

パーソナルコンピュータ上のBASICの乱数について

3C-9

森 博

小池 慎一

名古屋文理短期大学

はじめに

パーソナルコンピュータの登場とともにBASIC言語が広く普及した。BASIC言語には数学関数の一つとして乱数発生関数が組み込まれている。初期のBASICの乱数には、周期が202というとても乱数とは呼べないものもあった。現在のそれは、周期が 2^{24} を持つものもあり著しく改良されている。本報では乱数関数の変遷を、発生アルゴリズムと周期の点からたどり考察する。

周期の変遷

パーソナルコンピュータ上のBASICに組み込まれている乱数関数の周期はBASICの発展につれてしだいに長くなっている。以下にそれを列挙する。

BASIC	周 期
APPLE II	202, 37758, 32366
CBM3032	13379, 24114
MBASIC, NBASIC	1368, 13680, 90288
N88BASIC(86)	262144 ($=2^{16}$)
FBASIC	16777216 ($=2^{24}$)
2020BASIC	16777216 ($=2^{24}$)
QuickBASIC	16777216 ($=2^{24}$)

APPLE IIからMBASICまでは、シードの与え方によって複数の周期を有する。したがって、APPLE IIの202とか、MBASICの1368などの短い周期の乱数を不用意に利用してしまう可能性がある。実際は、周期に飛び込むまでにヒゲを持つので、これよりは長い乱数系列を利用できる。いずれにしても、利用したくない乱数である。

N88BASIC以降は周期は長くなっている。ほぼ、単精度実数の仮数部の大きさすなわち24ビットを周期にしている。

アルゴリズムの変遷

前節で述べたAPPLE IIからMBASICまでの乱数発生アルゴリズムでは線形合同法によらず、演算の途中でバイト単位で上位と下位を入れ換える⁽²⁾などの独自の方法が用いられている。乱数系列は、周期を持つばかりでなく、周期の系列に到達するまでのヒゲを持っている。したがって、利用可能な系列の長さは初期値からヒゲを通り周期上を一巡するまでである。そのために、任意のシードを与えた乱数系列間の関係を求めるのが面倒である。文献⁽²⁾に調べた結果が与えられている。

N88BASIC以降は、乱数発生アルゴリズムは線形合同法を基礎にしている。以下に調査した結果を示す。

N88BASIC	$X_{n+1} = 6012033 X_n + 5123127 \bmod 2^{24}$	(1)
FBASIC	$X_{n+1} = 4253261 X_n + 372837 \bmod 2^{24}$	(2)
2020BASIC	$X_{n+1} = 214013 X_n + 2531011 \bmod 2^{24}$	(3)
QuickBASIC	$X_{n+1} = 16598013 X_n + 12820163 \bmod 2^{24}$	(4)

これらの乱数はいずれもmodulus M を 2^{24} にとっているが、実際の周期が 2^{24} となるのはN88BASIC以外のBASICの乱数である。FBASICのそれは、整数値 X_{n+1} を区間 $[0, 1)$ の実数値に直す方法によって一見周期 2^{23} を持つようにみられる。

N88BASICは周期が 2^{16} と M の値に比べていちじるしく短くなっている。これは、式(1)の演算において、加算の定数5123127が下位2バイトと上位1バイトが分けて加えられる際、キャリーが足し忘れられているためである。意図的になされたものか、単なるバグかはわからない。具体的には、式(1)を16進数で表した場合次のようなことが生じる。

$$X_{n+1} = 5bbc81h X_n + 4e2c37h \bmod 1000000h \quad (5)$$

$x_n = 6c3ddbh$ とした場合、第一の項は、下位8バイトのみを書くとしてb0d0ff5bhとなる。これに定数4e2c37hを加えるのであるが、下位2バイトに続けて上位バイトを加えると、キャリーが生じて1f2b92hとなるべきである。実際は、キャリ

一が無視されるので、1e2b92hとなり65536だけ小さな値となる。

また、このために一周期ないの出現度数の均一性が損なわれている。たとえば、線形合同法では一周期すべての乱数の平均値は0.5になるはずであるが、RANDOMIZE -1で初期化された系列の場合の平均値は0.50011とわずかに片寄っている。

シードの与え方

線形合同法ではシードはmodulus Mの値を越えない整数値で与える。しかし、BASICのコマンドでは直接シードを与える関数またはコマンドは用意されていない。一般には、コマンドRANDOMIZE 引数を用いるかあるいは、乱数発生関数RNDの引数に負の値を与えることにより初期化をする。与えられるシードの値がいくつかはわからない。ただし、再現性はある。

BASICにはBASIC毎に式(1)-(4)で示される一個の乱数系列しか利用できない。初期化のコマンドは一個の乱数系列の一周期上のどこからかはじまるシードを与える。しかし、それがどこなのかはわからない。2020 BASICとQuick BASICのについては文献⁽³⁾に周期 2^{24} を128等分するシードを与えるテーブルが得られている。

N88 BASICの周期は前節で述べたように 2^{18} である。したがって、 $2^{24-18}(=128)$ 個の異なる系列の存在が予想される。調べた結果を表1に示す。

1	7	3	175	5	73	255	9	211	11	13	147	15	17	149	19
31	21	243	23	205	25	27	29	47	39	119	35	335	241	41	43
45	49	51	459	55	207	59	137	63	295	66	143	75	79	81	83
85	161	397	95	107	501	115	193	123	131	159	195	189	301	227	629

表1 N88BASICの異なる128個の系列を与えるRANDOMIZEの引数
RANDOMIZE -1のように表の値に-をつけて用いる

N88BASICではRANDOMIZEの引数に偶数を用いるとまったく同一の乱数系列を発生するものがあるので、偶数の引数は用いない方がよい。また、関数RNDの引数に負の値を与えても他のBASICのは異なり、新しい系列にはならない。

乱数関数組み込み上の問題点

コンピュータに組み込まれる乱数発生アルゴリズムは線形合同法によるのが一般的である⁽⁴⁾。線形合同法により乱数を発生する関数あるいはサブルーチンをBASICで組み込む場合、以下の障害がある。

- (1) 整数のビット長が短い(16ビット)
 - (2) 整数乗算のオーバーフローがチェックされる
- (1)からは、高々65536の周期の乱数しか作れない。(2)からは、オーバーフローを利用したmodulus演算がプログラムできないという制約を受ける。

倍精度を利用してたとえば次のようにすればプログラム可能であるが、実行時間が長いのが欠点である。

```
X#=69069*X#: X#=X# - FIX(X#/2^32) * 2^32 : RETURN
```

なお、Quick BASICでは4バイト長の長整数が用意されているが、(2)の理由により線形合同法のプログラミングには利用しがたい。

むすび

BASICに組み込まれた乱数は、周期も長くなり改良されてきているが、アルゴリズムが隠されていること、シードが利用者によって任意に与えられにくいことが問題である。また、利用者がプログラミングする上でも制約がある。

したがって、組み込まれたRND関数の性質をよく調べた上で利用すべきである。

参考文献

- (1) 小池慎一、マイコンに使われている乱数の例、数理科学 No.208, p57-61 (1980)
- (2) 小池慎一、BASICの乱数の性質、マイコンピュータ、P23-49(1984)
- (3) 森博、小池慎一、パーソナルコンピュータ用乱数のシードと周期について、名古屋文理短期大学紀要15号、P1-9(1990)
- (4) 伏見正則、乱数、東京大学出版局、(1989)