Random Numbers and Law of Large Numbers

モデル化とシミュレーション特論
2021 年度前期
佐賀大学理工学研究科　只木進一

Pseudo random number generators

- `AbstractRandom` class
  - `java.util.Random` class inside
  - Generate next random `getNext()`
- Transform method (変換法): `Transform` class
- Rejection method (棄却法): `Rejection` class

Sample Program

`https://github.com/modeling-and-simulation-mc-saga/Random`

Transform Method

- Probability density $f(x)$ ($x \in [a, b)$) and probability distribution F

$$F(x) = \int_a^x f(z)\mathrm{d}z \tag{1}$$

- Tranform method is available if the inverse of $F(x)$ is obtained.
- Process
    - Generate a random number $r \in [0, 1)$.
    - $x = F^{-1}(r)$
    - $\{x\}$ distribute with $f(x)$

Example: Exponential distribution

$$f(x) = Ae^{-x} \tag{2}$$

$$0 \leq x < 1$$

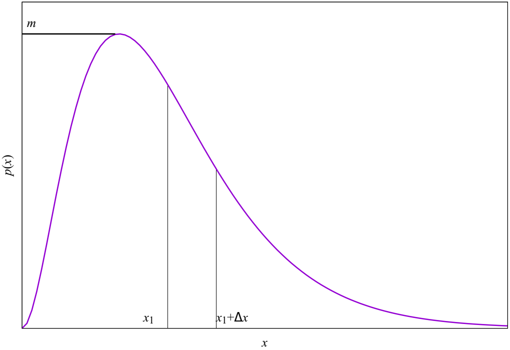
$$A = \frac{e}{e-1} \tag{3}$$

$$F(x) = \int_0^x f(z)\mathrm{d}z = A\left(1 - e^{-x}\right) \tag{4}$$

$$F^{-1}(r) = -\ln\left(1 - \frac{r}{A}\right) \tag{5}$$

Rejection Method

- Probability density $f(x)$ defined in $[a, b)$.
- Generate a random number pair $(x, y) \in [a, b) \times [0, \max f(x))$.
- Probability entering $[x_i, x_i + \Delta x]$ is proportional to $f(x)\Delta x$.

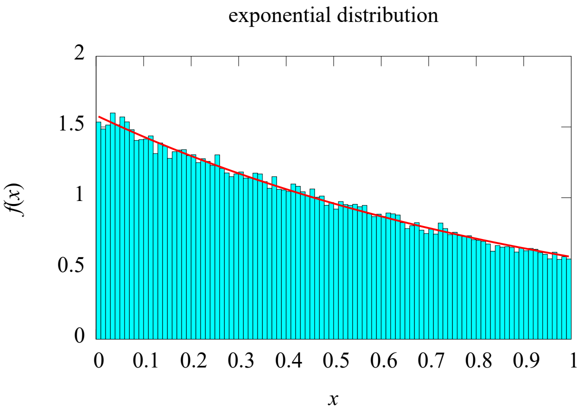


Classes for generating random numbers

- `randomNumbers` package
  - `AbstractRandom.java`
  - `Transform.java`
  - `Rejection.java`
- Using `java.util.Function.DoubleFunction`
  - Define the inverse of $F(x)$ for the Transform method.
  - Define $f(x)$ for the Rejection method.
  - Using lambda expressions

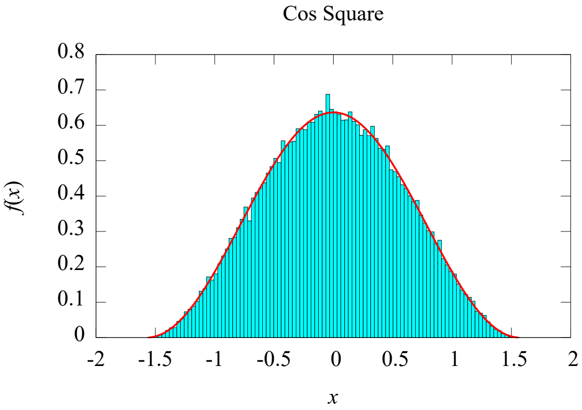
Example of Transform method: exponential distribution

```
指数分布に対応した分布関数の逆関数を定義
    //
    // A * exp (-x)
    double A = Math.E / (Math.E - 1);
    DoubleFunction<Double> invProDist = (x) -> {
        return -Math.log(1 - x / A);
    };変換法による乱数生成のインスタンス
    //
    AbstractRandom aRandom = new Transform(invProDist);
```


exponential distribution
2
1.5
1
0.5
0
f(x)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
x

Example of Rejection Method: Square of Cosine

$$f(x) = \frac{2}{\pi}\cos^2(x), \quad x \in [-\frac{\pi}{2}, \frac{\pi}{2}] \tag{6}$$



Probability and Law of Large Numbers (大数の法則)

- What does it mean that the probability of getting 1 on a dice is $1/6$?
- Consider the relative frequency of getting 1 on a dice.
    - It approaches $1/6$ with a large number of trials.
- Law of Large Numbers
- Let us take a closer look of this phenomenon.

Sample mean

- Consider a probabilistic variable X with the mean μ and deviation σ^2.
- Sample mean with size n.

$$\bar{X} = \frac{1}{n} \sum_{k=0}^{n-1} X_k \tag{7}$$

- Evaluate the population mean and deviation of $\bar{X}$.
  - Evaluate the mean and deviation of $\bar{X}$ with the probability of the population (**母集団**).
  - Equivalent to the mean of a large number of samples.

Population mean of sample means

The mean equals to the population mean μ.

$$
\begin{aligned}
E\left(\bar{X}\right) &= \frac{1}{n} E\left(\sum X_k\right) \\
&= \frac{1}{n} \sum E\left(X_k\right) = \frac{1}{n} n\mu \\
&= \mu
\end{aligned}
\tag{8}
$$

Population deviation of sample means

The deviation reduces with n^{-1}.

$$\begin{aligned}
V\left(\bar{X}\right) &= E\left(\left(\bar{X}-\mu\right)^2\right) = E\left(\frac{1}{n^2}\left(\sum_k \left(X_k-\mu\right)\right)^2\right) \\
&= \frac{1}{n^2}E\left(\sum_k \left(X_k-\mu\right)^2 + \sum_{i\neq j}\left(X_i-\mu\right)\left(X_j-\mu\right)\right) \\
&= \frac{1}{n^2}E\left(\sum_k \left(X_k-\mu\right)^2\right) + \frac{1}{n^2}E\left(\sum_{i\neq j}\left(X_i-\mu\right)\left(X_j-\mu\right)\right) \\
&= \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n} \qquad (9)
\end{aligned}$$

Confirm the law of large numbers by simulations

- Generate samples of size n.
- Instead evaluating the mean using the population distribution
  - Generate a large number m of samples with the same size.
  - Evaluate the mean and deviation for samples
- Changing n and observe n dependence.

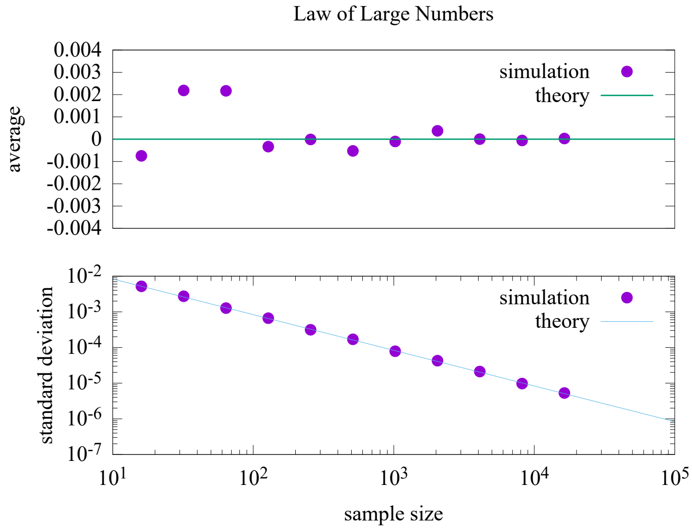
Example : uniform

$$f(x) = \begin{cases} 1 & -\frac{1}{2} \le x < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

$$\langle x \rangle = \int_{-1/2}^{1/2} x f(x) \mathrm{d}x = \int_{-1/2}^{1/2} x \mathrm{d}x = \left[\frac{1}{2}x^2\right]_{-1/2}^{1/2} = 0 \tag{11}$$

$$\langle x^2 \rangle = \int_{-1/2}^{1/2} x^2 f(x) \mathrm{d}x = \int_{-1/2}^{1/2} x^2 \mathrm{d}x = \left[\frac{1}{3}x^3\right]_{-1/2}^{1/2} = \frac{1}{12} \tag{12}$$

$$\sigma^2 = \langle x^2 \rangle - \langle x \rangle = \frac{1}{12} \tag{13}$$



$m = 1000$ for each sample size.

Pseudo Random Numbers (疑似乱数)

- Generate random numbers using a computer.
- Need some kind of algorithms.
    - Sequences are deterministic.
- In other word, you can generate an identical sequence as you need.

Linear Congruential Method (線形合同法)

- Recursion relation: $v_n = (av_{n-1} + c) \bmod m$
  - Property strongly depends on parameters a, c, m
  - Good parameter sets are known empirically.
- Unsigned integers are available in C/C++
  - No need to control overflow
- Languages, such as Java and Fortran, does not have unsigned integers.
  - Need some tips to suppress overflow

- Parameter example for 32 bit unsigned integers

$$m = 2416,\ c = 374441,\ m = 1771874$$

- Parameter example for 32 bit signed integers

$$m = 9301,\ c = 49297,\ m = 233280$$

Schrage's method

For 32 bit signed integers

- Set $m = 2^{31} - 1$

$$a = 17807,\ c = 0,\ m = 2^{31} - 1$$

$$\begin{aligned} q &= \lfloor m/a \rfloor \\ r &= m \bmod a \\ m &= aq + r \end{aligned}$$

- Need the condition $r < q$

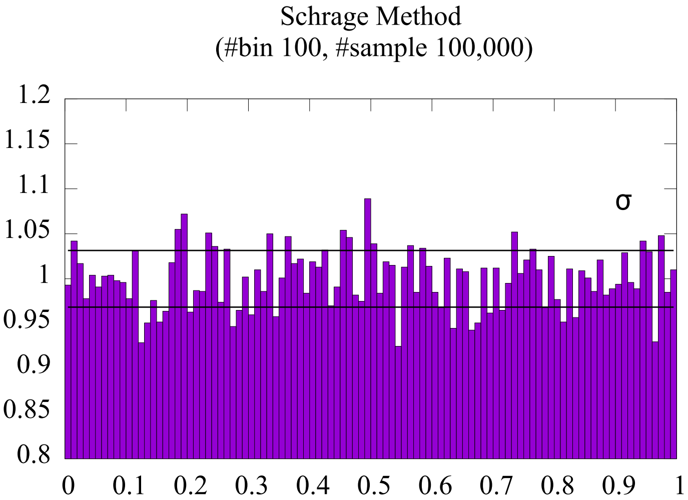
$$av_{n+1} \bmod m = \begin{cases} a\left(v_n \bmod q\right) - \lfloor v_n/q \rfloor r & \text{if not negative} \\ a\left(v_n \bmod q\right) - \lfloor v_n/q \rfloor r + m & \text{otherwise} \end{cases} \tag{14}$$

- Because, consider $v_n = xq + y$
  - LHS: $ay - xr$
  - And

$$\begin{aligned} av_{n+1} &= a\left(xq + y\right) = xaq + ay \\ &= x\left(m - r\right) + ay = xm + ay - xr \end{aligned} \tag{15}$$

$$av_{n+1} \bmod m = ay - xr \tag{16}$$

Example of Schrage's method



Difficulties in LCM

- The next value of a value a is determined.
- The period is limited by m.
  - Some kinds of simulation need a larger number of random numbers than m.
- Multidimensional sparse crystal (多次元粗結晶)
  - Consider consecutive n random numbers as a point in n dimensional space.
  - Those have crystal structure with some parameter set.