商学論集　第67巻第3号　　1999年2月

【 研究ノート 】

相互相関のある一様乱数の一発生法

星　野　珙　二

1. はじめに

在庫管理問題やスケジューリング問題などの解決の際に，互いに相関のある乱数を発生させてシミュレーションを行うことが必要になったり，ときにまたそれが便利であったりする場合がある。たとえば，筆者が今まで行ってきた在庫管理問題のシミュレーションにおいては，具体的な予測の方法はともかくとして，予測可能性の程度を表現する方法として，予測値と実現値との二つの系列間の関係を相関係数を用いて表わすことが大変便利であった。ジョブ・ショップ・スケジューリング問題においても，乱数列で加工時間を与える際に，ジョブの特性やショップの特性に応じて互いに相関を持たせることの方が現実的となる場合がある。

ところで，相互相関のある二つの乱数列の発生については，正規分布に従う場合についてのみ一般に良く知られた方法がある。しかしながら，それ以外の分布に従う場合については，現在までのところ筆者は確認していない。正規分布よりも少し幅の広い変化に富んだ状況を表現するために，一様分布に従う乱数で互いに相関を持たせた系列を作り出したいがどうしたらいいか，というのが本稿を認めた動機となっている。

2. 相互相関のある二つの一様乱数発生の試み

ごく自然に思いつく相互相関のある二つの一様乱数の発生方法としては，既に述べたように正規分布上で実現する方法が確立されているので，一旦正規分布の上で相互相関のある二つの乱数を発生しておき，それを一様分布上にそれぞれ変換していくやりかたである。しかしこの方法は，確率密度関数の積分値，すなわち累積分布の値を求めなければならず，変換の度に近似的に積分値を計算するか，もしくはメモリーに取り込んだ正規分布表の値を探索していかなければならないことになる。いずれの場合についても，計算精度と計算時間の双方の難点が認められる。

さて今回試みた方法は，計算精度の点では難を残すものの，計算時間については問題を解消することができるものである。考え方は，正規分布上で相関のある二つの乱数を発生する方法を準用しつつ，一様分布の特性に合わせた数値計算上の修正を加えていくものである。以下，その方法の概要を説明する。

区間 $[0, 1)$ 上の独立な二つの一様乱数を $R1, R2$ とし，それぞれ次のような変換を行う。それに伴って，区間はそれぞれ $[-0.5, 0.5)$ へ変わることになる。

$$X = R1 - 0.5$$
$$Y = R2 - 0.5$$

それらのX, Yおよび任意の相関係数ρを用いて次なるZを構成する。

$$Z = \rho X + \sqrt{(1-\rho^2)}\,Y \quad (-1 < \rho < 1) \tag{1}$$

このとき，X, Yの期待値は明らかに0であるから，Zの期待値は次のようになる。

$$\begin{aligned} E(Z) &= \rho E(X) + \sqrt{(1-\rho^2)}E(Y) \\ &= 0 \end{aligned} \tag{2}$$

また，X, Yが互いに独立であることに加えて $V(R) = V(X) = V(Y)$ と表すことができるので，Zの分散は次のように求められる。

$$\begin{aligned} V(Z) &= \rho^2 V(X) + (1-\rho^2)V(Y) \\ &= V(R) \end{aligned} \tag{3}$$

そこで，XとZの相関係数$\rho(X, Z)$を求めてみよう。まず，XとZの共分散$COV(X, Z)$を計算すると，$E(X)$, $E(Z)$は共に0でX, Yが互いに独立であることから次が得られる。

$$\begin{aligned} COV(X,\ Z) &= E[(X - E(X))(Z - E(Z))] \\ &= E[X \cdot Z] \\ &= E[\rho X^2 + \sqrt{(1-\rho^2)}\,YX] \\ &= \rho V(X) + \sqrt{(1-\rho^2)}E(X)E(Y) \\ &= \rho V(R) \end{aligned} \tag{4}$$

したがって，

$$\begin{aligned} \rho(X,\ Y) &= COV(X,\ Z)/(\sqrt{V(X)}\sqrt{V(Z)}) \\ &= \rho V(R)/(\sqrt{V(R)}\sqrt{V(R)}) \\ &= \rho \end{aligned} \tag{5}$$

となる。ここまでは相互相関のある正規分布の乱数発生と平行した論理で，一見したところ簡単に一様分布の乱数発生にも適用できそうにみえるが，次の二つの点で問題を残している。一つは，2つの一様乱数をρと$\sqrt{(1-\rho^2)}$でウエイト付けして加えていることからも明らかなように，一種の確率変数同士の和の分布となるため一様性が維持できないことである。もう一つは，(1) 式の構成からも明らかなようにZの範囲が$[-0.5, 0.5)$に納まらないことである。筆者は念のため数値実験を行って上記の (2)，(3)，(5) 式が成り立つことを確認し，その上で上述した二つの問題点があることも確かめた。

さて，少し強引なやり方であるが，二つの問題点に次のように対処してみた。まずZの範囲に関しては，(1) 式の作り方から，$-\rho - \sqrt{(1-\rho^2)} < Z < \rho + \sqrt{(1-\rho^2)}$ となるので，$|\rho + \sqrt{(1-\rho^2)}|$で除しておけば$Z$の範囲が$[-0.5, 0.5)$に納まることになる。その上で，一様性を回復させる調整を試みる。具体的には，一旦等分割した小区間毎に発生度数を求めておき，最終的に各区間の度数が一様

になるようにはじめの小区間の区切りを調整し，改めて発生した乱数を区間の区切りに対応して比例変換していく方法を採る。

3. アルゴリズム

上で述べた相互相関のある二つの一様乱数の発生方法を，ここでは改めてアルゴリズムとして整理しておく。

(1) 区間 $[0, 1)$ 上の二つの一様乱数(一般にプログラミング言語の関数として用意されており，連続的に乱数を発生させても独立性が保証されている) $R1, R2$ を発生させ，それぞれから 0.5 を減じ，X, Y とする。

(2) $Z=\rho X+\sqrt{(1-\rho^2)}\,Y$ を構成し，それをさらに $|\rho+\sqrt{(1-\rho^2)}|$ で割って改めて Z とおく。(計算を速めるために負の Z の値については一旦フラッグを立てて正の領域で処理し，後で負の値に戻すことにする)

(3) 区間 h の幅で k $(=1.0/hH_i=Q_i*h)$ 分割し，Z が含まれる区間の度数をカウントする。以上 (1)，(2)，(3) を適当に大きな数 n まで繰り返す。

(4) 期待度数 $ED=n/k$ と i 番目の区間の実現度数 D_i とから比率 $Q_i=ED/D_i$ および $H_1=Q_1*h$, $H_i=Q_i*h-(i/k-H_{i-1})Q_i/Q_{i-1}+i/k$ $(i=2, 3, \cdots, k)$ を求めておき，次の手順を踏む。

① $Z<H_1$ なら，$Z=Z/Q_1$ として④へ，そうでなければ $i=2$ として次へ

② $Z<H_i$ なら，$Z=(Z-H_{i-1})/(H_i-H_{i-1})+i/k$ とし④へ，そうでなければ i の値を 1 増やして $i=k-2$ まで繰り返す。$i=k-1$ となった場合は次へ

③ $Z=(k-1)/k+h*(Z-H_{k-1})/(0.5-H_{k-1})$

表 1 計算結果

所与の相関係数	結果の相関係数	Z の平均	Z の分散
0.1	0.09173	−0.00113	0.08084
	0.08917	−0.00109	0.08072
	0.09661	0.00038	0.08062
0.3	0.27654	−0.00166	0.07954
	0.27480	−0.00126	0.07945
	0.28156	0.00031	0.07954
0.5	0.47119	−0.00222	0.08054
	0.47009	−0.00149	0.08046
	0.47521	0.00019	0.08072
0.7	0.68848	−0.00263	0.08098
	0.68799	−0.00162	0.08090
	0.69048	0.00011	0.08125
0.9	0.90524	−0.00231	0.08011
	0.90490	−0.00070	0.07992
	0.90563	0.00024	0.08035

④　発生数が所定の数になるまで新たに Z を発生させ，① へ戻る

4.　発生した乱数の性質

上記のアルゴリズムに基づいて，相互相関のある二つの一様乱数 X, Z を発生させた結果を表1に示す。ただし，$n=20000$, $h=0.1$ とし，乱数の発生回数も 20000 として初期値を変えて3回くりかえしている。Z の頻度の結果については，本方式のねらいに照らして明らかなように一覧して一様性が保持されていることが確認でき，改めて示すことを省略した。また，Z の平均値についても問題のないことが表から確かめられる。結果の表から問題視される点は，一部を除いて相関係数 $\rho(X, Z)$ が所与の値をわずかに下回っていること，さらに，全体を通して Z の分散の値も理論値より若干小さ目になっていることであろう。ちなみに区間 $[-0.5, 0.5)$ の分散の理論値は 0.08333(=1/12) である。

計算時間については N88BASIC (PC-9801RX) で 20000 個の X, Y を発生させるのに5分程度であり，最新のプロセッサーでコンパイラー型のプログラミング言語を用いれば問題にならない時間で発生は可能であろう。

5.　ま　と　め

本稿では，提案の方法により，実用上差し障りない範囲で相互相関のある二つの一様乱数が発生できることを確かめた。この方法の計算は単純計算のみからなっており，計算時間はほとんど気にならない。しかしながら，この方法は一様性が崩れた分布を再度一様性を保持するようやや強引な調整を行っており，あまり洗練された方法とは言えない上，理論的にも曖昧さを残している。より優れた方法の開発が今後の課題となるが，どなたか良いアイディアをお持ちの方がおられたらご教示願いたいものである。

参考文献

[1]　伏見正則：UP応用数学選書12「乱数」，東京大学出版会，(1989)
[2]　林知己夫監修，脇本和昌著：「乱数の知識」，森北出版株式会社，(1970)
[3]　十代田三知男編著：「BASICによるシステム解析技法」，共立出版株式会社，(1988)
[4]　編集委員代表竹内啓：統計学辞典，東洋経済新報社，pp. 524-529, (1989)
[5]　Naylor, T.H., Balintfy, J.L., Burdick, D.S., Chu, K.: Computer Simulation Techniques, John Wiley & Sons, Inc., (1966)