Università
degli Studi di
Messina
DIPARTIMENTO DI INGEGNERIA

Dependable computing
modelling and symulation

Random number generation

Master degree in

Engineering in Computer Science

Simulation

- Usually events managed by a simulator are not deterministic
- Programming language should provide some tools for generating "random" number
  - Pseudo-random numbers are generated from uniform distribution over *[0, 1]*

Random numbers: characteristics

- Uniformity over *[0, 1]*
  - When an interval is spitted into N sub-intervals, *x/N* samples per sub-interval are expected
- Independence
  - A new sample does not depend on the previously generated one

Generation procedure

- Pseudo-randomness: properties of real sequences of random numbers are approximated
- Characteristics
  - Computational time (speed)
  - Long cycles
  - Reproducibility
  - Characteristics as close as possible to the ideal sequences

Linear congruential method

- A sequence of integer numbers X_1, X_2, … in the interval *[0, m-1]* is generated

$$X_{i+1} = (a\,X_i + c)\ mod\ m,\quad i \in \mathbb{N}$$

- X_0: seed
- *a* : multiplier
- *c* : increment (c=0, multiplicative congruential method)
- *m* : module

An example

$X_0=27,\ a=17,\ c=43,\ m=100$

$$X_0 = 27$$
$$X_1 = (17 * 27 + 43) \ mod \ 100 = 2$$
$$X_2 = (17 * 2 + 43) \ mod \ 100 = 77$$
$$X_3 = (17 * 77 + 43) \ mod \ 100 = 52$$
$$\vdots$$

Characteristics

- Generated numbers belongs to *[0, m-1]* interval
- Uniform distribution over *[0, m-1]*
- To generate numbers in [0, 1]:

$$R_i = \frac{X_i}{m}$$

- When *m* is very high the approximation is good
- Usually $m=2^{31}-1$ or $m=2^{48}$

Characteristics

- The period P depends on parameters
- $m=2^b$ and $c \neq 0$
  - $P=m$, if c is relatively prime to m and $a=1+4k$
- $m=2^b$ and $c=0$
  - $P=m/4$, if X_0 is odd and $a=3+8k$ or $a=5+8k$
- m is a prime number and $c=0$
  - $P=m-1$, if a is such that $k=m-1$ is the smallest integer number such that a^k-1 is a multiple of m

Linear combination

- Very long period (> 2×10^9)
- n generators with period equal to m_1-1, m_2-1, ..., m_n-1
- Sequences $X_{i,j}$, with $j=1, ..., n$

$$X_i = \left(\sum_{j=1}^{n} (-1)^{j-1} X_{i,j} \right) \bmod m_1 - 1$$

$$R_i = \begin{cases} \dfrac{X_i}{m_1} & X_i > 0 \\ \dfrac{m_1 - 1}{m_1} & X_i = 0 \end{cases}$$

$$P = \frac{(m_1-1)(m_2-1)\cdots(m_n-1)}{2^{n-1}}$$

Test

- χ^2
- Kolmogorov-Smirnov

Banks et al., 1996

Independent sequences

- Different generators
- **Different seeds**
- One sequence divided in sub-sequences

Cumulative random variable

Theorem

If X is a r.v. with continuous density function $f(x)$, the cumulative r.v. $C(x)$:

$$C(X) = \int_{-\infty}^{X} f(u)\,du$$

is uniformly distributed on [0, 1]

Application: continuous r.v.

- When the integral function is known the values of C(X) could be written as $c = F(x)$

- If $F(X)$ is invertible, $X = F^{-1}(C)$ has $f(c)$ as its pdf
- Samples generated by $F^{-1}(C)$, where C is a uniform r.v. over the support *[0, 1]* , have the same distribution of *X*
- Integral function is the CDF of X

Application: inverse transf.

- ξ r.v. with distribution $F_\xi(x)$
- $F_\xi(X)=C$
- ξ is a r.v. => C is a r.v.
- $X = F_\xi^{-1}(C)$, where C is a uniform r.v. on $[0,1]$

$$F_X(x) = P[X<=x] = P[F_\xi^{-1}(C)<=x] =$$

$$P[F_\xi(F_\xi^{-1}(C))<=F_\xi(x)] = P[C<=F_\xi(x)] = F_\xi(x)$$

Samples generation

- Computer systems are able to generate a sequence of uniformly distributed random number *R1, R2, …*

Algorithm: continuous r.v.

1. Given a cdf $F(X)$, set $F(X)=R$
2. Equation in 1. is solved assuming R as free variable, $X=F^{-1}(R)$
3. A sequence of uniformly distributed numbers on [0,1], R_1, R_2, R_3, ..., is generated
4. $X_i=F^{-1}(R_i)$ are computed using eq. in 3.

Application: discrete r.v.

- Let us assume C is a uniform r.v. over [0, 1]:

$$P\{c_1 \le C \le c_2\} = \frac{c_2 - c_1}{1 - 0}$$

- Then

$$P\{x_k < X \le x_{k-1}\} = F(x_k) - F(x_{k-1}) =$$
$$P\{F(x_{k-1}) < C \le F(x_k)\}$$

- Values of X can be generated according the following rule

$$X = x_k \Leftrightarrow F(X_{k-1}) < c \le F(x_k)$$

$$k = 1, F(x_0) = 0$$

Example

- $F(x)=1-e^{-\lambda x}$
- $R=1-e^{-\lambda x}$
- $x=-1/\lambda \ ln(1-R)$
- The $x_i=-1/\lambda \ ln(1-R_i)$, where R_i are the elements of the series given by a pseudo-random generator are samples of an exponential r.v.
- usually $x=-1/\lambda \ ln(R)$ is used

Application

- It is used each time it is possible to compute the inverse of F(x)
  - Exponential distribution
  - Uniform distribution
  - Weibull distribution
  - Triangular distribution
  - Empiric distribution

Gaussian distribution

- It is not invertible
- Sum of an infinite number of uniformly distributed r.v. is a Gaussian distribution
- Sum of n (with n>=12) uniformly distributed r.v. is a good approximation of a Gaussian distribution