数理情報学 1・講義ノート

木原 貴行

名古屋大学 大学院情報学研究科

Contents

1. 圧縮可能性とランダム性

ランダムとは何なのか，ということを理解することは,「ランダムでない」とは何なのか，というのを理解することと同等である．そういうわけで，まずは，ランダムではない列について学んでいこう．以下に挙げるのは，明らかにランダムでないバイナリ列である．

$$\underbrace{00000000000000000000000000\ldots\ldots00000000}_{1\text{ 億桁}},$$
$$\underbrace{01110111011101110111011101110111\ldots\ldots01110111}_{1\text{ 億桁}}.$$

上の列は，ただの「0 が 1 億個並ぶ列」であるから，明らかにランダムでない．下の列は，すぐには気付かないかもしれないが，よく見ると，0111 を繰り返しているだけであることが分かる．つまり，下の列とは「0111 が 2500 万回並ぶ列」であるから，ランダムではない．

それでは，これらのランダムでない列が共有する性質とは何であろうか．すぐに思いつくのは，これらの列が**規則的**である，という点であろうか．ランダム性のことを無秩序性や不規則性と呼ぶこともあるし，つまり，**規則性**というものは，ランダム性の対義語と考えられている．

それでは，そもそも規則的なバイナリ列とは何であろうか．その列を記述する法則か数式か論理式のようなものがある，という感じだろうか．ここでは，もう少し漠然と，

規則性とは，そのバイナリ列を「より短い言葉で説明できる」ということである

本ノートは，2017 年度春期に名古屋大学情報文化学部において開講されたオムニバス形式講義「数理情報学 1」の筆者担当回である第 11，12 回で講義した内容をまとめたものである．

と考えよう．たとえば，上の例では，長さ 1 億のバイナリ列を「0 が 1 億個並ぶ列」という 8 文字,「0111 が 2500 万回並ぶ列」という 14 文字で説明することができた．もう少し身近な言葉で説明すると，これは，数十メガバイトの容量を持つデータをほんの数十バイトのデータに**圧縮**したということである．

以上をまとめると，非ランダム性とは規則性であり，規則性とは圧縮可能性である．言い換えれば，

$$\text{ランダム性} = \text{不規則性} = \text{圧縮不可能性}$$

ということである．この発想を，もう少し数学的に厳密な形で定式化しよう．

1.1. **コルモゴロフ複雑性**. $\{0,1\}^*$ によって，バイナリ列全体の集合を表す．また，バイナリ列 $\sigma \in \{0,1\}^*$ の長さを $|\sigma|$ によって表す．$\{0,1\}^*$ の部分集合を定義域とし，$\{0,1\}^*$ を値域とする関数を $\{0,1\}^*$ 上の**部分関数** (*partial function*) と言って，$f :\subseteq \{0,1\}^* \to \{0,1\}^*$ のように書く．

定義 1. $\{0,1\}^*$ 上の部分関数 $f :\subseteq \{0,1\}^* \to \{0,1\}^*$ が**計算可能** (*computable*) とは，あるコンピュータプログラム P が存在して，入力 $x \in \mathrm{dom}(f)$ で P を実行したとき，有限時間で $f(x)$ を出力することを意味する．

以後，$\{0,1\}^*$ 上の部分計算可能関数を**マシン**と呼ぶ．

事実 2. 万能マシン $M :\subseteq \{0,1\}^* \to \{0,1\}^*$ が存在する．つまり，M はマシンであって，任意のマシン $f :\subseteq \{0,1\}^* \to \{0,1\}^*$ に対して，

$$(\exists e \in \mathbb{N})(\forall x \in \mathrm{dom}(f))\ M(0^e 1x) = f(x).$$

上の事実については，全ての人が触れたことのあるパーソナル・コンピュータなどが万能マシンの代表例である．つまり，我々は 1 つのコンピュータ M の内部で任意のプログラム P に σ を入力したものを実行できる．これが述べることは，どんなマシン f が与えられても，

$$(\exists P)(\forall x \in \mathrm{dom}(f))\ M(P,x) = f(x)$$

ということである．これに多少の修正を施せば，万能マシンを得る．

それではデータ圧縮の概念を数学的に定式化しよう．ここで考えるのは可逆圧縮であり，不可逆圧縮は考えない．可逆圧縮で重要なのは，解凍アルゴリズムである．圧縮されたデータから元のデータを復元できなければならない．上の例で言えば,「0 が 1 億個並ぶ列」という文字列が，実際に本物の 0 が 1 億個並ぶバイナリ列を意味しているという了解があるから,「0 が 1 億個並ぶバイナリ列は『0 が 1 億個並ぶ列』という文字列で説明された」と言えるのである．

つまり，σ という記述から τ というバイナリ列を実際に生成する方法 M があって初めて，τ は σ によって説明される，τ は σ に圧縮される，のように言うことができる．この M とは，マシン $M :\subseteq \{0,1\}^* \to \{0,1\}^*$ であり，$M(\sigma) = \tau$ のとき，τ は σ に圧縮された，と考える．そうすると，τ は M によってどれくらい小さいデータサイズまで圧縮できるか，の限界値は以下によって与えられる．

$$C_M(\tau) = \min\{|\sigma| : M(\sigma) = \tau\}.$$

この数値 $C_M(\tau)$ を τ の**平コルモゴロフ複雑性** (*plain Kolmogorov complexity*) と呼ぶ．もし $C_M(\tau) \geqslant |\tau|$ ならば，τ は決して圧縮できない列であり，したがって，ランダムな列と考えられる．まず，圧縮不可能列が存在することを見てみよう．

命題 3. どんな n についても，長さ n のバイナリ列 τ で，$C_M(\tau) \geqslant |\tau|$ を満たすものが存在する．

Proof. 長さ n 未満のバイナリ列の個数を数えよう．以下，$(\sigma)_2$ によって，σ を数を 2 進表記だと思ったときのその値を表す．たとえば，$(110)_2 = 6$ である．まず、長さ n のバイナリ列の種類はちょうど 2^n 個であることに注意しよう．すると，長さ n 未満のバイナリ列の個数は

$$\sum_{i=0}^{n-1} 2^i = (\underbrace{111\ldots111}_{n\text{ 個}})_2 = (1\underbrace{000\ldots000}_{n\text{ 個}})_2 - 1 = 2^n - 1$$

である．つまり，高々$2^n - 1$ 個のバイナリ列 τ だけが $C_M(\tau) < n$ となり得る．しかし，長さ n のバイナリ列は 2^n 種類存在するから，長さ n のバイナリ列 τ で $C_M(\tau) \geqslant n$ となるようなものが存在する．つまり，$C_M(\tau) \geqslant |\tau|$ である． □

ところで，上では M はマシンだったら何でも良いとしたが，以後は，入力文字列 σ がいつ読み込み終わるか分からないシチュエーションを考えたい．こういう状況では，M は，適当なタイミングで文字列の読み込みが終了したと判断して，出力を返す必要がある．たとえば，入力文字列側が自身のファイルサイズを最初に記述しているとか，あるいは入力文字列の最後にエンドマークが打たれているなどすれば，M はいつ文字列の読み込みが終了したかを判断できるであろう．ここでは，その判断方法は具体的には指定せず，単に，いつ終わるか分からない文字列を入力とする関数，という概念を以下のように定義する．

定義 4. マシン $M :\subseteq \{0,1\}^* \to \{0,1\}^*$ が**接頭** (*prefix-free*) とは，次の条件を満たすことを言う．

$$(\forall \sigma, \tau \in \{0,1\}^*)\ [\sigma \prec \tau\ \&\ \sigma \in \mathrm{dom}(M) \implies \tau \notin \mathrm{dom}(M)].$$

つまり，M は何らかの文字列を読み込み中に，ある時点 σ で出力 $M(\sigma)$ を返したならば，その時点で文字列の読み込みは打ち切っており，σ の拡張 τ については M はもはや反応を示さない．

例 5. $\Phi :\subseteq \{0,1\}^* \to \{0,1\}^*$ を任意のマシンとすると，

$$M(0^{|\sigma|}1\sigma) = \Phi(\sigma)$$

で定義される関数は接頭である．

定義 6. 接頭マシン R が**最適** (*optimal*) であるとは，任意の接頭マシン M に対して，次の条件を満たすものである．

$$(\exists c \in \mathbb{N})(\forall \tau \in \{0,1\}^*)\ C_R(\tau) \leqslant C_M(\tau) + c.$$

定理 7. 最適接頭マシンは存在する．

Proof. Φ を万能マシンとする．$R(0^e1\sigma)$ の値を決めるために，まず $\Phi(0^e1\sigma)$ をシミュレートする．もし，ある s ステップで $\Phi(0^e1\sigma)$ が出力を返したならば，$s' = \max\{s, |\sigma|\}$ とする．各 $\tau \prec \sigma$ について，$\Phi(0^e1\tau)$ を s' ステップだけ実行し，その間に出力を返すかどうかを確かめる．もし $\Phi(0^e1\tau)$ が出力を返したならば，$R(0^e1\sigma)$ は出力を返さないと宣言する．さもなくば，続いて，各 $\tau \succ \sigma$ で $|\tau| < s$ なるものについて，$\Phi(0^e1\tau)$ を $s' - 1$ ステップだけ実行し，その間に出力を返すかどうかを確かめる．もし $\Phi(0^e1\tau)$ が出力を返したならば，$R(0^e1\sigma)$ は出力を返さないと宣言する．さもなくば，$R(0^e1\sigma)$ は $\Phi(0^e1\sigma)$ を出力する． □

M が最適接頭マシンであるとき，$C_M(\tau)$ の代わりに $K(\tau)$ と書き，これを**接頭コルモゴロフ複雑性** (*prefix-free Kolmogorov complexity*) と呼ぶ．以下，$\leqslant^+$ によって，高々定数 c 程度の差を除いて，不等式 $\leqslant$ が成立することを意味する．

命題 8. $K(\tau) \leqslant^+ 2|\tau|$.

Proof. $M(0^{|\tau|}1\tau) = \tau$ と定義すれば，これは接頭マシンであり，$C_M(\tau) \leqslant 2|\tau| + 1$ を得る． □

自然数 $n \in \mathbb{N}$ が与えられたとき，$\mathsf{bin}(n)$ によって，n を 2 進展開したバイナリ列を表すものとする．このとき，$|\mathsf{bin}(n)| =^+ \log n$ である．ここで，対数関数 $\log$ の底は 2 である．

命題 9. $K(\tau) \leqslant^+ K(0^{|\tau|}) + |\tau| \leqslant^+ 2\log|\tau| + |\tau|$.

Proof. R を最適接頭マシンとする．先程のように，文字列の長さ $|\tau|$ の情報をヘッダに付与するが，それを $0^{|\tau|}1$ と記述するのは無駄が多い．代わりに，文字列の長さ情報を圧縮したデータをヘッダに付与しよう．つまり，$R(\sigma) = \mathsf{bin}(|\tau|)$ のときに限り $M(\sigma\tau) = \tau$ と定義すれば，これは接頭マシンであり，$C_M(\tau) \leqslant C_R(0^{|\tau|}) + |\tau|$ である．補題 8 より，$C_R(0^{|\tau|}) \leqslant^+ 2\log|\tau|$ である． □

次の補題は，様々なバイナリ列のコルモゴロフ複雑性の上界を求める際に便利である．

補題 10. M がマシンならば，次を満たす定数 $c \in \mathbb{N}$ が存在する．

$$(\forall \tau \in \{0,1\}^*)\ K(M(\tau)) \leqslant K(\tau) + c.$$

Proof. R を最適接頭マシンとする．このとき，$S(\sigma) = M(R(\sigma))$ によって定義すると，$\mathrm{dom}(S) \subseteq \mathrm{dom}(R)$ であるから，S は接頭マシンである．与えられた τ について，長さ $K(\tau)$ のバイナリ列 σ で，$R(\sigma) = \tau$ なるものが存在する．このとき，$S(\sigma) = M(R(\sigma)) = M(\tau)$ であるから，$C_S(M(\tau)) \leqslant K(\tau)$ である．よって，R の最適性より，ある c が存在して，任意の τ に対して，$K(M(\tau)) \leqslant C_S(M(\tau)) + c \leqslant K(\tau) + c$ となる． □

1.2. **チャイティンのオメガ**．それでは，具体的なランダム列を記述する準備を始めよう．接頭マシンにおいて，入力はいつ終わるか分からないバイナリ列であったことを思い出そう．いま，0 と 1 が同程度に出現する公平なコイン投げの繰り返しによってバイナリ列 $x_1x_2x_3\ldots$ を作る，という状況を考える．このような入力に対して，接頭マシン M が出力を返す確率 Ω_M はどれくらいだろうか？

たとえば，ある偶数 k について 0^k1 の形である入力のみ受理し，k を出力する接頭マシン M を考えよう．コインを $2n+1$ 回投げた結果が正確に $0^{2n}1$ となる確率は $0^{-(2n+1)}$ である．したがって，コイン投げを延々繰り返した結果，そのうち，ある偶数 k について 0^k1 の形となっている確率は以下によって与えられる．

$$\Omega_M = \sum_{n=0}^{\infty} 2^{-(2n+1)} = 2/3.$$

さて，どんなバイナリ列 $\sigma \in \{0,1\}^*$ が与えられても，コインを $|\sigma|$ 回投げた結果が正確に σ と一致する確率は $2^{-|\sigma|}$ である．したがって，一般の接頭マシン M についても，M が出力を返す確率 Ω_M は次によって計算できる．

定義 11 (停止確率). M が接頭マシンであるとき，M の**停止確率** (*halting probability*) とは，以下によって定まる値である．

$$\Omega_M = \sum_{\sigma \in \mathrm{dom}(M)} 2^{-|\sigma|}.$$

ここで，$\mathrm{dom}(M)$ は M の定義域を表す．

問題 12. M が接頭マシンであるとき，$0 \leqslant \Omega_M \leqslant 1$ であることを証明せよ．

接頭マシン M が与えられたとき，停止確率 Ω_M は，計算可能な方法で下から近似できる．これを確かめるために，M_s を M の長さ s 以下の入力に対する計算を s ステップだけ実行したものとする．つまり，$|\sigma| \leqslant s$ かつ $M(\sigma)$ が s ステップ以下で出力を返すならば，$M_s(\sigma) = M(\sigma)$ とし，さもなくば $M_s(\sigma)$ は出力を返さないものとする．すると，

$$\Omega_{M,s} = \sum_{\sigma \in \mathrm{dom}(M_s)} 2^{-|\sigma|}$$

であり，この値は有限種類の入力に対する M の計算を s ステップずつ実行することによって計算できる．さらに，

$$\Omega_{M,0} \leqslant \Omega_{M,1} \leqslant \Omega_{M,2} \leqslant \cdots \leqslant \Omega_M = \lim_{s \to \infty} \Omega_{M,s}.$$

あるバイナリ列 $\sigma \in \{0,1\}^*$ について $\sum_{i=0}^{|\sigma|-1} \sigma(i) 2^{-i-1}$ と書けるような実数を**二進有理数** (*dyadic rational*) と呼ぶ．つまり，$0.\sigma 0^\omega$ の形の実数のことである．たとえば，任意のマシン M および $s \in \mathbb{N}$ について，$\mathrm{dom}(M_s)$ は有限であるから，$\Omega_{M,s}$ は二進有理数である．一方，R が最適接頭マシンであれば，定義域は無限集合となり，したがって，二進有理数には成り得ない．つまり，Ω は二進有理数ではない．

問題 13. Ω が二進有理数でないことを証明せよ．

二進有理数でないような実数 $x \in [0,1]$ の二進表記は一意に定まるので，x を無限バイナリ列と同一視できる．無限バイナリ列 $z \in \{0,1\}^{\mathbb{N}}$ が与えられたとき，$z \upharpoonright n$ によって，長さ n の z の始切片を表すものとする．つまり，$z = z_0 z_1 z_2 \ldots z_i z_{i+1} \ldots$ ならば，$z \upharpoonright n = z_0 z_1 z_2 \ldots z_{n-1}$ のことである．

定理 14. 任意の最適接頭マシンの停止確率 Ω について，次が成立する．

$$(\exists c)(\forall n)\ K(\Omega \upharpoonright n) \geqslant n - c.$$

Proof. R を最適接頭マシンとする．$\Omega = \Omega_R$ かつ $\Omega_s = \Omega_{R,s}$ とする．次のような（接頭とは限らない）マシン M を考える．各入力 σ に対して，$0.\sigma \leqslant \Omega_t < 0.\sigma + 2^{-n}$ になるまで待つ．もし，そのようなステップ t が訪れたら，R_t の値域に入っていないバイナリ列 η を返す．つまり，どんな σ についても $R_t(\sigma) \neq \eta$ であるような η を選び，$M(\sigma) = \eta$ とする．

さて，実数 x について，$\sigma = x \upharpoonright n$ であるとは，$0.\sigma \leqslant x \leqslant 0.\sigma + 2^{-n}$ であるということに注意しよう．したがって，もし $\sigma = \Omega \upharpoonright n$ ならば，$0.\sigma \leqslant \Omega \leqslant 0.\sigma + 2^{-n}$ である．$(\Omega_s)_{s \in \mathbb{N}}$ は非減少であり，$\Omega = \lim_{s \to \infty} \Omega_s$ であることと，$\Omega = 0.\sigma$ では有り得ないことを利用すると，$0.\sigma \leqslant \Omega_t < 0.\sigma + 2^{-n}$ となるような t が存在することが分かる．つまり，マシン M は入力 $\sigma = \Omega \upharpoonright n$ に対しては，必ず R_t の値域に入っていない η を出力する．

一方，t ステップより後では，長さ $n-1$ 以下の入力に対して R が出力を返すことはない．なぜなら，さもなくば，停止確率の定義より $\Omega \geqslant \Omega_t + 2^{-n+1}$ となるが，t の選び方より，$0.\sigma + 2^{-n+1} \leqslant \Omega_t + 2^{-n+1} \leqslant \Omega \leqslant 0.\sigma + 2^{-n}$ となり，これは矛盾を導く．

さて，η は R_t の値域に入ってなかったため，もし $R(\sigma)=\eta$ となるような σ があったとしても，その計算は t ステップよりも長い時間がかかる．したがって，上で見た t の性質より，R が入力 σ に対して出力を返しているということは，$|\sigma|\geqslant n$ を導く．これより，$K(\eta)\geqslant n$ となる．ここで $\eta=M(\sigma)=M(\Omega\upharpoonright n)$ だったことを思い出すと，$K(M(\Omega\upharpoonright n))\geqslant n$ を得る．ところで，M はマシンであるから，補題 10 より，任意の $\sigma\in\{0,1\}^*$ に対して，$K(M(\sigma))\leqslant K(\sigma)+c$ である．以上をまとめると，

$$n\leqslant K(M(\Omega\upharpoonright n))\leqslant K(\Omega\upharpoonright n)+c$$

である．ここで，c はマシン M に依存する定数であって，n に依存しないことに注意する．以上より，任意の n について，$K(\Omega\upharpoonright n)>n-c$ を得る． □

定理 14 のような最適接頭マシンの停止確率は，**チャイティンのオメガ** (*Chaitin's Omega*) と呼ばれる．

2. エルゴード理論

2.1. **法 1 一様分布.** バイナリ列あるいは有限アルファベットの列は，実数と思うことができる．与えられた実数を b 進展開したときに，その中にどんな b 進列が現れるかについて考察しよう．たとえば，円周率 π を 10 進展開したときに 999999999999 という列がいつか現れることがあるだろうか，というものが典型的な問題である．そのような考察の下でよく現れるのは，正規数とよばれる概念である．実数 x が与えられているとしよう．自然数 $b\geqslant 2$ について，実数 x を b 進展開したとき，いかなる b 進有限列も極限的に等頻度で現れるならば，x は b **進正規** (*normal in base b*) であるといわれる．つまり，x の b 進展開した結果を $x_{(b)}$ と書けば，長さ k のどんな b 進列 $\sigma\in b^k$ が与えられても，

$$\lim_{N\to\infty}\frac{x_{(b)}\text{の小数点以下}N\text{桁までに}\sigma\text{が現れる回数}}{N}=\frac{1}{b^k}.$$

が成り立つことである．任意の自然数 $b\geqslant 2$ に対して b 進正規であるような実数は**絶対正規** (*absolutely normal*) であると言われる．

極限頻度には，**小数部** (*fractional part*) しか関わらないから，以後，実数 x の小数部のみに着目し，$[x]$ によって x の小数部を表すものとする．$[x]=[y]$ とは，ある整数 $k\in\mathbb{Z}$ について $x=y+k$ が成立することと同値である．具体的には，$[x]=|x|-\lfloor|x|\rfloor$ と書ける．

さて，簡単のために x は無理数であったとして，x の b 進展開の小数点以下 k 桁目から $k+\ell$ 桁目がちょうど $\sigma\in b^\ell$ であったという状況を考えよう．これは，$[b^kx]$ の b 進展開が σ で始まるということと同値である．b 進展開が σ で始まるような実数たちは区間をなすことに注意しよう．これより，x が b 進正規ということと，任意の実数 $p,q\in[0,1]$ に対して，次が成立することは同値である．

$$\lim_{N\to\infty}\frac{\#\{k<N:p\leqslant[b^kx]<q\}}{N}=q-p.$$

より一般に，次の概念を考えよう．

定義 15. 実数列 $(x_n)_{n\in\mathbb{N}}$ が**法 1 一様分布** (*uniformly distributed modulo* 1) であるとは，任意の実数 $p,q\in[0,1]$ に対して，次の式が成立することである．

$$\lim_{N\to\infty}\frac{\#\{n<N:p\leqslant[x_n]<q\}}{N}=q-p.$$

つまり，x が b 進正規であるとは，$(b^n x)_{n\in\mathbb{N}}$ が法 1 一様分布であることと同値である．集合 $S \subseteq [0,1]$ の特性関数を $\mathbf{1}_S$ によって表すと，$(x_n)_{n\in\mathbb{N}}$ が法 1 一様分布であるとは，任意の区間 $J \subseteq [0,1]$ に対して，次の式を満たすことと同値である．

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} \mathbf{1}_J([x_n]) = \int_0^1 \mathbf{1}_J(x)dx \tag{1}$$

法 1 一様分布に関する基本的な定理として，**ワイル規準** (*Weyl criterion*) と呼ばれるものがある．

定理 16 (ワイル規準)**.** 実数列 $(x_n)_{n\in\mathbb{N}}$ について，次の条件は同値である．

(1) (x_n) は法 1 一様分布である．

(2) 任意の連続関数 $f : [0,1] \to \mathbb{R}$ について，

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} f([x_n]) = \int_0^1 f(x)dx.$$

(3) 周期 1 の任意の複素数値リーマン可積分関数 $f : \mathbb{R} \to \mathbb{C}$ について，

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} f(x_n) = \int_0^1 f(x)dx.$$

(4) 任意の整数 $k \neq 0$ について，

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} e^{2\pi i k x_n} = 0.$$

Proof. (3)⇒(2) は自明である．また，整数 k について，$x \mapsto e^{2\pi i k x}$ は周期 1 の複素数値リーマン可積分関数であるから，(3)⇒(4) が成立することも導かれる．(1)⇒(3) について，法 1 一様分布性の等式 (1) による特徴付けを用いて，階段関数 f について (2) が成立することは容易に分かる．任意の実数値リーマン可積分関数は階段関数によって上下方向から近似できる．つまり，与えられた連続関数 f について，任意の $\varepsilon > 0$ について，$f_1(x) \leqslant f(x) \leqslant f_2(x)$ および $\int_0^1 (f_2(x) - f_1(x))dx < \varepsilon$ となる階段関数 f_1, f_2 を取れる．これより，(2) が成立することが導かれる．複素数値関数については，値を実部と虚部に分けて考えればよい．(2)⇒(1) については，同様に，階段関数を連続関数によって上下近似してやればよい．

(3)⇒(2) 周期 1 の複素数値リーマン可積分関数 f が与えられているとする．ワイエルシュトラスの近似定理より，f は複素三角多項式，つまり，$k \in \mathbb{Z}$ について $e^{2\pi i k x}$ 型の関数たちの有限線形結合によって任意に近似できる．与えられた $\varepsilon > 0$ について，複素三角多項式 Ψ で，任意の $x \in [0,1]$ について $|f(x) - \Psi(x)| < \varepsilon$ なるものを取る．このとき，

$$\begin{aligned}
&\left| \int_0^1 f(x)dx - \frac{1}{N} \sum_{n<N} f(x_n) \right| \\
&\leqslant \left| \int_0^1 (f(x) - \Psi(x))dx \right| + \left| \int_0^1 \Psi(x)dx - \frac{1}{N} \sum_{n<N} \Psi(x_n) \right| + \left| \frac{1}{N} \sum_{n<N} (\Psi(x_n) - f(x_n)) \right| \leqslant 3\varepsilon
\end{aligned}$$

であり，最初と最後の項は N に関わらず ε 以下の値であり，第 2 項は，十分大きい N について ε 以下となる． □

2.2. **バーコフのエルゴード定理.** 実数列 $(x_n)_{n\in\mathbb{N}}$ を時間経過によって，実数が $x_0, x_1, x_2, \ldots$ と刻々と変化していく過程を表しているものと考える．定理 16 (2) の式の左辺は，この変化していく入力に対する f の値の**時間平均** (*time average*) を表しており，右辺は f の値の**空間平均** (*space average*) を表す．つまり，実数列 $(x_n)_{n\in\mathbb{N}}$ が法 1 一様分布であるとは，この列が与える連続関数の値の時間平均が空間平均と一致するということである．

これを力学系の観点から見直そう．連続関数 $T : [0,1] \to [0,1]$ が与えられており，実数 $x \in [0,1]$ の T による軌道 $x, T(x), T^2(x), T^3(x), \ldots$ を考えよう．たとえば定理 16 より，x が b 進正規であるということは，$T(z) = [bz]$ および任意の連続関数 f について，以下の「時間平均 = 空間平均」が成立するということである．

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} f(T^n(x)) = \int_0^1 f(x)dx.$$

以下，λ で $[0,1]$ 上のルベーグ測度を表すものとする．関数 $T : [0,1] \to [0,1]$ が**保測** (*measure-preserving*) であるとは，任意の区間 $J \subseteq [0,1]$ について，$\lambda(J) = \lambda(T^{-1}[J])$ であることを意味する．

例 17. 自然数 $b \in \mathbb{N}$ について，$T(x) = [bx]$ によって定義される関数 $T : [0,1] \to [0,1]$ は保測である．なぜなら，区間 $J = [p,q]$ が与えられたとき，$T^{-1}[J] = \bigcup_{k<b}[(p+k)/b, (q+k)/b]$ であるが，各区間 $[(p+k)/b, (q+k)/b]$ の測度は $\lambda(J)/b$ であるから，$T^{-1}[J]$ の測度は $b \cdot \mu(J)/b = \lambda(J)$ である．

確率空間上の保測変換 T が**エルゴード的** (*ergodic*) とは，任意の T-不変可測集合が測度 0 または 1 であることを意味する．関数 f が $f \circ T = f$ を満たすとき，T-不変であるという．T がエルゴード的であることと，T-不変な自乗可積分関数は全て定数であることは同値である．

命題 18. 自然数 $b \geqslant 2$ について，$T(x) = [bx]$ によって定義される関数 $T : [0,1] \to [0,1]$ はエルゴード的である．

Proof. f を T-不変な自乗可積分関数とすると，$f \in L^2$ なので，フーリエ級数展開できる．これを $f(x) = \sum_{n\in\mathbb{Z}} c_n e^{2\pi inx}$ と表す．f の T-不変性より，$f(x) = f(T(x)) = \sum_{n\in\mathbb{Z}} c_n e^{2\pi ibnx}$ を得るから，任意の n について $c_{bn} = c_n$ を得る．一方，パーセバルの等式より，

$$\|f\|_2^2 = \int_0^1 |f(x)|^2 dx = \sum_{n\in\mathbb{Z}} |c_n|^2 < +\infty$$

を得るが，任意の n について $c_{bn} = c_n$ であることと最後の不等式を比較すると，$n \neq 0$ について，$c_n = 0$ であることを得る．したがって，$f(x) = c_0$ となり，定数であることが分かる．エルゴード性の自乗可積分関数による特徴付けにより，これは T がエルゴード的であることを導く． □

バーコフのエルゴード定理 (*Birkhoff's ergodic theorem*) とは，確率空間 X 上のエルゴード変換 T と可積分関数 $f \in L^1(X)$ が与えられたとき，殆ど全ての点 x は「時間平均 = 空間平均」を表す上式を成立させる，というものである．それでは，具体的には，一体如何なる実数が「時間平均 = 空間平均」を成立させるだろうか．実数 $x \in [0,1]$ の 2 進展開 $0.x_0x_1x_2\ldots$ の小数点以下 n 桁目までの部分 $x_0x_1x_2\ldots x_{n-1}$ を $x \upharpoonright n$ によって表していたことを思い出そう．

定理 19. 実数 $x \in [0,1]$ について，以下の条件は同値である．

(1) x は圧縮不可能である．つまり，

$$(\exists c \in \mathbb{N})(\forall n \in \mathbb{N})\ K(x \upharpoonright n) \geqslant n - c.$$

(2) 任意の計算可能[1]なエルゴード変換 $T : [0,1] \to [0,1]$ および下半計算可能[2]関数 $f : [0,1] \to [0,1]$ について，下式が成立する．

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} f(T^n(x)) = \int_0^1 f(x)dx.$$

定理 14 より，チャイティンの Ω は圧縮不可能であったことを思い出そう．したがって，Ω は上式の「時間平均 = 空間平均」を成立させる．特に，命題 18 より，Ω は絶対正規数である．このように，Ω は，あくまで最適マシンの停止確率というただの特殊な実数であったが，通常の意味でランダムとおぼしき性質は一通り満たしてしまう．

通常の確率論では “具体的なランダム列” に言及することはまず無いと言ってよいと思うが，アルゴリズム的ランダム性の理論では，このようにして Ω のような “具体的なランダム列” の理論を展開していく．

---

[1]関数 $g : [0,1] \to [0,1]$ が**計算可能** (*computable*) とは，それを任意精度で有理近似するコンピュータ・プログラムが存在することである．有理近似とは，我々のコンピュータは実数には直接アクセスできないが，たとえば，入力実数 x の任意精度情報を得ることができる，ということである．つまり，入力 x と要求精度 2^{-d} に対して，我々は $|x - p| < 2^{-d}$ なる何らかの有理数 p を受信できる．

もう少し詳細には，関数 g が計算可能とは，入力実数 x と出力精度 2^{-e} が与えられており，「入力値を精度 2^{-d} で近似する有理数を要求する」命令を自由に用いてよいとしたとき，$|g(x) - r| < 2^{-e}$ となるような有理数 r を返すプログラム P が存在する，ということである．

この計算可能性の定義を任意の可分距離空間に一般化できることは明らかであろうが，現代的には，関数の計算可能性は，距離化不可能空間を含む，より一般の空間における抽象的な方式で定義されている．

[2]関数 $h : [0,1] \to [0,1]$ が**下半計算可能** (*lower semicomputable*) とは，ある計算可能関数 $g : \mathbb{N} \times [0,1] \to [0,1]$ が存在して，任意の $x \in [0,1]$ について $f(x) = \sup_{n\in\mathbb{N}} g(n,x)$ となることを意味する．$[0,1]$ 上の任意の計算可能関数は連続であり，下半計算可能関数は下半連続である．