

# Uniform Distribution

Arabin Kumar Dey

Assistant Professor

Department of Mathematics  
Indian Institute of Technology Guwahati

14-th January, 2013

# Outline

Uniform  
Distribution

Arabin Kumar Dey

Alternative  
Algorithm for  
Uniform  
Distribution

Lattice Structure of  
Uniform  
Distribution

## 1 Alternative Algorithm for Uniform Distribution

## 2 Lattice Structure of Uniform Distribution

- Integer arithmetic is sometimes faster than floating point arithmetic
- If variable  $y$  and  $m$  are represented on integers in a computer, the integer operation  $y/m$  produces  $y/m$ .
- Therefore  $y \bmod m$  can be implemented as  $y - (y/m) \cdot m$ .  
Therefore  $y \bmod m$  can be implemented as  $y - (y/m) \cdot m$ .
- It can be shown that to avoid overflow a straight forward implementation of a linear congruential generator in integer variables must be restricted to small modulus.

- Let  $q = m/a$  and  $r = m \bmod a$ . Therefore  $m = aq + r$

- 

$$\begin{aligned} ax_i \bmod m &= ax_i - \left\lfloor \frac{ax_i}{m} \right\rfloor \cdot m \\ &= (ax_i - \left\lfloor \frac{x_i}{q} \right\rfloor m) + (\left\lfloor \frac{x_i}{q} \right\rfloor - \left\lfloor \frac{ax_i}{m} \right\rfloor)m \end{aligned}$$

- The term 1 of right hand side satisfy

$$\begin{aligned} (ax_i - \left\lfloor \frac{x_i}{q} \right\rfloor m) &= ax_i - \left\lfloor \frac{x_i}{q} \right\rfloor (aq + r) \\ &= a(x_i - \left\lfloor \frac{x_i}{q} \right\rfloor q) - \left\lfloor \frac{x_i}{q} \right\rfloor r \\ &= a(x_i \bmod q) - \left\lfloor \frac{x_i}{q} \right\rfloor r \end{aligned}$$

## ■ Substituting :

$$ax_i \bmod m = a(x_i \bmod q) - \lfloor \frac{x_i}{q} \rfloor r + (\lfloor \frac{x_i}{q} \rfloor - \lfloor \frac{ax_i}{m} \rfloor)m.$$

- To prevent overflow we need to avoid calculation of the potentially large term involving  $ax_i$  on the RHS. One can entirely avoid calculation of  $(\lfloor \frac{x_i}{q} \rfloor - \lfloor \frac{ax_i}{m} \rfloor)m$
- IF we can show that this expression takes only the values 0 and 1.
- In this case the last term  $(\lfloor \frac{x_i}{q} \rfloor - \lfloor \frac{ax_i}{m} \rfloor)m$  has value either 0 or m.
- But the final calculation must result in a value in  $\{0, 1, 2, 3, \dots, m-1\}$ .
- Therefore the last term is m precisely when  $a(x_i \bmod q) - \lfloor \frac{x_i}{q} \rfloor r < 0$
- One can check that  $(\lfloor \frac{x_i}{q} \rfloor - \lfloor \frac{ax_i}{m} \rfloor)$  takes only the values 0 and 1 only if  $a \leq \sqrt{m}$ .

# Algorithm

Uniform  
Distribution

Arabin Kumar Dey

Alternative  
Algorithm for  
Uniform  
Distribution

Lattice Structure of  
Uniform  
Distribution

- $m$  and  $a$  are integer constants.
- $q$  and  $r$  are pre-computed integer values with  $q = m/a$  and  $r = m \bmod a$ .
- $x$  is an integer variable holding the correct  $x_i$ .
- $k = \lfloor \frac{x}{q} \rfloor$
- $x \rightarrow a(x - k \cdot q) - k \cdot r$ .
- If  $x < 0$ , then  $x \rightarrow x + m$
- Finally convert uniform random number using  $u \rightarrow x \cdot h$ , where  $h = \frac{1}{m}$ .

# Outline

Uniform  
Distribution

Arabin Kumar Dey

Alternative  
Algorithm for  
Uniform  
Distribution

Lattice Structure of  
Uniform  
Distribution

## 1 Alternative Algorithm for Uniform Distribution

## 2 Lattice Structure of Uniform Distribution

- Random numbers  $N_i$  can be arranged in m-tuples  $(N_i, N_{i+1}, \dots, N_{i+m-1})$  for  $i \geq 1$ . Then the tuples or the corresponding points  $(u_i, u_{i+1}, \dots, u_{i+m-1}) \in [0, 1)^m$  are analyzed with respect to correlation and distribution. A sequence defined by  $N_0$  and for  $i = 1, 2, \dots$  and with generators  $N_i = (aN_{i-1} + b) \bmod M$ , lie on  $(m - 1)$  dimensional hyperplane.
- Analysis for case  $m = 2$  :

$$\begin{aligned} N_i &= (aN_{i-1} + b) \bmod M \\ &= (aN_{i-1} + b - kM), \quad kM \leq aN_{i-1} + b \leq (k+1)M \end{aligned}$$

where  $k$  is an integer.

- Let  $z_0$  and  $z_1$  be arbitrary. Then,

$$\begin{aligned} z_0 N_{i-1} + z_1 N_i &= z_0 N_{i-1} + z_1 (aN_{i-1} + b - kM) \\ &= M \left( N_{i-1} \frac{z_0 + az_1}{M} - z_1 k \right) + z_1 b \end{aligned}$$

The points calculated by the linear congruence generator line on these straight lines



- We divide by  $M$  and obtain the equation of a straight line in the  $(u_{i-1}, u_i)$  plane namely,

$$z_0 u_{i-1} + z_1 u_i = c + z_1 b M^{-1}$$

- The points calculated by the linear congruence generator line on these straight lines.
- If the tuple  $(z_0, z_1)$  such that only few of the straight lines cut the square  $[0, 1)^2$ , the wide areas of the square would be free of random points, which violates the requirements of uniform distribution.
- The minimum number of parallel straight lines cutting the square or equivalently the maximum distance between them serves as measure of equidistributiveness.

- We analyze for the worst case scenario. When we admit only integers  $(z_0, z_1)$  and require  $z_0 + az_1 \equiv 0 \pmod{M}$ . Now  $c = z_0 u_{i-1} + z_1 u_i - z_1 b M^{-1}$  and applying  $0 \leq u_i < 1$ , we obtain the maximal interval  $I_c$  such that for each integer  $c \in I_c$ , its straight line cuts or touches the square  $[0, 1)^2$ .
- $N_i = 2N_{i-1} \pmod{11}$
- Here  $a = 2, b = 0$  and  $M = 11$ . Need  $z_0 + az_1 = 0 \pmod{M} \rightarrow z_0 + 2z_1 = 0 \pmod{11}$ .  $z_0 = -2$  and  $z_1 = 1$  satisfies this relation.
- Now we investigate the family of straight lines  $-2u_{i-1} + u_i = c$ . For  $u_i \in [0, 1)$  we have  $-2 < c < 1$ .
- There are only two values which satisfy this namely,  $c = -1, 0$ . The two corresponding lines cut the interior of  $[0, 1)^2$ .

# Extended Fibonacci Generators

Uniform  
Distribution

Arabin Kumar Dey

Alternative  
Algorithm for  
Uniform  
Distribution

Lattice Structure of  
Uniform  
Distribution

- Fibonacci sequence:  $x_n = x_{n-1} + x_{n-2}$
- Fibonacci RNG:  $x_n = x_{n-1} + x_{n-2} \bmod m$
- Properties (a) not very good randomness (b) high correlation.
- Extended Fibonacci generator (Marsaglia 1983)

$$x_n = (x_{n-5} + x_{n-17}) \bmod 2^k$$

- Ring buffer with 17 values. We used to initialize the 17 values through LCG or usual Fibonacci RNG.
- Properties
  - (a) passes all statistical tests
  - (b) period =  $2^k(2^{17} - 1)$  [much longer than LCG]

# Seed Selection Guidelines

Uniform  
Distribution

Arabin Kumar Dey

Alternative  
Algorithm for  
Uniform  
Distribution

Lattice Structure of  
Uniform  
Distribution

- Don't use 0 [except few, not discussed here]
- Avoid even values
  - (a) seed should be odd for multiplicative LCG with  $m = 2^k$
  - (b) for full period generators, all non-zero values equally good
- Don't subdivide one stream - (a) don't use a single stream for all random variables
  - (b) might be a strong correlation between items in same stream
- Use non-overlapping streams - (a) each stream requires separate seed
  - (b) if seeds are bad, streams will overlap and not be independent.
  - (c)
    - example: need 3 streams of 20,000 numbers
    - pick  $u_0$  as seed for first stream
    - pick  $u_{20,000}$  as seed for second stream
    - pick  $u_{40,000}$  as seed for third stream

# Seed Selection Guidelines

Uniform  
Distribution

Arabin Kumar Dey

Alternative  
Algorithm for  
Uniform  
Distribution

Lattice Structure of  
Uniform  
Distribution

- Reuse seeds in successive replications – if simulation experiment is replicated several times. – can use seeds from end of previous replication in next one.
- Don't use random seeds
  - (a) simulation can't be reproduced
  - (b) impossible to guarantee multiple streams won't overlap.

# Two Candidate LCGs

Uniform  
Distribution

Arabin Kumar Dey

Alternative  
Algorithm for  
Uniform  
Distribution

Lattice Structure of  
Uniform  
Distribution

## ■ Which is better ?

$$x_n = ((2^{34} + 1)x_{n-1} + 1) \bmod 2^{35}$$

$$x_n = ((2^{18} + 1)x_{n-1} + 1) \bmod 2^{35}$$

## ■ Both must be full period generators

$m = 2k$ , for some integer  $k$

$a = 4c + 1$ , for some integer  $c$

$b$  is an odd integer

## ■ Check Autocorrelation !!

# Autocorrelation Function

Uniform  
Distribution

Arabin Kumar Dey

Alternative  
Algorithm for  
Uniform  
Distribution

Lattice Structure of  
Uniform  
Distribution

- The lag-1 sample autocorrelation function of  $r_t$  is defined as

$$\hat{\rho}_1 = \frac{\sum_{t=2}^T (r_t - \bar{r})(r_{t-1} - \bar{r})}{\sum_{t=1}^T (r_t - \bar{r})^2}$$

- The lag- $l$  sample autocorrelation function of  $r_t$  is defined as

$$\hat{\rho}_l = \frac{\sum_{t=l+1}^T (r_t - \bar{r})(r_{t-l} - \bar{r})}{\sum_{t=1}^T (r_t - \bar{r})^2}$$