Randomness and the Lebesgue Differentiation Theorem

Jason Rute

Department of Mathematical Science
Carnegie Mellon University

Southern Wisconsin Logic Colloquium
May 2011

Outline

The Lebesgue Differentiation Theorem

Theorem (Lebesgue Differentiation Theorem)

Let m be the Lebesgue measure on $[0,1]^n$. Let $f : [0,1]^n \to \mathbb{R}$ be an integrable (L^1) function. Define the average

$$A_r f(x) = \frac{\int_{B_r(x)} f(x)\,dx}{m(B_r(x))}.$$

Then $A_r f(x) \to f(x)$ as $r \to 0$ for almost every x.

- "Almost every x" intuitively means if x is "random enough" (relative to the parameters of the theorem) then the above theorem is true for x.
- Algorithmic randomness attempts to make rigorous this notion of "random".

An Effective Lebesgue Differentiation Theorem

Theorem (Lebesgue Differentiation Theorem)

Let m be the Lebesgue measure on $[0,1]^n$. Let $f : [0,1]^n \to \mathbb{R}$ be an integrable (L^1) function. Define the average

$$A_r f(x) = \frac{\int_{B_r(x)} f(x)\, dx}{m(B_r(x))}.$$

Then $A_r f(x) \to f(x)$ as $r \to 0$ for almost every x.

Theorem (Effective Lebesgue Differentiation Theorem, R.)

Let $f : [0,1]^n \to \mathbb{R}$ be an L^1-computable function. Then $A_r f(x)$ converges as $r \to 0$ for all Schnorr randoms x.

Two Effective LDT's

Pathak has already shown the following.

Theorem (Effective LDT for Martin-Löf Randoms, Pathak)

Let $f : [0,1]^n \to \mathbb{R}$ be an L^1-computable function. Then $A_r f(x)$ converges as $r \to 0$ for all Martin-Löf randoms x.

Theorem (Effective LDT for Schnorr Randoms, R.)

Let $f : [0,1]^n \to \mathbb{R}$ be an L^1-computable function. Then $A_r f(x)$ converges as $r \to 0$ for all Schnorr randoms x.

- Since ML randoms $\subset$ Schnorr randoms, the Schnorr result is better.
- Is Schnorr randomness the best we can do? Yes.

Characterization of Schnorr Randomness

Theorem ("The Reversal", R.)

Let $x \in [0,1]^n$ be not Schnorr random. Then there is an L^1-computable function f such that $A_r f(x)$ does not converge as $r \to 0$.

Therefore the Effective LDT and the reversal, together characterize Schnorr randomness by "differentiability". This is closely related to the field of Reverse Mathematics.

Related Work

- Pathak proved the Effective LDT for Martin-Löf randoms.
- Brattka, Miller, Nies characterize computable randomness, Martin-Löf randomness, and Π^0_2-randomness in terms of differentiability of absolutely-continuous computable functions on $[0, 1]$ (actually, functions of bounded variation). Martin-Löf case based on work of Demuth.
- Pathak, Rojas, and Simpson have a different proof of the Effective LDT for Schnorr randoms.
- Kenshi Miyabe also has a similar result.
- Freer, Kjos-Hanssen, Nies characterizes computable randomness and Schnorr randomness in terms of computable Lipschitz functions.
- The proof in this talk is based the tools of the Brattka, Miller, Nies result.

L^1-computable Functions

Any integrable function f can be approximated by a polynomial p_i with rational coefficients such that $\|f - p_i\|_1 = \int |f - p_i| \leq 2^{-i}$. In other words, rational polynomials are dense in L^1.

Definition

For any $f \in L^1$, a code for f is a sequence of rational polynomials (p_i) converging in the L^1-norm to f that are fast Cauchy, i.e. $\|p_{i+1} - p_i\|_1 \leq 2^{-i}$. We say f is L^1-computable if there exists such a computable code.

The following are all computable from (the codes for) $f, g \in L^1$ and (the codes for) any other parameters:

$$f+g, af, \max(f,g), \min(f,g), f^+, f^-, |f|, \|f\|_1, \int f\,dx, \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f\,dx.$$

Note that on $[0,1]$, if f is L^1-computable then $F(x) := \int_0^x f\,dx$ is computable, but the converse does not hold.

Functions of Absolute Continuity

Definitions

Take an a.e. differentiable, continuous function $F : [0,1] \to \mathbb{R}$. We say F is absolutely continuous if $F(x) = \int_0^x f(t)\,dt + F(0)$ where f is the derivative of F, i.e. F "satisfies the Fundamental Theorem of Calculus".

We say F is effectively absolutely continuous if its derivative f is L^1-computable.

effectively absolutely continuous

$\Rightarrow$ absolutely continuous and computable

But the converse doesn't hold.

Theorem (R.)

$x \in [0,1]$ is Schnorr random $\Leftrightarrow$
$A_r f(x)$ converges for all L^1-computable functions f.

Corollary

$x \in [0,1]$ is Schnorr random $\Leftrightarrow$
$F'(x)$ exists for all effectively absolutely continuous functions F.

Proof Sketch.
Let f be the derivative of F, hence $F(x) = \int_0^x f(x)\,dx$. Then

$$F'(x) = \lim_{r\to 0} \frac{F(x+r) - F(x-r)}{2r} = \lim_{r\to 0} \frac{\int_{x-r}^{x+r} f(x)\,dx}{2r} = A_r f(x). \quad \square$$

Translation Between L^1 and Abs. Cont.

Theorem (Brattka-Miller-Nies and Freer-Kjos-Hanssen-Nies)

$x \in [0,1]$ is Martin-Löf random (resp. computably random) $\Leftrightarrow$ $F'(x)$ exists for all absolutely continuous functions F s.t

- *F is computable*
- *(and resp. nondecreasing or Lipschitz).*

Corollary

$x \in [0,1]$ is Martin-Löf random (resp. computably random) $\Leftrightarrow$ $A_r f(x)$ converges for all integrable functions f s.t.

- *$\int_0^x f(t)\,dt$ is computable*
- *(and resp. nonnegative or bounded).*

Schnorr Random

Definitions

A set $U \subseteq [0,1]^n$ is Σ^0_1 (effectively open) if it is a union of a computable sequence of open cells $(a_1, b_1) \times \cdots \times (a_n, b_n)$ with rational endpoints.

A Martin-Löf test is a uniformly computable sequence (U_i) of Σ^0_1 sets such that

$$m(U_i) \leq 2^{-i}.$$

A Schnorr test is a uniformly computable sequence (U_i) of Σ^0_1 sets such that

$$m(U_i) = 2^{-i}.$$

$x \in [0,1]^n$ is Martin-Löf random if $x \notin \bigcap U_i$ for any ML test (U_i).
$x \in [0,1]^n$ is Schnorr random if $x \notin \bigcap U_i$ for any Schnorr test (U_i).

Borel-Cantelli-Solovay Test

Theorem (Borel-Cantelli Lemma)

Let (A_k) be a sequence of sets such that $\sum m(A_k) < \infty$. Then for almost every x, x is in at most finitely many A_k.

Theorem (Solovay Lemma for Martin-Löf Randoms)

Let (U_k) be a uniformly computable sequence of Σ^0_1 sets such that $\sum m(U_k) < \infty$. (Call (U_k) a Solovay test for Martin-Löf randomness.) Then for all Martin-Löf randoms x, x is in at most finitely many U_k. Further, if x is not Martin-Löf random, there is a Solovay test (U_k) such that $x \in U_k$ for infinitely many k.

Borel-Cantelli-Solovay Test

for Schnorr Randomness

Theorem (Borel-Cantelli Lemma)

Let (A_k) be a sequence of sets such that $\sum m(A_k) < \infty$. Then for almost every x, x is in at most finitely many A_k.

Theorem (Solovay Lemma for Schnorr Randoms, Hoyrup-Rojas)

Let (U_k) be a uniformly computable sequence of Σ^0_1 sets such that $\sum m(U_k)$ is finite and computable. (Call (U_k) a Solovay test for Schnorr randomness.) Then for all Schnorr randoms x, x is in at most finitely many U_k. Further, if x is not Schnorr random, there is a Solovay test (U_k) such that $x \in U_k$ for infinitely many k.

The Main Theorems to Prove

Theorem (Effective LDT for Schnorr Randoms)

Let $f : [0,1]^n \to \mathbb{R}$ be an L^1-computable function. Then

$$A_r f(x) := \frac{\int_{B_r(x)} f(x)\,dx}{m(B_r(x))}$$

converges as $r \to 0$ for all Schnorr randoms x.

Theorem (LDT Reversal)

Let $x \in [0,1]^n$ be not Schnorr random. Then there is an L^1-computable function f such that $A_r f(x)$ does not converge as $r \to 0$.

Structure of the Proof

Geometric part of Proof. Reduce the geometric complexity of the problem.

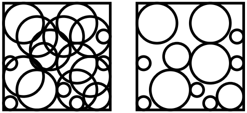
Martingale part of Proof. Use theorems about martingales to analyze the convergence.

Reversal. Construct a martingale that doesn't converge on a Schnorr test and read off an L^1-computable function.

Outline

Geometric Part

- The LDT is a very geometric theorem.
- For example, if the balls were replaced with arbitrary rectangles or ellipse, it would not be true. (But cubes are OK.)
- One approach is to replace overlapping balls with disjoint balls/cubes.



- Classical proofs use the Vitali Covering Theorem to do this.
- Brattka, Miller, Nies use a different approach, which this talk follows.

Work with Cubic Cells Instead

Work with cells instead of balls.

Definition

A closed cell Q on $[0,1]^n$ is a set of the form $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$. A cubic cell is a cell with equal lengths in all dimensions.

It is enough to approximate the balls with cubic cells. We wish to prove this equivalent version of the LDT.

Theorem

Let $f : [0,1]^n \to \mathbb{R}$ be an L^1-computable function. Let x be a Schnorr random. Let $(Q_k)_{k \in \mathbb{N}}$ be a decreasing sequence of cubes which converge to x. (x need not be center of Q_k's.) Then

$$\frac{\int_{Q_k} f(x)\,dx}{m(Q_k)}$$

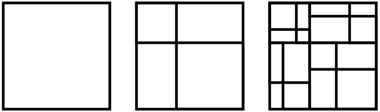
converges as $k \to \infty$.

Filtrations

Technically, a filtration is an increasing sequence of σ-algebras

$$\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \mathscr{F}_2 \subseteq \dots$$

For our purposes, we care only about filtrations made-up of finite partitions of $[0,1]^n$ into cells, e.g.



Call such a filtration of cells computable if it is "uniformly codable". Define $\mathscr{F}_\infty$ to be the minimal σ-algebra containing $\bigcup_i \mathscr{F}_i$. So $(\mathscr{F}_i)$ "converges to" $\mathscr{F}_\infty$.

Conditional Expectation

Let f be an L^1-function. Let $\mathscr{F}$ be a finite partition of $[0,1]^n$ into cells $Q_1, \ldots, Q_k$. Then the conditional expectation $E[f \mid \mathscr{F}]$ is a function from $[0,1]^n \to \mathbb{R}$ such that for $x \in Q_i$

$$E[f \mid \mathscr{F}](x) := \frac{\int_{Q_i} f\,dx}{m(Q_i)}.$$

Further, $E[f \mid \mathscr{F}]$ is L^1-computable from (codes for) $f, \mathscr{F}$, and $\frac{\int_{Q_i} f\,dx}{m(Q_i)}$ is computable from $f, \mathscr{F}, i$.

The LDT talks about simultaneous convergence over a "net of filtrations". We extract one filtration to work with.

Theorem (Insipred by Brattka-Miller-Nies)

Let $f : [0,1]^n \to \mathbb{R}$ be a non-negative L^1-computable function. For any x, if for some decreasing sequence of cubic cells (Q_i) (not necessarily "computable") converging to x,

$$\frac{\int_{Q_i} f(x)\,dx}{m(Q_i)}$$

diverges, then there is some computable filtration $(\mathscr{F}_k)$ (made up of finite partitions of cells) such that $\mathscr{F}_\infty = \mathscr{B}$ (Borel σ-algebra) and

$$E[f \mid \mathscr{F}_k](x)$$

diverges as $k \to \infty$.

f^+, f^- are L^1-computable

The requirement that f be non-negative in the previous lemma is OK, since the decomposition

$$f = f^+ - f^-$$

is computable from (the L^1 code for) f.

Outline

Levy 0-1 Law

By the Key Lemma, it is enough to show $E[f \mid \mathscr{F}_k](x)$ converges when x is Schnorr random and f is L^1-computable.

Theorem (Levy 0-1 Law)

Given a filtration $(\mathscr{F}_k)$ and an L^1-function f, then

$$E[f \mid \mathscr{F}_k] \to E[f \mid \mathscr{F}_\infty]$$

both in L^1-norm and pointwise almost-everywhere.
Therefore, if f is $\mathscr{F}_\infty$-measurable, then

$$E[f \mid \mathscr{F}_k] \to f$$

both in L^1-norm and pointwise almost-everywhere.

Effective Levy 0-1 Law

Theorem (Effective Levy 0-1 Law, R.)

Given a computable filtration $(\mathscr{F}_k)$ such that $\mathscr{F}_\infty = \mathscr{B}$ and an L^1-computable f, then

$$E[f \mid \mathscr{F}_k]$$

converges effectively in the L^1-norm, and converges pointwise on Schnorr randoms.

Step 1: Notice $M_k := E[f \mid \mathscr{F}_k]$ is a martingale.

Step 2: Approximate M_k to be rational-valued.

Step 3: Extract fast converging subsequence from M_k.

Step 4: Create Solovay test of points that don't converge "fast enough".

Step 1: Notice M_k is a Martingale

A martingale on a filtration $(\mathscr{F}_k)$ is a sequence of functions (M_k) such that M_k is $\mathscr{F}_k$-measurable and

$$E\left[M_{k+1} \mid \mathscr{F}_k\right] = M_k.$$

Let

$$M_k := E\left[f \mid \mathscr{F}_k\right].$$

It is well known that M_k is a martingale, by

$$E\left[M_{k+1} \mid \mathscr{F}_k\right] = E\left[E\left[f \mid \mathscr{F}_{k+1}\right] \mid \mathscr{F}_k\right] = E\left[f \mid \mathscr{F}_k\right] = M_k.$$

Martingales have very nice convergence properties!

Step 2: Approximate M_k to Be Rational Valued

By a straight-forward approximation argument, M_k can be assumed to be rational-valued. (Not entirely accurate, but close enough.)

Step 3: Extract Fast Converging Subsequence

Since $M_k \to f$ in the L^1-norm (Levy's 0-1 Law) and f is L^1-computable we can find a subsequence $\left(M_{k_j}\right)$ such that

$$\left\|M_{k_j} - f\right\|_1 \leq 2^{-(2j+1)}.$$

Therefore

$$\left\|M_{k_{j+1}} - M_{k_j}\right\|_1 \leq 2^{-2j}.$$

Actually, we can show $M_k \to f$ effectively in L^1-norm.

Step 4: Create Solovay Test

First notice that by Markov's Inequality we have

$$m\left\{x : \left|M_{k_{j+1}}(x) - M_{k_j}(x)\right| \geq 2^{-j}\right\} \leq \frac{\left\|M_{k_{j+1}} - M_{k_j}\right\|_1}{2^{-j}} \leq \frac{2^{-2j}}{2^{-j}} = 2^{-j}.$$

But we need something stronger. Since (M_k) is a martingale, we have this stronger inequality (Ville's Inequality/Doob's Submartingale Inequality).

$$m\left\{x : \max_{k_j \leq k \leq k_{j+1}} \left|M_k(x) - M_{k_j}(x)\right| \geq 2^{-j}\right\} \leq \frac{\left\|M_{k_{j+1}} - M_{k_j}\right\|_1}{2^{-j}} \leq 2^{-j}.$$

Step 4: Create Solovay Test

Let

$$U_j = \left\{ x : \max_{k_j \le k \le k_{j+1}} \left| M_k(x) - M_{k_j}(x) \right| \ge 2^{-j} \right\}.$$

Then

- (U_j) is uniformly Σ_1^0 (ignoring the boundaries of the cells).
- $m(U_j) \le 2^{-j}$ so $\sum m(U_j) \le 2$.
- $m(U_j)$ is computable (since U_j is a finite union of cells from the filtration and (M_k) is rational valued).

It follows that (U_j) is a Solovay test for Schnorr randomness.

If x is a Schnorr random, then x is in only finitely-many U_k. So, $M_k(x)$ is Cauchy.

Further, the rate of convergence is computable from x's "randomness deficiency", i.e. the highest k such that $x \in U_k$.

End of Proofs

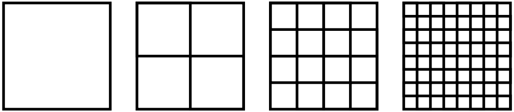
This ends the proof of the Effective Levy 0-1 Law,
and hence the proof of the Effective Lebesgue Differentiation Theorem.

Outline

Construction of L^1 Function

Remains to prove the reversal.

Let $\mathscr{F}_k$ be the partition of dyadic cubes on $[0,1]^n$.



It is enough for each non-Schnorr x to find an L^1-computable f such that $E[f \mid \mathscr{F}_k](x)$ diverges.

Uniformly Integrable Martingales

Theorem

A martingale M_k is of the form $E[f \mid \mathscr{F}_k]$ if and only if it is uniformly integrable, i.e.

$$\sup_n \int_{\{x:|M_k(x)|>C\}} |M_k|\, dx \to 0 \quad \text{as } C \to \infty.$$

This is the property that assures M_k converges in the L^1-norm.

Sketch of Construction

- Construct martingale M_k.
- Show M_k is uniformly integrable, hence $M_k = E[f \mid \mathscr{F}_k]$ for some L^1 function f.
- Show f is L^1-computable.

The Martingale

Martingales correspond to betting strategies. Bet as follows:

- Take Solovay test (U_i) for Schnorr randomness.
- Separate U_i into disjoint union of dyadic cubes.
- For every dyadic cube $Q \in \mathscr{F}_k$, ask if Q is in disjoint union of some U_i. (Pick disjoint unions such that this is decidable.)
  - If so, bet 1 (doesn't matter if win or lose)
  - If not, bet nothing.

Easy to see that the following are equivalent

- $M_k(x)$ does not converge
- $M_k(x)$ bets infinitely often
- x is in infinitely many U_i
- x passes the Schnorr test (U_i) (so x is not Schnorr random).

Uniformly Integrable

The martingale is uniformly integrable since $\sum m(U_i) < \infty$. (Hence same construction works for Martin-Löf randomness.)

The martingale converges to an L^1-computable function f since $\sum m(U_i)$ is computable.

Let f be the L^1 limit of the martingale. If in $[0,1]$ let F be the corresponding absolutely continuous function. F is a "saw-tooth function. (Hence this construction is similar to the one of Brattka, Miller, Nies.)

Outline

Measures v.s. L^1 functions

An extension of Lebesgue Differentiation theorem is as follows.

Theorem

Let ν be a finite signed measure on $[0,1]^n$. Then

$$\frac{\nu(B_r(x))}{m(B_r(x))} \to \frac{d\nu}{dm}$$

where $d\nu/dm$ is the Radon-Nikodym derivative.

On $[0,1]$,

- $F(x) = \int_0^x f\,dx + F(0)$ is an absolutely continuous function.
- $G(x) = \nu((0,x]) + G(0)$ is a function of bounded variation. (ν is Lebesgue–Stieltjes measure of G.)

Hence this theorem is equivalent to the fact that a function of bounded variation is differentiable almost-everywhere.

Computable Signed Measures and Functions of Effective Bounded Variation

Define computable finite signed measure and computable finite measure using the Riesz representation theorem (signed measures are isomorphic to bounded linear operators on continuous functions). This gives a computable description in the weak$*$ topology.

Define a function of effective bounded variation as one which corresponds to a computable signed measure.

- The term "effective bounded variation" is a bit misleading, since the total variation norm is not in general computable. Any suggestions for a different term?
- Computable functions of bounded variation are of effective bounded variation.

Measures v.s. L^1 functions

Theorem (R.)

Let ν be a finite signed measure (resp. measure) which is computable on $[0,1]^n$. Then

$$\frac{\nu(B_r(x))}{m(B_r(x))}$$

converges for Martin-Löf randoms x (resp. computable randoms x).

Corollary (R. extending Brattka-Miller-Nies)

If $F : [0,1] \to \mathbb{R}$ is a function (resp. a nondecreasing function) with effective bounded variation. Then

$$F'(x)$$

exists for Martin-Löf randoms x (resp. computable randoms x).

Decomposing a Measure

There are two decompositions for signed measures:
Han-Jordan Decomposition $\nu = \nu^+ - \nu^-$ (split into two positive measures).

Lebesgue Decomposition
$\nu \perp m$ (ν is mutually singular wrt m) if $d\nu/dm = 0$.
$\nu \ll m$ (ν is absolutely continuous wrt m) if

$$\nu(E) = \int_E \frac{d\nu}{dm}.$$

$\nu = \nu_1 + \nu_2$ where $\nu_1 \perp m$ and $\nu_2 \ll m$.

Decomposition of Bounded Variation Function

On $[0,1]$:

If ν is absolutely continuous with respect to m, then

$$F(x) := \nu((0,x]) = \int_0^x \frac{d\nu}{dm}$$

and F is absolutely continuous. (We have already investigated this case.)

If ν is mutually singular with respect to m, then $G(x) := \nu((0,x])$ has

$$G'(x) = \frac{d\nu}{dm} = 0 \quad \text{a.e.}$$

Call G singular.

Any function of bounded variation can be decomposed into absolutely continuous and singular functions.

Mutually Singular Results

Theorem (R.)

If ν is mutually singular (wrt m) computable measure (resp. computable signed measure), then

$$\lim_{r \to 0} \frac{\nu(B_r(x))}{m(B_r(x))} = 0$$

for Schnorr randoms (resp. ML randoms) x.

Corollary (R.)

If $F : [0,1] \to \mathbb{R}$ is a singular nondecreasing function (resp. a function) of effective bounded variation, then

$$F'(x) = 0$$

for Schnorr randoms (resp. ML randoms) x.

The reversals also hold.

Table of Results

For effective bounded variation functions on $[0, 1]$:

	Singular	Eff. Abs. Cont.	Abs. Cont.	No Decomp.
Nonneg.	Sch.	Sch.	Comp.	Comp.
Bdd Var.	ML	Sch.	ML	ML

For measures: Replace "Nonneg" and "Bdd Var." with "Measure" and "Signed Measure".

There are even further decompositions and results.

Convergence of Martingales

Levy's 0-1 law is related to Doob's Martingale Convergence Theorem.

Theorem (Doob's Martingale Convergence Theorem)

If (M_k) is an L^1-bounded martingale, then M_k converges almost everywhere as $k \to \infty$.

Effective versions of this have been given by Takahashi and Dean for Martin-Löf randomness and Brattka, Miller, Nies for computable randomness.

The techniques in this talk can be applied to get more effectivizations of Doob's Martingale Convergence Theorem.

What is the derivative?

Take f L^1-computable. Let

$$\hat{f} := \limsup_{k} \frac{\int_{B_r(x)} f\, dx}{m(B_r(x))}.$$

This is a unique (up to Schnorr randoms) representative for f.

Similarly, take a computable signed measure ν. Let

$$\hat{\frac{d\nu}{dm}} := \limsup_{r \to 0} \frac{\nu(B_r(x))}{m(B_r(x))}.$$

Again this is unique (up to randoms).

Can these "effective" representatives of the L^1-class be characterized?

This is related to layerwise computable functions of Hoyrup and Rojas (based on Lusin's Theorem), as well as previous work of Pathak (based on polynomial definition of L^1-computable).

Summary

- The standard randomness notions allow one to prove effective versions of "almost everywhere" convergence theorems in Analysis and Probability.
- These effectivizations can be reversed (similar to Reverse Mathematics) to characterize known (or new?) types of randomness.
- Martingales are a useful tool for the study of randomness, even outside their usual form in the computability community. What other tools from modern probability and analysis can be used to study randomness?
- A better understanding of the computable aspects of Measure Theory, Probability Theory, and Ergodic Theory may lead to a better understanding of randomness.
- And *vice-versa*.

Thank You!

These slides are available on my webpage:

`math.cmu.edu/~jrute`.

Or just Google me, "Jason Rute".