J. Plasma Fusion Res. Vol.96, No.7 (2020) 372-378

講座 シミュレーションのための乱数入門

4．乱数の品質と検定法

4. Quality and Verification Methods of Random Numbers

佐竹真介
SATAKE Shinsuke
自然科学研究機構　核融合科学研究所

（原稿受付：2020年3月17日）

疑似乱数には様々な発生法がありますが，良い疑似乱数，悪い疑似乱数の判断はどのように考えられているでしょうか．本章ではまず，結晶構造や多次元均等分布性について説明した後，乱数の満たすべき統計的な性質に基づくいくつかの検定テスト法と，それを2.1節と3章で紹介した疑似乱数の発生法に対して試した結果を紹介します．

Keywords:
pseudo random number, quality of random number, verification of pseudo random number

本章では，「良い乱数」「悪い乱数」とはどういうものか，疑似乱数の品質の判断基準と検定の仕方について概説します．第2章でも繰り返し述べてきたことですが，疑似乱数をシミュレーションに利用するにあたって最も気にするべきことの1つが，この乱数の「品質」です．疑似乱数の発達の歴史は，いかに品質の良い疑似乱数を効率よく決定論的なアルゴリズムから生み出すかの探求であると言っても過言ではないでしょう．また，疑似乱数に対する品質の不満や不信が物理乱数発生器の開発を促したとも言えます．

とは言え，何をもって「良い」疑似乱数であるとするかというと，これは非常に難しい問題でもあります．例えば偏微分方程式や数値積分の数値解法であれば，より高精度で誤差の蓄積が少ない，保存すべき量を保存する，数値不安定性を起こしにくい，といった明確な良し悪しの基準をその数値アルゴリズムの用途に応じて定められます．しかし乱数の場合，確率論的な統計量に対しての判定の議論が主となり，「この試験をパスした疑似乱数は真の乱数に近い」と言い切れるものがある訳ではありません．むしろ，発生した数列が，真の乱数であるならば満たす，考えうる限り多くの統計的な性質A，B，C，…を満たすならばそれは良い乱数だ，という少し漠然とした判断基準になります．また，数学的に厳密な乱数の品質の議論をするには高度な数学（統計学だけでなく整数論，Galois（ガロア）体等）の知識を必要とするため，この講座で全てを説明することはできません．そこで，ここではわかりやすい検定法の実例を挙げながら解説したいと思います．

・疑似乱数の結晶構造と均等分布性

「とても良い」乱数を定義することは難しいですが，「とても悪い」疑似乱数を見つけることはそれほど難しくありません．2.1節で線形合同法（混合合同法）の問題点は，周期長の短さと，多次元分布における結晶構造であることを触れました．この結晶構造とは，乱数列$\{r_i\}$を用いて，n次元空間に点列$(r_i, r_{i+1}, \cdots, r_{i+n})\,(i = 1, 2, \cdots)$を並べた時に，その点列が$n$次元空間を一様に埋め尽くすのではなく，ある決まった大きさのn次元格子の頂点にのみ点列が存在することを意味します．**図1**に2.1節で紹介した線形合同法による乱数列$r_{i+1} = 1103515245 r_i + 12345 \,(\mathrm{Mod}\, 2^{32})$を3次元空間に並べた場合の例を示します．見る角度を選ぶと，図のように点列(r_i, r_{i+1}, r_{i+2})がある平面上に整列していることがわかります．このようなパターンは，真に一様な乱数であれば出ないため，線形合同法は品質が劣る疑

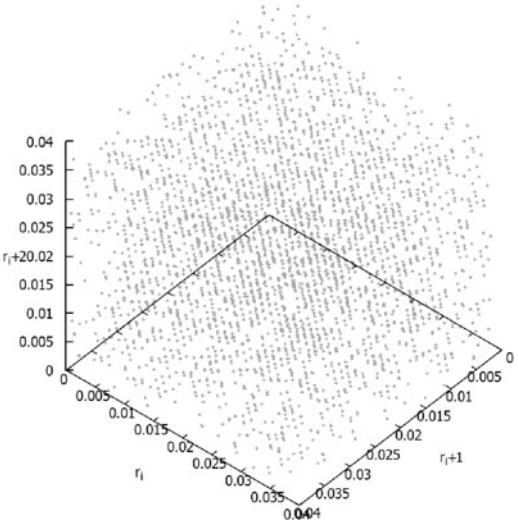

図1　線形合同法 $r_{i+1} = 1103515245\, r_i + 12345\ (\mathrm{Mod}\ 2^{32})$で発生した乱数から，3次元空間に点列$(r_i, r_{i+1}, r_{i+2})$を並べた場合に見られる結晶構造の例．

National Institute for Fusion Science, Toki, GIFU 509-5292, Japan
corresponding author's e-mail:satake@nifs.ac.jp

似乱数だと判断されるわけです．

線形合同法については，次元 n をどのように取っても必ず結晶構造を持つことが数学的に証明されています[1]．対照的に，2.1節で紹介したTausworthe（トーズワース）法やMersenne-Twister（メルセンヌ・ツイスター）法（MT法）は，長周期性だけでなく，多次元空間における分布の均等性が良いという特徴も持ちます．例えば，最もよく使われている，周期長が $2^{19937}-1$ の32ビット整数バージョンのMT法の場合，623次元空間に均等分布するとされています．しかし，実際にこのような超多次元空間の分布を，長い疑似乱数の一周期に渡って調べることは不可能です．そのため，これらの疑似乱数における多次元均等分布性は，乱数発生法の原理から以下のように説明されます[2]．

まずTausworthe法の場合，漸化式

$$\boldsymbol{x}_{i+p} = \boldsymbol{x}_{i+p} \oplus \boldsymbol{x}_i \qquad (1)$$

（$\boldsymbol{x}_i$ は$\{0,1\}$を元とするwビットの行ベクトル）に対応する特性多項式 $x^p + x^q + 1$ がGalois体GF(2)上で定義され，これが $t = 2^p - 1$ の時は多項式 $x^t - 1$ を割り切るが，$t < 2^p - 1$ に対しては割り切れない場合に，この特性多項式を p 次の原始多項式と呼び，生成される数列の各ビットの周期がそれぞれ $2^p - 1$ となると2.1節で説明しました．また，$2^p - 1$ が（1）式の形の漸化式が取りうる最大周期長でもあります．このように最大周期長を取っている場合，$\boldsymbol{x}_i$ の j 番目のビット $x_{i,j}$ に着目して，それを p 個ずつの組にして並べたもの $\{x_{i,j}, x_{i+1,j}, \cdots, x_{i+p-1,j}\}$ $(i = 1,2,\cdots,2^p-1)$ を考えると，これらの組の取りうるパターンは，全てのビットが0というパターンを除いた $2^p - 1$ 通りです．実は，特性多項式が p 次の原始多項式となる場合には，この $x_{i,j}$ を p 個並べた組 $\{x_{i,j}, \cdots, x_{i+p-1,j}\}$ が取りうる $2^p - 1$ 通りの全パターンが，乱数列の一周期の中にちょうど1回ずつ含まれることが証明されています．Tausworthe乱数の各ビットはこのような意味で p 次元に均等分布している，と言われます．

ところでTausworthe法では w ビットベクトル $\boldsymbol{x}_i$ の各ビットは独立に決まるため，$\boldsymbol{x}_i$ の初期値を上手く選ばないと w ビットベクトル全体としての周期を各ビットの最大周期長 $2^p - 1$ に一致させられないと2.1節で説明しました．では w ビットベクトル $\boldsymbol{x}_i$ に対する均等分布性に関してはどうでしょうか？その説明をするために，まず定義を明確にしましょう．疑似乱数の周期を P とした時，w ビットベクトル $\boldsymbol{x}_i$ を n 個並べた組 $\{\boldsymbol{x}_i, \boldsymbol{x}_{i+1}, \cdots, \boldsymbol{x}_{i+n-1}\}$ $(i = 1,2,\cdots,P)$ が v ビット（$v \le w$）の精度で n 次元均等分布するとは，$\{\boldsymbol{x}_i, \boldsymbol{x}_{i+1}, \cdots, \boldsymbol{x}_{i+n-1}\}$ の上位 v ビットが取りうる $2^{nv} - 1$ 個（全てゼロは除く）のパターンが，一周期中に同じ回数だけ現れることを意味します．Tausworthe法では，$\boldsymbol{x}_i$ ベクトルを p 個（ここで p は周期長ではなく，（1）式の階数）並べた組 $\{\boldsymbol{x}_i, \boldsymbol{x}_{i+1}, \cdots, \boldsymbol{x}_{i+p-1}\}$ が，全wビットの精度で均等分布しうる次元は最大で $\lfloor p/w \rfloor$ であることが知られています（$\lfloor a \rfloor$ は a 以下の最大整数の意味）．しかし，実際に何次元で均等分布するかは，周期長と同様に初期条件 $\boldsymbol{x}_i$ $(i = 1,2,\cdots,p-1)$ の各ビットの選び方に依存します．2.1節でも述べましたが，世の中に出回っているTausworthe法のプログラムは適切な初期化ルーチンと通常セットになっているのでユーザがあまり気に病まずともよいですが，周期長や均等分布性のよい初期値の設定の仕方について詳しく知りたい方は[3,4]を参考にして下さい．

ところでwビットベクトルの組 $\{\boldsymbol{x}_i, \boldsymbol{x}_{i+1}, \cdots, \boldsymbol{x}_{i+p-1}\}$ が取りうるビットの組み合わせは $2^{pw} - 1 > 2^p - 1$ 通りあるので，これらの組み合わせが一周期中に1回ずつ出てくるような数列を作ることができれば，より長周期でより高次元に均等分布する疑似乱数列を作ることができそうです．そこで，Tausworthe法では各ビットが（1）式から独立に決まっていたものを，ビット間でのシャッフルを取り入れることによってそれを実現したものがMT法です．MT法の漸化式は

$$\boldsymbol{x}_{i+n} = \boldsymbol{x}_{i+m} \oplus (\boldsymbol{x}_i^{w-r} | \boldsymbol{x}_{i+1}^{r})\boldsymbol{A} \qquad (2)$$

でした（式の詳細は2.1節を参照）．この場合，$\boldsymbol{x}_i$ ベクトルを w ビット乱数として先ほどと同様に漸化式の階数 n だけ並べた組 $\{\boldsymbol{x}_i, \boldsymbol{x}_{i+1}, \cdots, \boldsymbol{x}_{i+n-1}\}$ を考えると，取りうるビットの組み合わせは $2^{nw-r} - 1$ 通りになります．MT法ではこれらが一周期の中で1通りずつ出るようにパラメータ $(m, n, r, \boldsymbol{A})$ が上手く選択されています．例えば，2.1節で紹介したサンプルコードmt19937ar.fでは，$n = 624$，$m = 397$，$r = 31$，$w = 32$ であり，周期長 $P = 2^{624\times 32 - 31} - 1 = 2^{19937} - 1$ になっており，w ビット精度で $\lfloor (nw - r)/w \rfloor = 623$ 次元均等分布することになります．なお，MT法では w ビット精度の均等分布性だけでなく，各上位 v ビットに対する均等性についても工夫されています．上位 v ビットに関する均等分布性は，理論上 $\lfloor (nw - r)/v \rfloor$ 以下にしかなり得ませんが，MT法では $v = 1,2,\cdots,w-1$ のそれぞれについても極力最大値に近づくように，$\boldsymbol{x}_i$ ベクトルに追加の"Tempering"と呼ばれる操作を付け加えています（2.1節の（4）式を参照）．各 v ビット精度での均等分布の次元 $k(v)$ が具体的にどの様になっているかは，[5]に示されています．このように，MT法は単に周期長が長いだけでなく，多次元分布の均等性についてもTausworthe法より良好になるように作られています．

・統計的性質による疑似乱数の検定

ここまで多次元分布の均等性について見てきましたが，それさえ良ければ品質のよい疑似乱数であるという訳ではありません．実際に我々が疑似乱数を使う時には，w ビット整数乱数 $r_i \in [0, 2^w - 1]$ を $2^w - 1$ で割って $[0,1]$ 区間の一様乱数として使うことが一般的です．以下，$x_i = r_i/(2^w - 1)$ としてこの実数疑似乱数列 $\{x_i\}$ の品質について調べていきましょう．$\{x_i\}$ が「良い一様乱数」として満たすべき統計的な性質は，いくらでも挙げることができますが，例として以下のようなテストを考えてみましょう．

（ｉ）**連検定**：例えば，$x_{i-1} > x_i < x_{i+1} < x_{i+2} < \cdots < x_{i+m} > x_{i+m+1}$ となるような区間のことを，長さ m の上昇連と呼び

ます（逆は下降連）．このような連が全体の長さ N の数列の中に出現する数を l_m^N と表します．あるいは，平均値 1/2 以上（未満）の値が m 回続く連を考え，その出現数を f_m^N とします．$\{x_i\}$ が一様乱数であるなら，それらの期待値は以下のように求まります[1,6]．

$$\mathrm{E}[l_m^N]=\frac{2}{(m+3)!}[N(m^2+3m+1)-(m^3+3m^2-m-4)],$$

$$\mathrm{E}[f_m^N]=\frac{N+3-m}{2^{m+1}}.$$

十分な長さ N の乱数列を発生させ，その中の長さ m の連の出現回数と，上記の期待値との間に統計的に有意な差が生じていないかを確認します．

（ii）近接値の出現率：2 つの連続した乱数の組 $\{x_i, x_{i+1}\}$ $(i=1,2,\ldots)$ が，区間 $\Delta_{m-1}<|x_{i+1}-x_i|\le\Delta_m$ に現れる回数 c_m^N の期待値は，

$$\mathrm{E}[c_m^N]=(N-1)[(2\Delta_m-\Delta_m^2)-(2\Delta_{m-1}-\Delta_{m-1}^2)]$$

となります．$\Delta_m \ll 1$ に対して調べることで，近接した値が不自然に連続して出ていないかを確認します．

（iii）平均値と二乗平均値の分布：これは最も単純なテストの一つと言えます．一様乱数 $\{x_i\}$ の平均，二乗平均の期待値はそれぞれ $\mathrm{E}[x_i]=1/2$，$\mathrm{E}[x_i^2]=1/3$ ですが，大きさ N_s の標本を多数取った場合，$\{x_i\}$ が一様乱数であれば x_i，x_i^2 の平均値はそれぞれ，正規分布 $\mathrm{N}(1/2, 1/(12N_s))$，$\mathrm{N}(1/3, 4/(45N_s))$ に従って分布するはずです．生成された疑似乱数列の標本平均が，この分布に従うかを確認します．

なお，連検定と近接値の出現率の検定については，（i）の連の長さ，あるいは（ii）の近接値の区間 Δ_m のインデックス $m=1,2,\cdots$ に対する出現数分布の期待値が上の $\mathrm{E}[l_m^N]$，$\mathrm{E}[f_m^N]$，$\mathrm{E}[c_m^N]$ のように求められているので，長さ N の標本中の長さ m の連の出現数の分布が，期待値の分布とどの程度よく一致しているかをチェックします．このような「分布の当てはまりのよさ」のチェックには χ^2 検定が使われます．ある確率的事象が k 個の互いに背反なクラス $C_1, C_2, \cdots, C_k$ に分別される時，S 個の標本中の i 番目の事象 C_i の観測度数 n_i，その期待度数 E_i（$\sum_{i=1}^k n_i=\sum_{i=1}^k E_i=S$）に対して次式から求められる値を χ^2 値と呼びます．

$$\chi^2=\sum_{i=1}^k\frac{(n_i-E_i)^2}{E_i} \qquad (3)$$

もし観測度数の分布（例えば，疑似乱数列に含まれる長さ m の上昇・下降連の個数の分布）が，疑似乱数が一様乱数であると仮定して求めた期待値の分布に実際に従っているなら，（3）式に従って長さ N の標本を S 個用意して評価した χ^2 値は，自由度 $\nu=k-1$ の χ^2 分布に従うと期待されます．ここで自由度 ν の χ^2 分布（確率密度関数）は次式で定義されます．

$$f(\chi^2,\nu)=\frac{1}{2^{\nu/2}\Gamma(\nu/2)}(\chi^2)^{\nu/2-1}\mathrm{e}^{-\chi^2/2} \qquad (4)$$

今回の検定では，(i) の連検定については $m=1,2,\cdots,5$ と $m\ge 6$ の計 6 ケースの出現率，(ii) の近接値の出現率についても $\Delta_k=\{0.00, 0.01, 0.02, 0.05, 0.10, 0.20, 1.00\}$ とした場合の計 6 区間について調べるので，それぞれ自由度 = 6 − 1 = 5 の χ^2 分布に従うかどうかをチェックします．

一方，(iii) の平均値，二乗平均値の検定は Z 検定によって行います．ある確率変数 X の母集団の平均値と分散が (μ_0, σ^2) と仮定される時，大きさ N_s の標本 $\{x_i\}$ の標本平均を $\bar{x}$ とし，次式で定義される値 Z によって「母集団の平均値が μ_0 である」という仮定が正しいか否かを判定します．

$$Z=\frac{(\bar{x}-\mu_0)\sqrt{N_s}}{\sigma} \qquad (5)$$

仮説が正しければ，Z は平均 0，分散 1 の正規分布 $\mathrm{N}(0,1)$ に従うはずです．もし標本から評価した $|Z|$ が 1.96 以上の場合，仮説が正しければその確率は 5 %なので，この場合「有意水準 5 %で仮説（母平均が μ_0 である）は棄却された」ことになります．同様に，$|Z|>2.58$ となる場合は有意水準 1 %で棄却されます．ここでは，疑似乱数 $\{x_i\}$ の平均値，二乗平均値について，母集団が $[0,1]$ の一様乱数であるとの仮説から得られる平均と分散の期待値 (μ_0, σ^2) を用いて Z 検定します．

なお通常，Z 検定は N_s 個の標本のサンプリング 1 回について Z を評価し，仮説が正しいか否かを判定しますが，ここではそのようなサンプリングを S 回行い，評価した Z が実際に $\mathrm{N}(0,1)$ の正規分布をしているか，5 %，1 %の棄却域に入るサンプル数が実際に $0.05S$，$0.01S$ に近いかを調べます．

乱数検定の対象としては，2.1節で紹介した線形合同法 $(x_{i+1}=a_0x_i\,(\mathrm{Mod}\,P), a_0=1103515245, b=12345, P=2^{32})$，Tausworthe 法（$\boldsymbol{x}_{i+p}=\boldsymbol{x}_{i+p}\oplus\boldsymbol{x}_i$，$p=250$，$q=103$），Mersenne-Twister 法（mt19937ar.f）と，第 3 章で紹介した KMATH_RANDOM 版の SFMT の 4 つとします．なお，SFMT については Jump Method によって 2^{100} 飛ばしで作った初期値から開始した10本の乱数列それぞれについてテストしました．それぞれの疑似乱数から大きさ $N=10$万の標本を $S=1$ 万回サンプリングし，上記の検定を行いました．

まず，2 種類の連と，近接値の出現頻度の分布を見てみましょう．結果はどの疑似乱数もほぼ同じものになったため，代表として mt19937ar の例を**図 2** に示しました．この図では 3 つのテストそれぞれにおける，連や近接値の実際の出現数（図中点で表示）と期待値（線で表示）を比較していますが，一様乱数を仮定して得られる期待値とよい一致をしていることがわかります．この出現率の分布の期待値との適合度を見るために，（3）式で評価した χ^2 値の度数分布をグラフにしてみると，**図 3** のようになり，確かに自由度 5 の χ^2 分布に従っていることがわかります．紙面の都合上全ての結果について図は載せませんが，SFMT の10本の乱数列だけでなく，Tausworthe 法や，乱数としての品質が劣るとされている線形合同法についても**図 2** や**図 3** をプロットしても有意な差は見られませんでした．

次に疑似乱数の平均値，2 乗平均値の分布について見て

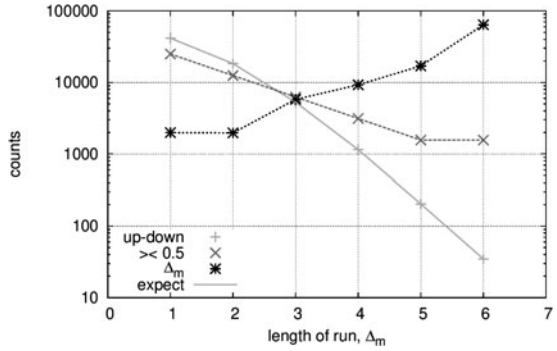

図 2　mt19937ar の疑似乱数に対する 2 種類の連と近接値出現回数の分布．“up-down”が上昇 - 下降連，“> < 0.5”が 0.5 以上又は以下の連，が $\Delta_{m-1} < |x_{i+1} - x_i| \leq \Delta_M$ の近接値の出現数であり，点がそれぞれの観測値，線が期待値を表します．長さ10万の乱数列を 1 万回サンプリングして評価しました．

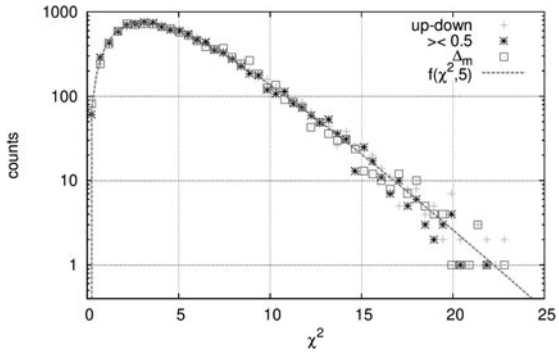

図 3　mt19937ar の疑似乱数に対する 2 種類の連と近接値出現回数の出現数分布に対する χ^2 値の度数分布（サンプル数＝1 万）．点がそれぞれの観測値，曲線が（4）式による自由度 5 の χ^2 分布を表します．

みましょう．**図 4** に線形合同法の結果を代表として示します．こちらも，一様乱数を仮定して求めた正規分布とよい一致を示していることがわかります．この場合も，乱数の発生法による有意な差は現れませんでした．（5）式による Z 検定の結果についても見てみましょう．**図 5** では，平均値と二乗平均値について，有意水準 5 %，1 %で棄却されるサンプルが 1 万個のうち実際にいくつあったかをプロットしてあります．また，**図 3** で見た χ^2 値の分布についても，χ^2 値が大きくて連などの分布が期待値によく一致していると言えない標本数（自由度 5 の場合 $\chi^2 > 11.07$ が 5 %，$\chi^2 > 15.09$ が 1 %の棄却域）も合わせて示しました．どの乱数発生法のどの検定結果も，実際にほぼ 5 %，1 %の標本がそれぞれの棄却域に入っています．

これらの結果から，ここで取り上げた 4 つの疑似乱数は，どれも今回紹介した検定をパスしたと言えそうです．驚くべきことに，線形合同法は明らかに結晶構造という乱数らしからぬ性質を持っているのに，これらの検定はパスできました．2.1節で整数の線形合同法乱数が偶・奇，偶・奇，のパターンを繰り返すことを紹介しましたが，このような乱数らしからぬ性質も今回のテストでは見逃されてしまいます．本章の冒頭でも述べた通り，「このテストさえ合格すれば OK」と言い切れるような，万能な乱数の検定法がある訳ではなく，考えうる限り多くのテストを合格で

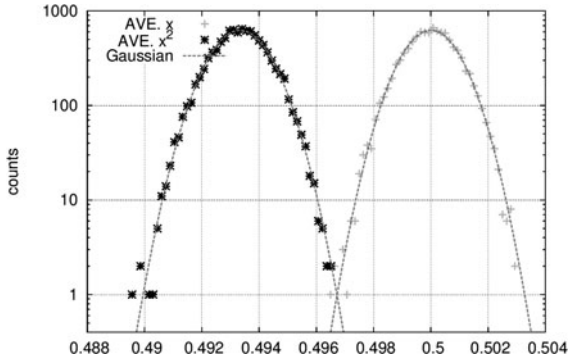

図 4　線形合同法で発生させた疑似乱数（標本の大きさ10万）の平均値と 2 乗平均値の分布（サンプル数 1 万）．なお 2 乗平均値は 0.16 右にオフセットしてプロットしてあります．“Gaussian”は一様乱数の場合に期待される正規分布を表します．

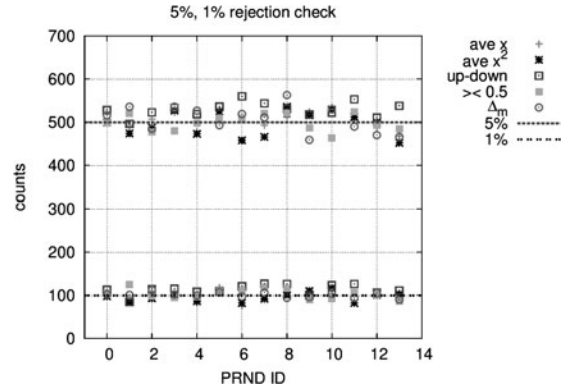

図 5　平均値，2 乗平均値に対する Z 検定と，2 種類の連と近接値の出現率に対する χ^2 検定で，それぞれ有意水準 5 %，1 %で棄却された標本数．横軸は乱数の発生法を表し，ID = 1～10 が10個の SFMT，11が線形合同法，12が Tausworthe 法，13が mt19937ar の結果です．また，ID = 0 は10個の SFMT の平均値を表します．

きる疑似乱数の発生法が「より乱数らしい」と言えるだけです．また，様々な検定を組み合わせても，「実際上この種の検定で棄却できるのは異常に不都合なアルゴリズムで，非常によいものを積極的に見出すことは，はなはだ難しい」（[1]の p. 32より抜粋）ことが以前から指摘されており，それゆえにより複雑な乱数検定法が多く考案され続けています．例えば，NIST SP 800-22[7]と呼ばれる乱数検定用のツール群には15種類もの検定法が含まれています．しかしながら，その検定法の中には問題点を指摘されているものも存在しているため[8]，利用には注意が必要です．

なお，NIST の検定法では疑似乱数に対する各テストの合否判定は，棄却域に入るサンプル数に対する Z 検定によって行っており，具体的には各テストにおいて 1 %棄却域に入った標本数が全体数 S に対して $S \times \left[0.01 \pm Z\sqrt{\frac{0.01(1-0.01)}{S}}\right]$，（$Z = 2.58$）の範囲に入っていれば，疑似乱数が「期待通りに有意水準 1 %の棄却域でテストに不合格した」ため「乱数検定のテストに合格した」という判断をします．ここでも，乱数の品質の検定はあくまで統計学的なスタンスでなされることが現れています．

NIST の合否判定を今回参考として行ったテストに当て

はめると，$S = 1$ 万なので合格の範囲は100±25ということになります．テスト結果の内訳を見ると，10個の SFMT の内の 2 つと線形合同法の上昇・下降連テストの χ^2 検定において126あるいは127個の標本が 1 %の棄却域に入った他は，全て 100±25 の範囲に収まりました．上昇連，下降連のテストについては筆者自身の経験として，一標本当たりの長さ N をかなり大きく取らないと，乱数の発生法に依らずテストから評価した χ^2 値が期待される χ^2 分布より若干上振れする傾向があります．これは長さ N の数列に含まれる連の総数自体が確率変数であることや，**図 2** からわかるように平均値以上・以下の連の分布と比較して，上昇・下降連の分布は長さ 5 や 6 以上の連の出現率が長さ 1， 2 の連に比べて極端に少ないため，χ^2 値の誤差（分散）が大きく出やすいためではないかと予想しています．いずれにしても，棄却されたサンプル数の 1， 2 程度の差をもって直ちに SFMT の内の 2 つと線形合同法がテストに不合格だとみなすのは早計でしょう．実際，1 サンプルの長さを10万から100万に延ばすと，全ての Z 検定，χ^2 検定において 1 %の棄却域に入った標本数が 100±25 の合格圏に収まりました．このように，検定法によって適切な 1 サンプルの大きさが異なったり，χ^2 値の出方に癖があったりするので，NIST やその他，世の中に出回っている乱数の検定法を使う際，テストのサイズ設定や合否判定の解釈には注意が必要です．

検定プログラム群としては他にも，TESTU01[9]と呼ばれる乱数の品質検定プログラム群が知られています．こちらはテストの規模に応じて SmallCrush，Crush，BigCrush いう 3 つの検定プログラム群が用意されており，最も詳細な検定をする BigCrush では106種類もの検定法による160通りのテストが行われます．文献[9]の Table.I には多くの乱数発生法に対する検定テストの結果が載っており，棄却域 10^{-8}% で不合格だったテストの数が示されています．残念ながら本節で試したものと同じ線形合同法はありませんが，パラメータの異なる様々な線形合同法の結果が載っており，おしなべて多くのテストで不合格になっています．Tausworthe 法（本節と同じものは文献[9]の Table.I 中の GSFR (250, 103)），MT 法（同 Table 中の MT19937）の結果も載っています．Tausworthe 法に比べて MT 法の方が不合格になったテストの数が少ないことがわかりますが，MT 法でも BigCrush の160個のテスト中 2 つで不合格になっています．これを「たった 2 つ」とみなすか，「 2 つも不合格」とみなすかは，先ほども述べたように万能な乱数のテストというものがないため難しいところです．なお MT 法については初期のバージョンには，初期化の seed の値を変えても初期ベクトルの状態があまり変わらない，一旦漸化式中のベクトルのビットが 0 が過多な状態に陥るとそこからなかなか抜け出せない，といった問題点が指摘されていました．これらが 2 つのテストで不合格になったことと関係があるかについては，筆者が調べた限りではっきり述べられたものはありませんでした．初期化の問題については本講座で紹介した mt19937ar において既に改善されており，また 0 超過状態からの回復についても並列化 MT 法として紹介した SFMT において改善されています．もし，BigCrush テストの結果を重視するのであれば，MT 法を改良して省メモリ化し，かつ BigCrush テストも全て通るように出力を工夫した TinyMT[10] が MT 法の作者らによって発表されているので，それを使うという選択肢もあります．ただし均等分布性の次元は元の MT 法より低くなっています．TinyMT はまだ新しく一般的に普及していないため本講座では紹介しませんでしたが，Dynamic Creator や Jump Method も備えているので，最新の疑似乱数を試したい方は利用してみてはいかがでしょうか．

・並列乱数の品質について

ここまで，1 つの疑似乱数発生法に対する品質の検定について説明してきましたが，最後に，第 3 章で紹介したような並列疑似乱数についての品質について少し述べたいと思います．乱数を並列化する上で重要な性質は，それらの乱数列同士が互いに独立で相関関係を持たないことです．第 3 章では MT 法の並列化の方法として，過去に本学会誌[11] で紹介した Dynamic Creator の他に，1 本の MT 法乱数列のそれぞれ遠く離れた点を出発点として，並列計算において異なる部分区間の乱数列を利用するという Jump Method を新たに取り上げました．並列計算によるシミュレーションで，これら並列疑似乱数を利用する時には，それらの乱数列が互いに独立であるという前提で利用するわけですが，その保証はどこにあるのでしょうか？

Dynamic Creator の場合は，MT 法を特徴づける漸化式の形が異なるものを必要な数だけ用意するというものでした．その独立性については，作者自身も厳密な数学的証明はなく，漸化式を特徴づける特性方程式が互いに素な疑似乱数同士は独立である，という仮説に基づいていると説明しています[12]．一方，Jump Method では 1 本の MT 法の疑似乱数列を例えば 2^{100} 飛びに使うという方法であり，それぞれの部分区間が統計学的な意味で独立とみなせるかどうかについて，筆者が調べた限り具体的な説明は見当たりませんでした．しかし，MT 法が持つ一般的な意味での疑似乱数の品質の良さや高次元均等分布性を踏まえると，部分区間 A と B から取った数列 $\{a_j\}$ と $\{b_j\}$ に対して，例えば $a_j b_{j+k}$ $(k = 0, 1, \dots)$ に対する相関係数を具体的に評価したところで，何かおかしな相関が現れることはないと期待して良さそうです．このように，MT 法による疑似乱数並列化に関しては，まず間違いなく大丈夫だろう，という感じで実応用に供されているというのが実情のように思われます．

筆者自身は15年くらい前から MT 法（当時は Dynamic Creator，現在は Jump Method）を並列シミュレーションに応用してきましたが，当時から並列乱数列の独立性について議論や検証の話が見当たらないことに不安を感じつつ使い始めました．最終的には，160本（後に1024本）の Dynamic Creator の MT 法乱数列について，${}_{160}C_2$ (${}_{1024}C_2$) 通りの全ての組み合わせに対して，相関チェックを自ら試すという実力行使に出ました．この際に考えた Checker-

board test と名付けた検定法[13]では，2つの乱数列 $\{a_j\}\{b_j\}\in[0,2^n]$ に対し，$(2^n\times 2^n)$ の2次元のマス目を考え，(a_j,b_j) を順次打っていき，iステップ目まで進んだ時に，全部で $N=2^{2n}$ 個のマス目のうちまだ1度も点を打たれていない数 $m(i)$ を観測します．$\{a_j\}\{b_j\}$ が無相関な一様乱数だと仮定した場合の期待値とその分散は

$$\mathrm{E}[m(i)]=N\left\{1-\left(\frac{N-1}{N}\right)^i\right\}\simeq N\exp\left(-\frac{i}{N}\right),$$

$$\sigma^2[m(i)]=N\left\{N+(N-1)\left(\frac{N-2}{N}\right)^2-(2N-1)\left(\frac{N-1}{N}\right)^i\right\}$$

と与えられます．これに対し，観測した $m(i)$ が $E[\mathrm{m}(i)]\pm(1\sim 3)\times\sigma(i)$ の範囲に現れる割合が，上式の期待値と分散を持つ正規分布に従うかどうかをチェックしました．もしテストの結果がそうなっていなければ $\{a_j\}\{b_j\}$ が無相関な一様乱数だという仮定が棄却される，即ち相関があると考えられるわけです．MT 法は32ビットの精度を持っていますが，その精度でチェックすることは時間的に不可能だったので，$\{a_j\}\{b_j\}\in[0,2^{14}]$ にビットを落として $i=20N$ まで追跡しました．その結果は[13]に示すように，独立な乱数として期待される通りに $m(i)$ が i と共に減少するということが，Dynamic Creator で作った全ての MT 法の乱数列の組み合わせについて確認できました．

このようなテストをわざわざするほど並列乱数の独立性について心配する必要は今にして思えばなかったかも知れません．また，このテストに限らず乱数の検定全般についてよく指摘される点ですが，いくら乱数のテストを行ったとしてもそれはテストを行った部分区間における品質をチェックしたに過ぎないということは否めません．しかしながら，並列乱数の独立性については自らテストをしなければ何も分からないという程に，検定法や検定結果が発表されていないというのが実情です．並列計算に疑似乱数を応用するプログラマは，その事実を頭の片隅に留めておいてもよいのではないか，と私は考えます．よりシビアに乱数の独立性を求めたいユーザは，第2.2節で紹介したような物理乱数の利用を検討するのも一つの手段だと思います．

・おわりに

最後に本章の補足として，実際にシミュレーションを行う際に乱数を検定してから使うべきか否か，という問題について少し述べたいと思います．例えば粒子シミュレーションで1空間セル当たり N 個のテスト粒子の初期速度を乱数で与える際に，大きさ N の疑似乱数に対して本節で例を示したような統計学的な検定テストを行えば，5％，1％など任意の有意水準で「一様乱数らしくない」疑似乱数のセットを排除することも可能です．しかし，疑似乱数が実際に真の乱数に近い挙動をしている場合，1％の棄却域に入るような N 個の疑似乱数のセットは，実際に約1％の確率で生じて然るべきものでもあります．空間セルの数が例えば10万個ある場合，統計的には1000個のセルで疑似乱数が1％の棄却域に落ちる可能性がある，ということです．

この例のように，初期分布を作るのに乱数を使う場合は1回きりのことなので，検定を掛けて棄却域に入った場合は乱数を作りなおしたとしても，総計算時間に対する計算時間の増加は無視できるでしょう．また，テスト粒子の平均初期速度がなるべく仮定した初期条件通りの状態から計算をスタートさせたいといった要請があるかも知れません．しかしそれならば，平均値が棄却域に入るような乱数でもそのまま N 個のテスト粒子の初速の決定に使い，後で粒子の速度を全体的にずらして仮定した初期条件に粒子の平均速度を合わせてもよいわけです．また，時間発展シミュレーションの中で，ランダムウォークや衝突計算を乱数で模擬する場合，1％の棄却域に入る乱数を一々排除していると計算時間の増大につながりかねませんし，何より本当に1％の確率で起こるはずの「レアイベント」を人為的に排除してしまう可能性があります．

筆者自身は初期化においても時間発展計算の最中においても，疑似乱数の検定はしてもいいが，それは乱数の品質の確認のためだけであり，棄却域に入る乱数もそのまま使うというスタンスを取りますが，これはシミュレーションコードが扱う問題の性質や研究者個々人の乱数に対する考え方によって変わりうるものだと思います．ここで強調しておきたいことは，乱数の検定法が統計的な性質に基づくものである以上，真の乱数でも疑似乱数でも，ある確率で検定の棄却域に入る乱数の標本は必ず生じ，むしろそれは乱数として自然な振る舞いであるということです．

本章では筆者自身の過去の経験を交えながら，疑似乱数の品質について概説してみました．乱数を使うシミュレーションを行う研究者の方々が，ふと「自分が使っているこの疑似乱数は大丈夫なのだろうか？」という疑問を感じた時にこの記事で紹介した内容が参考になれば幸いです．

参考文献

［1］津田孝夫：モンテカルロ法とシミュレーション（三訂版）（培風館，1995）．

［2］松本 眞，栗田良春：Twisted GFSR:新しい乱数発生法，京都大学数理解析研究所講究録 **85**, 86-95 (1993).

［3］伏見正則：乱数（東京大学出版会，1989）．

［4］B.J. Collins, G. Barry Hembree, J. Association for Computing Machinery 33, 706-711 (1986).

［5］松本 眞：有限体の擬似乱数への応用，http://www.math.sci.hiroshima-u.ac.jp/~m-mat/TOKYOHOMEPAGE/TEACH/1011.pdf

［6］H. Levene and J. Wolfowitz, Ann. Math. Stat. **15**, 58-69 (1944).

［7］https://csrc.nist.gov/projects/random-bit-generation/documentation-and-software

［8］吉田等明 等：電子情報通信学会 信学技報 **112**, 13 (2012).

［9］P. L'ecuyer and R. Simard, ACM T. Math. Software, 33, Article 22 (2007).

［10］M. Saito *et al.,* RFC 8682, DOI 10.17487/RFC8682 (https://www.rfc-editor.org/info/rfc8682).
［11］講座「核融合プラズマシミュレーションの技法 5．粒子シミュレーションのコーディング技法」プラズマ・核融合学会誌 **89**, 245 (2013).
［12］M. Matsumoto and T. Nishimura, "Dynamic Creation of Pseudorandom N umber Generators", Monte-Carlo and Quasi-Monte Carlo Methods pp. 56-69 (Springer, Berlin, Heidelberg, 1998).
［13］S. Satake *et al.,* Lecture Notes in Computer Science book series Vol. 4759, pp. 344-357 (Springer-Verlag, Berlin, Heidelberg, 2006).

なお、本講座の執筆のために統計数理研究所の研究者に意見を聞きに伺ったことがきっかけで、朝日新聞の記者から乱数のシミュレーション研究への応用について取材を受けました。その記事は2020年 6 月 3 日の朝日新聞夕刊および朝日新聞デジタルに『現場へ！「乱数の世界へようこそ」』というタイトルで掲載されました。

佐竹真介（さたけしんすけ）

自然科学研究機構 核融合科学研究所 ヘリカル研究部 核融合理論シミュレーション研究系 准教授，2003年 総合研究大学院大学 博士（学術）．
モンテカルロ法を使った 3 次元磁場配位中の新古典輸送現象，新古典粘性のシミュレーションや最適化配位の研究が主なテーマ．乱数については深い思い入れがありますが，特にギャンブル好きというわけではありません．