

研究計画

釜江哲朗

本研究ではランダムの意味とこの度合いを表す量を考察するとともに、逆に、ランダムでないこと、広い意味での周期性の考察も同時に行いたい。「ランダムとは何か？」という哲学的な問いに正面から答えるには論理的観点が必要となる。事実 Kolmogorov-Chaitin 計算量を基礎にして、アルゴリズム的観点から乱数が定義されるが、このようにして定義される乱数は「存在するが作れない」という致命的な欠陥をもつ。しかしながら、コンピュータ・シミュレーションにおいてはアルゴリズムによって生成される疑似乱数は欠かせない道具となっている。本研究では哲学的問いは排し、実用的観点からランダム、非ランダムに迫りたい。

正規列は独立確率変数列に対する大数の法則を実現する無限列とし導入された。これがどれだけランダムなものであるかは、1970年代に必勝戦略の非存在 (von Mises の collective) という立場から考察されてきた。すなわち、次の記号の選択をそれまでの列の有限オートマトンによる採否にしたがって行う場合は正規列から得られる無限部分列は再び正規列となることは知られていた。私はオートマトンの有限性を緩める形でこれを発展させた。最近中国での共同研究では、この十分条件をさらに拡張したが、必要十分条件には至ってない。これは本研究の課題の一つである。

正規列であることは乱数列がみたすべき最低限の要件であるが、乱数列とはほど遠いものも含まれている。このため、より強い概念が必要となる。このため私はアルファベット \mathbb{A} ($2 \leq d := \#\mathbb{A} < \infty$) 上の有限列 $x = x_1 \cdots x_n \in \mathbb{A}^n$ にランダムの度合いを示す数値 $\Sigma(x)$ を導入した。これは \mathbb{A} 上の有限列 ξ が x に出現する回数 $|x_1 \cdots x_n|_\xi$ の 2 乗をすべての有限列に対して加えたもの、 $\Sigma(x) = \sum_{\xi} |x_1 \cdots x_n|_\xi^2$ である。同じ長さ n の有限列どおしの比較ではこの値が小さいほどランダムである。 \mathbb{A} 上一様な i.i.d. 確率変数列 $X_1 X_2 \cdots$ に対して確率 1 で $\lim_{n \rightarrow \infty} (1/n^2) \Sigma(X_1 \cdots X_n) = \frac{d+1}{2(d-1)}$ が成立する。これを満たす無限列 $x_1 x_2 \cdots$ は Σ 乱数列と呼ばれ研究されている。どのような初期有限列からも、この延長として Σ 乱数列を作るアルゴリズムが存在する。さらに、性質の異なる初期有限列からこのアルゴリズムによって作られる 2 つの Σ 乱数列は互いに独立であるかのように振舞うことが観測されている。 Σ 乱数列は正規列となるが逆は成立しない。また、この Σ 値によって、ランダムとは逆の Sturmian 列や

eventually periodic な列の特徴付けも可能となった。本研究では Σ 乱数列の生成とその実際問題への応用を考察したい。

ランダムネスの大きさは不確定性の度合いであり，これを解消することによって得られる情報量とあわせてエントロピーとして量的に把握される。ただし，無限の対象に対しては1文字当たりのエントロピーとして，不確定性の度合いの対数を文字数で割った極限が議論される。これは不確定性の度合いが文字数 n とともに指数的に c^n という形で増える場合の $\log c$ を表す。これは不確定性あるいは複雑さの度合いを表す量として非常に粗い。例えば多項式的に増える場合，エントロピーは0であるが，多項式の次数は machine learning の分野で VC-dimension として重要視される。このようにエントロピーより詳細な複雑さの度合いの量的把握を試みた。この量についての研究は始まったばかりで，本研究ではこれを発展させたい。