

予測限界と ランダムネス

京都大学 数理解析研究所 計算機科学セミナー

2012年4月26日

宮部賢志

京都大学 数理解析研究所

概要

- 微分可能性とランダムネス
- ランダムネスの向かうべき方角
- アルゴリズム的確率における確率・ランダムとは

微分可能性と ランダムネス

ランダムネスの階層

弱 2 ランダムネス

$\cap$

Martin-Löf ランダムネス

$\cap$

計算可能ランダムネス

$\cap$

Schnorr ランダムネス

$\cap$

Kurtz ランダムネス

微分可能性

定理 (Lebesgue 1904)

すべての単調増加関数 $f : [0, 1] \rightarrow \mathbb{R}$ はほとんど至る所で (almost everywhere) 微分可能.

関数 f が有界変動 (bounded variation) であるとは

$$\sup \sum_{i=1}^{n-1} |f(t_{i+1}) - f(t_i)| < \infty.$$

ただし, $0 \leq t_1 \leq t_2 \leq \dots t_n \leq 1$.

Jordan の定理より有界変動関数は 2 つの単調増加関数の差なので, すべての有界変動関数 f はほとんど至る所で微分可能.

ほとんど至る所とは？

具体的にどの点で微分可能だろうか？

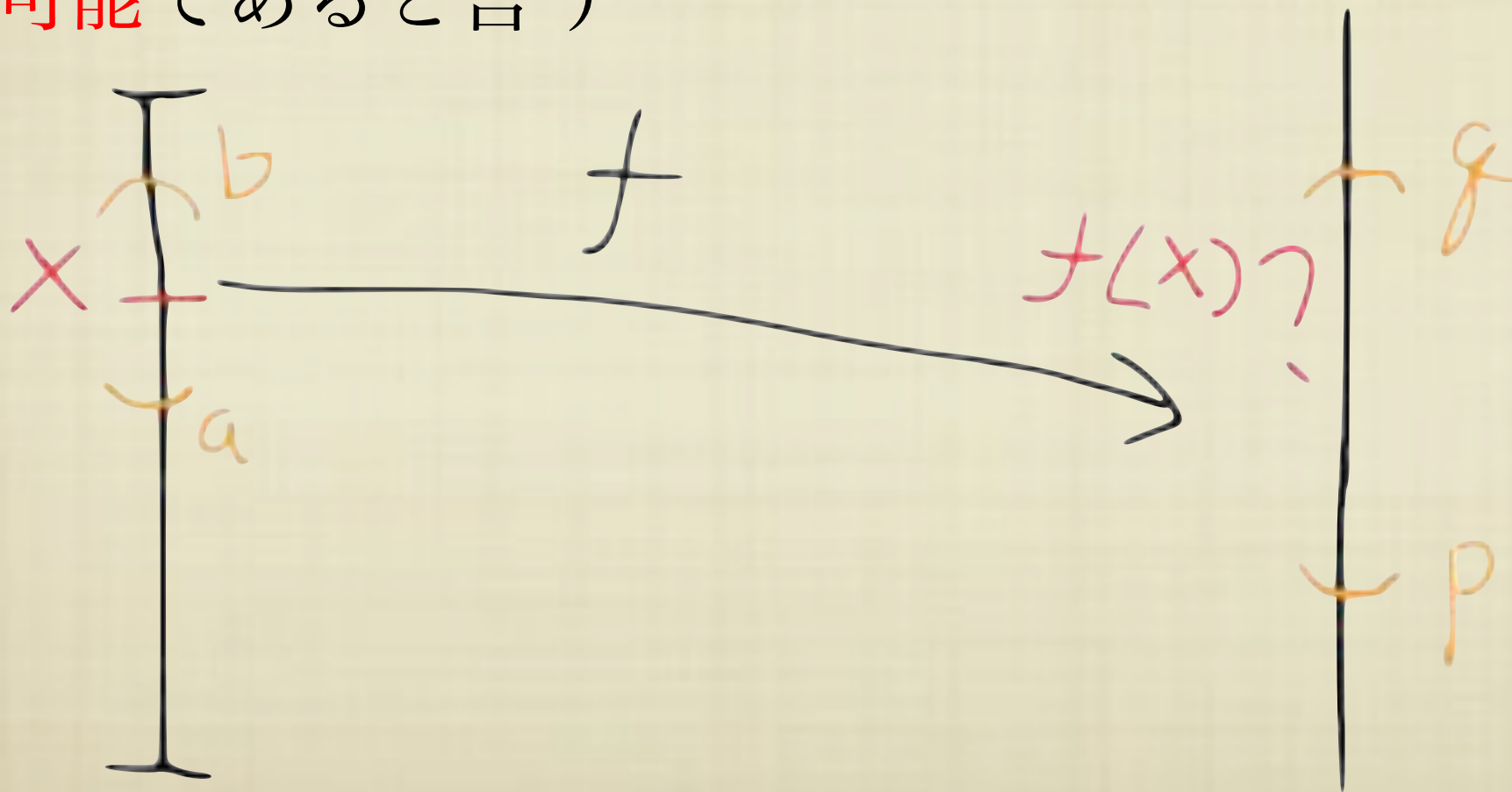
- $N(f) = \{x \mid f \text{ は } x \text{ で微分可能} \}$ とおく.
- すべての f で $\mu(N(f)) = 1$.
- すべての点 x である f が存在して、 $x \notin N(f)$.

よって古典的な立場では意味の無い問題.

しかし，可算の加法族 C に対しては $\mu(\bigcap_{f \in C} N(f)) = 1$.
そこで計算可能な関数の族を考える.

計算可能性

実数上の可算開基 (countable base) として $\{(p, q) \mid p, q \in \mathbb{Q}\}$ を固定する. 開集合が開基の計算可能な和であるとき **c.e.** であると言ひ, 各開基の f による逆像が一様に c.e. であるとき、 f は**計算可能**であると言ふ



MLランダムネス

C を計算可能な有界変動関数の集合とすると

$$\mu\left(\bigcap_{f \in C} N(f)\right) = 1$$

定理 (Demuth 1975, Brattka, Miller, Nies to appear)

$x \in [0, 1]$ が Martin-Löf ランダム

$\iff$ すべての $f \in C$ に対し f は x で微分可能

その他のランダムネス

定理 (Brattka, Miller, Nies to appear)

x が弱 2 ランダム $\iff$

すべての計算可能で至る所微分可能な関数 f が x で微分可能

定理 (Brattka, Miller, Nies to appear)

x が計算可能ランダム

$\iff$ すべての計算可能な単調関数 f が x で微分可能

定理 (Freer, Kjos-Hanssen, Nies to appear)

x が計算可能ランダム

$\iff$ すべての計算可能な Lipschitz 関数 f が x で微分可能

Schnorr ランダムネス

定理 (Freer, Kjos-Hanssen, Nies to appear)

x が Schnorr ランダム

$\iff$ すべての「変動ノルムで計算可能」な Lipschitz 関数 f
が x で微分可能

注意

「変動ノルムで計算可能」は「計算可能」よりも強い条件

Kurtz ランダムネス

定理 (M.)

K を以下を満たす関数の族とする

1. 計算可能
2. 微分が拡張実数への関数として計算可能

このとき、 x が Kurtz ランダム

$\iff$ すべての $f \in K$ が x で微分可能

階層の対応

弱2ランダムネス	至る所微分可能	古典的性質
MLランダムネス	有界変動	
計算可能 ランダムネス	単調 or Lipschitz	
Schnorr ランダムネス	変動ノルムで 計算可能	+計算可能性
Kurtz ランダムネス	微分が拡張実数 への計算可能	

微分定理

定理 (微分定理)

積分可能な f に対しほとんど至る所

$$f(x) = \lim_{r \rightarrow 0} \frac{\int_{B_r(x)} f d\mu}{\mu(B_r(x))}$$

ここで $B_r(x) = (x - r, x + r)$

この等式が成り立つ点を **Lebesgue 点** という

L^1 計算可能性

定義 (L^1 計算可能性; Pour-El & Richards 1989)

ある関数 f に対し，計算可能な有理数係数多項式の列 f_n があって，

$$\|f - f_n\|_1 \leq 2^{-n}$$

となる時. f は L^1 計算可能であると言う

また $\hat{f} = \lim_n f_n$ とおき，実効的 L^1 計算可能関数と呼ぶ.

Schnorr ランダムネス

定理 (Pathak-Rojas-Simpson, Rute)

以下は同値

- x は Schnorr ランダム
- $\hat{f}(x)$ がすべての L^1 計算可能関数に対して定義される
- すべての L^1 計算可能関数 $\hat{f}$ の Lebesgue 点である

積分テスト

積分テスト とは積分可能な下側半計算可能関数 $f : [0, 1] \rightarrow \overline{\mathbb{R}}$
 x が ML ランダム $\iff$ すべての積分テスト f で $f(x) < \infty$.

定理 (M.)

計算可能な積分値を持つ積分テストを

Schnorr ランダムネスに対する積分テスト と呼ぶ

x が Schnorr ランダム

$\iff$ すべての Schnorr ランダムネスに対する積分テスト f
に対して $f(x) < \infty$

自然な関数族の発見

定理 (M.)

以下の関数を Schnorr ランダムな点に制限した関数の族はそれぞれ互いに一致する

1. 2 つの Schnorr ランダムに対する積分テストの差
2. 実効的 L^1 計算可能関数 $\hat{f}$
3. 計算可能な L^1 ノルムを持つ Schnorr layerwise 計算可能関数

Kurtz ランダムネス (1/3)

定理 (M.)

以下は同値

1. z は Kurtz ランダム
2. すべての非負拡張計算可能関数でほとんど至る所有限な関数 f に対し $f(z) < \infty$
3. すべての非負拡張計算可能関数で積分可能な関数 f に対し $f(z) < \infty$
4. すべての非負拡張計算可能関数で積分値が計算可能な関数 f に対し $f(z) < \infty$

Kurtz ランダムネス (2/3)

定義

部分計算可能関数 f の定義域の測度が 1 の時, **a.e. 計算可能**であると言う.

命題 (M.)

関数 f が a.e. 計算可能であることと, 2 つの Kurtz ランダムネスに対する積分テストの差であることは**”ほとんど同じ”**

Kurtz ランダムネス (3/3)

命題 (M.)

以下は同値.

1. z は Kurtz ランダム.
2. すべての a.e. 計算可能な関数 f で $f(z)$ が定義される.
3. すべての a.e. 計算可能な関数 f に対して, z は Lebesgue 点である.

「2つの積分テストの差」が重要な関数族である

弱いランダムネスの場合

R を Kurtz もしくは Schnorr ランダムな点の集合とする.

この時, 以下が同値になっている.

1. $z \in R$.
2. すべての R に対する積分テスト f で $f(z) < \infty$.
3. すべての R に対する積分テストの差 g で $g(z)$ が定義される.
4. すべての R に対する積分テストの差 g に対して, z が Lebesgue 点である.

一般の場合

R を弱 2, ML, Schnorr もしくは Kurtz ランダムな点の集合とする. この時, 以下が同値になっている.

1. $z \in R$.
2. すべての R に対する積分テスト f で $f(z) < \infty$.
3. 以下を満たすすべての g で z が Lebesgue 点である.
 - (a) R に対する積分テストの差.
 - (b) 各区間の積分値が一様に計算可能.

積分テストの階層

弱2ランダムネス	至る所有限	区間ごとの 積分値が 計算可能でない
MLランダムネス	積分可能	
計算可能 ランダムネス	?	
Schnorr ランダムネス	積分値が 計算可能	区間ごとの 積分値が 計算可能
Kurtz ランダムネス	拡張実数 への計算可能	

階層の対応

弱2ランダムネス	至る所微分可能	古典的性質
MLランダムネス	有界変動	
計算可能 ランダムネス	単調 or Lipschitz	
Schnorr ランダムネス	変動ノルムで 計算可能	+計算可能性
Kurtz ランダムネス	微分が拡張実数 への計算可能	

ランダムネスの 向かうべき方角

ランダムネスの今まで

■ ロシアの研究者

Kolmogorovらの後を継ぎ，哲学的議論も好む。

若い研究者が現れていないように見える。

Demuth programmeなど。

■ ヨーロッパ+ニュージーランドの研究者

計算可能性理論出身の人が多く，

ランダムネスとTuring degreeとの関連を好む。

計算可能解析との融合

- 今，取り組まれている方向としては．．．
- 微分可能性（Niesら）
 - 関数の階層を明らかにしたという大きな成果.
- 力学系（Hoyrup, Avigadのグループ）
 - 特に重要． 解釈が変わる． 計算可能性との関連を調べる事でシミュレーションに影響する？

その他重要な方向性

- フーリエ解析への応用→デジタル信号処理

- 新しい確率論の構築.

Solomonoffのアルゴリズム的確率など.

確率哲学(1/2)

確率に対する立場		その数学的定式化
客観確率	傾向説	Kolmogorov
	頻度説	von Mises
主観確率		--

- 通常, Kolmogorovによる公理主義的確率論はどの確率にも当てはまると言われる
- 実際には傾向説で, 主観確率の研究者はKolmogorovの枠組みではない確率論を構築している

確率哲学(2/2)

確率に対する立場		その数学的定式化	
傾向説	实在論的	Kolmogorov	
頻度説	認識論的	von Mises	Solomonoff?
主観説		--	

- 1919年に始まる von Mises による頻度説の定式化
“First collective, then probability.”
- アルゴリズム的確率は von Mises の理論の拡張？
主観説とも相性が良い (Solomonoff 2009)

統計の立場から

一般的に考えられている立場	中心人物	ランダムネスから見ると？
頻度説	Neyman, Pearson	計算可能性として頻度のみを考慮している
相対頻度に基礎を置く信頼性の尺度	Fisher	
主観説	Bayesian	計算可能性が考慮されていない

予測理論としての立場

予測に利用する性質	理論
典型性	統計・情報理論
圧縮不可能性	赤池情報量規範・最小記述長
予測不可能性	ゲーム論的確率論

- アルゴリズム的信息理論では、これら3つの性質は本質的に同じもの
- 予測理論として統一的な扱いができるはず

アルゴリズム的確率における 確率・ランダムとは

アルゴリズム的確率入門

定義 (Levy 1937)

関数 $M : 2^{<\omega} \rightarrow \mathbb{R}^+$ が, $M(\epsilon) = 1$,

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

を満たす時, M を (半) 測度と呼ぶ.

半測度 M が以下を満たす時, 最適 (optimal) であると言う:
任意の半測度 N に対して, ある定数 C が存在し, すべての σ
で, $C \cdot M(\sigma) \geq N(\sigma)$.

定理 (Schnorr 1971, Levin)

最適な半測度 M は存在する.

Solomonoffの結果

ある $x \in 2^\omega$ に対して以下のようにおく：

$$M(x_n | x_{<n}) = \frac{M(x_{1:n})}{M(x_{<n})}.$$

定理 (Solomonoff 1978)

計算可能な測度 μ に対して,

$$\frac{M(x_n | x_{<n})}{\mu(x_n | x_{<n})} \rightarrow 1 \quad (n \rightarrow \infty)$$

が μ 確率 1 で成り立つ.

自然に解釈すると

- ある測度でランダムなデータが与えられた時,
- 最適な予測 M に依れば,
- その測度を予め知らなかったとしても,
- やがてその測度に収束する.

具体例

- 偏りのあるコインを投げる.
- 長い間投げ続ければ, そのコインの偏りのだいたいの値は計算できる.
- 次のコインが表である確率を計算できる.
- 独立同分布を仮定できない場合でも
計算可能でありさえすれば可能!

ベイズ確率との比較

	アルゴリズム的確率	ベイズ確率
設定	文字列のみ	任意の測度空間
予測する確率	次の文字のみ	その後の事象
最初に与える確率	最適	任意
予測関数	semimeasure	測度
計算可能性	計算不可能	計算可能

利点と欠点

- 利点 — 「確率の求め方」を提案
- 欠点 — 設定の限定, 計算不可能性
→ 計算可能解析 + Schnorr ランダムネス

微分と条件付き確率

- 微分定理
- Radon-Nikodym微分（測度の微分）
- 条件付き確率

予測する確率の違い

Solomonoff によれば、最適な確率 M があって、 $x_{<n}$ が与えられたときの、次の文字の確率は、

$$M(x_n | x_{<n}) = \frac{M(x_{1:n})}{M(x_{<n})}.$$

ベイズ確率によれば、ある確率 μ が与えられていて、 $x_{<n}$ が与えられたときの、その後の確率は、

$$M(\sigma | x_{<n}) = \frac{M(x_{<n}\sigma)}{M(x_{<n})}.$$

条件付き確率

μ を 2^ω 上の一様測度とする.

g を非負実効的 L^1 計算可能関数で $\mu(g) = 1$ を満たすとする.

N を g を確率密度関数とする測度とする: すべての $\sigma \in 2^*$ で

$$N(\sigma) = \int_{[\sigma]} g d\mu$$

$X \in 2^\omega$ を固定する. $n \in \mathbb{N}$ に対し, N_n を

$$N_n(\sigma) = \frac{N((X \upharpoonright n)\sigma)}{N(X \upharpoonright n)}$$

で定義し, n 番目の条件付き確率と呼ぶ.

長期予測

予想

g_i を非負一様実効的 L^1 計算可能関数で $\mu(g_i) = 1$ とする. N^i をそれらを密度に持つ測度とする.

$$f = \sum_i 2^{-i-1} g_i + \frac{1}{2}$$

として, M を f を密度に持つ測度とする.

X が $N = N^i$ に対して Schnorr ランダムならば,

$$H^2(M_n, N_n) \rightarrow 0 \quad (n \rightarrow \infty).$$

ここで H^2 は Hellinger 距離を表す.

注意

- Solomonoffが次の文字列を予測しているのに対し、
この場合は次以降の文字列すべてを予測している。
- 証明にはLevy 0-1 lawが本質的に使われる(はず)。
- 予測する測度は計算可能よりも制限されている。
- 新たな解釈ができるようになる！！

確率とは何か？

- von Misesは「ランダムな列」(Collective)の頻度であると主張
- Solomonoffは「最適な予測」から導かれる掛け金の配分の割合であると主張
→ 「ランダム」の言葉を使うとどうなる？

確率とランダムネス

図式

データ →

測度（確率）＋その測度でランダム

注意

- 「ランダム」の意味が厳密でない古典的枠組みでは、厳密には正当化できない主張。
だが、本質的にはよく使われる。
- 統計で「数値の組」を「適当な関数で近似」する時、最小二乗法を使って「残差」を小さくする。
- 情報理論で「届いたデータ」から「情報」と「エラー」を識別する。

MLランダムでは？

長い間未解決であったこの問題は，最近否定的に解決された．

定理 (Hutter & Muchnik 2007)

最適な半測度 M と計算可能な測度 μ と μ ML ランダムな列 x があって，

$$M(x_n | x_{<n}) - \mu(x_n | x_{<n}) \not\rightarrow 0 \quad (n \rightarrow \infty).$$

選択は適切に

- すべてのMLランダムな列で正しい測度に収束する半測度が存在する. (Hutter & Muchnik 2007)
- 最適な半測度はが正しい測度に収束するためには, 1回ジャンプから計算可能ランダムな列であれば十分. (M.)
- 例の設定の場合, すべてのSchnorrランダムな列で収束する.

今後の重要な問題

- Schnorr ランダムでない点では？
- データから N^i のインデックス i の計算可能性は？
- 「最適な予測」の特徴付けになるか？
- 距離空間でも成立するか？
- 他の解釈・特徴付けはあり得るか？

ではランダムとは？

- ランダムネスの定義で主要な3つ：
- 典型性（統計的検定）
- 予測不可能性（マルチンゲール）
- 圧縮不可能性（Kolmogorov複雑性）

測度は何を意味するか？

- 典型的 → 確率を想定している？
- 予測不可能性 → 胴元の予測
- 圧縮不可能性 → 見つけられるだろう規則

確率とランダムネス2

図式 2

データ →

見つけられた規則

+ その測度（規則）でランダム

参考文献

- [1] H. Lebesgue. *Leçons sur l'Intégration et la recherche des fonctions primitives*. Gauthier-Villars, Paris, 1904.
- [2] O. Demuth. The differentiability of constructive functions of weakly bounded variation on pseudo numbers. *Comment. Math. Univ. Carolin.*, 16(3):583–599, 1975.
- [3] V. Brattka, J. S. Miller, and A. Nies. Randomness and differentiability. Submitted.
- [4] C. Freer, B. Kjos-Hanssen, and A. Nies. Computable aspects of Lipschitz functions. In preparation.

参考文献

- [5] N. Pathak, C. Rojas, and S. G. Simpson. Schnorr randomness and the Lebesgue Differentiation Theorem. To appear in Proceedings of the American Mathematical Society.
- [6] K. Miyabe. Characterization of Kurtz randomness by a differentiation theorem. Submitted.
- [7] K. Miyabe. An integral test for Schnorr randomness and its application. Submitted.
- [8] K. Miyabe. Weak L^1 -computability and limit L^1 -computability. In preparation.

参考文献

- [9] R. J. Solomonoff. A formal theory of inductive inference I, II. *Information and Control*, 7:1–22,224–254, 1964.
- [10] R. J. Solomonoff. Complexity-based induction systems: Comparisons and convergence theorems. *IEEE Transaction on Information Theory*, IT-24:422–432, 1978.
- [11] M. Hutter and A. Muchnik. On semimeasures predicting Martin-Löf random sequences. *Theoretical Computer Science*, 382:247–261, 2007.
- [12] K. Miyabe. An optimal superfarthingale and its convergence over a computable topological space. To appear in *Lecture Notes in Artificial Intelligence*.