# デタラメさの効用と、１＋１＝０の世界

松本 眞（東京大学数理科学研究科）

2012年1月25日　東京大学学術俯瞰講義

**第壱話**　デタラメさの効用：意外なところでランダムネスのお世話に

**第弐話**　デタラメさを生み出すのは意外に難しい：（デタラメ禅問答）

**第参話**　近現代数学によるデタラメさの生成（メルセンヌ・ツイスター）

**第参話 近現代数学によるデタラメさの生成**（人之章）

前回の復習：乱数列の数学的定義はあった。
が、「計算機で効率よく発生できる疑似乱数」
という観点から満足行くものではない。

現在広く使われている疑似乱数：

- 周期が長い
- 高次元空間内での均等分布性が保証されている
  （高い一様性と独立性）
- 使用してみて経験的に問題がなかった
- 高速に生成可能

といった、妥協の産物として現在の疑似乱数発生法はある。
小学校で習う「循環小数」の発展形が主流。

## 分数と循環小数

$1 \div 7$ の 10 進小数展開の余りの列

$x_1, x_2, \ldots$

は、$x_1 = 1$ として以下の法則で求まる。

$$
\begin{aligned}
1 \times 10 &= 10,\ \div 7 = 1 \text{ 余り } 3 =: x_2 \\
3 \times 10 &= 30,\ \div 7 = 4 \text{ 余り } 2 =: x_3 \\
2 \times 10 &= 20,\ \div 7 = 2 \text{ 余り } 6 =: x_4 \\
6 \times 10 &= 60,\ \div 7 = 8 \text{ 余り } 4 =: x_5 \\
4 \times 10 &= 40,\ \div 7 = 5 \text{ 余り } 5 =: x_6 \\
5 \times 10 &= 50,\ \div 7 = 7 \text{ 余り } 1 =: x_7
\end{aligned}
$$

$$
\begin{array}{r}
0.\dot{1}42857\dot{1} \\
7\,\overline{)\,10} \\
\underline{7} \\
30 \\
\underline{28} \\
20 \\
\underline{14} \\
60 \\
\underline{56} \\
40 \\
\underline{35} \\
50 \\
\underline{49} \\
10 \\
\underline{7} \\
30
\end{array}
$$

**定義**　整数$a, N$に対し、

$$a \mod N$$

で、

$a$を$N$で割ったあまり（$0, 1, \ldots, N-1$のいずれか）

を表す。先の余りの列は

$$\begin{aligned}
x_1 &= 1 \\
x_2 &= x_1 \times 10 \mod 7 = 10^1 \mod 7 = 3 \\
x_3 &= x_2 \times 10 \mod 7 = 10^2 \mod 7 = 2 \\
&\cdots \\
x_{n+1} &= x_n \times 10 \mod 7 = 10^n \mod 7 = \ldots
\end{aligned}$$

## 定理

$N$を10と互いに素な自然数とする。
$1/N$の小数展開の周期は$N-1$以下で、
$1 \;=\; 10^P \mod N$となる最小の自然数
$P \geq 1$となる。

**証明**：余りの種類は$N$種。そのうち0は出てこないから、周期は$N-1$以下。
（余りに0が小数点$n$桁目に出てきたら、$10^n$を$N$が割り切ることが筆算の形から分かる)。
後半については、循環するのは
$10^n \mod N$が1となったとき。
筆算を見よ。

$0.\dot{1}42857\dot{1}$

```
7 ) 10
     7
    30
    28
     20
     14
      60
      56
       40
       35
        50
        49
         10
          7
         30
```

**線形合同法** Linear Congruential Generator
(LCG, Lehmer '60)

- ある整数 $x_1$ を初期シードとして選ぶ。
- 次の例のような漸化式により $x_2, x_3, \ldots$ を次々に生成：

$$x_{n+1} = x_n \times 1103515245 + 12345 \bmod 2^{32}.$$

例 $x_1 = 3$ ならば

$$\begin{aligned}
3 \times 1103515245 + 12345 &= 3310558080 \ (\mathrm{mod} 2^{32}) \to 3310558080 = x_2 \\
3310558080 \times 1103515245 + 12345 &= 3653251310737941945 \ (\mathrm{mod} 2^{32}) \to 465823161 = x_3 \\
465823161 \times 1103515245 + 12345 &= 514042959637601790 \ (\mathrm{mod} 2^{32}) \to 679304702 = x_4 \\
679304702 \times 1103515245 + 12345 &= 749623094657194335 \ (\mathrm{mod} 2^{32}) \to 2692258143 = x_5
\end{aligned}$$

擬似乱数のメリット:

- 漸化式と初期シードを記録しておけば、
  誰でも同じ数列を再現できる
- 高速で低コスト

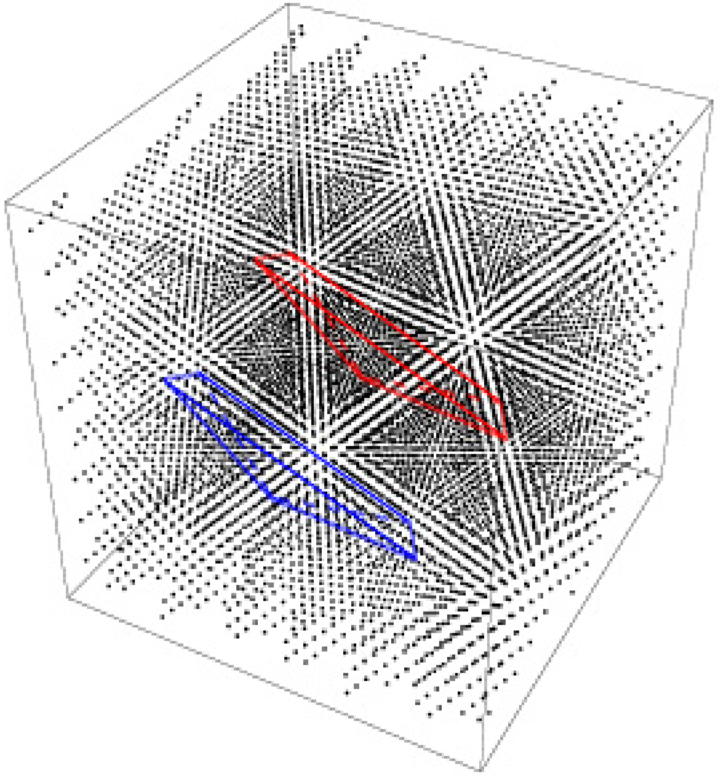
問題点:「乱数と呼んでいいのか」...$\Leftarrow$（実用的）定義の不在

von Neumann 「漸化式で乱数を作るのはある種の罪」

"Anyone who considers arithmetical methods of
producing random digits is, of course, in a state of sin"

たとえば:

- 先の線形合同法による数列の周期は、
  初期シードの選び方によらず$2^{32}$。
  (前回の平方採中法の例では、初期値により周期が1や4にもなりえた)
- この生成法は、70年代から80年代にかけて
  ANSI-Cなどの標準擬似乱数であった。
- 現代のパソコンは数分で$2^{32}$個の乱数を使ってしまう
- 生成される数列はかなり乱数に見えるが、
  数千万個の出力を使うと、非乱数性が現れてくる

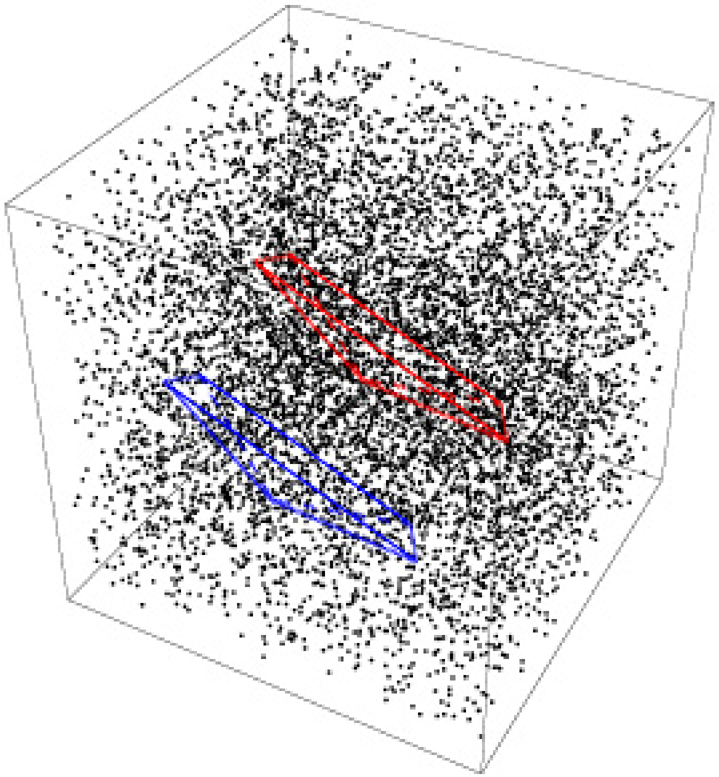


この線形合同法による$2^{32}$個の
全周期3次元ランダム点プロット
（70倍拡大図）

赤領域体積$\simeq \frac{62}{7253} = 8.5 \times 10^{-3}$.

青領域体積$\simeq \frac{45}{7253} = 6.2 \times 10^{-3}$.

(真の値: $8.333 \cdots \times 10^{-3}$)



$1+1=0$の数学に基づく

Mersenne Twister(松本-西村 '98)

による3次元ランダム点プロット

赤領域 $\simeq \frac{61}{7248} = 8.4 \times 10^{-3}$.

青領域 $\simeq \frac{64}{7248} = 8.8 \times 10^{-3}$.

(真の値: $8.333\cdots \times 10^{-3}$)

## $1+1=0$の世界での多項式

### 二元体$\mathbb{F}_2$

$\mathbb{F}_2 := \{0,1\}$とおく。

0,1の掛け算は普通に定義して、$\mathbb{F}_2$からはみ出ない。
足し算は$1+1=2$だけが$\mathbb{F}_2$からはみ出してしまうので、

$$1+1=0$$

と定義する（2で割ったあまりを見ている）。

1+1=0から1を移項して

$$1=-1.$$

## $\mathbb{F}_2$ 多項式 $\mathbb{F}_2[t]$

$$\mathbb{F}_2[t] := \left\{ \sum_{i=0}^{n} a_i t^i \mid a_i \in \mathbb{F}_2,\ n \in \mathbb{N} \right\}$$

を考える。係数が $0$ または $1$ の多項式のことである。
掛け算・足し算は通常の多項式同様

$$\begin{aligned} &(t+1) \times (t+1) = t(t+1) + 1(t+1) \\ &= t^2 + t + t + 1 = t^2 + 1 \end{aligned}$$

といった具合に計算できる。
（$1+1=0$ より $t+t=(1+1)t=0$。）

係数のみを表記することにして、「$t$進くらい取り」で

$$t^3 + t^2 + 1 = 1t^3 + 1t^2 + 0t + 1 = 1101$$

と表わすことにする。

$\mathbb{F}_2$多項式での和差積商は、

「繰り上がり・繰り下がりのない世界での計算」

になる。

$$\begin{array}{rrr} & 1 & 1 \\ \times & 1 & 1 \\ \hline & 1 & 1 \\ 1 & 1 & \\ \hline 1 & 0 & 1 \end{array} \qquad \begin{array}{rrr} & t & +1 \\ \times & t & +1 \\ \hline & t & +1 \\ t^2 & +t & \\ \hline t^2 & +0t & +1 \end{array}$$

## 形式べき級数 ($1+1=0$ での無限小数)

$\mathbb{F}_2[t]$ の世界で $1 \div (t^3+t^2+1) = 1 \div 1101$ を小数展開すると

$$0.\dot{0}0111010\dot{0}$$

```
              0.0 0 1 1 1 0 1 0
1 1 0 1 ) 0 0 0 1
          0 0 0 0
            0 0 1 0
            0 0 0 0
              0 1 0 0
              0 0 0 0
                1 0 0 0
                1 1 0 1
                  1 0 1 0
                  1 1 0 1
                    1 1 1 0
                    1 1 0 1
                      0 1 1 0
                      0 0 0 0
                        1 1 0 0
                        1 1 0 1
                          0 0 1 0
                          0 0 0 0
                            0 1 0 0
```

検算：

$$\begin{array}{rr}
 & 0.00111010011101\cdots \\
\times & 1101 \\
\hline
 & 0.00111010011101\cdots \\
 & 00.00000000000000\cdots \\
 & 000.11101001110100\cdots \\
 & 0001.11010011101001\cdots \\
\hline
 & 0001.00000000000000\cdots
\end{array}$$

$1 \div 1101 = 0.00111010011101 \cdots$ は次の省略形である。

$$1 \div (t^3 + t^2 + 1) =$$
$$0 + 0t^{-1} + 0t^{-2} + 1t^{-3} + 1t^{-4} + 1t^{-5} + 0t^{-6} + 1t^{-7} \cdots$$

この無限小数のような式を$\mathbb{F}_2$**形式べき級数**と言い、
右辺を左辺の形式**べき級数展開**という。

**定理**：定数項が$1$で次数$n$の$\mathbb{F}_2$多項式$f(t) = t^n + \cdots + 1$に対し、$1/f(t)$のべき級数展開の係数は循環し、
周期は$2^n - 1$以下。
周期は$1 = t^P \mod f(t)$となる最小の自然数$P \geq 1$。

証明：余りの種類が$2^n$で、そのうち$0$はあらわれないから。
後半は、筆算を見よ。

$1 \div 1101 = 0.x_1x_2x_3x_4\cdots$ とおくと$\mathbb{F}_2$での漸化式

$$x_{n+3}+x_{n+2}+x_n = 0 \quad (n \geq 1) \quad \text{を満たす}$$

理由：

$$\begin{array}{cccccccccccc}
 & & & & 0 & . & x_1 & x_2 & x_3 & x_4 & x_5 & \cdots \\
\times & 1 & 1 & 0 & 1 & & & & & & & \\
\hline
 & & & & 0 & . & x_1 & x_2 & x_3 & x_4 & x_5 & \cdots \\
 & & & 0 & 0 & . & 0 & 0 & 0 & 0 & 0 & \cdots \\
 & & 0 & x_1 & x_2 & . & x_3 & x_4 & x_5 & x_6 & x_7 & \cdots \\
 & 0 & x_1 & x_2 & x_3 & . & x_4 & x_5 & x_6 & x_7 & x_8 & \cdots \\
\hline
 & 0 & 0 & 0 & 1 & . & 0 & 0 & 0 & 0 & 0 & \cdots
\end{array}$$

したがって、漸化式

$$x_{n+3} = x_{n+2} + x_n \; (n \geq 1), \quad x_1 = 0, x_2 = 0, x_3 = 1$$

を解くことで上の循環「小数」が得られる。

$$x_{n+3} = x_{n+2} + x_n \ (n \geq 1), \quad x_1 = 0, x_2 = 0, x_3 = 1$$

で $x_1 x_2 x_3 x_4 \cdots$ を計算すると

0011

00111

001110

0011101

00111010

001110100

0011101001

00111010011

001110100111

## 周期の最大性と均等分布性

上の循環「小数」を三つずつ組にしてみると、000 以外の $2^3 - 1$ 通りのパターンを一回ずつ一周期にとる。

$$0011101001110 1$$

$$001, 011, 111, 110, 101, 010, 100, 001$$

証明：漸化式が３階だから、連続する３個の数のパターンにより以後の数列は決まってしまう。000 はあらわれないから、周期が $2^3 - 1$ ならば他の全パターンが一個ずつ現れる。

## 定理

1. 任意の自然数 $n$ に対し、次数 $n$ の $\mathbb{F}_2$ 多項式 $f(t)$ であって $1/f(t)$ のべき級数展開の係数の周期が $2^n-1$ となるものがたくさん存在する。

2. このとき、連続する $n$ 個の0-1の並びは、

   $000\cdots0$ を除いてすべて一回ずつ一周期に現れる。

性質1を満たす $f(t)$ を $\mathbb{F}_2$ **原始多項式**という。
このとき $f(t)$ で割り切れない任意の $\mathbb{F}_2$ 多項式 $g(t)$ に対して、
$g(t)/f(t)$ の「小数部分」は周期 $2^n-1$ となる。

性質2は均等分布性と呼ばれ、数列のバランスの良さを示している。疑似乱数に用いるのに適している。

原始多項式の例：
$t^3+t^2+1$(上でみた、周期$2^3-1=7$)
$t^{31}+t^3+1$(周期$2^{31}-1=2147483647$)
$t^{607}+t^{273}+1$(周期$2^{607}-1=5.3\times10^{182}$)
など多数が知られている。

したがって、たとえば

$$x_{n+607}=x_{n+273}+x_n$$

を使って周期が$2^{607}-1$で、連続する607項が均等分布する数列を、きわめて高速に作り出すことができる。

**実験では確かめられない**が、**証明できる**。数学の強み。
このような擬似ランダムビット列生成法を**Tausworthe法**('65)という。

# ベクトル化：$\mathbb{F}_2$多項式から$\mathbb{F}_2$線形変換へ

**GFSR** (Lewis-Payne '73) 計算機ワード長の$\mathbb{F}_2$ベクトル列を

$$\vec{x}_{n+p} := \vec{x}_{n+q} + \vec{x}_n$$

で生成($+$は$\mathbb{F}_2$ベクトルとしての和)
整定数$p, q$をうまく選ぶと周期$2^p - 1$にできる

- 各桁はTausworthe法で生成される数列に一致
- 高速だが、各桁の間に情報のやり取りがない
- 乱数性に問題あり（特にランダムウォークで）

**Twisted GFSR** (松本-栗田良春 '92, '94):
Twister と呼ぶ$\mathbb{F}_2$係数正方行列$A$を導入する：

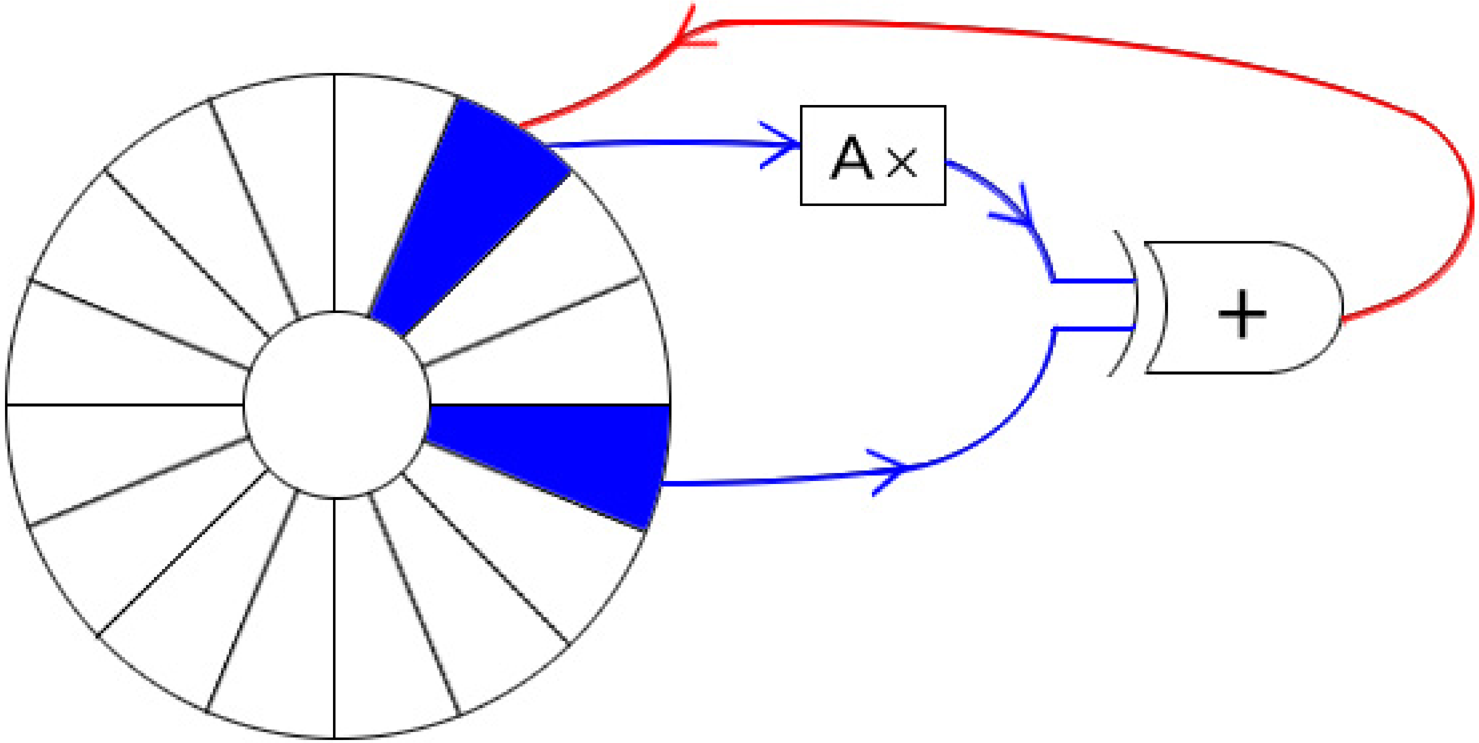
$$\vec{x}_{n+p} = \vec{x}_{n+q} + \vec{x}_n A.$$

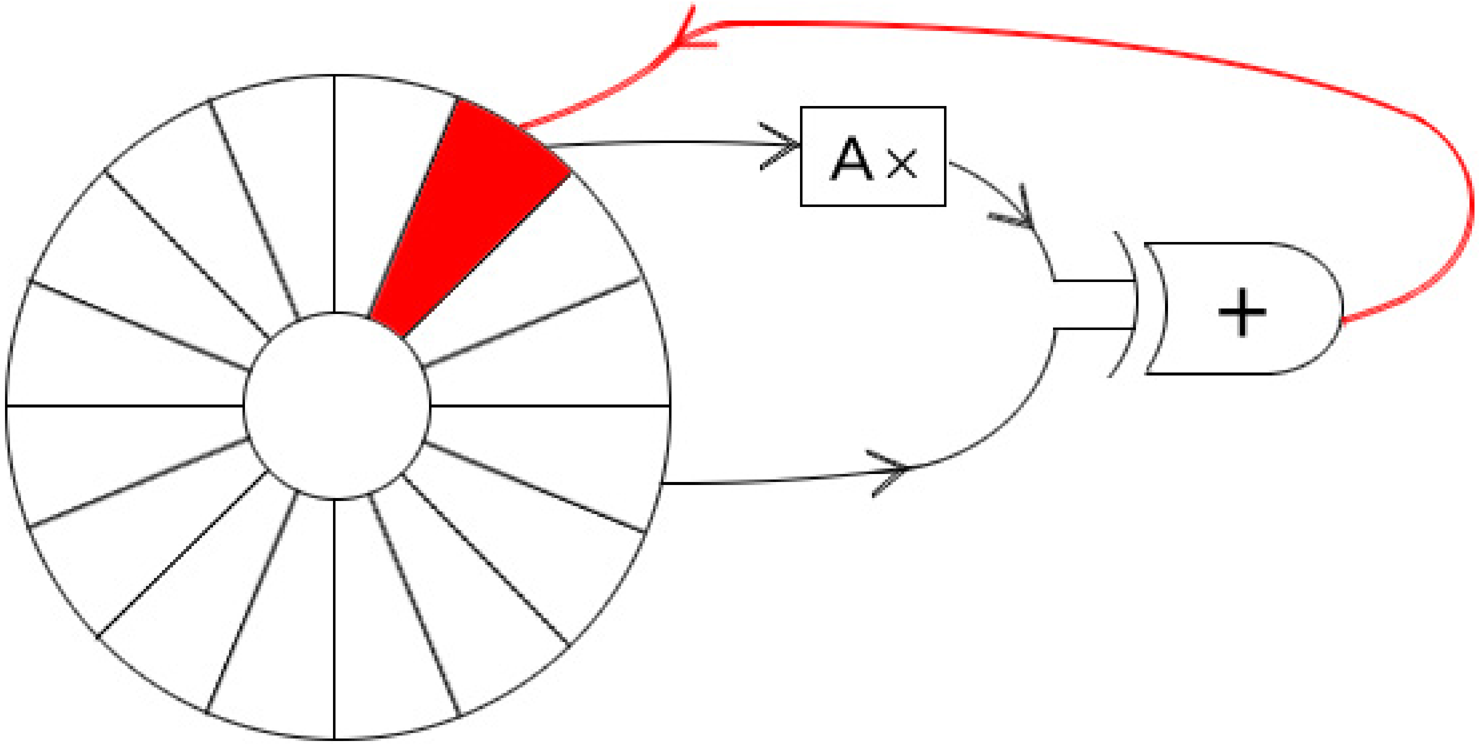
- $A$ は桁の間の情報を混ぜる
  $\Rightarrow$ より長周期: $2^{32p} - 1$が達成可能(32:ワード長)
- $A$は次のようなものを選ぶ:定数ベクトル$\vec{a}$により

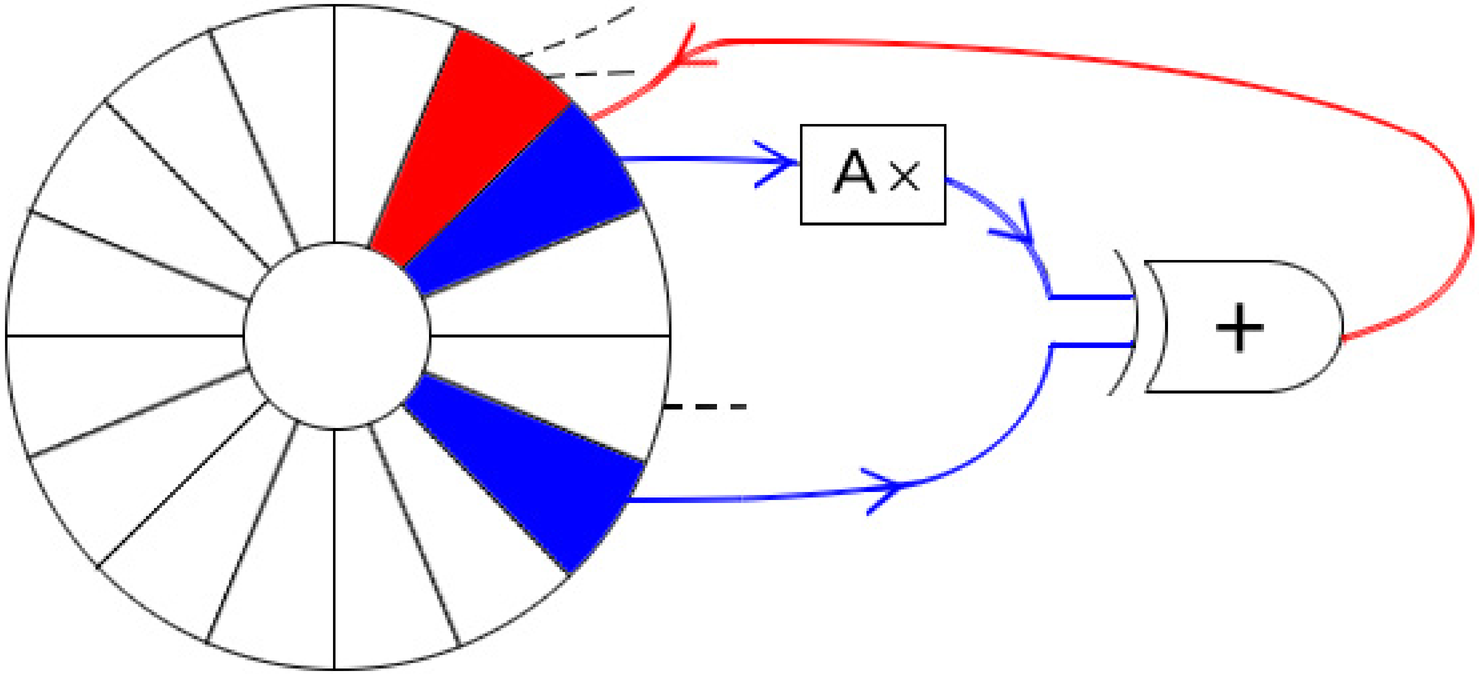
$$\vec{x}A = \begin{cases} \mathrm{shiftright}(\vec{x}) & (\vec{x}\text{の最下位ビットが}0\text{の場合}) \\ \mathrm{shiftright}(\vec{x}) + \vec{a} & (\vec{x}\text{の最下位ビットが}1\text{の場合}) \end{cases}$$

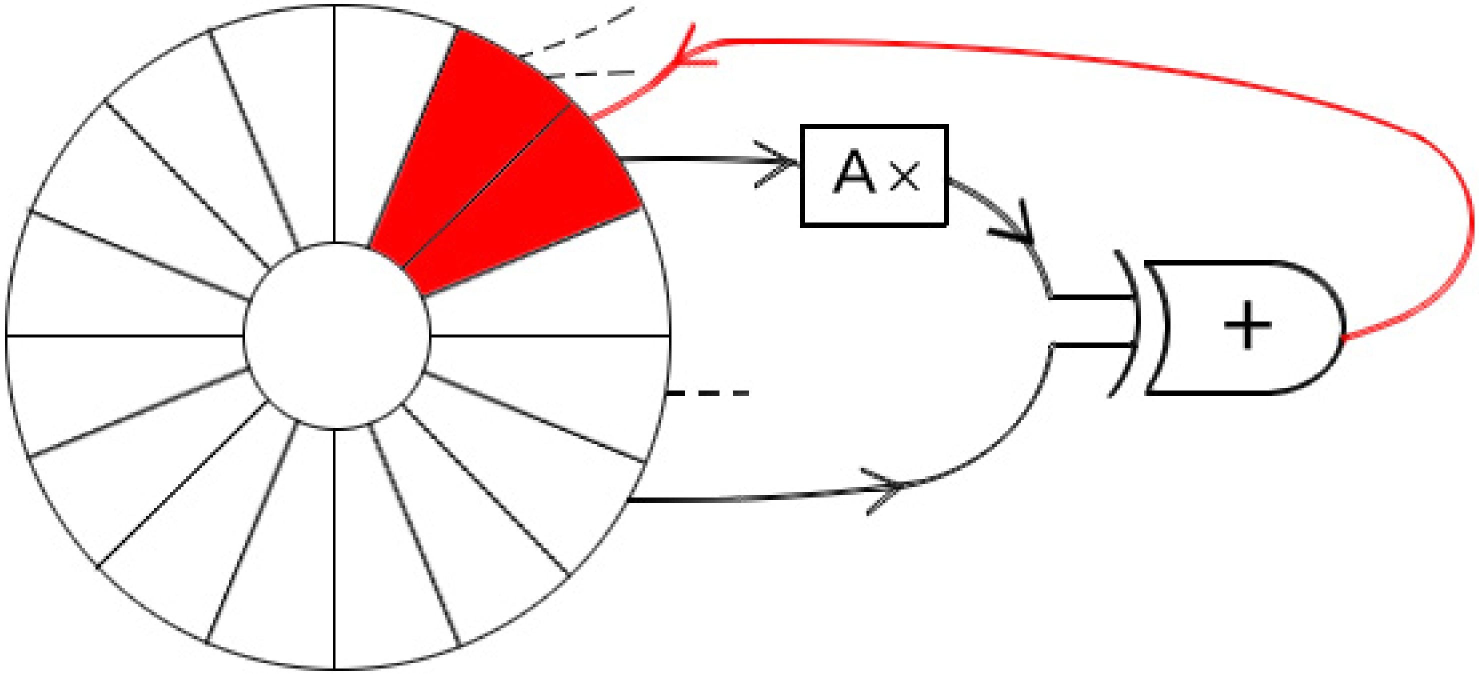
  $\Rightarrow$高速に計算可能,　十分一般: $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ a_3 & a_2 & a_1 & a_0 \end{pmatrix}.$
- 高次元均等分布性を改善するため$\vec{x}_n T$を出力('94)

# なぜMT(TGFSR)は高速か？ $\Leftarrow$ ラウンドロビン実装




A×
+


A×
+


A×
+

## メルセンヌツイスター疑似乱数(MT法、松本-西村1998)

$\mathbb{F}_2$成分の32次元ベクトルの列を次の漸化式で生成し、

$$\vec{x}_{n+624} = \vec{x}_{n+q} + B\vec{x}_{n+1} + C\vec{x}_n$$

$T\vec{x}_n$を出力列とする。ここに$B, C, T$はうまく選ばれた
３２次正方行列。

1. 周期は$2^{19937} - 1 > 10^{6000}$(24番目のメルセンヌ素数)

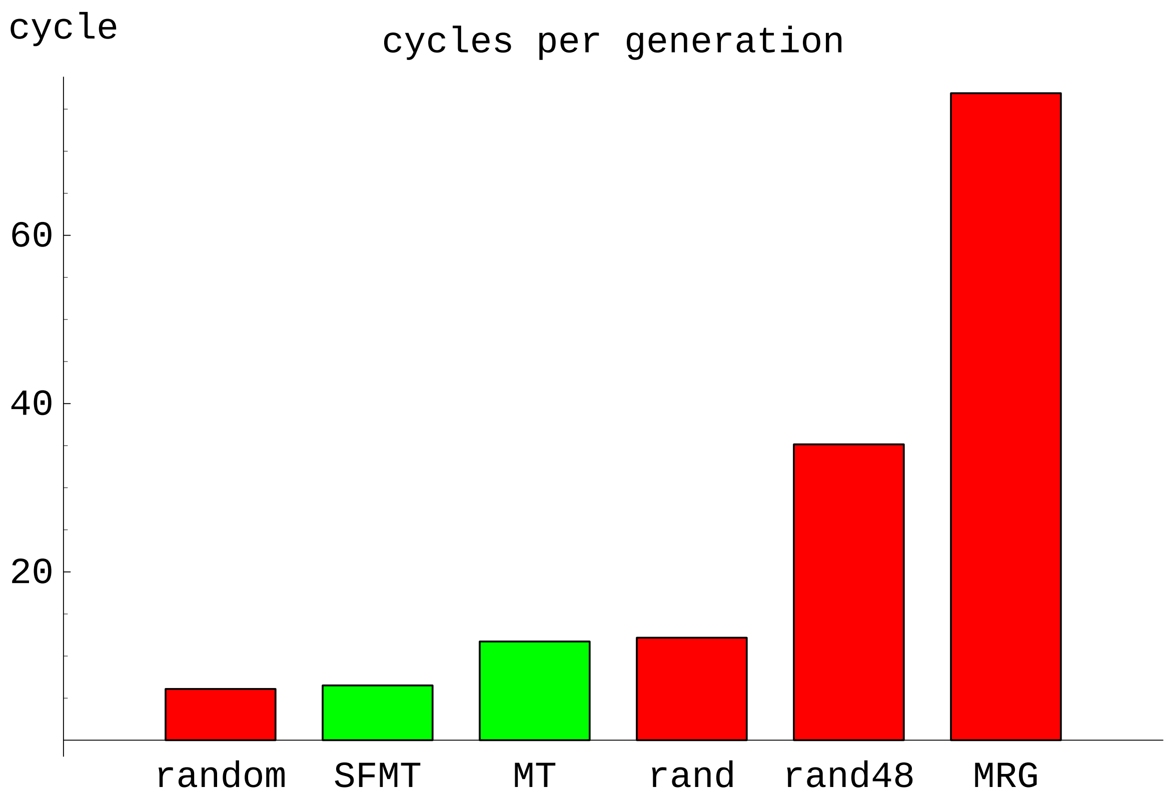
2. 出力列は、623次元空間内で均等分布している

3. 高位ビットはさらに高次元
（例えば上位3ビットは6240次元まで均等分布）

4. それまでの生成法よりも数倍高速に生成

高次元の互除法（格子簡約）など、$1 + 1 = 0$の世界
での代数・幾何を利用して周期や均等分布の次元を求めた。

# 速度比較




**random**: ラグ付きフィボナッチ周期 $\sim 2^{63}$

**rand**: LCG 周期 $2^{32}$

**SFMT** SIMD Fast MT

**MT**: Mersenne Twister
周期 $2^{19937}-1$

**rand48**: LCG 周期 $2^{48}$

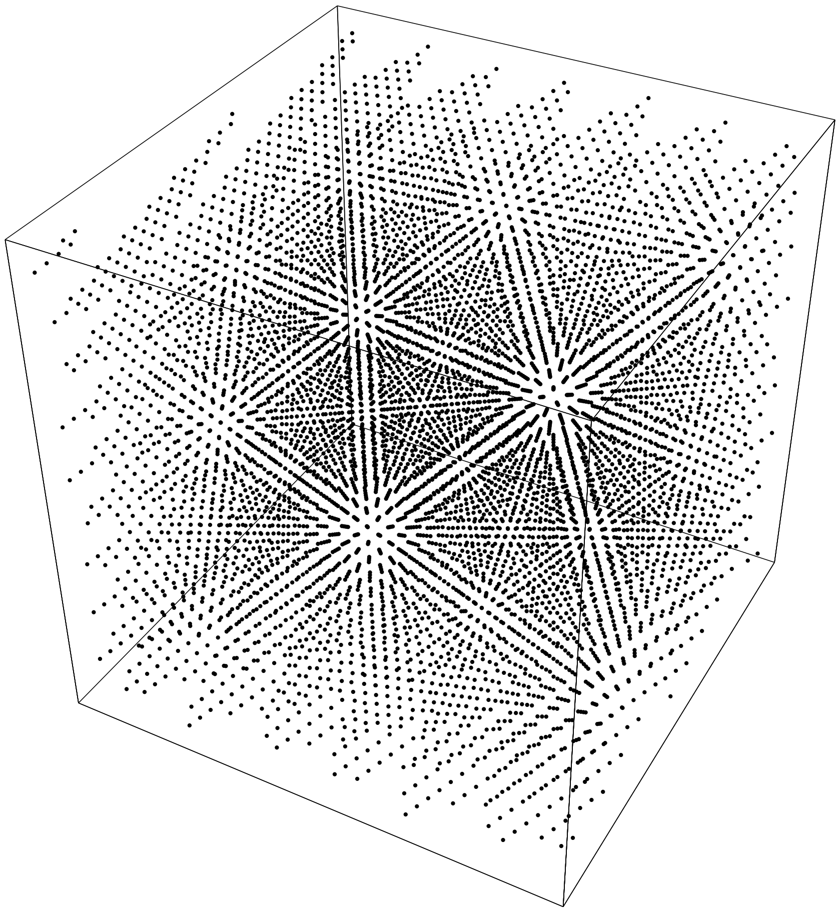
**MRG**: L'Ecuyer
周期 $\sim 2^{186}$

## 数学の予期せぬ効用

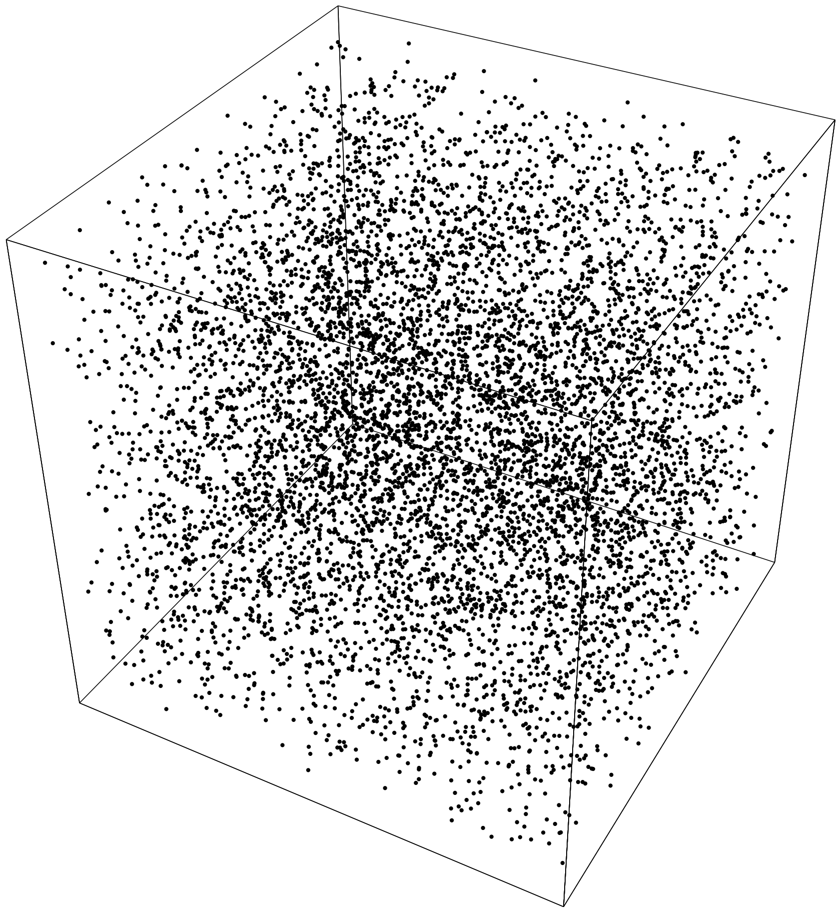
- $1+1=0$ の数学の研究は、ガロア (1830 ごろ) に遡る
- 当時は応用の見えなかった純粋数学が、
  現在実用されている。
- $\mathbb{F}_2[t]$ と整数は良く似ており、代数・幾何が展開できる。
  前者が扱いやすい（現代整数論の指導原理の一つ）

# 終わりに：注意喚起

現在も、品質の悪い擬似乱数が広く使われている



ANSI-C標準擬似乱数('70-'90)



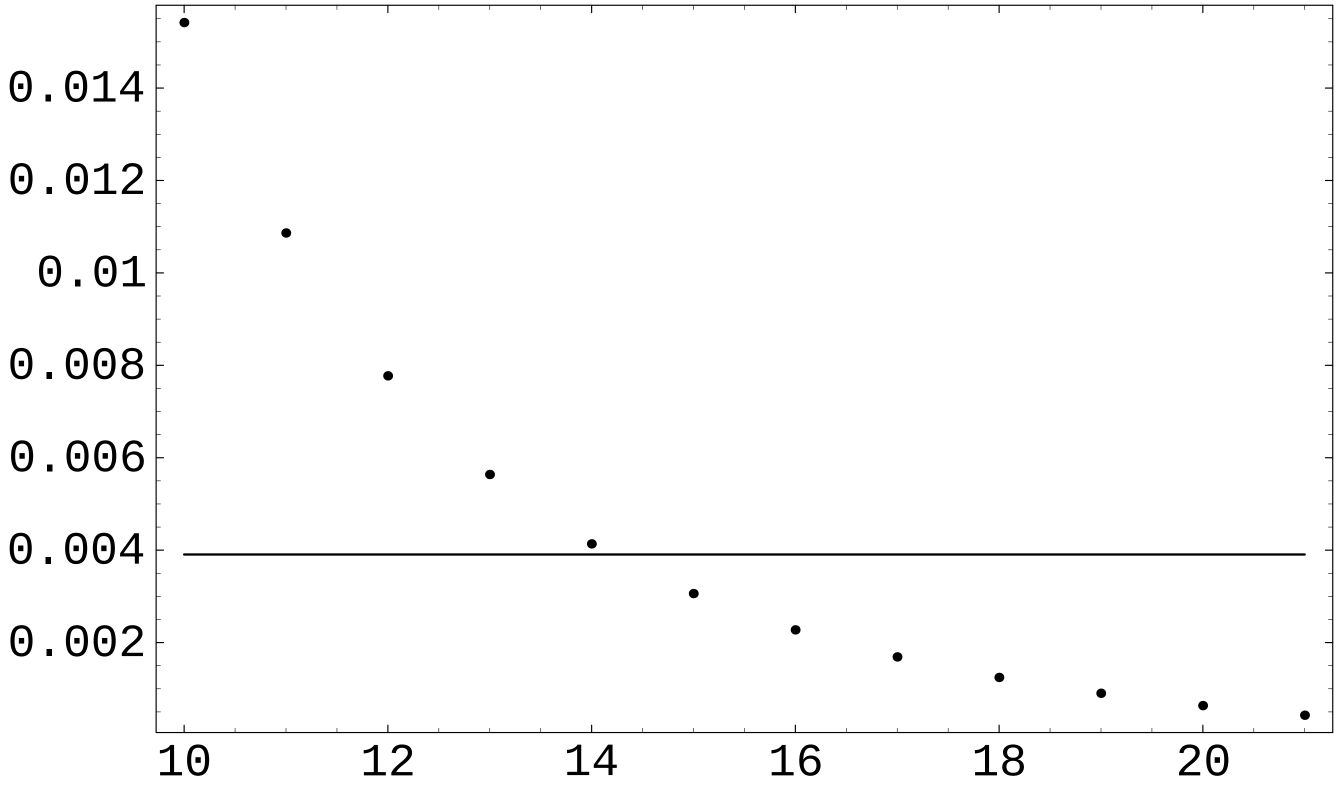
mt19937('98)

## 新しく提唱されたものの中にも悪いものが多い

`random`: ’90-現在UNIX系C言語での標準的擬似乱数の最下位ビット0-1を見る。(原本博史-M ’07)

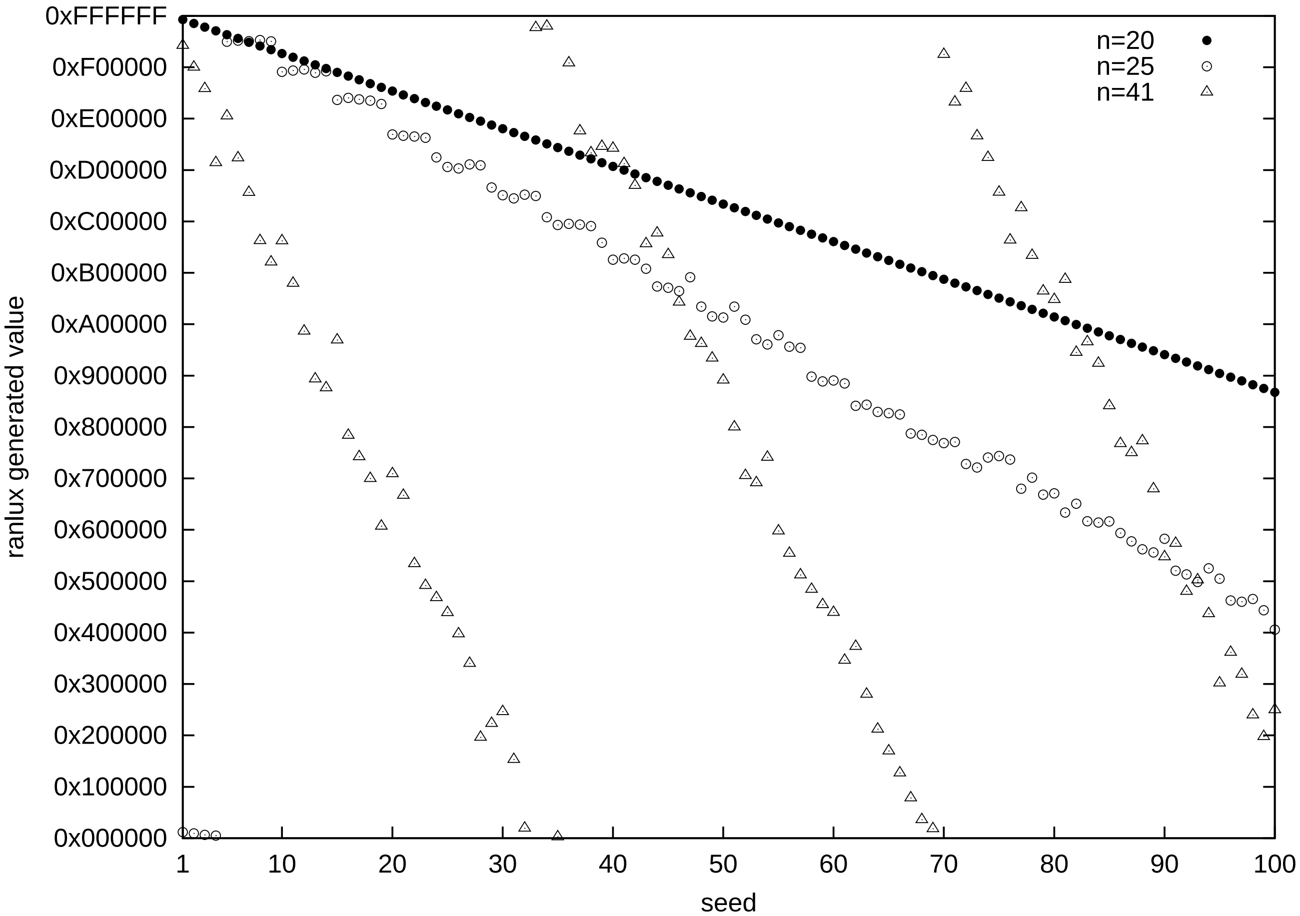
横軸：過去31回中の1の個数

縦軸：その条件下で、次の8回が全て0の確率：



## 初期値を系統的に選んだときの、非乱数性

ranlux(カオス理論に基づく擬似乱数, Lüscher ’94 )で、
20, 25, 41番目の出力（縦軸）を、初期シードを1, 2, 3, . . . , 100（横軸）と動かしてプロット



GNU Scientific Libraryに入っている
58個の擬似乱数発生法のうち、最新のものも含めて
45個にこのような問題が観測された。
Mersenne Twisterでは観測されなかった。
（M, 和田維作、倉本愛、芦原評 2007）

## 総まとめ

擬似乱数は極めて大量に用いられる。微細な統計的偏りや、初期値へのわずかな依存性が、計算機の高速化・大規模化に伴いシミュレーションを狂わせる可能性がある。

この際、使用者は擬似乱数が原因だと中々気づけない。

$\Rightarrow$ 精密なデタラメさを高速生成する必要性

それに答えるのが、「1+1=0の数学」

講演者は、

「擬似乱数をMT系に変えたらうまく動いた」

というメールをたくさんもらっている。

注：MTは、そのままではストリーム暗号に使えない。複雑な関数により出力を圧縮すれば使える。

## 擬似乱数研究の混迷

$\Leftarrow$ 理論的かつ実用的な「擬似乱数の定義」がないため

**「擬似乱数の定義」へのまるで異なる３アプローチ：**

1. 記録しておかない限り再生不能なものを乱数という (Kolmogorov-Chaitin, '60末)
2. 数列の一部から、他の部分が計算量的に計算できないものをいう (Blum-Blum-Shub, '86)
3. 周期や高次元分布性といった指標を用いて、良い漸化式を探す (古典的, '45-)

MT は古典的な3番だが、漸化式の「良さ」の評価に $\mathbb{F}_2((t^{-1}))$ など現代数学の手法を用いた。

終わり

## 周期保証：Mersenne素数の利用

**定理** $f(t)$ を定数項が1の $n$ 次 $\mathbb{F}_2$ 多項式とする $(n \geq 2)$。
$2^n - 1$ が素数とする。
$1/f(t)$ のべき級数展開の周期が最大値 $2^n - 1$ を満たす必要十分条件は、

$$t = t^{2^n} \mod f(t)$$

となること。

$n$ 回 $\mod f(t)$ で二乗すればよい。
$n$ が１００万くらいまでなら計算機で数分でチェックできる。

必要性：周期が $2^n - 1$ ならば $1 = t^{2^n - 1} \mod f(t)$.

十分性：$t = t^{2^n} \mod f(t)$ と $f(t)$ が $t$ と互いに素なことから

$$1 = t^{2^n - 1} \mod f(t).$$

周期を $P$ とすると

$$1 = t^P \mod f(t).$$

$2^n - 1$ を $P$ で割ったあまりを $r < P$ とすると

$$1 = t^{2^n-1} \mod f(t) = t^{Pq+r} \mod f(t) = t^r \mod f(t).$$

周期 $P$ の最小性から $r = 0$。
すなわち $P$ は $2^n - 1$ の約数である.
$2^n - 1$ が素数であることを使うと $P = 1$ または $P = 2^n - 1$。
$P = 1$ は $1 = t \mod f(t)$ となり $n \geq 2$ より不可能。
よって $P = 2^n - 1$。

**注**　$2^n - 1$が素数となるとき、メルセンヌ素数という。
2012年1月現在で47個知られており、既知の最大は

$$2^{43,112,609} - 1。(2008/8/23発見)$$

無限個あるかどうかは分かっていない。

メルセンヌツイスターに使った$2^{19937} - 1$は
２４番目のメルセンヌ素数(1971年発見)。