条件付き確率の計算可能性

宮部賢志*

1 条件付き確率は如何なる意味で計算可能か

確率という概念が科学において広範囲に使われていること，また計算機がその科学の研究および応用において広範囲に使われていることを考えれば，条件付き確率を計算する場面がいかに多いかは想像に難くない．しかし，数学における測度論では非構成的な証明が多用されているため，その計算可能性は自明ではない．昨今の計算可能解析学とランダムネスの理論の発展を使って，これらの問題に対して統一的な理論を与えよう．

高校で習う条件付き確率の定義は以下である．A および B を事象とし，$P(B) > 0$ とすると，B における A の 条件付き確率 は，

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

により定義される．この定義から明らかに，$P(A|B)$ は $P(A \cap B)$ と $P(B)$ から計算可能であることが分かる．問題は $P(B) = 0$ の場合である．

具体的な問題を設定しよう．今，$2^\omega \times 2^\omega$ 上に確率測度 μ が与えられているとする．第一成分が $x \in 2^\omega$ である時の，第二成分の確率測度を求めたい．もう少し正確には，x からその測度への計算可能性を考えたい．

*京都大学数理解析研究所

2 準備

2.1 素朴な方法

$\mathcal{B}$ を 2^ω 上の Borel 集合族としよう．測度 μ_1 を，すべての $A \in \mathcal{B}$ に対して $\mu_1(A) = \mu(A \times 2^\omega)$ で定義する．ある $\sigma \in 2^{<\omega}$ に対して，第一成分 x が $\sigma \prec x$ を満たすときの条件付き確率は，すべての $\tau \in 2^{<\omega}$ に対して，

$$\mu_2^\sigma([\tau]) = \frac{\mu([\sigma] \times [\tau])}{\mu([\sigma] \times 2^\omega)}$$

であるから，その極限として条件付き確率 μ_2^x を定義しよう．

定義 2.1. μ_2^x を以下で定義する．すべての $A \in \mathcal{B}$ に対して，

$$\mu_2^x(A) = \lim_{n\to\infty} \frac{\mu([x \upharpoonright n] \times A)}{\mu([x \upharpoonright n] \times 2^\omega)}.$$

果たしてこの μ_2^x は測度であろうか？どんな x ならば，すべての $A \in \mathcal{B}$ で定義されるであろうか？以下の結果が知られている．

定理 2.2 (folklore; [2])**.** μ を計算可能な測度とする．x が計算可能ランダム[1]ならば，μ_2^x は定義され，測度になっている．

しかし，一般には x から μ_2^x は計算不可能である．

定理 2.3 (Ackermann-Freer-Roy [1])**.** 以下を満たす μ が存在する．μ は *Lebesgue* 測度に対して絶対連続な計算可能確率測度で，すべての x に対して，

$$\mu_2^x \equiv_T x'.$$

ここで，x' は x の *Turing* ジャンプである．

[1] 2.2 節で定義する．

これより μ_2^x は xk あら計算可能ではない．では，μ がどんな測度であれば，μ_2^x は x から計算可能となるだろうか？更に言えば，$x \mapsto \mu_2^x$ はどんな関数だろうか？

2.2 マルチンゲールの極限の計算可能性

この素朴な方法の失敗は以下のことを思い起こさせる．

アルゴリズム的ランダムネスの理論 [8, 15] において，マルチンゲール とは，関数 $d : 2^{<\omega} \to \mathbb{R}^+$ であって，

$$d(\sigma) = \frac{d(\sigma 0) + d(\sigma 1)}{2}$$

を満たすものを言う [12]．すべての計算可能マルチンゲール d に対して，$\sup_n d(x \upharpoonright n) < \infty$ である時，列 $x \in 2^\omega$ は 計算可能ランダム であると言われる [18, 19]．一般に $\lim_n d(x \upharpoonright n)$ は x から計算可能ではないが，$\lim_n d(x \upharpoonright n)$ が存在するならば，x' から Turing の意味で計算可能である．

それに対して，Pathak et al. [16] の結果と筆者らの最近の結果 [13, 14] を組み合わせることで，以下のことが分かる．f を $L^1(\lambda)$ 計算可能関数[2]とすると，すべての $\sigma \in 2^{<\omega}$ で $d(\sigma) = \int_{[\sigma]} f d\lambda$ で定義される d は計算可能マルチンゲールとなる．そのような d に対しては，すべての Schnorr ランダム[3]な列 x に対して，$\lim_n d(x \upharpoonright n)$ は存在して，x から計算可能である．

そこで μ に対して，似たような制限を与えれば，μ_2^x が x から計算可能になってくれるのではないかという期待ができる．

2.3 正則条件付き確率

条件付き確率の存在は基本的な結果として確率論でよく知られているが，その証明は（いずれも）非構成的であり，どうやって計算したら良いかは明らかではない．そこで正則条件付き確率を考えるのは 1 つの方法であろう．

定義 2.4 (Probability kernel [11], Transition measure [3])**.** 関数 $\kappa : 2^\omega \times \mathcal{B} \to [0,1]$ が以下を満たす時，κ を *probability kernel* と呼ぶ．

1. すべての $x \in 2^\omega$ に対して，$\kappa(x, \cdot)$ が 2^ω 上の確率測度．

2. すべての $B \in \mathcal{B}$ に対して，$\kappa(\cdot, B)$ は可測．

定義 2.5. *probability kernel* κ が以下を満たす時，κ を 正則条件付き確率 *(*の変形*)* と呼ぶ．すべての $A, B \in \mathcal{B}$ に対して，

$$\mu(A, B) = \int_A \kappa(x, B) d\mu_1(x).$$

μ は確率測度だから，κ は少なくとも積分が計算可能であるような関数である．しかし一般に関数が計算可能であっても積分は計算可能とは限らないし，積分が計算可能であっても計算可能とは（もちろん）限らない．すなわちこの文脈においては正則条件付き確率に計算可能であることを求めるのは不自然で，むしろ積分が計算可能であるような条件を求めるべきである．そのような条件の一つが L^1 計算可能性である．

3 L^1 計算可能な密度を持つ測度

3.1 L^1 計算可能性

L^1 計算可能性についてこれまで知られていることを復習しよう．

定義 3.1 (計算可能距離空間 [20, 6, 7])**.** *3* つ組 (X, d, α) が 計算可能距離空間 であるとは以下を満たすことを言う．

1. $d : X \times X \to \mathbb{R}$ は X 上の距離*,*

2. $\alpha : \mathbb{N} \to X$ は X で稠密な列*,*

3. $d \circ (\alpha \times \alpha) : \mathbb{N}^2 \to \mathbb{R}$ は計算可能な実数列．

[2] 3.1 節で定義する．
[3] 3.1 節で定義する．

$\alpha_i = \alpha(i)$ と書く．すべての $i \in \mathbb{N}$, $p \in \mathbb{Q}$ に対して，$B(\alpha_i, p) = \{x : d(\alpha_i, x) < p\}$，$\overline{B}(\alpha_i, p) = \{x : d(\alpha_i, x) \leq p\}$ とする．計算可能距離空間上の測度 μ が計算可能であるとは，開球の有限和が一様に下側半計算可能であることを言う [4, 10].

命題 3.2 (Bosserhoff [4], Hoyrup and Rojas [10])**.** μ を計算可能距離空間上の計算可能確率測度とする．この時，計算可能な稠密な列 $\{r_n\}$ が存在して，すべての i, j に対して $\mu(\overline{B}(\alpha_i, r_j) \setminus B(\alpha_i, r_j)) = 0$.

今，そのような $\{r_i\}$ を固定し，$B(\alpha_i, r_j)$ を basic set，$\overline{B}(\alpha_i, r_j)$ を co-basic set と呼ぶ．$\mathcal{I}$ を basic set もしくは co-basic set の有限和の集合とする．

定義 3.3. 計算可能距離空間上の関数 $s : X \to \mathbb{R}$ が<u>有理階段関数</u>*(rational step function)* であるとは，

$$s = \sum_{i=1}^{k} q_i \mathbf{1}_{B_i}$$

となる $q_i \in \mathbb{Q}$, $B_i \in \mathcal{I}$ が存在することを言う．

定義 3.4 ([17, 16, 13])**.** 関数 $f :\subseteq X \to \mathbb{R}$ が<u>実効的 L^1 計算可能</u>であるとは，有理階段関数の計算可能な列 $\{s_n\}$ が存在して，すべての n で $||s_n - s_{n-1}||_1 \leq 2^{-n}$ かつ $f(x) = \lim_n s_n(x)$.

L^1 計算可能性は Schnorr ランダムネスと相性が良い．

定義 3.5 ([18])**.** 一様に *c.e.* 開集合の列 $\{U_n\}$ が，すべての n で $\mu(U_n) \leq 2^{-n}$ かつ $\mu(U_n)$ が一様に計算可能であるとき，$\{U_n\}$ を<u>*Schnorr* テスト</u>と呼ぶ．点 $x \in X$ が<u>*Schnorr* ランダム</u>であるとは，すべての *Schnorr* テスト $\{U_n\}$ に対して，$x \notin \bigcap_n U_n$ であることを言う．

命題 3.6 ([16])**.** μ を計算可能距離空間 X 上の計算可能な測度とする．$x \in X$ に関して以下は同値．

1. x は *Schnorr* ランダム．
2. すべての実効的 L^1 計算可能関数 f に対して，$f(x)$ が定義される．

定義 3.7 ([13, 9])**.** ある関数 $f :\subseteq X \to \mathbb{R}$ が<u>*Schnorr* 各層計算可能</u>であるとは，*Schnorr* テスト $\{U_n\}$ が存在して，$f|_{X \setminus U_n}$ が一様に計算可能．

Schnorr テスト $\{U_n\}$ に対し $\bigcap_n U_n$ を Schnorr null と呼ぶ．

定理 3.8 ([13])**.** 関数 $f :\subseteq X \to \mathbb{R}$ に関して以下は同値．

1. ある実効的 L^1 計算可能関数と *Schnorr null* を除いて一致．
2. *Schnorr* 各層計算可能かつ計算可能な L^1 ノルムを持つ．

3.2 Banach 空間上への L^1 計算可能性

前節の結果を Banach 空間への関数に拡張しよう．

定義 3.9 (Brattka [5])**.** 表現付き空間 *(represented space)*(X, δ) が<u>計算可能ベクトル空間</u>であるとは，$(X, +, \cdot, 0)$ がベクトル空間で以下を満たすことを言う．

1. $+ : X \times X \to X$, $(x, y) \mapsto x + y$ が計算可能，
2. $\cdot : \mathbb{R} \times X \to X$, $(a, x) \mapsto a \cdot x$ が計算可能，
3. $0 \in X$ が計算可能な点．

定義 3.10 (Brattka [5])**.** *3* つ組 $(X, ||\cdot||, e)$ が<u>計算可能ノルム付き空間</u>であるとは，

1. $||\cdot|| : X \to \mathbb{R}$ が X 上のノルム，
2. $e : \mathbb{N} \to X$ が基本列，すなわちその線型包が X で稠密，
3. (X, d, α_e) が計算可能距離空間．ここで

$$d(x, y) := ||x - y||,$$

$$\alpha_e\langle k, \langle n_0, \cdots, n_k \rangle\rangle := \sum_{i=0}^{k} \alpha_{\mathbb{R}}(n_i) e_i.$$

4. (X, δ_X) は計算可能ベクトル空間，

であることを言う．更に (X, d) が完備距離空間であれば，計算可能 *Banach* 空間 と呼ぶ．

実数の場合と同様に L^1 計算可能性を定義する．

定義 3.11. μ を計算可能距離空間 (X, d, α) 上の計算可能測度とし，$(N, ||\cdot||_N, e)$ を計算可能ノルム空間とする．関数 $f :\subseteq X \to N$ が 実効的 L^1 計算可能 であるとは，X から N への階段関数の計算可能な列 $\{s_n\}$ が存在して，すべての n で $\int ||s_{n+1}(x) - s_n(x)||_N d\mu \leq 2^{-n}$ かつ $f(x) = \lim_n s_n(x)$ であることを言う．

L^1 空間は以下のように定義されていることを思い出そう．

$$L_1 := \{[f] \ : \ \text{可測関数}\ f : X \to \mathbb{R},\ ||[f]||_1 < \infty\},$$
$$[f] := \{g : X \to \mathbb{R} \ : \ \text{可測関数}\ g,\ ||[f-g]||_1 = 0\}.$$

ここでの距離は，$d([f], [g]) = ||[f] - [g]||_1$ で定めされていて，$||\cdot||_1$ がノルムになっている．

今，$f : X \times Y \to \mathbb{R}$ を $L^1(\mu_X \times \mu_Y)$ 関数とすると，f は X から $L^1(Y, \mu_Y)$ への関数と見ることができる．この計算可能性を見てみよう．

命題 3.12. μ を計算可能距離空間 X 上の計算可能測度とする．以下で定義される $(L_1(X, \mu), ||\cdot||_1, e)$ は計算可能ノルム空間である．

1. $e_i = e(i) = [\mathbf{1}_{B_i}]$,
2. $\{B_i\}$ はすべての *basic* または *co-basic set* の計算可能な数え上げ．

定理 3.13. X, Y を計算可能距離空間とし，μ_X, μ_Y は X, Y 上の計算可能測度とする．$f :\subseteq X \times Y \to \mathbb{R}$ を非負の実効的 $L^1(\mu_X \times \mu_Y)$ 計算可能関数とし，$f_1(x) = \int_Y f(x, y) d\mu_Y(y)$ とする．また，$h :\subseteq X \to L_1(Y, \mu_Y)$ を次のように定める．もし $x \in X$ が $0 < f_1(x) < \infty$ を満たすならば $h(x)(y) = [f(x, y)/f_1(x)]$．この時，$h$ は実効的 $L^1(f_1\mu_X)$ 計算可能．

また，定理 3.8 に相当する結果が少なくとも一方向は成り立つ．

定理 3.14. 計算可能距離空間から計算可能ノルム空間への関数 $f :\subseteq X \to N$ が実効的 L^1 計算可能ならば，*Schnorr* 各層計算可能．

$\mathbb{R}$ の場合の逆方向の証明は，ノルム空間に一般化できないが，逆方向に相当する結果が成り立つ可能性はあり，今後の研究課題である．

3.3 条件付き確率の計算可能性

本題に戻ろう．今，$2^\omega \times 2^\omega$ 上に計算可能な測度 μ が与えられているとしよう．更に，μ は Lebesgue 測度に対して実効的 L^1 計算可能な密度 f を持つとしよう．この時，前節の結果を使うと，$x \mapsto \mu_2^x$ が μ_1 に関して Schnorr 各層計算可能であることが分かる．

更に，

$$\kappa(x, B) = \int_B f(x, y) d\lambda(y)$$

で定義される κ は，μ_1-Schnorr null を除いて，probability kernel の条件を満たしている．よって，probability kernel の定義はこのように弱められるべきであり，実際弱い形での定義が採用されている教科書も存在する [21]．

参考文献

[1] N. L. Ackerman, C. E. Freer, and D. M. Roy. On the Computability of Conditional Probability. arXiv:1005.3014.

[2] L. Bienvenu, M. Hoyrup, and A. Shen. In Preparation.

[3] V. Bogachev. *Measure theory.* Springer, 2007.

[4] V. Bosserhoff. Notions of probabilistic computability on represented spaces. *Journal of Universal Computer Science*, 14(6):956–995, 2008.

[5] V. Brattka. *Computability of Banach Space Principles.* Fernuniv., Fachbereich Informatik, 2001.

[6] V. Brattka. Computability over topological structures. In S. B. Cooper and S. S. Goncharov, editors, *Computability and Models*, pages 93–136. Kluwer Academic Publishers, New York, 2003.

[7] V. Brattka, P. Hertling, and K. Weihrauch. A tutorial on computable analysis. *New Computational Paradigms*, pages 425–491, 2008.

[8] R. Downey and D. R. Hirschfeldt. *Algorithmic Randomness and Complexity.* Springer, Berlin, 2010.

[9] M. Hoyrup and C. Rojas. An Application of Martin-Löf Randomness to Effective Probability Theory. In *CiE*, pages 260–269, 2009.

[10] M. Hoyrup and C. Rojas. Computability of probability measures and Martin-Löf randomness over metric spaces. *Information and Computation*, 207(7):830–847, 2009.

[11] O. Kallenberg. *Foundations of modern probability.* Springer, New York, 2nd ed. edition, 2002.

[12] P. Lévy. *Théorie de l'Addition des Variables Aléatoires.* Gauthier-Villars, 1937.

[13] K. Miyabe. L^1-computability, layerwise computability and Solovay reducibility. Submitted.

[14] K. Miyabe and J. Rute. Van Lambalgen's Theorem for uniformly relative Schnorr and computable randomness. Submitted.

[15] A. Nies. *Computability and Randomness.* Oxford University Press, USA, 2009.

[16] N. Pathak, C. Rojas, and S. G. Simpson. Schnorr randomness and the Lebesgue Differentiation Theorem. To appear in Proceedings of the American Mathematical Society.

[17] M. B. Pour-El and J. I. Richards. *Computability in analysis and physics.* Springer, 1989.

[18] C. Schnorr. A unified approach to the definition of a random sequence. *Mathematical Systems Theory*, 5:246–258, 1971.

[19] C. P. Schnorr. *Zufälligkeit und Wahrscheinlichkeit*, volume 218 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin-New York, 1971.

[20] K. Weihrauch. *Computable Analysis: an introduction.* Springer, Berlin, 2000.

[21] D. Williams. *Probability with Martingales.* Cambridge University Press, 1991.