でたらめを利用する

逆瀬川　浩孝

さいころは振るたびに出る目の数が違い，その数を予想することはできない．また普通はどの目も同等に出やすくできているので，古今東西，その特徴を利用したゲームが数限りなく発明され，数えきれない人々の楽しみのために奉仕してきた．そのさいころが，ゲームに限らず，学問の世界で大きな役割を果たしている，と聞かされたら意外に思う読者が多いに違いない．さいころを何回も振った結果を使えば円周率が推定できる．あるいは，スーパーのレジの混雑が，レジを１台増やすことによってどれくらい改善されるのか予想できる．ほんとかしら．

乱数とは

さいころを振ってみたら最初の10回は次のような目が出た．

1，5，3，4，3，3，6，2，1，6

さいころをさらに振り続ければ，このようにでたらめに並んだ数列がいくらでも続くだろう．「でたらめ」とは何かと聞かれれば，「規則的でないこと」あるいは「次に何が来るか予想がつかないこと」というような答えが返ってくることが予想される．さいころの数列は，それに加えてもう１つの重要な性質がある．それは，十分に多くの数字を集計してみると６つの数字がほぼ均等に（一様に，ともいう）出現している，ということである．このような性質を持った数列を一様乱数列，あるいはただ単に乱数ということにしよう．１から６までの数字ではなく０から９までの10個の数字をでたらめに並べたものは乱数表として知られている．乱数の性質から乱数表はどこから読み始めても，どのような順番で読んでいってもでたらめである．

さかせがわ　ひろたか　早稲田大学　理工学部工業経営学科

〒169 新宿区大久保３－４－１

表１　乱数表（部分）

78161	88382	71490	66940	14852
88849	38473	78485	19881	76785
62814	95306	26635	0035	
7053	83497	0861		

ここで簡単な実験をしてみよう．

実験１　数をでたらめに並べれば良いのだから，自分でも乱数表ができそうである．自作の乱数表ができたとして，最初の10個の数字を書いてみてください．

この実験で多くの人は０，３，２，８，７，４，６，９，１，５というような数字を書く傾向にある．このような10個組みを並べたものは乱数表と言えるだろうか．この数列の特徴は０から９までの数字が１回ずつ現われるということである．したがって，たとえばこの場合，最後の数字は５と決まっていてでたらめではないことになる．

逆に１を10個書いた人は相当ひねくれている人である．乱数表は特定の規則がない数字の列だから，０から９までの数字がほぼ均等に現われているのはもちろんのこと，隣どうしの２つの数字を組み合わせてできる２桁の数00から99までの数字も均等に出現しなければならないし，３つ，４つ，... を組み合わせた数字も均等でなければならない．そうすると10桁の数である1111111111も大きさが10^{10}の乱数表の中に含まれていてもおかしくはないということになり，それがたまたま先頭に来ただけのことかもしれない．ふつうはそのような数の並びは乱数とは言わないが，もし10^{20}個の数字を並べたときに1111111111が１回も現われなかったとしたら，それは乱数とはいえないと言ってよいだろう．

このように乱数かどうかを決めるのは，たくさんの数を集めたときにその数が全体として満たすべき性質を持っているかどうかを調べる必要があるので，ある数が乱数かどうか，という質問は意味を持たない．

でたらめな数は，なにもさいころを振らなくても見い出すことができる．たとえば缶ジュースの中身の量を0.01 ccまで量れる器具を使って量ったとき，同じ銘柄のものでも結果はばらつくはずである．コンビニの来客数を5分ごとに数えれば，その数字はばらついている．公衆電話の1人の通話時間は一定ではない．

これらの数字がばらつくということは，いずれも制御できない要因が絡んでいる，ということで納得できるが，完全に規則的な計算手順によってもでたらめな数列ができる．次の数列はその1つの例だが，さて，どのようにして作ったものだろう．

0.3141，0.5926，0.5358，0.9793，0.2384，...

少数点以下の数字を順番に並べて書けば 31415926535897932384...となり，これは3のあとに少数点を補えば言うまでもなく円周率になる．円周率は単位円の周の長さという決まった値で，公式を使って導くことができるが，その数字はでたらめに並んでいるようにみえる．計算手順が分からなければでたらめにみえるという数列を作ることは円周率を持ち出さなくても，もっと簡単に作ることができる．

乱数を作る

最初の数をたとえば1111と決める．それにマジックナンバー179を掛けて末尾4桁の数を取り出すと8869が得られる．8869に再び179を掛けて末尾4桁の数を取り出すと7551が得られる．このように，前の数に179を掛けて末尾4桁の数を取り出す，という計算を続けると下のような数列が得られる．

(*)　1111，8869，7551，1629，1591，4789，7231，4349，8471，6309，9311，...

ここで，これらの数の千の位の数字だけを並べると

1，8，7，1，1，4，7，4，8，6，9，...

となる．さて，179という数字を知らなければ3の次に来る数字は何だろうと聞かれてすぐに答えられる人はいないに違いない．このように計算によってでたらめそうに並んだ数列ができるとき，その数列を擬似乱数という．

上の方法で末尾4桁の代わりに末尾2桁とした場合でも

11，69，51，29，91，89，31，49，71，09，...

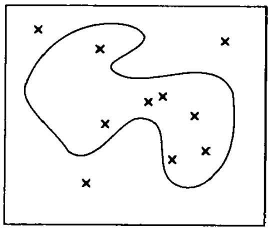
のようにでたらめに並んでいるようにみえるが，09のあとは11となり，そのあとはこの10個の数字が周期的に繰り返される．これではとてもでたらめとはいえない．周期的であるという性質は末尾を4桁取ったとしても事情は同じであるが，人が覚えられないくらい十分に長い周期であれば，乱数として認めてもよいだろう．パソコンゲームの意外性のもとはこのような擬似乱数なのである．

末尾の桁をどれくらい取れば周期の長い数列が得られるのか，マジックナンバーの179はどうやって決めたのか，という問題は数学を使って解くことができるし，擬似乱数がでたらめといえるのかということは統計を使って調べることができる．

図形の面積を測る

図形の面積は図形を細かな領域に分割してから，各々の面積を積み上げることによって求めることができる．定積分の考え方である．ここにはでたらめに並んだ数列の入り込む余地はないように思われる．ところが...

実験2　図形を紙に書いて，雨が降り始めたとき地面に置いて，雨粒がどこに落ちたか見える程度の状態で紙を引っ込める．紙に落ちた雨粒の数n，そのうち図形の内部に落ちた雨粒の数mを数えて，図形の内部に落ちた雨粒の相対頻度m/nを計算したとき，それはどういう意味があるだろう．



図形の内部に落ちた雨粒の相対頻度は紙全体の面積に対する図形の相対的面積にほぼ近いことが期待されるので，紙を1メートル四方に切っておけば，図形の面積はほぼm/n平方メートルと言えるだろう．

さて，雨粒はでたらめに落ちてくると考えてよいから，紙に直角座標系を設定すれば，雨粒の落下点のx座標，y座標は乱数と言ってよい．ここで面積と乱数が結び付いたことになる．

たとえば，前に書いた円周率の数字を使った擬似乱数列を先頭から2つずつ区切って((0.3141，0.5926)，(0.5358，0.9793)，...)それを1つ1つの点のx-y座標とする．そして1メートル四方の紙にその座標を持つ点を描いてゆく．これと実際に雨粒が落ちた点を描いたものと比較しても，どっちがどっちと言えないものができるはずである．

そこで計算機を使って仮想的に雨を降らせ，図形の

面積を計算することができる．このように，実際のでたらめで予測不可能な現象を計算機のなかで仮想的に実現し，その実験にもとづいて問題を解く方法をシミュレーションと呼んでいる．

円周率を計算する

図形を一辺1メートルの紙に内接する4分円とすれば，その面積は $\pi/4$ なので，今述べた方法で円周率を「計算する」ことができる．でたらめに選んだ点がその4分円の内部にあるかどうかを判定するためには，その点の座標を (x, y) としたとき，x^2+y^2 が1より小さいかどうかを確かめればよい．

点の数を増やすほど相対頻度 $4m/n$ は円周率に近づくと思われる．n を変えて実際に計算してみたのが表2である．

表2　円周率の計算実験

点の数	100	1000	10000	100000	1000000
1回目の実験	3.20	3.144	3.1432	3.1400	3.1399
2回目の実験	3.00	3.016	3.1520	3.1421	3.1409

この表から，実際 n を増やすと円周率の数字に近づいているようにみえること，しかしその近づき方は非常に緩やかで，しかも実験ごとに結果が異なり，円周率が大体3.14くらいだということはわかるが，この方法を使ってたとえば小数点以下7桁目が6であることを知ることはかなりむずかしいこと，などがわかる．このように，理論的に式が立てられていて計算ができる問題に対して乱数を使って答えを求める方法を特にモンテカルロ・シミュレーションという．モンテカルロは地中海にある賭博が公認の国モナコにある地区の名前である．でたらめ＝さいころ＝賭博という図式からこのような名前がつけられた．

実験の精度と信頼性

シミュレーション実験は物理や化学の実験と異なり，乱数を替えるたびに違った実験結果が得られる．したがって，そのさまざまな結果から真の値をどのように「推定」すればよいのか，あるいは実験結果が真の値とどれくらい近いのかを知る必要がある．たとえば表2のような実験を繰り返すと，点の数が100の場合，実験結果は0.3くらいの誤差があるが，点の数を1000000に増やすと誤差は0.003くらいになる，というようなことがわかる．でたらめに落とす点の数が増えれば増えるほど，推定の誤差が小さくなることは直観的に理解できるが，どれくらいの点をとればどれくらい正確なのか，少数点以下3桁までわかるためにはどれくらいの点をとる必要があるのか，というような量的な関係について考えてみよう．

正確さとは，いいかえれば同じ実験を乱数を替えて実行したとき，2つの結果がどれくらい食い違う可能性があるか，ということである．

でたらめに点を落としたとき，その点が図形の中に含まれるチャンス，これを確率というが，それは図形の面積と紙全体の面積の比で決まってくる．紙の面積を1としたときの図形の面積を p とすると，図形の内部に落ちる確率は p と考えてよい．そうすると，上の実験は，たとえば100円玉を投げて表が出れば点が図形に落ちた，裏が出たら図形の外に点が落ちた，というように解釈して，表の出た相対頻度によって表の出る確率 p（この場合は0.5に近い数と期待されるが，歪んでいればどうなるかわからない）を推定する問題と同じになる．

そこで，100円玉の実験を少し詳しく調べてみよう．

実験3　100円玉を10回投げたとき表がちょうど5回出る可能性はどれくらいあるだろうか．最初は直観的に結果を予想し，次に手元の100円玉を使って実験して見てください．

100円玉を2回投げたとき，表裏の出るパターンは，表を○，裏を×とすると，次の4通りにまとめられる．

○○，○×，×○，××

表の出た相対頻度で $p(=0.5)$ を推定しようとすると，正しい結果を推定しているのは全体の50％に過ぎない．

投げる回数を増やしていくとどうなるだろう．3回投げたとき，あるいは奇数回投げたとき相対頻度が0.5になることはありえないから，正しい推定をする可能性は0である．4回投げたときは，2回の時と同様に○と×を使って確かめてみると，4つの中に○がちょうど2つ入る並び方は6通りあるが，これは順列組合せの考え方を使って求めることができる．全体で4つの記号の組合せは16通りあるので，正しい推定をする可能性（これも確率という）は6/16となる．6回投げたときは，6つの記号の組合せは64通りあり，○がちょうど3回出る並び方は ${}_6C_3=20$ 通りなので，正しく推定される確率は20/64となる．一般に100円玉を n 回（n は偶数）投げたとき，n 個の記号の組

合せは 2^n 通りあり，○がちょうど $n/2$ 回出る並び方は ${}_nC_{n/2}$ 通りなので，正しく推定される確率 r は ${}_nC_{n/2}/2^n$ となる．これをいくつかの n について計算したのが下の表である．

n	4	10	20	50	100	200
r	0.375	0.246	0.176	0.112	0.080	0.056

この表から，実験を繰り返せば繰り返すほど表の出る相対頻度が真の値に等しい可能性が0になるということがわかる．実験3で，はじめに4分の1と予想した人はどれくらいいますか．

この結果を見て，このような実験で真の値を調べることは不可能だ，という結論を出すのはまだ早い．実験3を繰り返し試した人は，表の出る回数が毎回ばらつき，4回，5回，6回となることが多いが，0回，1回，9回，10回となることはほとんどなかった，という結果を得たことだろう．このようなでたらめな現象を相手にした実験では乱数と同様，多くの実験結果を集めたとき，それらが集団として持つ性質を論じなければいけない，ということが言える．この場合は，相対頻度は0.5に近い値になる可能性が大きい，ということが重要である．

相対頻度と真の値との差を誤差という．たとえば上の実験で10回投げたときの誤差が0.1以下になる確率は表が4回，5回，6回となる確率を加えたものに等しい．これらの確率は上の計算同様，順列組合せの考え方を利用して計算することができる．その結果は約0.66である．同様に20回，50回，100回，200回と実験を増やしていった場合の誤差が0.1以下になる確率 q は下の表のようになる．

n	10	20	50	100	200
q	0.656	0.737	0.881	0.965	0.996

50回コインを投げるという実験で誤差が0.1以内に収まる確率が0.881ということは，たくさんの人が同じ実験をしたとき9人中8人までが表の出る回数が20回から30回の間になる，ということを意味している．そして200回のコイン投げではほとんど確実に(250人中249人)相対頻度は0.4と0.6のあいだにあるということが言える．したがって，上の表の q は実験結果の信頼度を表わしていると言ってよい．この表から，投げる回数を増やせば信頼度が上がることがわかる．

一方，上の表の q が一定になるような誤差の範囲は，投げる回数を増やせば小さくなることが予想される．$q=0.95$ に近くなるようにすると誤差 e は次の表のようになる．

n	50	100	200
e	0.14	0.1	0.07
$1/\sqrt{n}$	0.141	0.1	0.071

この表は，たくさんの人がたとえば200回コインを投げたとすると，20人中19人までが表の出る回数が86回から114回の間になる，ということを表わしている．この表から，コイン投げの回数を増やせば増やすほど，相対度数が真の値に近くなる可能性が大きくなるということがわかる．さらにデータを増やして n と e の関係を図に表わすと下のようになり，簡単な関係がありそうなことがわかる．実際，上の表の3段目の数字は $1/\sqrt{n}$ を計算したものであるが，それらは e にほとんど等しい．この結果，信頼性が同じならばコインを投げる回数の平方根に反比例して誤差が小さくなることがわかる．このことは，誤差を半分にするためには回数を倍ではなく4倍の実験を実施しないと駄目，ということを意味している．p が0.5と限らない一般の場合でも事情は同じである．

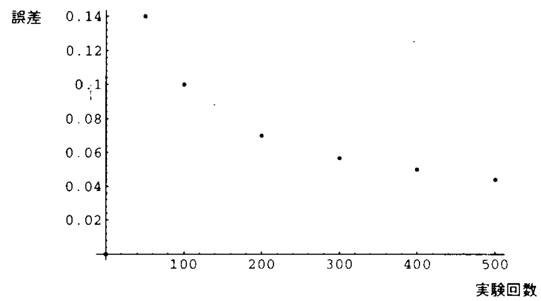


図1 実験回数と誤差

このように考えると，表1の実験結果で，点を10000個取ったときは3.1までわかり，1000000個取ったときは3.14までわかったことが説明できる．昔から試みられていた円周率の計算法である内接正多角形と外接正多角形の周の長さを計算するという方法では，10000角形の場合を計算すると3.1415926くらいまで正確に計算できることを考えると，この方法はあまり効率が良いとは言えないが，もっと複雑な問題になった時，乱数を使ってある程度の結果が得られるという点で有効な方法である．

参考文献

森戸晋ほか「システム・シミュレーション入門」の第1章，共立出版，1993