アルゴリズム的ランダムネスと学習可能性(2)

宮部賢志(明治大学)

2024年7月9日(火) -- 12日(金)
RIMS共同研究（公開型）
数理論理学の最近の進展
Symposium on Advances in Mathematical Logic

目次

Martin-Löfランダム性

定義(Martin-Löf 1966)

Martin-Löf検定(ML検定): 一様c.e.開集合の列$(U_n)_{n\in\omega}$で，すべての$n \in \omega$で$\mu(U_n) \leq 2^{-n}$となるもの.

$X \in 2^\omega$とML検定$(U_n)_n$に対して，$X \notin \bigcap_n U_n$となるとき，XはML検定$(U_n)_n$に**合格する**(pass)という.

$X \in 2^\omega$がすべてのML検定に合格するとき，Xは**MLランダム**であるという.

マルチンゲール

予測不可能性

気持ち：ランダム=予測不可能

公平な賭けであれば借金せずに元手を無限大に増やすことはできない.

$X \in 2^\omega$の列に対して賭けを行う.

$X \upharpoonright (n-1)$を見て，次のビットが0か1かに資金の一部を賭け，当たればその分増えて，外れればその分減る.

その資金課程(capital process)をマルチンゲールと呼ぶ.

マルチンゲール

マルチンゲールとは，関数$d : 2^{<\omega} \to [0, \infty)$で，

$$2d(\sigma) = d(\sigma 0) + d(\sigma 1)$$

を満たすものである．

公平な賭けの資金過程を表す．

戦略dは列の接頭辞がσであったとき，次のビットが0であることに，$d(\sigma 0) - d(\sigma)$だけ賭けている．
当たればその金額だけ資金が増え，外れたらその金額だけ資金が減って，資金は

$$d(\sigma) - (d(\sigma 0) - d(\sigma)) = d(\sigma 1)$$

となる．
賭ける金額は負でも良いが，当たっても外れても変化したあとの資金は非負である必要がある．

マルチンゲール

Martingale (tack) from Wikipedia

A martingale is any of several designs of tack that are used on horses to control head carriage.

日本語訳

マーチンゲールとは、馬の頭の位置を制御するために使用される馬具のいくつかのデザインのことを指します。

注意：日本語版は存在しない！！

Martingale (betting system) from Wikipedia

A martingale is a class of betting strategies that originated from and were popular in 18th-century France.

日本語訳

マーチンゲールとは、18世紀のフランスで起源を持ち、人気を博した一連の賭け戦略のことを指します。

特に負けた金額分賭け続けるとやがて勝てるという戦略のことを指す.

注意：日本語版は存在しない！！

Villeのマルチンゲール

Ville(1939)のcollectiveへの批判の中で，

戦略=資金過程

として，数学の関数としてマルチンゲールが導入された．

Martingale (probability theory) from Wikiepedia

In probability theory, a martingale is a sequence of random variables (i.e., a stochastic process) for which, at a particular time, the conditional expectation of the next value in the sequence is equal to the present value, regardless of all prior values.

日本語訳

確率論において、マルチンゲール（英: martingale）とは確率過程の性質の一つであり、過去の情報に制限して計算した期待値と未来の期待値が同一になる性質である。

注意：日本語版も存在する！！

マルチンゲールによる特徴づけ

定理 (Schnorr 1971)
$X \in 2^\omega$がMLランダムであることと，
すべての下側半計算可能(lower semicomputable)なマルチンゲールdに対して，

$$\sup_n d(X \upharpoonright n) < \infty$$

となることは同値．

ランダム性は公平な賭けで資金を無限に増やせないことを意味している．

注意

実数に対してleft-c.e.といったことに対応して，関数に対してはlower semicomputableと呼ぶ.

マルチンゲールに対してはc.e. マルチンゲールと呼ぶこともある.

証明

マルチンゲールから検定を作る.

$d(\lambda) \leq 1$として良い.

Kolmgorovの不等式より，$k \geq 1$に対して，

$$\mu(\{X \in 2^\omega \ : \ (\exists n) d(X \upharpoonright n) > 2^k\}) \leq 2^{-k}$$

この集合は一様にc.e.開集合なので，$\sup = \infty$となる列はML検定で覆うことができる.

証明2

検定からマルチンゲールを作る．$\mu(U_n) \leq 2^{-2n}$と仮定．

$$U_n = \bigcup_k [\sigma_k^n]$$

として，排反になるようにσ_k^nを選ぶ．

B_σはσに全賭けする戦略: $\sigma \in 2^{<\omega}$に対して，

$$B_\sigma(\tau) = \begin{cases} 2^{|\tau|} & \text{if } \tau \preceq \sigma \\ 2^{|\sigma|} & \text{if } \sigma \preceq \tau \\ 0 & \text{otherwise.} \end{cases}$$

証明3

$$d(\tau) = \sum_n \sum_k 2^{n-|\sigma_k^n|} B_{\sigma_k^n}(\tau)$$

$$d(\lambda) = \sum_n \sum_k 2^{n-|\sigma_k^n|} \leq \sum_n 2^{-n} < \infty$$

よりdはマルチンゲール．また，lower semicomputable.

$X \in \bigcap_n U_n$とすれば，任意の$n \in \omega$に対して，ある$k \in \omega$が存在して，
$\sigma_k^n \prec X$となり，その時点で資金が1確保される．
任意のnについて成立するから，$\sup_n d(X \upharpoonright n) = \infty$.

万能マルチンゲール

証明から明らかなように，万能なc.e.マルチンゲールが存在する．

置き換え

supは$\limsup$に置き換えることができる．当然．

上の証明から$\liminf$に置き換えても同値であることが分かる．

優マルチンゲールとは，関数$d : 2^{<\omega} \to [0, \infty)$で，

$$2d(\sigma) \geq d(\sigma 0) + d(\sigma 1)$$

を満たすものである．

定理の主張において，マルチンゲールを優マルチンゲールに置き換えても良い．

ただし，$\lim$の存在には置き換えることができない．

計算可能ランダム性

定義
$X \in 2^\omega$が**計算可能ランダム**(computably random)とは，任意の計算可能マルチンゲールdに対して，$\sup_n d(X \upharpoonright n) < \infty$となることをいう．

MLランダム$\Rightarrow$計算可能ランダム$\Rightarrow$Schnorrランダム(演習問題:示せ)

万能な計算可能マルチンゲールは存在しない(演習問題:示せ)

有理数値マルチンゲール

ランダムな列の構成において，マルチンゲールは有用な道具となる．

命題
dを実数値の計算可能なマルチンゲールとする．
このとき，ある有理数値の計算可能なマルチンゲールfが存在して，すべての$\sigma \in 2^{<\omega}$に対し，

$$|d(\sigma) - f(\sigma)| \leq 2$$

とできる．

証明

$\alpha(\sigma) = d(\sigma 0) - d(\sigma)$とし，
$\beta(\sigma)$を有理数値で$|\alpha(\sigma) - \beta(\sigma)| \leq 2^{-|\sigma|-2}$となるものとする．
$f(\sigma)$を，空文字列λに対して$d(\lambda) + 1/2 \leq f(\lambda) \leq d(\lambda) + 1$で，

$$f(\sigma a) = f(\sigma) + (-1)^a \beta(\sigma)$$

で定める．fは有理数値のマルチンゲールである．

$$|d(\sigma) + 1 - f(\sigma)| \leq |d(\lambda) + 1 - f(\lambda)| + \sum_{\tau a \prec \sigma} |(-1)^a||\alpha(\tau) - \beta(\tau)|$$

$$\leq \frac{1}{2} + \sum_{n < |\sigma|} 2^{-n-2} \leq 1$$

ランダム列の構成

ランダムな列の構成において，マルチンゲールは有用な道具となる．

命題
有理数値計算可能なマルチンゲールdに対し，すべてのnで

$$d(X \upharpoonright (n+1)) \leq d(X \upharpoonright n)$$

となる$X \in 2^\omega$が存在する．

つまり，Xはdに対してランダムに見える列．

証明
有理数値なので$d(\sigma 0), d(\sigma 1)$を比較して大きくない方を選べば良い．

有限個の計算可能マルチンゲールに対しては同じことができる.

万能なlower semicomputableなマルチンゲールに対して上記のように構成した列はMLランダムになる. 列は計算可能ではない.

演習問題:
任意の実数値計算可能マルチンゲールdに対して,

$$\sup_n d(X \upharpoonright n) \leq d(\lambda)$$

となる計算可能な列$X \in 2^\omega$が存在することを示せ.
ここでλは空文字列.

ランダム性の分離

定理 (?)

計算可能ランダムだが，MLランダムでない列が存在する．

証明

十分速く発散する計算可能な関数$h(n)$を固定する．

$(d_n)_n$: 有理数値の部分計算可能マルチンゲールの計算可能な数え上げ

$$d(\sigma) = \sum_n 2^{-n} d_n(\sigma)$$

が発散しないようにXを作れば，Xは計算可能ランダム．

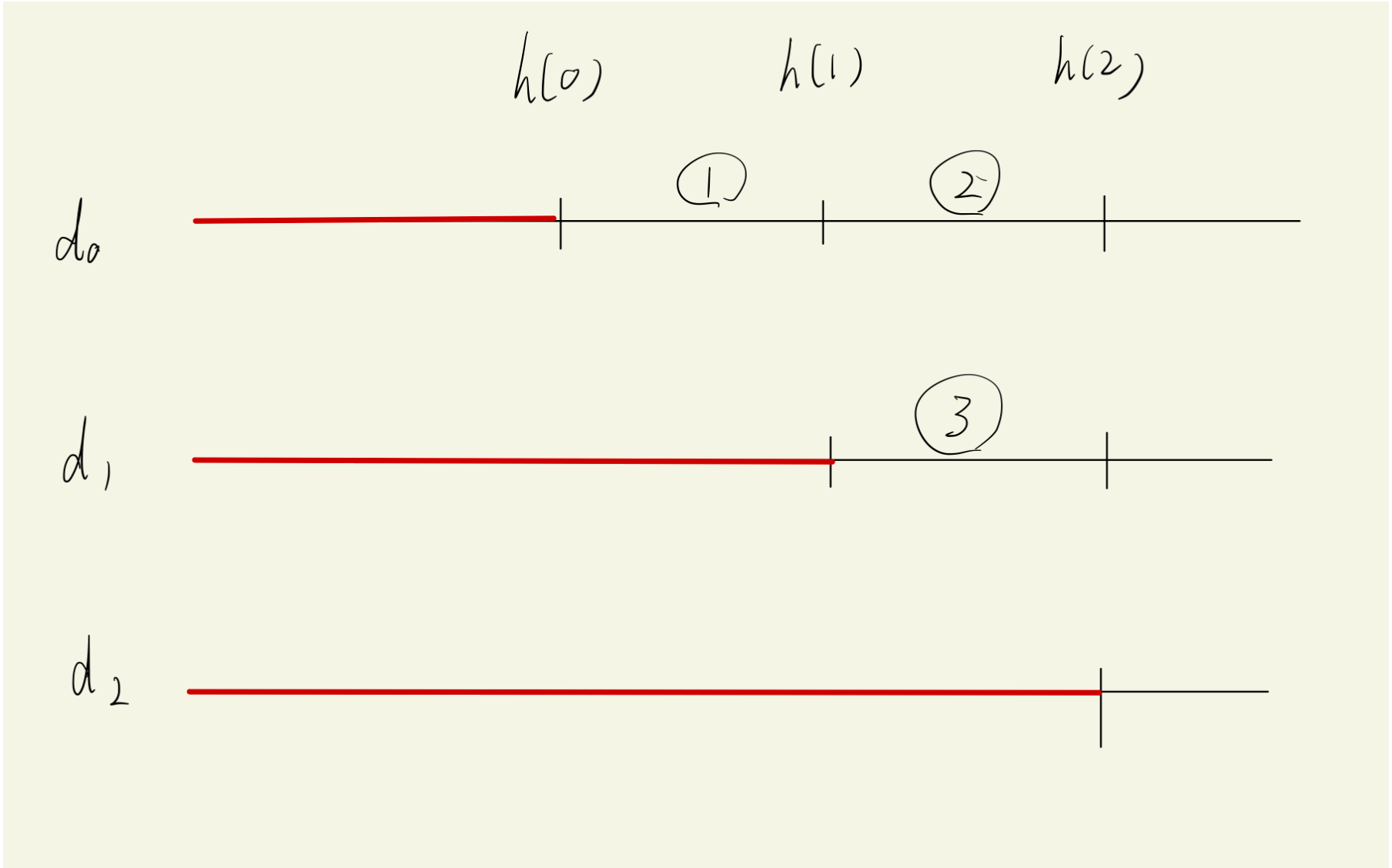
$|\sigma| \leq h(n)$ならば$d_n(\sigma) = 1$であると修正しても良い.

ある$|\sigma| \leq h(n)$に対して$d_k(\sigma)\downarrow$ならば，すべての$|\tau| \leq h(n)$に対して$d_k(\tau)\downarrow$と修正しても良い.

dはlower semicomputableな優マルチンゲールで，$X \in 2^\omega$をこのdに対して小さくなるように構成する.

$k \in \omega$を固定すると，$|\sigma| \leq h(k)$および$n \geq k$に対し$d_n(\sigma) = 1$.
また，$X \upharpoonright h(k)$が変化するのは，高々$\frac{k(k+1)}{2}$回.

hを十分大きく取れば，XはMLランダムでない.


h(0)
h(1)
h(2)
1
2
3
d_0
d_1
d_2

ランダム性の分離

定理 (Nies, Stephan, and Terwijn 2005)

任意のhigh次数に，計算可能ランダムだがMLランダムでない列が含まれる．

逆にhighでなければ存在しない．

Kolmogorov複雑性

圧縮不可能性

測度もマルチンゲールも確率論でよく使われる．圧縮不可能性との関係はランダムネスの理論の特に面白いところ．

定義
文字列$\sigma \in 2^{<\omega}$に対して，**非接頭Kolmogorov複雑性**(prefix-free Kolmogorov complexity)を以下で定義する：

$$K(\sigma) = \min\{|\tau| \ : \ U(\tau) = \sigma\}$$

ここで，$U :\subseteq 2^{<\omega} \to 2^{<\omega}$は万能非接頭機械である．

τはσを出力するための情報を持っている．
$K(\sigma)$が小さければσは単純，大きければσは複雑．

非接頭

$2^{<\omega}$の部分集合Sが**非接頭**(prefix-free)であるとは，任意の2つの異なるSの元に対し，どちらももう一方の接頭辞にならないことをいう．

例：電話番号の集合

機械Uの定義域$\mathrm{dom}(U) = \{\sigma \in 2^{<\omega} \ : \ U(\sigma)\downarrow\}$が非接頭のとき，**非接頭機械**と呼ぶ．
非接頭機械に対しては，入力を前から順に見ていったときに，定義域に含まれる文字列を見つけることができるという意味で，自己終端判別可能(self-deliminating)である．

任意の非接頭を模倣できる非接頭機械を万能非接頭機械と呼ぶ．

注意

$K(\sigma)$はUに依存するので，$K_U(\sigma)$とも書く．
ただし，Uの万能性から，高々定数しか変化しない．

Kolmogorov複雑性はその具体的な値に興味があるわけではなく，長い文字列に対する漸近的な振る舞いを明らかにする道具である．

Levin-Schnorrの定理

定理 (Levin-Schnorrの定理 1973)
$X \in 2^\omega$がMLランダムであることと，
$K(X \upharpoonright n) > n - O(1)$であることは同値．

ランダムな列は最初のn文字を高々定数分しか圧縮できない列として特徴づけられる．
ここで，機械に対する非接頭の条件を外すと，上記の同値性は成り立たなくなる．

歴史的コメント

Kolmogorov複雑性は，Kolmogorov (1965) により導入されていた．

Solomonoff (1964) は人工知能の文脈でほぼ同じ概念をKolmogorovより早く導入していた．

Levin, Schnorrが1973年に独立にわずかに異なる複雑性でMLランダム性との同値性を示した．

現在よく知られている非接頭Kolmogorov複雑性による特徴付けは Chaitin (1975)による．

記号コメント

単純複雑性(plain Kolmogorov complexity)はCで，
非接頭Kolmogorov複雑性(prefix-free Kolmogorov complexity)はKで表す．

古い文献ではKとHで表記されることがあるので注意．

Cの性質

命題

$$C(\sigma) \leq |\sigma| + O(1)$$

証明

$M(\sigma) = \sigma$とすると，Mは計算可能関数．

$$C_M(\sigma) \leq |\sigma|$$

Cの万能性から，

$$C(\sigma) \leq C_M(\sigma) + O(1) \leq |\sigma| + O(1)$$

Cの性質(続き)

命題

任意の$n \in \omega$に対し，$C(\sigma) \geq n$となる長さnの文字列σが存在する．

証明

長さnの文字列は2^n個ある．$C(\sigma) < n$となるためには，σを出力する長さn未満の文字列τが存在する必要がある．1つの入力に対しては1つの出力しか持ち得ない．長さn未満の文字列は，

$$1 + 2 + \cdots + 2^{n-1} = 2^n - 1$$

個しかないので，少なくとも1個は$C(\sigma) \geq n$となる必要がある．

Cの性質(続き)

命題

$$C(n) := C(0^n) \leq \log n + O(1)$$

注意：logの底は2

証明

$M(\sigma) = 0^n$とする．ここでnは1σが2進数で表す自然数である．
万能性から，

$$C(0^n) \leq C_M(0^n) + O(1) = \log n + O(1)$$

MLランダム列の特徴づけに向けて

命題

$C(X \upharpoonright n) > n - O(1)$となる列$X \in 2^\omega$は存在しない．

証明アイディア

n桁の文字列σには，長さnの情報に加えて，nという情報も持っている．任意の$X \in 2^\omega$に対して，$X \upharpoonright k$はその2進桁表現により2^kくらいの大きさの自然数から復元できる．すなわちこの分だけ圧縮が可能である．

Kの性質

命題 (Chaitin 1975)

$$K(\sigma) \leq |\sigma| + K(|\sigma|) + O(1)$$

証明

Kに使われる万能非接頭機械Uを固定し，

$$M(\rho\tau) = \tau \text{ if } U(\rho) = |\tau|$$

によりMを定義する．Mは計算可能である．

Uが非接頭であることから，入力のσ, τへの分割は一意である．

Kの性質(続き)

そのことから，Mは非接頭機械であることが導かれる．

これより，

$$K(\sigma) \leq K_M(\sigma) + O(1) \leq |\sigma| + K(|\sigma|) + O(1)$$

Kの性質(続き)

命題 (Counting theorem; Chaitin 1975)

$$|\{\sigma \ : \ |\sigma| = n \wedge K(\sigma) \leq n + K(n) - r\}| \leq 2^{n-r+O(1)}$$

特に,

$$\max\{K(\sigma) \ : \ |\sigma| = n\} = n + K(n) \pm O(1)$$

Levin-Schnorrの定理の証明1

機械から検定を作る.

$$U_d = \bigcup \{[\sigma] \ : \ K(\sigma) < |\sigma| - d\}$$

とおくと，$(U_d)_d$は一様c.e.開集合である.
さらに，Counting theoremより

$$\mu(U_d) \leq \sum_n 2^{-n} \cdot 2^{n-K(n)-d} \leq 2^{-d}$$

より$(U_d)_d$はML検定である.
もしAがMLランダムであれば，ある$d \in \omega$が存在して$A \notin U_d$なので，
すべてのnで$K(A \upharpoonright n) \geq n - d$である.

Kraftの不等式

定理(Kraftの不等式)
非接頭集合$S \subseteq 2^{<\omega}$に対し，

$$\sum_{\sigma \in S} 2^{-|\sigma|} \leq 1$$

Sが非接頭なので，$[\sigma]$が互いに素になり，それぞれの一様測度が$2^{-|\sigma|}$であるから．

特に，

$$\sum_{\sigma \in 2^{<\omega}} 2^{-K(\sigma)} \leq \sum_{\sigma \in \mathrm{dom}(U)} 2^{-|\sigma|} \leq 1$$

KC定理

定理 (KC定理, Machine existence theorem)
$(d_i, \tau_i)_{i\in\omega}$を自然数$d_i \in \omega$と有限列$\tau_i \in 2^{<\omega}$の計算可能な組の列で
$\sum_i 2^{-d_i} \leq 1$を満たすものとし**要請**(request)と呼ぶ.
この要請から接頭機械Mと文字列σ_iの列を計算可能に作れて,

- すべてのiで$\left|\sigma_i\right| = d_i$,
- $M(\sigma_i) = \tau_i$,
- $\mathrm{dom}(M) = \{\sigma_i \ : \ i \in \omega\}$

とできる.

計算可能性を忘れれば簡単だが，計算可能性を要請すると難しい.

Levin-Schnorrの定理の証明2

万能ML検定$(U_n)_n$に対し，一様c.e.な非接頭集合V_nが存在して，

$$U_n = \bigcup \{[\sigma] \ : \ \sigma \in V_n\}$$

すべての$\sigma \in V_{2n} (n \in \omega)$に対し，$(|\sigma| - n + 1, \sigma)$を要請するKC集合を考える．この重さは

$$\sum_n \sum_{\sigma \in V_{2n}} 2^{-|\sigma|+n-1} \leq \sum_n 2^{-2n+n-1} = 1$$

KC定理より非接頭機械Mが存在して，

$$\sigma \in V_{2n} \Rightarrow K_M(\sigma) \leq |\sigma| - n + 1$$

AがMLランダムでないとする．
すべてのdで$A \in U_{2n}$であり，$A \upharpoonright m \in V_{2n}$となる$m$が存在し，

$$K_M(A \upharpoonright m) \leq m - 2n + 1$$

Kの最小性より，これは$K(A \upharpoonright k) > k - O(1)$を満たさないことを意味する．

MLランダム性の特徴づけ

Levin,Schnorr,Chaitinの定理: MLランダム性のKによる特徴づけ

定理 (Miller-Yu 2008)
$X \in 2^\omega$がMLランダムであることと,

$$C(X \upharpoonright n) > n - K(n) - O(1)$$

となることが同値.

2-ランダム性の特徴づけ

停止問題$\emptyset'$に対するMLランダム性を2-ランダム性と呼ぶ.

定理 (Miller 2004, Nies, Stephan, and Terwijn 2005)
$X \in 2^\omega$が2-ランダムであることと，無限に多くのnで

$$C(X \upharpoonright n) > n - O(1)$$

となることは同値.

2-ランダム性の特徴づけ(続き)

定理 (Miller 2010)
$X \in 2^\omega$が2-ランダムであることと，無限に多くのnで

$$K(X \upharpoonright n) > n + K(n) - O(1)$$

となることは同値.

まとめ

- MLランダム性はマルチンゲールによる特徴づけがある.
- MLランダム性はKolmogorov複雑性による特徴づけがある.