Vol. 26 No. 3 情報処理学会論文誌 May 1985

部分列の多次元均等分布が保証された擬似乱数の発生法†

今井 徹†† 伏見正則††

擬似乱数列はその原系列のみならず，部分列も乱数列として用いるのにふさわしい性質を有していることが望ましい．そこで，M系列に基づく一様乱数列の部分列の性質を多次元均等分布の観点から評価することによって，部分列の性質もすぐれた一様乱数の発生法を設計する方法が提案されている．本論文では，この方法を適用して得られたいくつかの実用的な擬似乱数の発生法と，その性質を紹介する．

1. はじめに

真の乱数列の任意の部分列はまた乱数列であるが，アルゴリズムによって決定論的に発生される擬似乱数列については，その部分列が乱数列として用いるのに適当であるとは限らないので，部分列の性質は一般に別途検討する必要がある．

部分列の性質を吟味することの重要性は，Kolmogorov[8], Knuth[7], Tootill et al.[12] らによって論じられている．なかでも Kolmogorov[8]は，与えられた数列 $X=\{x_t;\ t=0,1,2,\cdots\}$ が乱数列として見なせるかどうかを，種々の部分列 $\{x_{t_1}, x_{t_2}, x_{t_3}, \cdots\}$ の頻度分布が原系列の頻度分布に近いかどうかによって判定するという立場で議論を展開している．そして部分列の抽出規則の例をいくつか挙げているが，そのなかでも等間隔抽出（系統サンプリング）が，実用上とくに重要であろう．たとえば待ち行列のシミュレーションでは，$Y=\{x_0, x_2, x_4, \cdots\}$ を客の到着間隔の決定に使い，$Z=\{x_1, x_3, x_5, \cdots\}$ をそれらの客に対する所要サービス時間の決定に使うという方法がよく使われる．この場合，原系列Xのみならず，部分列Y，Zも乱数列と見なせることが必要である．前記の Knuth, Tootill et al. が論じているのも，等間隔抽出による部分列のランダムネスの重要性についてである．

このように，等間隔抽出によって得られる部分列もまた乱数列と見なせるような原系列の重要性は，比較的よく認識されているが，そのような性質をもった系列を簡単なアルゴリズムによって具体的に発生させる問題は，これまであまり論じられてこなかったように思われる．そこで，Fushimi[4]は，M系列を用いてこのような性質をもった系列を発生させるアルゴリズムを設計する方法を提案している．この趣旨に沿って検討した結果得られたいくつかの実用的なアルゴリズムと，それらによって生成される系列の性質を紹介するのが，本論文の目的である．

2. 基本的概念

2.1 M 系列および Tausworthe 型系列

ガロア体 GF(2)上のp次の原始多項式

$$f(D)=1+c_1D+c_2D^2+\cdots+c_pD^p,\quad c_p=1 \tag{2.1}$$

を特性多項式とする漸化式

$$a_t=c_1a_{t-1}+c_2a_{t-2}+\cdots+c_pa_{t-p}\pmod 2 \tag{2.2}$$

を任意の初期条件 $(a_0, a_1, \cdots, a_{p-1})\neq(0,0,\cdots,0)$ の下に解いて得られる系列 $\{a_t\}$ のことをM系列と呼ぶ．これは $T=2^p-1$ を周期とする周期列である[6]．

$\{a_t\}$ を用いてlビット $(2\leq l\leq p)$ の2進小数の系列 $\{x_t\}$ を次のように構成する．

$$x_t=0.a_{\sigma t}a_{\sigma t+1}\cdots a_{\sigma t+l-1} \tag{2.3}$$

ここでσはTと互いに素な自然数である．$\{x_t\}$ も周期がTの周期列であり，Tausworthe 型系列と呼ばれている[11]．$\{x_t\}$ からの等間隔抽出によって得られる系列 $\{x_{nt};\ t=0,1,2,\cdots\}$ は，抽出間隔nが周期Tと互いに素であれば，やはり Tausworthe 型系列である．

2.2 多次元均等分布

一般に，lビットの2進数値をとる確率変数の列 $\{X_t\}$ に対して，k次元確率変数ベクトル$(X_t, X_{t+1}, \cdots, X_{t+k-1})$ の列が l ビットのあらゆる k-ベクトル値を等確率でとるとき，$\{X_t\}$ はk次均等分布をするという．Tausworthe 型系列 $\{x_t\}$ については，$\{(x_t, x_{t+1}, \cdots, x_{t+k-1})\}$ が一周期 $0\leq t\leq T-1$ の間に $(0, 0, \cdots, 0)$ 以外のlビットのあらゆる k-ベクトル値を

† Pseudorandom Number Generators Whose Subsequences are Multidimensionally Equidistributed by TORU IMAI and MASANORI FUSHIMI (Department of Mathematical Engineering and Instrumentation Physics, Faculty of Engineering, University of Tokyo).

†† 東京大学工学部計数工学科

2^{p-kl} 回ずつとるとき，$\{x_i\}$ は k 次均等分布をするという．この場合，$(0,0,\cdots,0)$ をとる回数は，$2^{p-kl}-1$ 回である．また $p-kl\geq 0$ であることが必要であるから，Tausworthe 型系列が達成できる均等分布の最大の次元は，

$$m=\lfloor p/l \rfloor^{*} \tag{2.4}$$

である．

任意の自然数 k について k 次均等分布をする系列は ∞ 次均等分布をするといい，乱数列としてきわめて好ましい性質をもっていることが知られている[7]．しかし，∞ 次均等分布をする系列を実際に発生させる簡単なアルゴリズムは得られていない．そこで，ある程度大きな k に対して k 次均等分布をする系列を擬似乱数列として使うことにする．

Tausworthe 型系列等，M系列を用いて構成される l ビットの擬似乱数列 $\{x_i\}$ が k 次均等分布をするための必要十分条件は，この系列の連続する k 個の要素，たとえば $x_0, x_1, \cdots, x_{k-1}$ を構成するM系列 $\{a_i\}$ の $k\times l$ 個の要素が一次独立であることである[1),4)]．ここで，一次独立の意味は次のとおりである．

M系列 $\{a_i\}$ の任意の要素は，必要ならば漸化式 (2.2) を繰り返し用いて，$a_0, a_1, \cdots, a_{p-1}$ の一次結合

$$e_0a_0+e_1a_1+\cdots+e_{p-1}a_{p-1}$$

の形に表現できる．特性多項式 (2.1) は原始多項式であるから，$\{a_i\}$ の各要素に対して定まる 0-1 ベクトル $(e_0, e_1, \cdots, e_{p-1})$ は一意である．これらの 0-1 ベクトルが GF(2) 上で一次独立のとき，対応するM系列の要素は一次独立であるという．

$k\times l$ 個の p 次元 0-1 ベクトルが与えられたとき，これらが一次独立かどうかは，たとえば，これらを並べて得られる $k\times l$ 行 p 列の行列に対してガウスの消去法の前進部分を行うことによって，$(kl)^2p$ のオーダの手間で判定できる[1)]．

多次元均等分布については，次の性質が成立する．

l ビットの擬似乱数列 $\{x_i\}$ が k 次均等分布をしているとき，

(1) $\{x_i\}$ は k 以下の任意の次元の均等分布をしている．

(2) $\{x_i\}$ の上位 $l'(l'\leq l)$ ビットに着目したときの均等分布の次元を k' とすると，$k'\geq k$ である．

とくにM系列を用いて構成される擬似乱数列については，

(3) $\{a_i\}$ の各要素に対する 0-1 ベクトルは (2.1) のみによって定まり，初期条件 $(a_0, a_1, \cdots, a_{p-1})$ の与え方には依らない．したがって均等分布の次元は，特性多項式および (2.3) の σ, l によって定まり，初期条件に依らない．

2.3 部分列の多次元均等分布の改善法[4)]

われわれが発生する乱数のビット数 l は，通常使用する計算機の語長によって定められる．たとえば，最近の大型計算機では，$l=32$ あるいは 31 とすることが多い．しかし実際に使用する場合，これだけの精度を必要とすることはまれで，多くの場合は上位のビットについて多次元均等分布をしていれば十分であろう*．

そこで 2.2 節の性質 (2) を利用し，上位ビットに着目して，Tausworthe 型系列の部分列 $\{x_{ni}\}$ を多次元均等分布の意味で改善する方法を考える．ここで，n としては実用上重要な，小さな自然数のみを考えることとし，それらを要素とする集合を N とする．おのおのの $n\in N$ に対し，$\{x_{ni}\}$ が何次元までの均等分布をしているか調べ，m 次均等分布**をしていない系列がある場合は，原系列を以下のように改善する．

上位何ビット目までに着目するなら，すべての $n\in N$ に対し $\{x_{ni}\}$ が m 次均等分布をするかという指標 l' を考え，(2.5) のように $\{x_i\}$ のビットを入れ替えて，l' ができるだけ大きい $\{x_i'\}$ を作る．

$$x_i'=0.a_{ni+j(0)}a_{ni+j(1)}\cdots a_{ni+j(l-1)} \tag{2.5}$$

ここで，$\{j(0), j(1), \cdots, j(l-1)\}$ は，$\{0,1,\cdots,l-1\}$ の置換であり，i には依らない．

すべての置換のなかで l' を最大にするものを求めるという問題は，数理計画法の分野で著名な集合被覆問題 (set covering problem) として定式化できる[4)]．なお，実際にビットの入れ換えを行う必要があるのは，系列 $\{x_i'\}$ の初期値設定の段階，すなわち $0\leq t\leq p-1$ に対応する値を構成する段階だけであり，乱数発生段階では $\{x_i\}$ と同じ漸化式を用いるので，余分な時間はかからないことを注意しておく．

3. 具体例

前章の方法を適用して得られるいくつかの擬似乱数発生アルゴリズムを示す．3.1 節および 3.2 節に示すものは，語長が 32 ビット以上からなる大型計算機向きであり，前者は原始 3 項式を使う方法，後者はきわ

* $\lfloor x \rfloor$ は x を超えない最大の整数を表す．

* 実際，現在よく使われている合同法乱数については，下位ビットは大変規則的なふるまいをしており，ランダムとは言いがたいが，実用上は支障ない場合も多いことが経験的に知られている．

** (2.4) より，n と周期 T が互いに素ならば，$\{x_{ni}\}$ は $m=\lfloor p/l \rfloor$ 次までの均等分布が可能である．

めて多数の項からなる原始多項式を使う方法である．また3.3節に示すものは，15, 16ビットの擬似乱数を発生させる方法で，語長の短いパソコン等で使うのに適している．いずれの方法においても，擬似乱数列の周期 T は素数であり，$N=\{n|1\leq n\leq 16\}$ とした．

表1から表4は，何次元までの分布が均等となっているかをおのおのの n について調べたものである．具体的には，$\{x_{nt}\}$ にたいして，$k=\lfloor p/l\rfloor$ として，2.2節で述べた方法を適用し，$x_0, x_n, \cdots, x_{n(r-1)}$ を構成するM系列の要素が一次独立であるような，r の最大値として求められる．

3.1 $f(D)=1+D^{32}+D^{521}$, $l=32$

この原始多項式は多くの文献で取り上げられており，$\sigma=32$ として得られる系列は文献5)で提案されている．$\sigma=32$ とした系列 $\{x_t\}$ と $\sigma=512$ とした系列 $\{y_t\}$ について，$\{x_{nt}\}$, $\{y_{nt}\}$ が何次元までの均等分布をしているか調べたところ，**表1**のようになった．なお，実際に使用する場合は上位ビットについての多次元均等分布が重要であるという観点から，$l=16$ とした場合についても付記した．この結果，$\sigma=32$ よりも512とした場合のほうがより高い次元までの分布が均等となっており，ここで注目している性質に関する限り，$\sigma=512$ のほうがよいという結論が得られた．

表1 $f(D)=1+D^{32}+D^{521}$ に基づく系列の部分列 $\{x_{nt}\}$ および $\{y_{nt}\}$ が均等分布をする最大次元
Table 1 Maximum orders of equidistribution of decimated sequences $\{x_{nt}\}$ and $\{y_{nt}\}$ generated by $f(D)=1+D^{32}+D^{521}$.

・$l=32$ ビット，$\lfloor p/l\rfloor=16$ 次元までの可能性がある

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\{x_{nt}\}$	16	16	11	16	13	13	15	16	16	13	12	15	10	15	13	16
$\{y_{nt}\}$	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16

・$l=16$, $\lfloor p/l\rfloor=32$

$\{x_{nt}\}$	16	24	21	20	13	27	27	22	26	29	28	30	28	32	32	32
$\{y_{nt}\}$	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32

l: The number of bits.
$\lfloor p/l\rfloor$: Maximum possible order of equidistribution.

そこで $\sigma=512$ とした原系列 $\{y_t\}$ について，上位から任意番目のビットまで着目したとき，何次元までの均等分布をしているかを調べたところ，**表2**のようになった．いずれの場合も，そのビット数で達成可能な最大の次元をほぼ達成しているので，Tootill et al.[13]の定義した asymptotic randomness という性質をほぼ満たしている．この意味からも，$\{y_t\}$ は好ましい系列ということができる．

2.3節の改善法を $\{x_t\}$ に適用すると，ある置換を行うことにより，$l'=19$ とすることができることが文献4)に記されているが，$\{y_t\}$ については，$l=32$, 16 の双方とも改善の必要はない．

最後に $\{y_t\}$ の発生法について述べる．M系列を用いて生成される Tausworthe 型系列および Lewis & Payne 型系列[9]に関する相反定理[3]により，$\{y_t\}$ は次式で定義される系列 $\{z_t\}$ と位相のずれを除いて一致する．

$$z_t=0.a_ta_{t+\tau}a_{t+2\tau}\cdots a_{t+31\tau} \tag{3.1}$$

ここで，τ は $2^{521}-1$ を法とする乗法に関する $\sigma=512$ の逆元，すなわち

$$\tau=2^{521}/512=2^{512} \tag{3.2}$$

である．したがって $\{Y_t\}=\{2^{32}y_t\}$ は漸化式

$$Y_t=Y_{t-521}\oplus Y_{t-32} \tag{3.3}$$

によって発生できる*．ただし⊕はビットごとの繰り上りなしの加法（排他的論理和）を示す．

この漸化式を用いるための初期値 $Y_0, Y_1, \cdots, Y_{520}$ の設定は，文献5)あるいは2)に記されている $\{x_t\}$ の初期値設定の方法を利用して行えばよい．なぜなら，$\{y_t\}$ は $\{x_t\}$ から16番目ごとに等間隔抽出することによって得られる系列だからである．

3.2 521次の原始279項式を用いる方法

一般に，項数の多い原始多項式を用いると乱数発生に時間を要するため，多くの場合は原始多項式として

* まず整数の乱数列 $\{Y_t\}$ を漸化式によって発生し，Y_t を正規化して y_t を求める．

表2 上位 s ビットに着目したときの $\{y_t\}$ の均等分布の最大次元
Table 2 Maximum orders of equidistribution of the leading s bits of $\{y_t\}$.

ビット数 s	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
実際の最大次元	260	170	130	102	81	72	64	57	49	41	40	37	33	32	32	30	26	26	24	22	22	22	19	18	17	16	16	16	16	16	16
理論上の最大次元	260	173	130	104	86	74	65	57	52	47	43	40	37	34	32	30	28	27	26	24	23	22	21	20	20	19	18	17	17	16	16

First row: The number of bits s.
Second row: Actual order of equidistribution.
Third row: Maximum possible order of equidistribution.

表 3　521 次 279 項の原始多項式に基づく系列[10]の部分列が均等分布をする最大次元

Table 3　Maximum orders of equidistribution of decimated sequences generated by a 521st-degree primitive polynomial with 279 terms.

$l=32$, $\lfloor p/l \rfloor=16$

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
次元	16	16	16	16	16	16	16	16	16	16	16	16	16	16	13	16

表 4　$f(D)=1+D^{15}+D^{127}$ に基づく系列の部分列が均等分布をする最大次元

Table 4　Maximum orders of equidistribution of decimated sequences generated by $f(D)=1+D^{15}+D^{127}$.

(1)　$l=\sigma=16$, $\lfloor p/l \rfloor=7$

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
次元	7	7	7	7	7	7	1	7	7	7	7	7	7	4	7	7

(2)　$l=\sigma=15$, $\lfloor p/l \rfloor=8$

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
次元	8	8	6	8	7	8	7	8	8	8	8	8	8	8	8	8

3 項式を用いる．しかし，項数の多い原始多項式に基づきながらも 3 項式と同じ時間で発生できる方法が文献 10)で提案されている．このなかに記されている 521 次の原始 279 項式に基づく方法を用いて，$l=\sigma=32$ として得られる系列の部分列が何次元までの均等分布をするかを調べた結果，表 3 のようになった．$n=15$ の系列のみが 16 次均等分布をしていないが，2.3 節の改善法を試みたところ，ビットの入れ替えを行わないで上位 27 ビットに着目すると 16 次均等分布をしていることが判明したので，実用上は差し支えないであろう．

3.3　$f(D)=1+D^{15}+D^{127}$

(1)　$l=\sigma=16$ の場合

部分列が何次元までの均等分布をするかを調べたところ，表 4 (1)のようになった．$n=7$ の系列は均等分布の次元が極端に低いが，2.3 節の改善法を試みると，ビットの入れ替えを行わないで上位 15 ビットに着目すれば任意の $n\in N$ について 8 次均等分布をすることが判明した*．したがって乱数の精度が 15 ビット以下でよい場合には，好ましい系列であるといえる．

系列 $\{X_t\}=\{2^{16}x_t\}$ は漸化式

$$X_t=X_{t-127}\oplus X_{t-16} \tag{3.4}$$

を用いて発生できる．初期値 $X_0, X_1, \cdots, X_{126}$ は次のように設定すればよい．

1°　$X_0, X_1, \cdots, X_7$ を任意に与える（ただしすべてが 0 にはならないようにする）．

2°　X_7 を次式により更新する．

$$X_7=L^1(R^1X_7)\oplus R^{15}(X_0\oplus X_7) \tag{3.5}$$

3°　X_t $(8\le t\le 126)$ を次式により計算する．

$$\begin{aligned} W&=L^1(M^{15}(X_{t-1}\oplus X_{t-8}))\\ X_t&=R^{15}(X_{t-7}\oplus W)\oplus W \end{aligned} \tag{3.6}$$

ここで，L^j，R^j はそれぞれ j ビットの左，右論理シフトを表し，M^{15} は下位 15 ビットを取り出すマスク演算を表す．

(2)　$l=\sigma=15$ の場合

これは Tootill et al.[12] が推奨している系列のひとつであり，$\{X_t\}=\{2^{15}x_t\}$ は漸化式

$$X_t=X_{t-128}\oplus X_{t-127}\oplus X_{t-16} \tag{3.7}$$

により発生できる（導出方法は付録参照）．

部分列が何次元までの均等分布をするかを示すのが表 4 (2)である．置換 $\{j(0), j(1), \cdots, j(14)\}=\{0, 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 3, 12\}$ を行うと，上位 13 ビットに着目したとき任意の $n\in N$ について 8 次均等分布をさせることができる．しかし(1)の系列と比べると，均等分布の次元および発生速度の点で劣っている．

4. ま と め

擬似乱数列を使って行う実験において，原系列の性質だけではなく，等間隔抽出によって得られる部分列の性質も実験結果に大きな影響を及ぼす場合がしばしばある．このような場合でも，擬似乱数の利用者は部分列の性質を吟味しないで使うことが多い．したがって実用上しばしば使用される部分列の性質があらかじめ理論的に保証されているような擬似乱数列の必要性は高いと考えられる．

そこで本論文では，部分列の性質もすぐれた擬似乱数列の発生法として，語長の長い計算機向きの方法と，短い計算機向きの方法を紹介し，その均等分布の次元を示した．

参 考 文 献

1) Fushimi, M. and Tezuka, S.: The k-Distribution of the Generalized Feedback Shift Register Pseudorandom Numbers, *Comm. ACM*, Vol. 26, No. 7, pp. 516-523 (1983).

* $l=16$ のとき，$m=\lfloor p/l \rfloor=7$ であるが，$l=15$ のときは，$m=8$ となる．

2) 伏見正則：一様乱数の発生法，情報処理，Vol. 24, No. 4, pp. 367-371 (1983).
3) 伏見正則：M系列に基づく乱数発生法に関する相反定理とその応用，情報処理学会論文誌，Vol. 24, No. 5, pp. 576-579 (1983).
4) Fushimi, M.: Designing a Uniform Random Number Generator Whose Subsequences are k-Distributed, Technical Reports, METR 83-7, Department of Mathematical Engineering and Instrumentation Physics, Faculty of Engineering, University of Tokyo (1983).
5) 伏見正則，手塚 集：多次元分布が一様な擬似乱数列の生成法，応用統計学，Vol. 10, No. 1, pp. 151-163 (1981).
6) Golomb, S. W.: *Shift Register Sequences*, 224 pp., Holden-Day, San Francisco (1967).
7) Knuth, D. E.: *The Art of Computer Programming*, Vol. 2, *Seminumerical Algorithms*, 2nd ed., Addison-Wesley, Reading, Mass. (1981).
8) Kolmogorov, A. N.: On Tables of Random Numbers, *Sankhya*, Vol. 25 A, pp. 369-376 (1963).
9) Lewis, T. G. and Payne, W. H.: Generalized Feedback Shift Register Pseudorandom Number Algorithms, *J. ACM*, Vol. 20, No. 3, pp. 456-468 (1973).
10) 斎藤隆文，伏見正則，今井 徹：多数項の原始多項式に基づくM系列乱数の高速発生法，情報処理学会論文誌，Vol. 26, No. 1, pp. 148-152 (1985).
11) Tausworthe, R. C.: Random Numbers Generated by Linear Recurrence Modulo Two, *Math. Comput.*, Vol. 19, pp. 201-209 (1965).
12) Tootill, J.P.R., Robinson, W. D. and Adams, A. G.: The Runs Up-and-Down Performance of Tausworthe Pseudo-Random Number Generators, *J. ACM*, Vol. 18, No. 3, pp. 381-399 (1971).
13) Tootill, J. P. R., Robinson, W. D. and Eagle, D. J.: An Asymptotically Random Tausworthe Sequence, *J. ACM*, Vol. 20, No. 3, pp. 469-481 (1973).

付録 (3.7)式の導出方法

$$a_t = a_{t-127} + a_{t-15} \pmod 2 \qquad (1)$$

$Da_t = a_{t-1}$ なる演算子 D を用いると(1)は，

$$a_t = D^{127}a_t + D^{15}a_t \qquad (2)$$

と書ける (mod 2 は省略する)．(2)式は任意の t に対して成立するから，

$$1 = D^{127} + D^{15} \qquad (3)$$

ここで $\sigma=15$ であるから，$\{X_t\}$ の従う漸化式は $\{a_t\}$ を15番目ごとに等間隔抽出した系列の従う漸化式に等しい．$Dx = D^{15}$ として，(3)より

$$(D^{127})^{15} = (D^{15}+1)^{15}$$

$$Dx^{127} = (Dx+1)^{15} = Dx^{15} + Dx^{14} + \cdots + 1 \qquad (4)$$

(4)の両辺に Dx をかけて，

$$Dx^{128} = Dx^{16} + Dx^{15} + \cdots + Dx \qquad (5)$$

(4)，(5)を加えて，

$$Dx^{128} + Dx^{127} + Dx^{16} + 1 = 0$$

よって(3.7)式が得られる．

(昭和59年 6 月20日受付)
(昭和59年10月18日採録)