a real x by a Cauchy name $(p_i)_{i\in\mathbb{N}}$. $p_{ß\in\mathbb{Q}}$, and $\forall k > i |p_i - p_k| \le 2^{-i}$.
- g is { polytime, computable, Δ^1_1 } if there is a procedure at the right level turning every Cauchy name for x into a Cauchy name for $g(x)$.
- Same notion of "computable" as before.
- For poly time, need to restrict to a compact set of Cauchy names (signed digit representation does it).

Functions like e^x, x^2, $\sin x$ are polynomial time.
Δ^1_1 means number of steps is a recursive ordinal, such as ω. This gives time to look at the whole real x to see whether $x \ge 1/2$. So Δ^1_1 functions can be discontinuous.

Tests-to-functions, again

Recall that for a martingale M, μ_M is the measure, and $g_M(x) = \mu_M[0, x)$. M has the savings property if $M(\sigma) \geq M(\tau) - 2$ whenever $\sigma \succeq \tau$. This implies $M(\sigma) = O(|\sigma|)$ so M grows slowly. In particular, μ_M has no atoms.

- If M is poly time and has savings property, then g_M is poly time.
- If M is computable and μ_M atomless then g_M is computable.
- If M is Δ^1_1 then so is g.

Characterization via the lower Dini derivative

Theorem (N + others)

A real z is NOT $\{$ polytime, computable, Δ^1_1 $\}$ random $\Leftrightarrow$ some nondecreasing continous function g at the same level of effectivity satisfies $\underline{D}g(z) = +\infty$.

We can develop the theory of martingales with bases other than 2. We get the same connections with nondecreasing functions. Since the right side of the theorem is base invariant, we obtain

Corollary

$\{$ polytime, computable, Δ^1_1 $\}$ randomness of a real is base invariant.

Figueira + student Javier have direct proof for polytime.

Questions on polytime randomness

Let z be a polynomial time random real.

- Does $f'(z)$ exist for each nondecreasing polytime computable f?
- Easier: does $f'(z)$ exist for each Lipschitz polytime computable f?
- Does the Denjoy alternative hold for each polytime computable f?