CPSC 531: System Modeling and Simulation

UNIVERSITY OF CALGARY

Carey Williamson
Department of Computer Science
University of Calgary
Fall 2017

Recap and Terminology

- (Pseudo-) Random Number Generation (RNG)

- A fundamental primitive required for simulations
- Goal: Uniform(0,1)
- Uniformity
- Independence
- Computational efficiency
- Long period
- Multiple streams
- Common approach: LCG
- Careful design and seeding
- Never generates 0.0 or 1.0
- Covered in guest lecture (JH)
- Readings: 2.1, 2.2

- Random Variate Generation (RVG)

- Builds upon Uniform(0,1)
- Goal: **any** distribution
- Discrete distributions
- Continuous distributions
- Independence (usually)
- Correlation (if desired)
- Computational efficiency
- Common approach: the inverse transform method
- Straightforward math (usually)
- Might generate 0.0 or 1.0
- Covered in today's lecture
- Readings: 6.1, 6.2

UNIVERSITY OF CALGARY

Outline

- Random variate generation
  - Inverse transform method
  - Convolution method
  - Empirical distribution
  - Other techniques

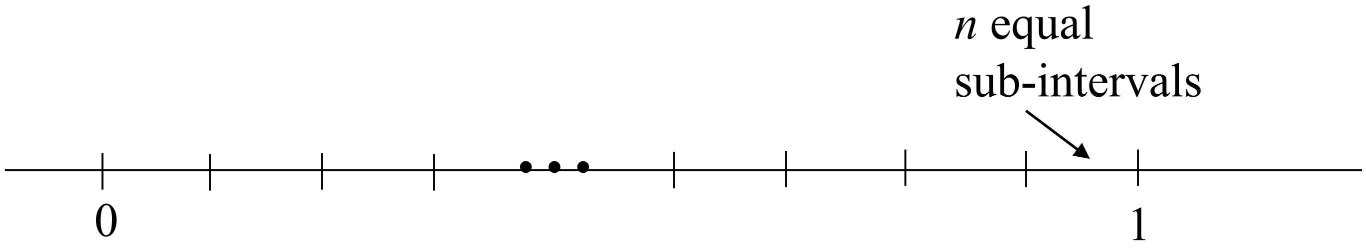
- Input parameters such as inter-arrival times and service times are often modeled by random variables with some given distributions

- A mechanism is needed to generate variates for a wide class of distributions

This can be done using a sequence of random numbers that are independent of each other and are uniformly distributed between 0 and 1

Uniform Random Numbers

- Uniformly distributed between 0 and 1
  - Consider a sequence of random numbers $u_1, u_2, \ldots, u_N$



  - Uniformity: expected number of random numbers in each sub-interval is N/n
  - Independence: value of each random number is not affected by any other numbers

Bernoulli Variate

A Bernoulli variate is useful for generating a binary outcome (0 or 1) to represent "success" (1) or "failure" (0)

Example: wireless network packet transmission

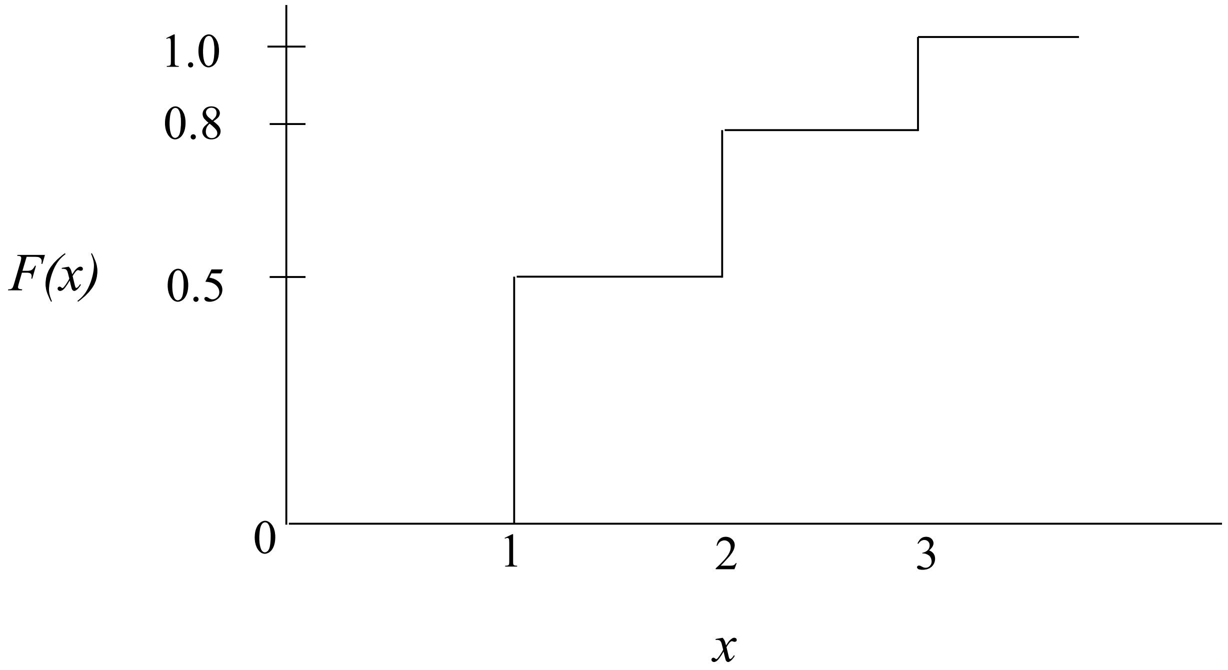
Example: coin flipping to produce "heads" or "tails"

Bernoulli trial (with parameter p)

$$p(1) = p, \qquad p(0) = 1 - p$$

- Random variate generation
  - Generate u
  - If $0 < u \leq p$, $x = 1$;
  - Otherwise $x = 0$

- Consider a tri-modal discrete distribution
  - Example: size of an email message (in paragraphs, or KB)
  - Example: $p(1) = 0.5$, $p(2) = 0.3$, $p(3) = 0.2$
- Cumulative distribution function, $F(x)$



- Algorithm
  - Generate random number u
  - Random variate $x = i$ if $F(i-1) < u \leq F(i)$

- Example: $F(0) = 0, F(1) = 0.5, F(2) = 0.8, F(3) = 1.0$
  - $0 < u \leq 0.5$ variate $x = 1$
  - $0.5 < u \leq 0.8$ variate $x = 2$
  - $0.8 < u \leq 1.0$ variate $x = 3$

Discrete Uniform Variate

- Discrete uniform (with parameters *a* and *b*)

$$p(n) = 1/(b - a + 1) \quad for\ n = a, a + 1, \ldots, b$$

$$F(n) = (n - a + 1)/(b - a + 1)$$

- Random variate generation
  - Generate u
  - $x = a + floor(u * (b - a + 1))$ OR
  - $x = (a - 1) + ceiling(u * (b - a + 1))$

Geometric Variate

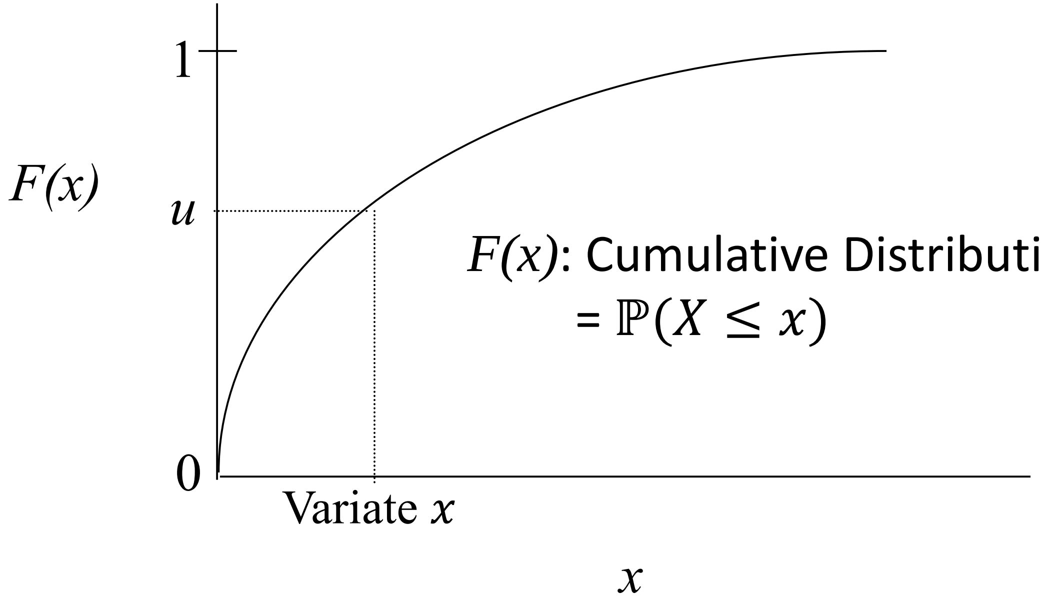
- Geometric (with parameter p)

$$p(n) = p(1-p)^{n-1}, \mathrm{n} = 1,2,3, \ldots$$

- Gives the number of Bernoulli trials until achieving the first success

- Random variate generation
  - Generate u
  - Geometric variate $x = \left\lceil \frac{\ln(u)}{\ln(1-p)} \right\rceil$

- Algorithm
  - Generate uniform random number u
  - Solve $F(x) = u$ for random variate x

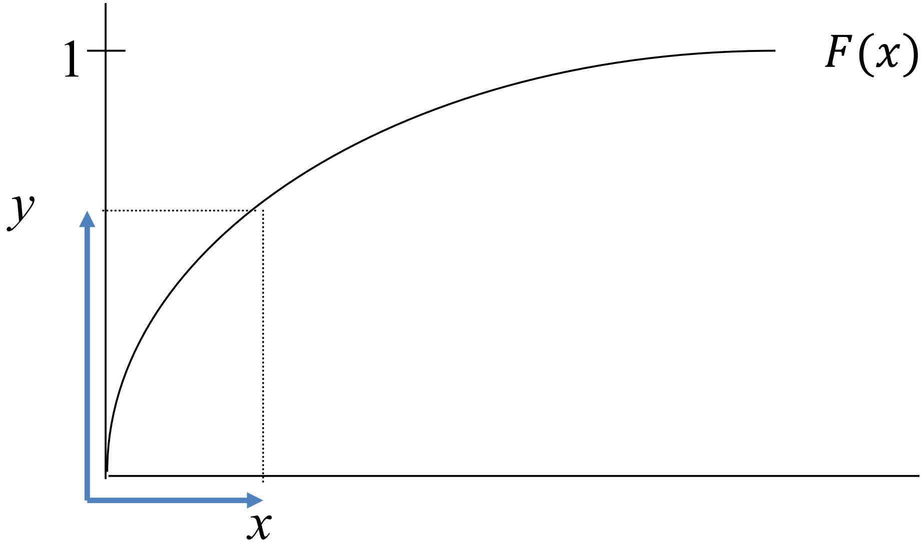


$F(x)$: Cumulative Distribution Function of X
$= \mathbb{P}(X \leq x)$

Proof

- Define the random variable Y as:

$$Y = F(X)$$



$$\mathbb{P}(Y \leq y) = \mathbb{P}(X \leq x) = y$$

Therefore,

$$Y \sim U(0,1)$$

Continuous Uniform Variate

- Uniform (with parameters a and b)

$$f(x) = \begin{cases} 1/(b-a) & a \le x \le b, \\ 0 & \text{otherwise.} \end{cases}$$

$$F(x) = (x - a)/(b - a), a \le x \le b$$

- Random variate generation
  - Generate u
  - $x = a + (b - a)u$

Exponential Variate

- Exponential (with parameter λ)

  $f(x) = \lambda e^{-\lambda x}$

  $F(x) = 1 - e^{-\lambda x}$

- Random variate generation
  - Generate u
  - $x = -\left(\frac{1}{\lambda}\right) \cdot \ln(u)$
- Can also use $x = -\left(\frac{1}{\lambda}\right) \cdot \ln(1 - u)$

Note: If u is Uniform(0,1), then $1 - u$ is Uniform(0,1) too!

Convolution Method

- Sum of n variables: $x = y_1 + y_2 + \cdots + y_n$

1. Generate n random variate y_i's
2. The random variate x is given by the sum of y_i's

Example: the sum of two fair dice that are rolled

P(x=2) = 1/36; P(x=3) = 2/36; P(x=4) = 3/36;

P(x=5) = 4/36; P(x=6) = 5/36; P(x=7) = 6/36;

P(x=8) = 5/36; P(x=9) = 4/36; P(x=10) = 3/36;

P(x=11) = 2/36; P(x=12) = 1/36

Geometric Variate

- Geometric (with parameter p)

$$p(n) = p(1-p)^{n-1}, \mathrm{n} = 1,2,3, \dots$$

- Gives the number of Bernoulli trials until achieving the first success
  - let $b = 0$, $n = 0$
  - while ($b == 0$)
    - Generate Bernoulli variate b with parameter p
    - Geometric variate $n = n + 1$

Inefficient!!

Binomial Variate

- Binomial (with parameters p and n)

$$p(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \; k = 0,1,\ldots,n$$

Random variate generation

- Generate n Bernoulli variates, $y_1, y_2, \ldots, y_n$
- Binomial variate $x = y_1 + y_2 + \cdots + y_n$

Poisson Variate

- Poisson (with parameter λ)

$$p(k) = \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda},\ k = 0,1,2,\ldots$$

- Random variate generation (based on the relationship with exponential distribution)
  - let $s = 0$, $n = 0$
  - while ($s \leq 1$)
    - Generate exponential variate **y** with parameter λ
    - $s = s + y$
    - $n = n + 1$
  - Poisson variate $x = n - 1$

Other Techniques: Normal Variate

- Normal (with parameters μ and σ^2)

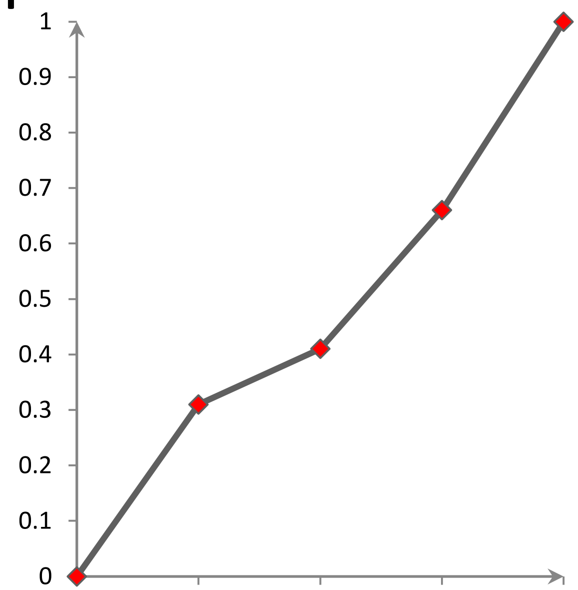
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad \text{for } -\infty \leq x \leq +\infty$$

- Random variate generation using approximation method
  - Generate two random numbers u_1 and u_2
  - Random variates x_1 and x_2 are given by:

$$x_1 = \mu + \sigma\sqrt{-2\ln(u_1)} \cdot \cos(2\pi u_2)$$
$$x_2 = \mu + \sigma\sqrt{-2\ln(u_1)} \cdot \sin(2\pi u_2)$$

Could be used if no theoretical distributions fit the data adequately

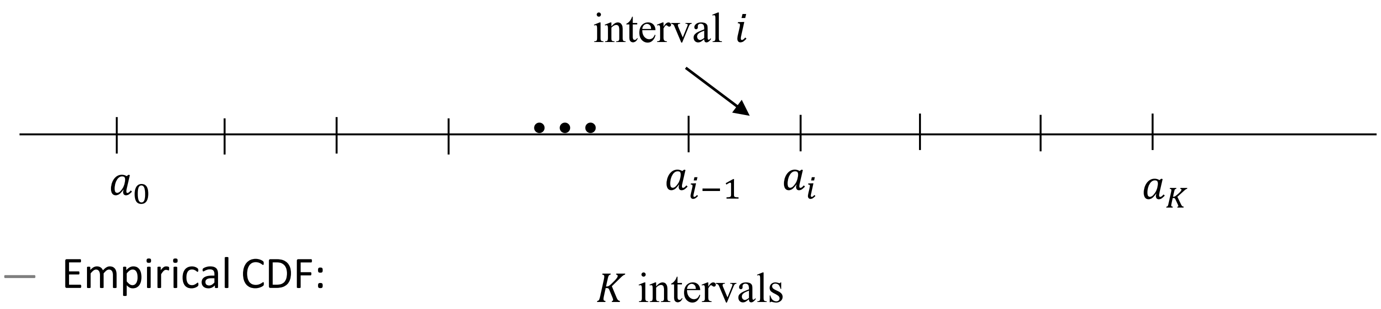
- Example: Piecewise Linear empirical distribution
  - Used for continuous data
  - Appropriate when a large sample data is available
  - Empirical CDF is approximated by a piecewise linear function:
    - the 'jump points' connected by linear functions



Piecewise Linear Empirical CDF

Empirical Distribution

- Piecewise Linear empirical distribution
  - Organize X-axis into K intervals
  - Interval i is from a_{i-1} to a_i for $i = 1,2, \ldots, K$
  - p_i: relative frequency of interval i
  - c_i: relative cumulative frequency of interval i, i.e., $c_i = p_1 + \cdots + p_i$



  - Empirical CDF:
    - If x is in interval i, i.e., $a_{i-1} < x \leq a_i$, then:

$$F(x) = c_{i-1} + \alpha_i (x - a_{i-1})$$

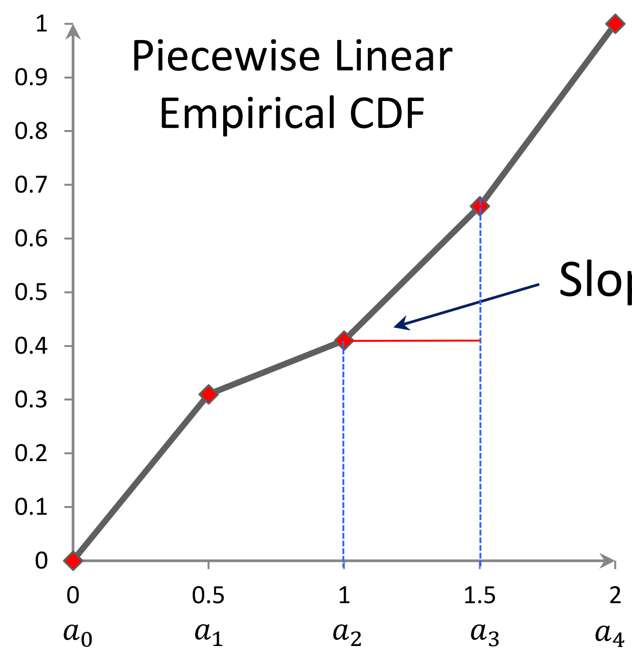
where, slope α_i is given by

$$\alpha_i = \frac{c_i - c_{i-1}}{a_i - a_{i-1}}$$

Example Empirical Distribution

- Suppose the data collected for 100 broken machine repair times are:

i	*Interval (Hours)*	*Frequency*	*Relative Frequency*	*Cumulative Frequency*	*Slope*
1	0.0 < x ≤ 0.5	31	0.31	0.31	0.62
2	0.5 < x ≤ 1.0	10	0.10	0.41	0.2
3	1.0 < x ≤ 1.5	25	0.25	0.66	0.5
4	1.5 < x ≤ 2.0	34	0.34	1.00	0.68



Slope $\alpha_3 = \frac{c_3 - c_2}{a_3 - a_2} = 0.5$

Empirical Distribution

- Random variate generation:
  - Generate random number u
  - Select the appropriate interval i such that

$$c_{i-1} < u \le c_i$$

  - Use the inverse transformation method to compute the random variate x as follows

$$x = a_{i-1} + \frac{1}{\alpha_i}(u - c_{i-1})$$

Example Empirical Distribution

- Suppose the data collected for 100 broken machine repair times are:

i	Interval (Hours)	Frequency	Relative Frequency	Cumulative Frequency	Slope
1	0.25 < x ≤ 0.5	31	0.31	0.31	1.24
2	0.5 < x ≤ 1.0	10	0.10	0.41	0.2
3	1.0 < x ≤ 1.5	25	0.25	0.66	0.5
4	1.5 < x ≤ 2.0	34	0.34	1.00	0.68

- Suppose: $u = 0.83$

$$c_3 = 0.66 < u \leq c_4 = 1.00 \Rightarrow i = 4$$

$$x = a_3 + \frac{1}{\alpha_4}(u - c_3)$$

$$= 1.5 + \frac{1}{0.68}(0.83 - 0.66)$$

$$= 1.75$$