Computable analysis of martingales and related measure-theoretic topics, with an emphasis on algorithmic randomness

Jason Rute

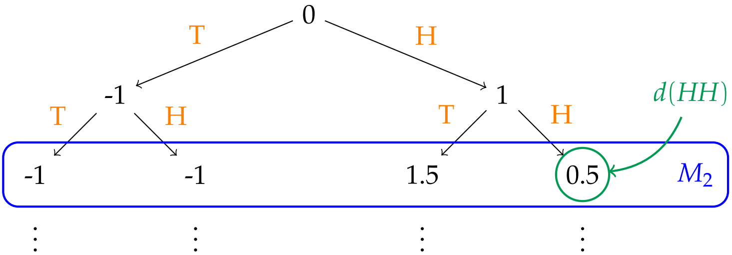
Carnegie Mellon University

PhD Defense
August, 8 2013

What is a martingale?

Informally: A martingale is a **gambling strategy**.

Example: Bet on a fair coin.



For logicians: If $d(\sigma)$ is a martingale in the computability theory sense and $x \in 2^{\mathbb{N}}$, then we are considering

$$M_n(x) = d(x \upharpoonright n).$$

What is a martingale in general?

Formally:

A martingale is a sequence of integrable functions $M_n \colon (\Omega, P) \to \mathbb{R}$ such that

$$\mathbb{E}[M_{n+1} \mid M_0, M_1, \ldots, M_n] = M_n.$$

Slogan:

The expectation of the future is the present (conditioned on the past & present).

An a.e. convergence theorem

Example (Doob's martingale convergence theorem)

Let (M_n) be a martingale. Assume $\sup_n \|M_n\|_{L^1} < \infty$. Then M_n converges a.e. In particular, $\lim_n M_n < \infty$ a.s.

4 Questions

1. (Computable analysis) Is the rate of convergence of (M_n) computable (from the martingale (M_n))?
2. (Computable analysis) If not, what other information is needed to compute a rate of convergence?
3. (Algorithmic randomness) For which points x does $M_n(x)$ converge for all "computable" L^1-bounded martingales?
4. (Algorithmic randomness) Which assumptions are needed to characterize convergence on (insert favorite randomness notion)?

Computable analysis

Computable reals

A real number a is **computable** if it can be effectively approximated by rationals.

Example

π is computable. We can compute a sequence of rationals q_n such that $|q_n - \pi| \leqslant 2^{-n}$. (This sequence is called a **name** for x.)

Computable points, functions

This definition extends to any complete separable metric space with a "nice" countable set of simple points $\{q_n\}_n$.

Examples

Computable L^1-functions, computable sequences of reals, computable sequences of L^1-functions, computable Borel measures, etc.

Computable analysis

Computable maps

We say that y is computable from x if there is an algorithm which computes a name for y uniformly from a name for x, more formally

- takes in $\varepsilon > 0$
- keeps reading the name for x: $q_0, q_1, q_2, \ldots$
- when it has a close enough approximation q_n,
  it returns r_n such that $|r_n - y| \leqslant \varepsilon$.

Note

Total computable maps are continuous. Most continuous maps in practice are computable.

Question 1

Question 1.

Is the rate of a.e. convergence computable?

A **rate of a.e. convergence** is some $n(\varepsilon,\delta)$ such that

$$\mu\left\{x \,\middle|\, \sup_{n \geqslant n(\varepsilon,\delta)} |f_n(x) - f_\infty(x)| \geqslant \varepsilon\right\} \leqslant \delta,$$

i.e. $f_n \to f_\infty$ with an ε-uniform rate of convergence outside a set of measure $\leqslant \delta$.

Theorem

The rate of a.e. convergence of a martingale M_n is not necessarily computable.

Proof sketch.

Code in the halting problem. Enumerate the programs $\{e_n\}$ that halt and bet 3^{-e_n} dollars that the e_nth program halts. Any rate of convergence would compute the halting problem.

□

Question 2

Question 2.

What other information is needed to compute a rate of convergence?

Theorem (R.)

The rate of convergence of $M_n \to M_\infty$ is computable uniformly from

- (M_n) *(as a sequence of L^1 functions),*
- M_∞ *(as an L^1 function), and*
- $\sup_n \|M_n\|_{L^1}$.

Note

This is not just because we know the limit M_∞. Without $\sup_n \|M\|_{L^1}$ being computable, the limit could be 0 but the rate of convergence not computable.

Question 2

Question 2.

What other information is needed to compute a rate of convergence?

Theorem (R.)

The rate of convergence of $M_n \to M_\infty$ is computable uniformly from

- (M_n) *(as a sequence of L^1 functions),*
- M_∞ *(as an L^1 function), and*
- $\sup_n \|M_n\|_{L^1}$.

Theorem (R.)

The rate of convergence (a.e. and L^2) of $M_n \to M_\infty$ is computable uniformly from

- (M_n) *(as a sequence of L^2 functions) and*
- $\sup_n \|M_n\|_{L^2} = \|M_\infty\|_{L^2}$.

Question 3

Question 3.

For which points x does the following hold?

$$M_n(x) \text{ converges for all "computable" } L^1\text{-bounded martingales } (M_n)$$

Theorem

Every L^1-bounded martingale converges on almost every point.

Corollary

For almost every point, every "computable" L^1-bdd martingales converges.

There are countably many "computable" martingales.

Question

What is this measure-one set of points?

Algorithmic randomness

- π
  0010010000111111011010101000100010000101101000110000100011010011...
- `random.org`
  1010110101001101011100001001101001010010100001000000010000111100...

Are either of these random? How can we check?

1. Are they normal?
2. Do they satisfy the law of the iterated logarithm?
3. Is the number not π?
4. It did not come from `random.org`?

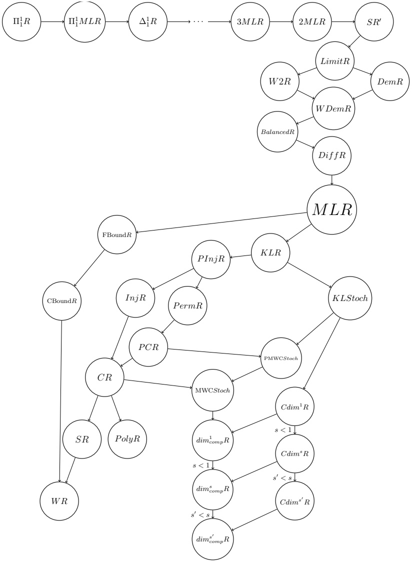
Wait!

Is any bit sequence random? They all fail some statistical test!

A bit sequence is **"algorithmically random"** if it passes all **"computable" statistical tests.**

The randomness zoo

- There are many randomness notions.
- Most start out on $2^{\mathbb{N}}$ (coin flipping).
- Some of the more natural ones are:
  - 2-randomness
  - Weak 2-randomness
  - Difference randomness
  - **Martin-Löf randomness**
  - **Computable randomness**
  - **Schnorr randomness**
  - Kurtz randomness
- The natural ones have connections with computable analysis.
- The natural ones can be extended to other computable probability spaces.



Randomness Zoo (Antoine Taveneaux)

Martin-Löf and Schnorr randomness

Definition

- A **Martin-Löf test** is a computable sequence (U_n) of effectively open sets (uniform sequence of Σ^0_1 sets) such that $\mu(U_n) \leqslant 2^{-n}$.
- A **Schnorr test** is a Martin-Löf test, where $\mu(U_n)$ is uniformly computable.
- x is **Martin-Löf/Schnorr random** (for the measure μ) if $x \notin \bigcap_n U_n$ for each ML/Schnorr tests.

Question 3 Again

Question 3.

For which points x does the following hold?

$$M_n(x) \text{ converges for all computable } L^1\text{-bounded martingales } (M_n)$$

Theorem (Takahashi; Merkle-Mihalovic-Slaman; Dean; R.)

The answer is Martin-Löf randomness, even if the martingales are dyadic or nonnegative (but not both).

Corollary

Doob's martingale convergence theorem characterizes Martin-Löf randomness!

Question 4

Question 4.

Which assumptions characterize convergence on **Schnorr randoms**?

Lemma (R.)

If (f_n) and f are L^1-computable, and $f_n \to f$ effectively a.e., then $\widetilde{f}_n(x) \to \widetilde{f}(x)$ on Schnorr randoms x.

Here

$$\widetilde{f}(x) = \lim_n p_n(x)$$

where (p_n) is a sequence of simple functions $\|f - p_n\|_{L^1} < 2^{-n}$. (We need this since f is an equivalence class.)

Question 4

Question 4.

Which assumptions characterize convergence on **Schnorr randoms**?

Lemma (R.)

If (f_n) and f are L^1-computable, and $f_n \to f$ effectively a.e., then $\widetilde{f}_n(x) \to \widetilde{f}(x)$ on Schnorr randoms x.

Theorem (R.)

Assume (M_n) is L^1-computable, M_∞ is L^1-computable, and $\sup_n \|M_n\|_{L^1}$ is computable. Then

$$M_n \to M_\infty \quad \textit{effectively a.e.}$$

*Therefore **(for free!)***

$$\widetilde{M}_n(x) \to \widetilde{M}_\infty(x) \quad \textit{on Schnorr randoms } x.$$

Question 4

Question 4.

Which assumptions characterize convergence on **Schnorr randoms**?

Lemma (R.)

If (f_n) and f are L^1-computable, and $f_n \to f$ effectively a.e.,
then $\widetilde{f}_n(x) \to \widetilde{f}(x)$ on Schnorr randoms x.

Theorem (R.)

Assume (M_n) is L^1-computable, M_∞ is L^1-computable, and $\sup_n \|M_n\|_{L^1}$ is computable. Then $\widetilde{M}_n(x) \to \widetilde{M}_\infty(x)$ on Schnorr randoms x.

Theorem (R.)

If x is not Schnorr random, there is some L^1-computable martingale (M_n), with an L^1-computable limit M_∞, and $\sup_n \|M_n\|_{L^1} = 1$ such that

$$\lim_n M_n(x) = \infty.$$

Martingale convergence

Theorem

Assume (M_n) is L^1-computable, M_∞ is L^1-computable, and $\sup_n \|M_n\|_{L^1}$ is computable. Then

$$M_n \to M_\infty \quad \textit{effectively a.e.}$$

Hence,

$$\widetilde{M}_n(x) \to \widetilde{M}_\infty(x) \quad \textit{on Schnorr randoms } x.$$

Proof sketch.

Decompose $M_n = N_n + L_n$ where

- N_n converges in L^1 and
- L_n converges to 0.

Handle each case individually.

□

Martingale convergence in L^1

Theorem

Assume (M_n) is L^1-computable, M_∞ is L^1-computable and $M_n \to M_\infty$ in L^1. Then

$$M_n \to M_\infty \quad \textit{effectively a.e. and in } L^1.$$

Proof sketch.

- Fix k, and find n_k such that $\|M_\infty - M_{n_k}\|_{L^1} \leqslant 2^{-2k}$.
- Facts: $\|M_n - M_{n_k}\|_{L^1}$ is increasing and $(M_n - M_{n_k}) \xrightarrow{L^1} (M_\infty - M_{n_k})$.
- $M_n \to M_\infty$ effectively in L^1 since

$$\forall n \geqslant n_k \; \|M_n - M_{n_k}\|_{L^1} \leqslant \|M_\infty - M_{n_k}\|_{L^1} \leqslant 2^{-2k}.$$

- $M_n \to M_\infty$ effectively a.e. since (by Kolmogorov's inequality)

$$\mu\left\{\sup_n |M_n - M_{n_k}| \geqslant 2^{-k}\right\} \leqslant \frac{\sup_n \|M_n - M_{n_k}\|_{L^1}}{2^{-k}} \leqslant \frac{2^{-2k}}{2^{-k}} \leqslant 2^{-k}.$$

Similar results

All these theorems can be used to characterize Schnorr randomness

- Differentiability of bounded variation functions
- "Lebesgue differentiation theorem" for signed measures
- Ergodic theorem (Avigad-Gerhardy-Towsner; Gács-Hoyrup-Rojas; Galatalo-Hoyrup-Rojas)
- Sub/supermartingale convergence theorem (nonnegative)
- Backwards martingale convergence theorem
- Monotone convergence theorem
- Strong law of large numbers
- De Finetti's theorem

An observation

Observation

In most common a.e. convergence theorems, the rate of convergence is computable from

- the sequence (f_n),
- the limit f_∞, and
- the bounds $\sup_n \|f_n\|_{L^1}$ and $\inf_n \|f_n\|_{L^1}$.

- There are easy, but contrived, counterexamples.
- Can this observation be made into a theorem with a few more assumptions?

Sub/supermartingales

What about sub/supermartingales? This is one of the only cases I have not been able to work out. It is also one of the only cases where $\|f_n\|_{L^1}$ is not monotone (or nearly monotone).

Lebesgue differentiation theorem

Theorem (R.; Pathak-Rojas-Simpson)

Assume f is L^1-computable on $[0,1]$. Then

$$\frac{1}{2r}\int_{x-r}^{x+r} f(x)\,dx \xrightarrow[r\to\infty]{} f(x) \quad \text{effectively a.e.}$$

and

$$\frac{1}{2r}\int_{x-r}^{x+r} f(x)\,dx \xrightarrow[r\to\infty]{} \widetilde{f}(x) \quad \text{on Schnorr random } x.$$

Theorem (R.; Pathak-Rojas-Simpson)

If x is not Schnorr random, there is some L^1-computable f such that

$$\frac{1}{2r}\int_{x-r}^{x+r} f(y)\,dy \to \infty.$$

Backwards martingales with an application

Theorem (R.)

*Assume (M_{-n}) is an L^1-computable **backwards martingale**, $M_{-\infty}$ is L^1-comp.*
Then $M_{-n} \to M_{-\infty}$ effectively a.e. and in L^1.
Hence $\widetilde{M}_{-n} \to \widetilde{M}_{-\infty}$ on Schnorr random x.

Corollary (Variation on Kučera's theorem, R.)

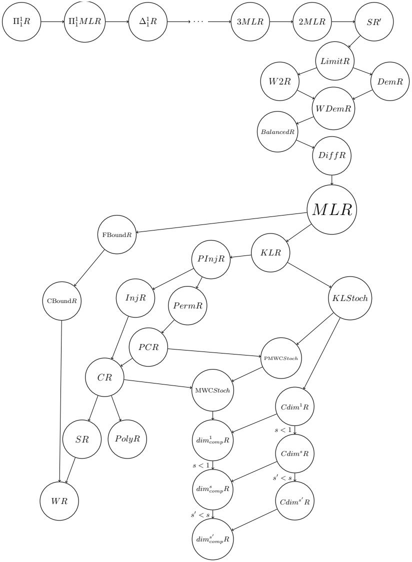
Let C be a closed set of positive computable measure $\mu(C)$. Let x be Schnorr random. There is some $y \in C$ such that y is the same as x but with finitely many bits permuted.

Proof sketch.

- Let M_{-n} be the average of $\mathbf{1}_C$ under all permutations of the first n bits.
- It turns out M_{-n} is a reverse martingale with limit $\mu(C)$.
- Then $M_{-n}(x) \to \mu(C)$ by the above theorem.
- Hence $M_{-n}(x) > 0$ for some n.
- Hence some $y \in C$ where y is a permutation of first n bits of x.

The randomness zoo

- There are many randomness notions.
- Most start out on $2^{\mathbb{N}}$ (coin flipping).
- Some of the more natural ones are:
  - 2-randomness
  - Weak 2-randomness
  - Difference randomness
  - Martin-Löf randomness
  - **Computable randomness**
  - Schnorr randomness
  - Kurtz randomness
- The natural ones have connections with computable analysis.
- **The natural ones can be extended to other computable probability spaces.**



Randomness Zoo (Antoine Taveneaux)

Computable randomness

Definition

A **test for computable randomness** is a nonnegative dyadic martingale $M\colon 2^{<\omega} \to \mathbb{R}_+$ such that

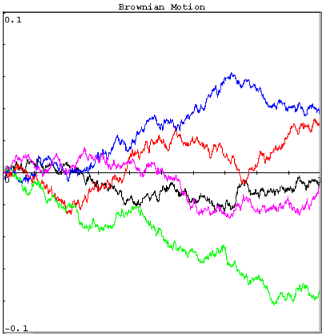
$$\mu(\sigma 0)M(\sigma 0)+\mu(\sigma 1)M(\sigma 1)=\mu(\sigma)M(\sigma)$$

and $M(\sigma)$ is computable from σ, provided that $\mu(\sigma)>0$.

Definition

We say that $x \in (2^{\mathbb{N}}, \mu)$ is **computably random** if $\limsup_n M(x \upharpoonright n) < \infty$ for all martingale tests M.

Computably random Brownian motion?



- B a "computably random Brownian motion".
- Are these also computably random?
  - $B(1)$ (Gaussian distribution)
  - Last hitting time before $t = 1$ (arcsin distribution)
  - Maximum/minimum values for $t \in [0,1]$
  - argmax/argmin
  - set of zeros $\{t : B(t) = 0\}$
- **We need a good definition of computable randomness** for Brownian motion and for reals (with, say, Gaussian distribution).
- **Which maps preserve computable randomness?** For example $B \mapsto B(1)$?

Computable randomness on $[0,1]$.

Base invariance

Say that x is "random" on $[0,1]$ (with Lebesgue measure) if

- its binary digits are "random" on $2^{\mathbb{N}}$.
- its decimal digits are "random" on $10^{\mathbb{N}}$.

Are these the same? What about other bases?

Easy

2-randomness, weak 2-randomness, difference randomness, Martin-Löf randomness, Schnorr randomness, and Kurtz randomness are base invarient!

Brattka, Miller, Nies; Silveira

Computable randomness is base invariant.

The proofs for comp. randomness are not trivial. (The Brattka, Miller, Nies proof uses differentiability and does not even work in multiple dimensions.)

Computable randomness on Polish space $\mathbb{X}$.

- Let $\mathbb{X}$ be a computable Polish metric space.
- Let μ be a computable measure on $\mathbb{X}$.
- Break up $\mathbb{X}$ into cells (Gács; Hoyrup-Rojas; Bosserhoff)
- **Now $(\mathbb{X}, \mu)$ looks like a measure $(2^{\mathbb{N}}, \nu)$ on Cantor space**
- Say $x \in \mathbb{X}$ is computably random if the corresponding point in $(2^{\mathbb{N}}, \nu)$ is computably random.

Theorem (R.)

It does not matter how we break up $\mathbb{X}$ into cells.

Example

We then have computably random Brownian motion, Gaussian distributed reals, etc.

Computing randoms from randoms

Almost-everwhere computable maps

Let $T : (2^{\mathbb{N}}, \mu_1) \to (2^{\mathbb{N}}, \mu_2)$ come from an algorithm which

1. Takes in a sequence of coin flips with distribution μ_1.
2. Outputs a sequence of coin flips with distribution μ_2.
3. Almost every input has an output.

However, there is a problem with computable randomness!

Theorem (Bienvenu-Porter; R.)

There is an a.e. computable map $T : (2^{\mathbb{N}}, \lambda) \to (2^{\mathbb{N}}, \lambda)$ where x is computably random, but $T(x)$ is not computably random!

Preservation of computable randomness

Theorem (R.)

Assume $T : (2^{\mathbb{N}}, \mu_1) \to (2^{\mathbb{N}}, \mu_2)$ and $T^{-1} : (2^{\mathbb{N}}, \mu_2) \to (2^{\mathbb{N}}, \mu_1)$ are a.e. computable such that $T \circ T^{-1} = id$ and $T^{-1} \circ T = id$. Then T preserves computable randomness.

Proof sketch.

- Take $y \in (2^{\mathbb{N}}, \mu_2)$ **not** computably random.
- There is a martingale M which succeeds on y ($\lim_n M_2(y \upharpoonright n) = \infty$).
- Slow down the martingale by saving some of your money (savings trick).
- This gives an **absolutely continuous** measure $\nu_2(\sigma) = M_2(\sigma)\mu_2(\sigma)$.
- **Since $\nu_2 \ll \mu_2$, then T^{-1} is a.e. computable on ν_2.**
- Let $\nu_1 = \nu_{2*}T^{-1}$ (pushforward of ν_2 along T^{-1}).
- **Since T^{-1} is ν_2-a.e. comp., then ν_1 is comp.**
- Let $M_1(\sigma) = \frac{\nu_1(\sigma)}{\mu_1(\sigma)}$ and $x = T^{-1}(y)$.
- It can be shown that $M_1(x \upharpoonright n) \to \infty$. So x is not computably random.

□

More recent work (not in the dissertation)

Theorem (R.)

Assume $T:(\mathbb{X},\mu_1)\to(\mathbb{Y},\mu_2)$ is ***effectively measurable*** *with a* ***"computable" conditional probability*** $\mu_1[\,\cdot\mid T]$*. Then $\widetilde{T}(x)$ preserves computable randomness.*

Theorem (R.)

TFAE:

- $(\mathbb{X},\mu)$ *is computable.*
- *There exists a map $T:(2^{\mathbb{N}},\lambda)\to(\mathbb{X},\mu)$ as above.*

TFAE:

- *x is computably random on $(\mathbb{X},\mu)$.*
- *$x=\widetilde{T}(\omega)$ for some computably random $x\in(2^{\mathbb{N}},\lambda)$.*

Theorem (R.)

If $(M_n,\mathcal{F}_n)$ is an L^1-comp. martingale, $\sup_n\|M_n\|_{L^1}$ is comp., and $f\mapsto\mathbb{E}[f\mid\mathcal{F}_\infty]$ is $(L^1\to L^1)$-computable, then $\widetilde{M}_n(x)$ converges on computable randoms.

Thank You!

These slides will be available on my webpage:

math.cmu.edu/~jrute

Or just Google™ me, “Jason Rute”.