ｒ２５０アルゴリズムによる疑似乱数発生について

1X−6

小池慎一　山住富也
愛工大　名古屋文理短大

はじめに

パーソナルコンピュータ上のBASIC言語等にインプリメントされている乱数は、合同法によるものが主である。その周期は N88BASIC で 2^{16}、Quick BASIC で 2^{24}である(1)．Turbo Pascalでは 2^{32}であり、周期の点では改良されている。

合同法は

$$X_n = a*X_{n-1} + b \mod P \qquad (1)$$

で示され、演算中に乗算が含まれている。また、数値の精度は法Ｐで定まる。

一方、Ｍ系列として知られる乱数発生法がある。この特徴は演算が基本的に排他的論理和のみですむこと、合同法に比べて周期が長くとれることなどの特徴がある。

本報では、文献(2)でＣ言語上にインプリメントされたｒ２５０アルゴリズムによる乱数発生法を、Quick BASIC に乗せ、その性質について実験的に調査した。なお、ｒ２５０アルゴリズムはＭ系列による乱数発生法である。

Ｍ系列乱数

Ｍ系列乱数は一般に

$$X_i = c_1*X_{i-1}+c_2*X_{i-2}+\cdots +c_p*X_{i-p} \mod 2 \qquad (2)$$

の線形漸化式で生成される数列である。係数ciおよび数X_iは０または１をとる。この演算はフィードバックつきシフトレジスタで実現できる。ハードウェア的に乱数を発生するのに適している。

係数$c_1, c_2, \ldots, c_p$は生成多項式の係数である。すなわち

$$c_p*X^p+c_{p-1}*X^{(p-1)}+\cdots+c_1*X+1 \qquad (3)$$

ソフトウェアで実現する場合には、２個の係数のみがゼロでなく、他の係数はゼロである３項多項式、すなわち$c_p=c_q=1$, $c_i=0$ $(i\neq p,q)$ であり

$$X^p + X^q + 1 \qquad (4)$$

を用いる。この場合の乱数生成の式は

$$X_i = X_{i-p} + X_{i-q} \mod 2 \qquad (5)$$

となる。

Ｍ系列で得られる乱数列は１ビットの列で、すなわち０と１のみからなる数列である。有効精度ｍビットの乱数列を得るには

$$b_1 b_2 \ldots b_{m-1}$$
$$b_i = 0, 1 \quad i = 0, \ldots, m-1 \qquad (6)$$

として、各ビットをＭ系列で発生する。実際は初期系列を適当に選んであたえる。この方法によれば、任意の精度の乱数列を得ることができる。しかし、実用上は16ビット程度あれば十分であろう。

Ｍ系列の周期は 2^p-1 である。ｐが大きければ、合同法に比べて大きな周期の乱数列を得ることができる。ｐとｑの具体的な値については文献(3)にある。

p=521, q=32 のＣによる別のインプリメント例は文献(5)にある。

On Pseudo Random Number Generator by r250 Algorithm
Shin-ichi Koike, Tomiya Yamazumi
Aichi Institute of Technology, Nagoya Bunri College

r２５０アルゴリズム

r２５０アルゴリズムでは、p=103, q=1を採用している。周期は$2^{103}-1$、精度は16ビットである。初期系列はrand関数で発生された系列に、対角要素を11個ごとに1とするなどの処理をしたものを用いる。

一度初期化をすると以下、それをもとにして逐次新しい乱数を生成する。最小103個の系列は常にメモリの上に常駐しており、発生された乱数を保存してゆく。そのために、式(5)のX_iに相当する値の位置を常に保存しておく必要がある。

BASICにインプリメントしてみると、発生に要する時間は組み込み関数RNDに比較して、３倍程度かかるが、周期が長いのが利点である。原理的にはRNDより速くなると予想されるが、アセンブラで組み込まないと比較できる速さにはならないのかもしれない。

統計的性質

r２５０アルゴリズムを用いて発生させた乱数について、平均値、標準偏差、自己相関関数を求めた結果を表１に示す。シードは１～10、サンプル数は１千万で、乱数の値は０～１の範囲である。

表１　統計処理の結果

seed	平均値	標準偏差	自己相関関数
1	0.5002	0.2886	-0.00000
2	0.4996	0.2888	-0.00001
3	0.5002	0.2889	-0.00006
4	0.4999	0.2886	-0.00006
5	0.5003	0.2887	-0.00005
6	0.4995	0.2885	0.00001
7	0.4999	0.2885	-0.00003
8	0.4997	0.2890	0.00008
9	0.4996	0.2893	0.00003
10	0.5002	0.2883	-0.00010

平均値は0.5、標準偏差は0.288、自己相関関数は10^{-5}のオーダーが得られた。他の特性に付いても検討中である。

むすび

FORTRAN, BASIC, Pascal, Cなどの言語にインプリメントされている乱数はほとんど合同法によるものである。この理由としては、容易にプログラム可能なこと、特別な初期化プログラムが不要なことなどが考えられる。しかも、基本言語長が８ビット、16ビットと大きくなるにしたがい、法Pの値も大きくなり、周期も長くなった。ある意味では実用的には問題がないと言える。

しかし、より長周期の乱数が要求されるような場面ではM系列の長周期は利点となる。たとえば、ある種のゲーム機あるいはゲームソフトでは簡単なハードウェアで、長周期の乱数が発生できる必要がある。そのような応用面の研究にはM系列が有利である。

文献

(1) 森、小池　"パーソナルコンピュータ上のBASICの乱数について",情報処理学会第４１回全国大会（1992.後期）
(2) W.L.Moier "A Fast Pseudo Random Number Generator", Dr.Dobb"s Journal, (1991.May) p.152
(3) 伏見政則　"乱数",
東京大学出版会　(1989)
(4) Knuth "準数値算法／乱数"
サイエンス社 (1985)
(5) 奥村晴彦　"C言語による最新アルゴリズム辞典"　技術評論社 (1991)