

第5章 乱数

第III部では、主に Excel を活用したシミュレーション問題を取り上げている。本章では、次章以降の本題に入る前にシミュレーション問題で共通に利用する乱数についてふれておく。

5.1 はじめに

自然現象や社会現象の多くは確率的な要素を含む。仮に系の時間発展が原理的に決定論的なものであったとしても、実際には我々は全ての要素を知ることができず、確率的な議論をしなければならぬことは多い。例えば、サイコロを振って次にどの目が出るか、は原理的には初期条件がわかれば運動方程式を解けば求まるのであろうが、実際にサイコロを振る時の初期条件を正確に知ることは困難である。そこで確率的な議論をするわけである。系の時間発展が単純な規則で記述される場合も、初期条件の微小な差が時間とともに拡大される場合がある、ということは「カオス」の基本原則である。

逆に、確率的な議論を行なうことで、決定論的に扱うのは非常に困難な系を簡単に扱うことができる場合がある。いろいろな現象をコンピュータで数値的に解析する場合も、確率的な要素を導入すると便利ことが多い。しかし、コンピュータの計算は決定論的であるため、普通にやると確率的な要素は入ってこない。

コンピュータで確率的な計算をするには、「乱数」つまりランダムな値を取る確率変数が必要である。実際には何らかの決定論的な計算で乱数を作るため、真の乱数は得られない。そこで、正確にはコンピュータで用いる乱数を「疑似乱数」と呼ぶ。「真の乱数」とはどういうもので疑似乱数はそれとどう異なるか、というのは難しい問題である。

Excel には、組み込み関数として疑似乱数 `rand()` が用意されている。`rand()` は一様な確率で 0 から 1 までの間の実数の値を取る。例えば任意のセルで次のようにすれば乱数を表示することができる。

```
= rand()
```

この関数を利用すると、他の様々な分布に従う乱数を生成することができる。例えば、5.3 で取り上げる正規分布や 8 章で必要となる指数分布に従う乱数である。指数分布の例をあげると、その密度分布は、

$$f(t) = \lambda e^{-\lambda t} \quad (5.1)$$

であり、累積分布は

$$F(x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x} \quad (5.2)$$

であるから、一様乱数 y に対して、 $F(x) = y$ を解いて

$$x = -\frac{\log(1-y)}{\lambda} \quad (5.3)$$

を求めれば、 x は指数乱数になる。これを Excel 内で記述すると次のようになる。

$$= -\log(1 - \text{rand}()) / [\lambda \text{の値を格納したセル}]$$

5.2 演習 — 乱数のランダム性について

情報処理演習では、モンテカルロシミュレーション、待ち行列など乱数を用いたシミュレーション課題を取り上げる。乱数はランダム (出現確率が同じ割合で規則性を持たない) な一様分布の性質を満たさなければならない。この性質を完全に満たす乱数の生成は難しく、実際には疑似乱数と呼ばれる実用上乱数とみなしうる数列が用いられている。演習で用いる関数 `rand()` も、この疑似乱数を生成する。

この疑似乱数のランダム性を調べるには、通常、

- 等確率性の検証：度数検定
- 不規則性の検証：間隔検定，ポーカー検定 (χ 二乗検定がベース)

が行われる。ここでは、ヒストグラムと相関係数を使って簡単に調べてみる。

- `rand()` を使って、乱数を多数生成し、ヒストグラムを作成し、一様分布になるかどうか確認せよ。
- 2 個の乱数からなる組を多数生成し、その相関係数を調べてみよ。

5.3 演習 — 正規分布に従う乱数の生成に関して

平均 μ 、標準偏差 σ の正規分布に従う乱数 z は、標準正規分布に従う乱数 x から、

$$z = \mu + \sigma x$$

と求めることができるので、標準正規分布に従う乱数さえ生成できればよいことになる。

標準正規分布に従う乱数を生成する簡単な方法は、 $0 \leq U < 1$ の一様乱数 U を 12 個加えて 6 を引くという方法である。一様乱数に限らず、どんな乱数でも、多数加えると中心極限定理により正規分布に近づく。一様乱数 U は平均 $\frac{1}{2}$ 、分散 $\frac{1}{12}$ なので、ちょうど 12 個加えて 6 を引くと平均 0、分散 1 になる。これは標準正規分布を意味する。したがって、Excel では、

$$=\text{rand}()+\text{rand}()+\text{rand}()+\text{rand}()+\text{rand}()+\text{rand}()+\text{rand}()+\text{rand}()+\text{rand}()+\text{rand}()+\text{rand}()+\text{rand}()+\text{rand}()+\text{rand}()-6$$

と数式を任意のセルに入力すると、標準正規分布に従う乱数を生成できることになる (上記の例は 2 行にわたっているが、実際は 1 行)。

また、Excel では乱数生成ツールが準備されている。「ツール」→「分析ツール」→「乱数発生」を選択し、分布や発生数を指定すると目的の乱数を生成してくれる。標準正規分布の場合、図 5.1 のように、「乱数発生数」を指定し、「分布」として「正規」を、「平均」として「0」を、「標準偏差」として「1」を指定し、「出力先」を指定すれば正規乱数を生成してくれる。

この Excel の乱数生成機能と上記の方法を用いて、それぞれ標準正規分布に従う乱数を多数生成し、その平均値、標準偏差などを比較することにより、正確な乱数が生成されているかどうかを調べることができる。

ただし、Excel 内部では、乱数生成にどのような方法を使っているか把握することはできないが、おそらく Box-Muller(極座標)法のような正確な正規分布の乱数を発生させる原理を採用しているものと考えられる。

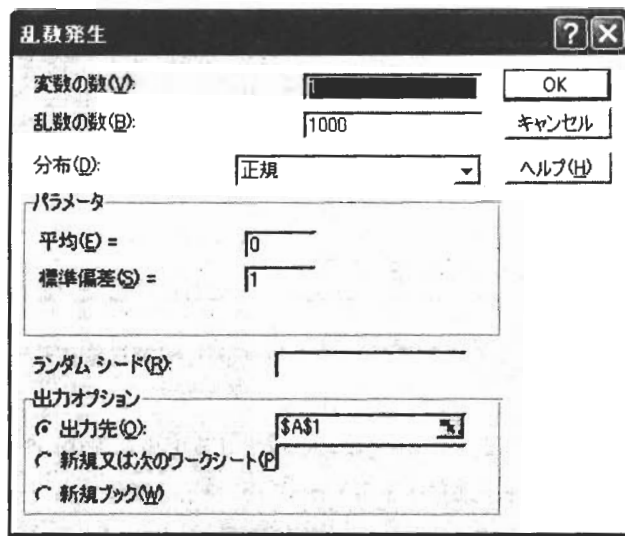


図 5.1: 正規乱数の発生

余裕があるならば、12.10 に示した通りに独自に VBA を用いて正規乱数生成のための関数を作成し、上記方法と比較してみよ。