## 第10回: 擬似乱数の生成法

確率的な事象を、計算機上でシミュレーションするには、どうしたら良いか？

ある確率分布に従って生起する事象を、計算機上で実現するには、どうしたら良いか？

⇒何らかの方法で、「サイコロ振り」をプログラムによって実現する必要がある。

⇒ 乱数 を生成する。

## 乱数

以下の2つを満たす数の系列を乱数という。

(1)不規則性: 数字の並び方がでたらめ(不規則)

(2)一様性: 各数字が同じ割合で出現する

乱数表:

乱数サイ: 正二十面体のサイコロ

## 一様乱数

区間$0 \leqq x < 1$の一様乱数

⇔

連続な確率変数$x$が、以下の確率密度関数をもつ

$$f(x)=\begin{cases}1 & (0\le x<1)\\ 0 & (x<0, x\ge 1)\end{cases}$$

一様乱数の生成法

- 乱数表から
- 計算機の利用

## 生成法: 合同式法 (congruence method)

### 区間 $0 \leqq x < 1$の一様乱数の生成法

手続き:

$a, c, m \in N$(自然数)。このとき、$r_0 \in N$を出発点として、以下の漸化式により、自然数の数列 $\{r_0, r_1, r_2, \ldots\}$を作る。

$$r_{n+1} \equiv a \cdot r_n + c \pmod m$$

$0 \leqq r_n < m$ より、$x_n = r_n/m$ は、$0 \leqq x_n < 1$

mod付きの等式を「合同式」と呼ぶ

---

例えば $a=1229,\ c=351750,\ m=1664501$ を使うと

$r_0 = 1, (x_0 = 0.000001504)$

$r_1 = (1229\times 1 + 351750)\,\%1664501 = 352979\%1664501 = 352979$

$x_1 = 352979/1664501 = 0.212062954$

$r_1 = 352979,\ x_1 = 0.212063$

$r_2 = (1229\times 352979 + 351750)\,\%1664501 = 434162941\%1664501 = 1392681$

$x_2 = 1392681/1664501 = 0.836695802$

$r_2 = 1392681,\ x_2 = 0.836695802$

$r_3 = (1229\times 1392681 + 351750)\,\%1664501 = 1711956699\%1664501 = 849671$

$x_3 = 849671/1664501 = 0.510465899$

$r_3 = 849671,\ x_1 = 0.510465899$


## 計算機で生成できるのは擬似乱数

× 見かけは一見「でたらめ」な数列だが、先の合同式法の規則に基づいて(決定論的に)生成されているので、決して「不規則」とは言えない。

○ 一様性は、おおむね成立している(ことが知られている)。

⇒ 「真の」乱数ではなく、擬似乱数

しかも、合同式法で生成すると、周期的になる宿命。ただ、その周期が非常に長いことを期待。

→ 周期が長いような、$a$, $c$, $m$ を選ぶ。

## 正規乱数の生成法

正規乱数：連続な確率変数 $x$ が、標準正規分布$N(0,1)$に従うような乱数

区間 $0 \leqq x < 1$ の（擬似）一様乱数から、（擬似）正規乱数を生成する方法：

生成法1：中心極限定理を用いる簡便法

手続き：

一様乱数を、12個加えて6を引く

つまり、$0 \leqq x < 1$の一様乱数として12個の$x_1, x_2, \ldots, x_{12}$を生成し、それから、$y = x_1 + x_2 + \ldots + x_{12} - 6$ を求めて、1つの正規乱数とする。

## 復習　中心極限定理

とはどんな定理か…

● 中心極限定理 (central limit theorem)

$n$個の確率変数$x_1, x_2, \ldots, x_n$ が、互いに独立に同じ確率分布に従うとする。その分布の平均を$\mu$、分散を$\sigma^2$とする；

$$E(x_i) = \mu, \quad E\left((x_i - \mu)^2\right) = \sigma^2 \quad (i = 1, 2, \cdots, n)$$

この$n$個の変数の平均 $\bar{x} = \frac{1}{n} \cdot (x_1 + x_2 + \cdots + x_n)$

（これも確率変数）は、$n$が大きいときには、正規分布$N(\mu, \sigma^2/n)$に従う。

## 中心極限定理

- $n$ 個の確率変数の平均（'中心'）は、
- $n$ が十分大きな極限においては、
- （もとの各$x_i$が従う確率分布がどうであれ）

正規分布$N(\mu, \sigma^2/n)$に従う

…という定理。

平均は$\mu$のまま（分布の中心は同じところ）、分散は$1/n$で小さくなってゆく（バラツキは小さくなる）。

同じ対象を、（念のために）何度も繰り返し測定して、その平均をとることにする。すると、より多くの測定値を平均するほうが、値のバラツキが小さくなり、精度が上がる…

## 元に戻って　$y = x_1 + x_2 + x_3 + \cdots + x_{12} - 6$ が $N(0,1)$に従うこと

- 中心極限定理が保障していること：どんな確率分布でも（よって、$0 \leqq x < 1$の一様分布でも）、$n$ 個の確率変数の平均は、$n$ が十分大きいとき、正規分布$N(\mu, \sigma^2/n)$に従う。
- いま、$0 \leqq x < 1$の一様分布では、平均$\mu = 1/2$, 分散$\sigma^2 = 1/12$（計算！）であり、

$n=12$で考えると、$\bar{x} = \frac{1}{12} \sum_{i=1}^{12} x_i$ は $N(1/2, 1/12^2)$ に従う。つまり、

$y = 12 \cdot \bar{x} - 6 = \frac{\bar{x} - 1/2}{1/12}$ は、$N(0, 1)$ に従う。

## 生成法2：ボックス・ミュラー法 (Box-Muller's method)

手続き：

$0 \leqq x < 1$の一様乱数として2つの$x_1$, $x_2$を生成し、それから、

$$y_1 = \sqrt{-2\log x_1} \cdot \cos(2\pi x_2)$$
$$y_2 = \sqrt{-2\log x_1} \cdot \sin(2\pi x_2)$$

を求めて、2つの正規乱数 $y_1, y_2$ とする。

## $y_1, y_2$ が$N(0,1)$に従うこと

$$dx_1 dx_2 = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} dy_1 dy_2 = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_1^2}{2}\right) dy_1 \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_2^2}{2}\right) dy_2$$

ヤコビアン(Jacobian)

計算は次で…

$$\underbrace{\mathrm{P}(x_1)}_{=1} dx_1 \cdot \underbrace{\mathrm{P}(x_2)}_{=1} dx_2 = \underbrace{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_1^2}{2}\right)}_{\substack{= f(y_1) \\ = N(0,1)}} dy_1 \cdot \underbrace{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_2^2}{2}\right)}_{\substack{= f(y_2) \\ = N(0,1)}} dy_2$$

$$x_1, x_2 : 0 \to 1 \quad \Rightarrow \quad y_1, y_2 : -\infty \to +\infty$$

## ヤコビアン $\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$ の計算（1 /2ページ目)

$$\begin{cases} y_1 = \sqrt{-2\log x_1} \cdot \cos(2\pi x_2) \\ y_2 = \sqrt{-2\log x_1} \cdot \sin(2\pi x_2) \end{cases}$$

$$\therefore \quad y_1^2 + y_2^2 = -2\log x_1, \quad \frac{y_1}{y_2} = \tan(2\pi x_2)$$

$$\therefore \quad x_1 = \exp\left(-\frac{y_1^2 + y_2^2}{2}\right), \quad x_2 = \frac{1}{2\pi}\tan^{-1}\left(\frac{y_1}{y_2}\right)$$

$$\therefore \quad \frac{\partial x_1}{\partial y_1} = -y_1 x_1, \quad \frac{\partial x_1}{\partial y_2} = -y_2 x_1$$

$$\frac{\partial x_2}{\partial y_1} = \frac{1}{2\pi} \cdot \frac{1}{1+\left(\frac{y_1}{y_2}\right)^2} \cdot \frac{1}{y_2}, \quad \frac{\partial x_2}{\partial y_2} = \frac{1}{2\pi} \cdot \frac{1}{1+\left(\frac{y_1}{y_2}\right)^2} \cdot \left(-\frac{y_1}{y_2^2}\right)$$

$$\frac{d}{dx}\left(\tan^{-1} x\right) = \frac{1}{1+x^2}$$

## ヤコビアン $\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$ の計算（2 /2ページ目)

$$\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \frac{1}{2\pi} \cdot \frac{1}{1+\left(\frac{y_1}{y_2}\right)^2} \cdot x_1 \cdot \begin{vmatrix} -y_1 & -y_2 \\ \frac{1}{y_2} & -\frac{y_1}{y_2^2} \end{vmatrix}$$

$$= \frac{1}{2\pi} \cdot \frac{1}{1+\left(\frac{y_1}{y_2}\right)^2} \cdot x_1 \cdot \left(1 + \frac{y_1^2}{y_2^2}\right) = \frac{x_1}{2\pi} = \frac{1}{2\pi} \cdot \exp\left(-\frac{y_1^2 + y_2^2}{2}\right)$$

$$\underbrace{\mathrm{P}(x_1)}_{=1} dx_1 \cdot \underbrace{\mathrm{P}(x_2)}_{=1} dx_2 = \underbrace{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_1^2}{2}\right)}_{\substack{=f(y_1) \\ =N(0,1)}} dy_1 \cdot \underbrace{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_2^2}{2}\right)}_{\substack{=f(y_2) \\ =N(0,1)}} dy_2$$

## レポート課題（配布）

一様乱数から、2つの方法で正規乱数を生成するプログラムを作成し、実際に10万個の乱数を発生させて、その性質を調べなさい。

- 課題1：「中心極限定理を用いる簡便法」
- 課題2：「ボックス・ミュラー法」
- 課題3：課題1，2の結果から、2つの方法で生成された正規乱数の質の違いを考察しなさい。

テキスト138ページの図8．7は、実行結果の1例

## ポアソン乱数の生成法

ポアソン乱数：離散的な確率変数 $n$ が、 ポアソン分布 $P(\lambda)$に従うような乱数

$$p(n) = \frac{\lambda^n}{n!} \cdot e^{-\lambda}, \quad n = 0,1,2,\cdots$$

区間 $0 \leqq x < 1$の（擬似）一様乱数から、（擬似）ポアソン乱数を生成する方法：

一様乱数法

手続き：

$0 \leqq x < 1$の一様乱数の系列$x_1, x_2, x_3, \ldots$から、

⇒続く

## ポアソン乱数の生成法（続き）

$$y_0 = e^{\lambda} \cdot x_0, \; y_1 = e^{\lambda} \cdot x_0 x_1, \; y_2 = e^{\lambda} \cdot x_0 x_1 x_2, \; y_3 = e^{\lambda} \cdot x_0 x_1 x_2 x_3, \; \cdots$$

によって

- 系列$x_0, x_1, x_2, \ldots$から系列$y_0, y_1, y_2, \ldots$を作る。
- $x_i < 1$ より、$y_0 > y_1 > y_2 > y_3 > \ldots$である。
- そこで、初めて $y_n \leqq 1$ となる $n$ を求める。
- この$n$ を、ポアソン乱数として得る。

（このようにして得た $n$ がポアソン分布に従うことの証明は略）