Computer Simulation

Dr. Bardia Safaei

Chapter Seven: Random-Variate Generation

Purpose & Overview

SHARIF UNIVERSITY OF TECHNOLOGY

- Understanding the generating process of samples from a **specified distribution** as input to our simulation model
  - Unlike the previous lecture, where the goal was to generate random numbers who follow the uniform distribution
- Explaining some widely-used techniques for generating random variates
  - Inverse-transform technique
  - Acceptance-rejection technique
- Many well-known simulators provide random-variate generators, but some may not
  - In this lecture you learn the essentials for developing routines for generating random variates if the simulator does not support it

Assumptions

SHARIF UNIVERSITY OF TECHNOLOGY

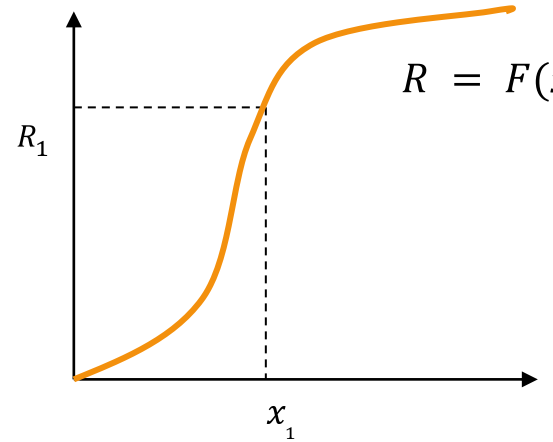
- All of the techniques in this lecture consider the two following assumptions:
  - R_i is a previously generated random number readily available
  - R_i is uniformly distributed on [0,1]
    - Therefore, the PDF, and CDF of these numbers are defined as follows:
- With these R_is in one hand, and a specific statistical distribution in another, we generate variates, which support that intended distribution

$$f_R(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

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

Inverse-Transform Technique

- This technique is generally applicable to distributions that their CDF is **reversable**
  - You can simply obtain F^{-1} from $\mathrm{F(x)}$
- Steps to take:
  - Produce $\mathrm{R_i} \sim \mathrm{U}[0,1]$
  - Use:
  - Find $x\mathbf{R_i = F(x)} \Rightarrow \mathrm{x} = \mathrm{F}^{-1}(\mathrm{R_i})$
    - Is it correct to put R, and F(x) equal?



- This technique enables us to sample data with various distributions, and generate similar numbers $(\mathrm{x_i})$ with the same distribution
  - Exponential, Weibull, Uniform, triangular, and even empirical

Exponential Distribution (1)

- Recall: the PDF, and CDF for the exponential distribution are defined as follows:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, x \geq 0 \\ 0, \qquad x < 0 \end{cases}$$

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 1 - e^{-\lambda x}, x \geq 0 \\ 0, \qquad x < 0 \end{cases}$$

  - Where λ indicates the rate of events occurred in a time unit, and the average time between two consequent events is λ^{-1}
- Our goal here is to generate a set of random variates $\{x_1, x_2, x_3, \ldots\}$, who follow the exponential distribution

Exponential Distribution (2)

SHARIF UNIVERSITY OF TECHNOLOGY

- The step by step guide for exponential variate generation via inverse-transform:
  - **1st step:** Specify the CDF of the ultimate variates (x_i), which you want to generate: $F(x) = 1 - e^{-\lambda x}, x \geq 0$
  - **2nd step:** Place R equal to $F(x)$: $F(x) = 1 - e^{-\lambda x} = R$
    - Since x is a random variable, $1 - e^{-\lambda x}$ is also a random variable, which we have named it as R
      - As we illustrate later, R follows the uniform distribution over [0,1] (as expected)
  - **3rd step:** Solving the equation to obtain x based on R

$$1 - e^{-\lambda x} = R \rightarrow e^{-\lambda x} = 1 - R \rightarrow -\lambda x = Ln(1 - R) \rightarrow \boldsymbol{x = -\frac{1}{\lambda} Ln(1 - R)}$$

- This is called **exponential variate generator function**
- Independent from the distribution, the **variate generator function** is denoted with $x = F^{-1}(R)$

Exponential Distribution (3)

- **4th step:** Based on how many variates you want to generate, you must produce R_is first
  - You can use the techniques you studied in lecture 6
  - Then apply them to the variate generator function to obtain x_is

$$x_i = F^{-1}(R_i) = -\frac{1}{\lambda} Ln(1 - R_i)$$

- To simplify the exponential variate generation, we can replace $1 - R_i$ with R_i
  - Why?

$$x_i = -\frac{1}{\lambda} Ln(R_i)$$

Exponential Distribution (4)

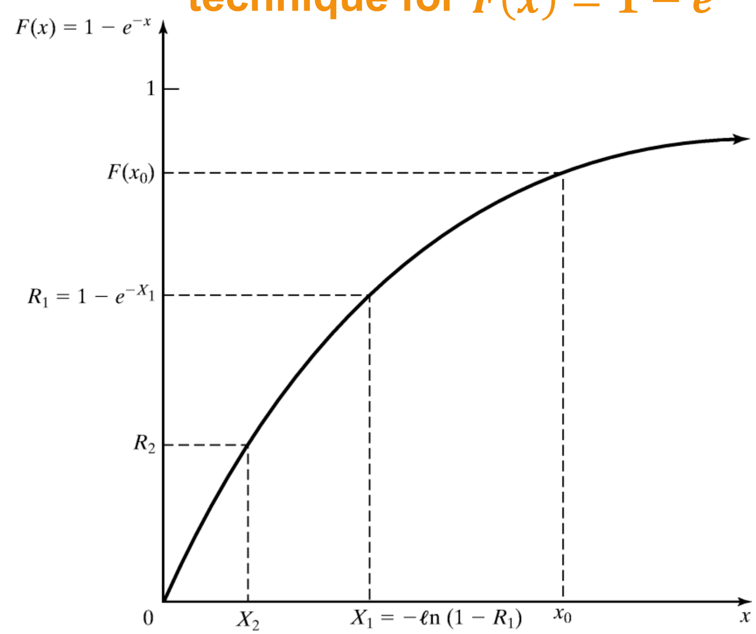
SHARIF UNIVERSITY OF TECHNOLOGY

- Does the generated $\mathrm{x_i}$ really follows the exponential distribution with using this equation?
  - Consider the CDF of an exponential distribution with $\lambda = 1$
  - Assume x_0 as a prespecified number
- Let's obtain the CDF for $\mathrm{x_i}$ $(\mathrm{i} = 1)$

$$\overbrace{P(x_1 \le x_0) = P\big(R_1 \le F(x_0)}^{A}\underbrace{\big) = F(x_0)}_{B}$$

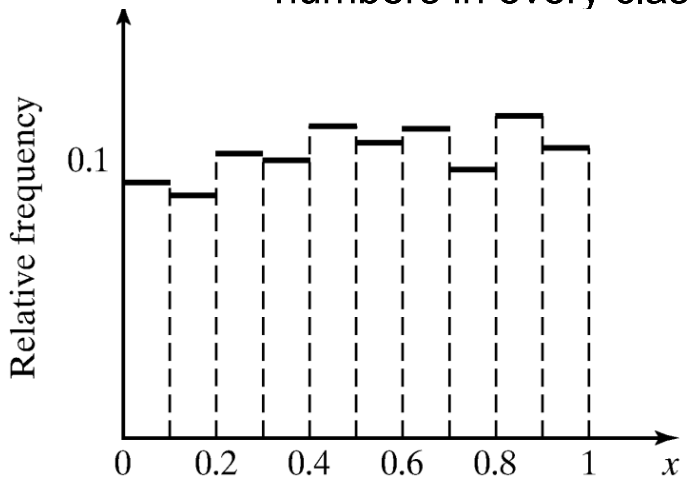
- How A is applicable?
  - $x_1 \le x_0$ if and only if $R_1 \le F(x_0)$
- How B is applicable?
  - Because $0 \le F(x_0) \le 1$, and R_1 is uniformly distributed on [0,1]

Figure: Inverse-transform technique for $F(x) = 1 - e^{-x}$

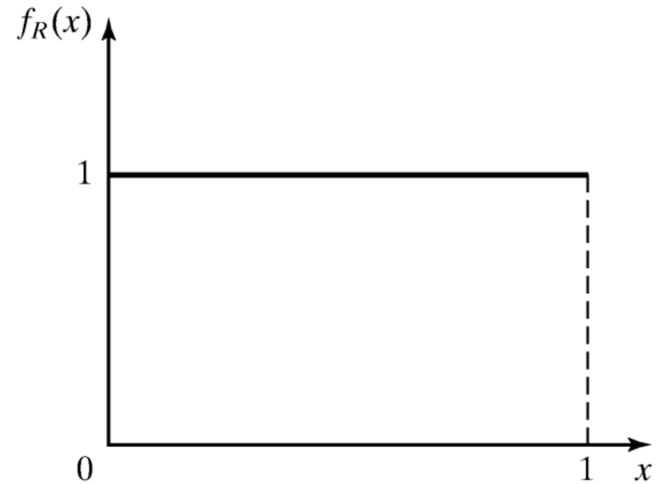


Example (1)

- Generate 200 variates x_i with distribution $\exp(\lambda = 1)$
  - First, we produce 200 $R_i \sim U(0,1)$
    - These random numbers are shown in the following histogram
      - 10 classes have been used for representing the frequency of numbers in every class



Histogram for 200 empirically generated R_i Random numbers $\sim U(0,1)$



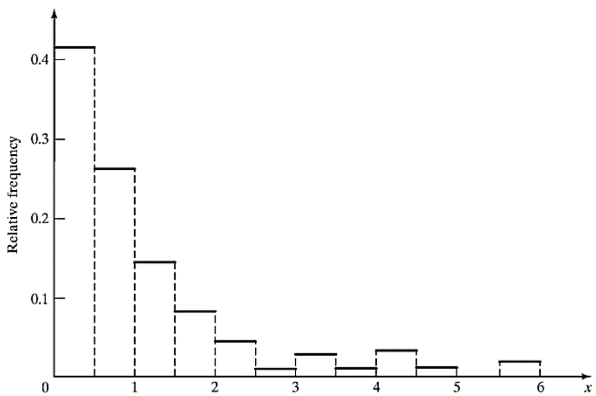
Theoritical representation of R_i Random numbers $\sim U(0,1)$

Example (2)

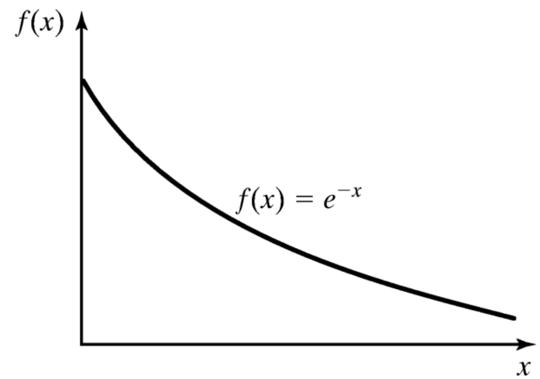
- Then we utilize the following equation:

$$x_i = -\frac{1}{\lambda} Ln(R_i) = -Ln(R_i)$$

- The generated variates are illustrated in the following histogram



Histogram for 200 empirically produced $\mathbf{x_i}$ exponential variates $\sim\mathbf{exp}(\lambda = 1)$



Theoritical representation of $\mathbf{x_i}$ variates$\sim\mathbf{exp}(\lambda = 1)$

Other Distributions (1)

- Now let's talk about producing variates for other distributions with employing the **inverse-transform**
  - Recall: The CDF of the targeted distribution must be **reversible**
- Uniform distribution:
  - The random variable X is uniformly distributed on [a,b] if:

$$x_i \sim U(a,b) \Rightarrow x_i = a + (b-a)R_i$$

    - Where a, and b can get any values $(a < b)$
    - Based on this equation, if R_i is selected from [0,1], x_i resides in [a,b]
- Weibull distribution $(v = 0)$:

$$f(x) = \begin{cases} \dfrac{\beta}{\alpha^\beta} x^{\beta-1} e^{-(x/a)^\beta}, & x \geq 0 \\ 0, & otherwise \end{cases} \longrightarrow F(x) = 1 - e^{-\left(\frac{x}{\alpha}\right)^\beta}, x \geq 0$$

$$\downarrow$$

$$x_i = \alpha[-Ln(1 - R_i)]^{\frac{1}{\beta}}$$

Other Distributions (2)

SHARIF UNIVERSITY OF TECHNOLOGY

- Triangular distribution:
  - The same process must be applied to its PDF:

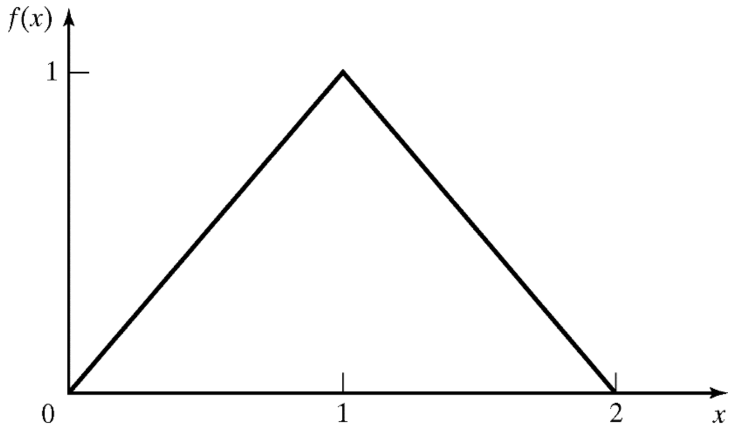
$$f(x) = \begin{cases} x, & 0 \le x \le 1 \\ 2 - x, & 1 < x \le 2 \\ 0, & \text{otherwise} \end{cases}$$

$$\downarrow$$

$$F(x) = \begin{cases} 0, & x \le 0 \\ \dfrac{x^2}{2}, & 0 < x \le 1 \\ 1 - \dfrac{(2-x)^2}{2}, & 1 < x \le 2 \\ 1, & x > 2 \end{cases}$$

$$\downarrow$$

$$x_i = \begin{cases} \sqrt{2R_i}, & 0 \le R_i \le 1/2 \\ 2 - \sqrt{2(1-R)}, & 1/2 < R_i \le 1 \end{cases}$$



Obtain the x_i equation in triangular distribution with inverse-transform technique

Empirical Continuous Distributions (1)

- When there isn't sufficient data, known theoretical distributions cannot be used for **modeling** and also **generating random variates**
- What approaches could be used here?
  - Resampling more data (if possible), and reaching a better understanding from the collected data
    - Then, try to suggest values for those areas with no data available based on the graphical presentation of the collected data
    - Also known as **discrete** approach
  - Interpolate between observed data points to fill in the gaps in the empirical CDF
    - This is where we want to focus in this section and produce random variates based on empirical **continuous** distribution

Empirical Continuous Distributions (2)

- Problem statement:
  - There is a **small sample set (size n)** composed of values you have observed
  - We want to know, **if** supposed to be other observed values in this set, what would they like be?
- The process:
  - Arrange your limited collected data from smallest to largest

$$y_1 \leq y_2 \leq y_3 \leq \cdots \leq y_n$$

  - Consider the gap between two numbers as an interval, and assign the probability 1/n to each interval
  - Plot these values and draw a line between every two consequent points
    - This is your empirical CDF composed of segmented lines

Empirical Continuous Distributions (3)

SHARIF UNIVERSITY OF TECHNOLOGY

- The process (Cont.):
  - Produce R_i, which is uniformly distributed on [0,1]
  - Determine the value of j according to $\frac{j-1}{n} < R_i \leq \frac{j}{n}$
    - You can do this with either the CDF or a table composed of your gathered information
  - Calculate the value of x_i according to the following equation:

$$x_i = F^{-1}(R_i) = y_{j-1} + a_j \left[R_i - \frac{j-1}{n} \right]$$

    - x_i is the value you want to produce, and y_j is the value you have already collected from the system
    - a_j is indicating the slope of the jth segmented line in the CDF

$$a_j = \frac{y_j - y_{j-1}}{\frac{j}{n} - \frac{j-1}{n}} = \frac{y_j - y_{j-1}}{1/n}$$

Example (1)

- The speed of preparation for a group of firefighters have been measured and reported in minutes
  - 2.76 1.83 0.80 1.45 1.24
- We sort these data and insert them in a table as follows:

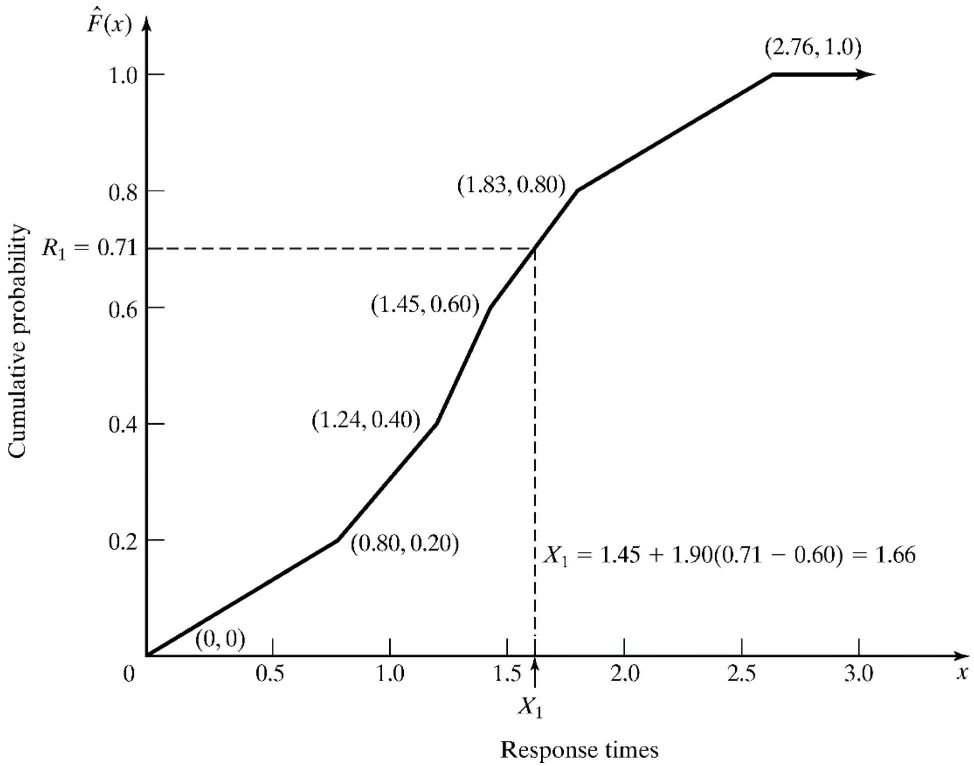
j	Interval (Hours)	Probability 1/n	Cumalative Probability, j/n	Slope, aj
1	$0.0 \le x \le 0.80$	0.2	0.20	4.00
2	$0.8 < x \le 1.24$	0.2	0.40	2.20
3	$1.24 < x \le 1.45$	0.2	0.60	1.05
4	$1.45 < x \le 1.83$	0.2	0.80	1.90
5	$1.83 < x \le 2.76$	0.2	1.00	4.65

  - Since there are only 5 values, the probability of every interval is 1/5
  - The slope of every segmented line is also calculated
    - These values are then used to plot the empirical CDF

Example (2)

- The CDF is now obtained
- How to generate a random variate based on this empirical CDF?
  - Assume $R_1 = 0.71$
  - R_1 lies between 0.6, and 0.8

$$0.6 = \frac{j-1}{n} < R_1 \leq 0.8 = \frac{j}{n}$$



- Therefore, $j = 4$ and $a_4 = 1.90$

$$x_1 = y_{j-1} + a_j\left[R_i - \frac{j-1}{n}\right] = 1.45 + a_4\left[0.71 - \frac{3}{5}\right] = 1.45 + 1.90 \times (0.71 - 0.6) = 1.66$$

Empirical Continuous Distribution with Large Sample Set (1)

SHARIF UNIVERSITY OF TECHNOLOGY

- What if the number of samples is high?
  - We need a classification approach
    - Consider a number of intervals with identical length
    - Every interval embraces a number of samples
      - The **cumulative frequency** of every interval is indicated with c_j
- So, the only difference with the previous approach is in the frequency of observations in every interval
  - Previously, every interval was contained of only one sample (check out the table in slide 16)
    - Therefore, the probability of every interval was identical to others (1/n), and the cumulative frequency of the jth interval was j/n
  - In the new approach, intervals contain different number of samples
    - So, we use the actual cumulative frequency of every interval, which may not be equal to other intervals

Empirical Continuous Distribution with Large Sample Set (2)

SHARIF UNIVERSITY OF TECHNOLOGY

- The process:
  - Produce R_i, which is uniformly distributed on [0,1]
  - Determine the value of j according to $c_{j-1} < R_i \leq c_j$
  - Calculate the value of x_i according to the following equation:

$$x_i = F^{-1}(R_i) = y_{j-1} + a_j\left[R_i - c_{j-1}\right]$$

    - x_i is the value you want to produce, and y_j is the value you have already collected from the system
    - a_j is indicating the slope of the jth segmented line in the CDF

$$a_j = \frac{y_j - y_{j-1}}{c_j - c_{j-1}}$$

Example (1)

- Assume that the repairing time of a device has been measured in n=100 consecutive observations
  - These values are classified into 4 intervals and added in the following table

j	Interval (Hours)	Frequency	Relative Frequency	Cumulative Frequency, cj	Slope, aj
1	$0.25 \le x \le 0.5$	31	0.31	0.31	0.81
2	$0.5 < x \le 1.0$	10	0.10	0.41	5.0
3	$1.0 < x \le 1.5$	25	0.25	0.66	2.0
4	$1.5 < x \le 2.0$	34	0.34	1.00	1.47

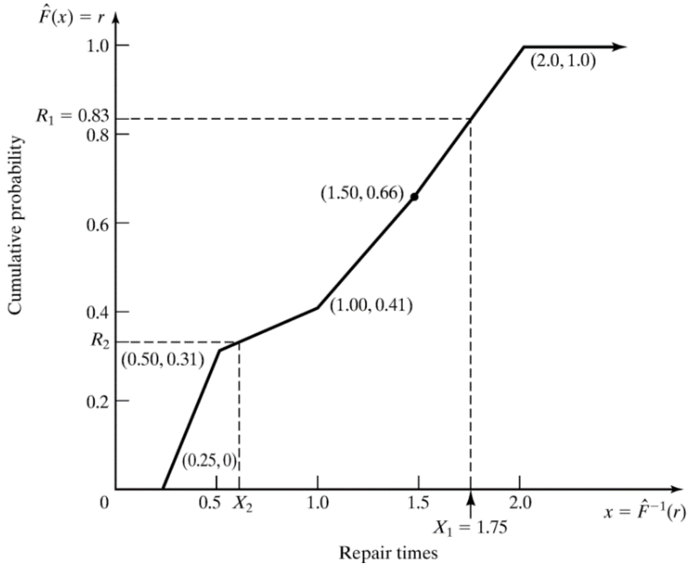
- One of the differences between this table and the previous version is the **lower bound** of the 1^{st} interval
  - Previously, since we built the intervals from the samples themselves, we put 0, but here we use the sample with minimum value

Example (2)

- Now we could depict the CDF from the table with drawing the segmented lines
- How to generate a random variate based on this empirical CDF?
  - Assume $R_1 = 0.83$
  - R_1 lies between 0.66, and 1.0

$$0.66 = c_3 < R_1 \leq 0.8 = c_4$$

  - Therefore, $j = 4$, and $a_4 = 1.47$



$$x_1 = y_{j-1} + a_j\left[R_i - c_{j-1}\right] = 1.5 + 1.47[0.83 - 0.66] = 1.75$$

- Calculate x_2, if $R_2 = 0.33$

Continuous Distributions without a Closed-Form Inverse

- A number of useful continuous distributions do not have a closed-form expression for their CDF or its inverse
  - Example: Normal, Gamma, and Beta distribution
- We are willing to **approximate** the inverse of CDF
  - Example: Simple approximation for the inverse of CDF in the standard normal distribution:

$$x_i = F^{-1}(R_i) \approx \frac{R_i^{0.135} - (1 - R_i)^{0.135}}{0.1975}$$

  - In the following table, the approximated inverse has been compared with the exact inverse based on their corresponding R_i values

Only 0.5% error

R	*Approximate Inverse*	*Exact Inverse*
0.01	−2.3263	−2.3373
0.10	−1.2816	−1.2813
0.25	−0.6745	−0.6713
0.50	0.0000	0.0000
0.75	0.6745	0.6713
0.90	1.2816	1.2813
0.99	2.3263	2.3373

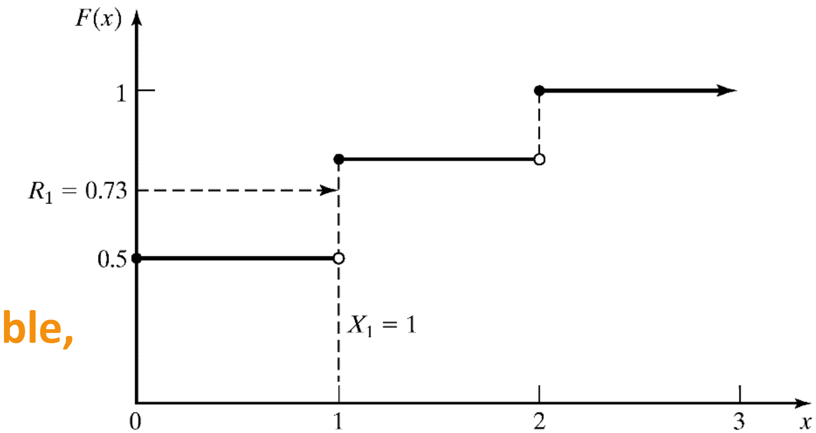
Discrete Distribution (1)

- Now let's concentrate on **discrete distributions**
- Random variates supporting all kinds of discrete distributions can be generated via inverse-transform technique
  - This could be done either numerically with table-lookup procedures, algebraically, or a formula
- Examples for distributions:
  - Empirical
  - Discrete uniform
  - Gamma

Discrete Distribution (2)

- Example:
  - Suppose that at the end of a working day, the number of shipments of a sailing company (X), on the loading dock of a port is either 0, 1, or 2
  - The probability distribution of the gathered data is as follows
    - Accordingly, the cumulative distribution is also calculated, and indicated in the 3rd column

x	p(x)	F(x)
0	0.50	0.50
1	0.30	0.80
2	0.20	1.00



The CDF could be plotted based on this table, and according to $F(X) = P(X \leq x)$

  - In order to take financial decisions, we have been asked to model the number of shipments (X) based on what we have observed

Discrete Distribution (3)

SHARIF UNIVERSITY OF TECHNOLOGY

- Assume we want to find the random variate corresponding to $\mathrm{R}_1 = 0.73$
  - First, find R_1 on the y-axis of F(X)
  - Draw a line to cut the 1st jump step
  - Find the corresponding x_1 to R_1, which is 1
- You can also use a lookup table to find the corresponding x_i values
  - This table could be simply obtained from the previous CDF

$$x_i = \begin{cases} 0, & if\ R_i \le 0.5 \\ 1, & if\ 0.5 < R_i \le 0.8 \\ 2, & if\ 0.8 < R_i \le 1.0 \end{cases}$$

i	Input ri	Output xi
0	0.50	0
1	0.80	1
2	1.00	2

- Since $r_1 = 0.5 < R_1 = 0.73 \le r_2 = 0.8, \rightarrow i = 2 \rightarrow x_i = 1$

Geometric Distribution

SHARIF UNIVERSITY OF TECHNOLOGY

- Consider a geometric distribution with a success probability of P
  - The PMF of this distribution is:

$$p(x) = p(1-p)^x, \quad x = 0, 1, 2, \ldots \qquad 0 < p < 1$$

- The CDF is given by: $F(x) = 1 - (1-p)^{x+1}$
- Using the Inverse-Transform Technique:

$$X = \left\lceil \frac{\ln(1-R)}{\ln(1-p)} - 1 \right\rceil \qquad R \sim U[0,1]$$

  - This variate generator gives us geometric variates ≥ 0, but sometimes we need geometric variates ≥ q:

$$X = q + \left\lceil \frac{\ln(1-R)}{\ln(1-p)} - 1 \right\rceil \qquad x_i \in \{q, q+1, q+2, \ldots\}$$

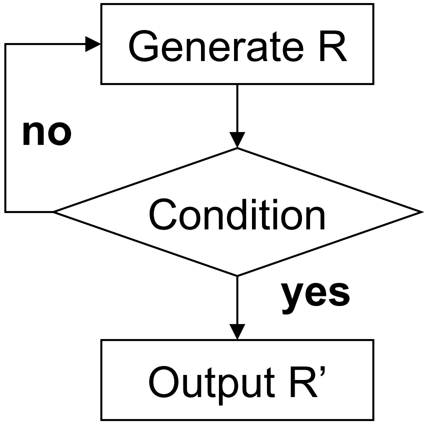
Acceptance-Rejection Technique (1)

SHARIF UNIVERSITY OF TECHNOLOGY

- Useful particularly when inverse CDF **does not exist in closed form**
- Let's discuss this technique using an example
  - Assume we want to generate random variates $X \sim U(1/4, 1)$
  - Follow these steps:
    - 1st : Generate a random number $R \sim U[0,1]$
    - 2nd : If $R \geq 1/4$, it lies in the acceptable range
      - So, we **accept** R and assign it to x
      - Unless, if $R < 1/4$, it will be **rejected**
    - 3rd : Go back to step 1, and produce an other random number R
      - Of course, if you need another variate
- In acceptance-rejection, the number of accepts is always less or equal to the total number of generated Rs

Acceptance-Rejection Technique (2)

- For instance, in our example, with producing high enough random numbers R, approximately 1/4 of the total number of produced Rs will be rejected
  - Because they were less than $1/4$
  - Only 3/4 will be accepted
- The process of variate generation using the AR technique is shown in the flowchart
  - While all of Rs do not meet the condition, a fraction of them (R') does
    - So, R' has the desired distribution



- **Efficiency:** heavily depends on the ability to minimize the number of rejections in the algorithm

Poisson Process (1)

- Recall: The PDF of a random variable N supporting the Poisson distribution with $\alpha > 0$:

$$P(N = n) = \frac{e^{-\alpha}\alpha^{n}}{n!}, n = 0,1.2, \ldots \qquad \boldsymbol{\alpha = \lambda t}$$

  - Where n indicates the number of arrivals in the intended observation period
- Consider $\mathrm{A_i}$ as the interarrival between customer i, and $i - 1$
  - These interarrivals follow the exponential distribution with rate λ
  - We know how to generate exponential variates with $x_i = -\frac{1}{\lambda} LnR_i$
- Now, we want to use this relation between Poisson and exponential distributions, to generate Poisson variates

Poisson Process (2)

SHARIF UNIVERSITY OF TECHNOLOGY

- Assume $\mathrm{N} = \mathrm{n}$ arrivals are supposed to happen in our observation period t
  - If we assume t to be a time unit ($\lambda = \alpha$), then:

$$A_1 + A_2 + A_3 + \cdots + A_n \leq 1 < A_1 + A_2 + A_3 + \cdots + A_n + \mathrm{A}_{\mathrm{n+1}} \quad \mathbf{I}$$

  - This indicates that the n+1th customer arrives after our observation period (t=one time unit)
  - So, n **MUST** meet this temporal condition
- Hence, let's generate a number of exponential $\mathrm{A_i}$ interarrivals with accordance to the above condition:

$$A_i = -\frac{1}{\lambda} LnR_i \quad \mathbf{II} \quad \Rightarrow \quad \sum_{i=1}^{n} -\frac{1}{\lambda} LnR_i \leq 1 < \sum_{i=1}^{n+1} -\frac{1}{\lambda} LnR_i$$

Poisson Process (3)

SHARIF UNIVERSITY OF TECHNOLOGY

- To simplify the unequal relation, multiply the sides by $-\lambda$

$$\boxed{\sum_{i=1}^{n} LnR_i} \geq -\lambda > \sum_{i=1}^{n+1} LnR_i$$

We know that sum of logarithms equals to logarithm of their multiplication

$$Ln\prod_{i=1}^{n} R_i \geq -\lambda > Ln\prod_{i=1}^{n+1} R_i$$

- Take e^x from both sides → $e^{Ln\prod_{i=1}^{n} R_i} \geq e^{-\lambda} > e^{Ln\prod_{i=1}^{n+1} R_i}$
- We know $e^{Ln\boldsymbol{x}} = \boldsymbol{x}$:

$$\prod_{i=1}^{n} R_i \geq e^{-\lambda} > \prod_{i=1}^{n+1} R_i$$

Poisson Process (4)

- Finally, according to what we have obtained, based on the question that **is it possible to have n arrivals in a time unit**, we can have an acceptance/rejection criterion
- Therefore, follow these steps to produce Poisson variates:
  - 1st step: Assign n=0 (no arrivals yet), and P=1 (a probability)
  - 2nd step: Generate a random number R_{n+1}, and replace the old P value with $P \times R_{n+1}$
  - 3rd step: If $P < e^{-\lambda}$, the condition is met (P is actually the multiplication of previous R_is)
    - n is accepted, and N is assigned with n
    - Otherwise, n is rejected, increase n by one, and return to step 2
- To produce a **single Poisson variate**, the above algorithm must be iteratively executed until an n is accepted

Example

- Generate 3 Poisson variates with $\lambda = 0.2$
  - First calculate $e^{-\lambda} = 0.8187$
  - Then generate a sequence of random numbers $\mathrm{R} \sim \mathrm{U}[0,1]$
    - You can use table A.1 in the textbook
  - Then, follow the algorithm:

Step 1. Set $n = 0, P = 1$.

Step 2. $R_1 = 0.4357, P = 1 \cdot R_1 = 0.4357$.

Step 3. Since $P = 0.4357 < e^{-\alpha} = 0.8187$, accept $N = 0$.

Step 1–3. ($R_1 = 0.4146$ leads to $N = 0$.)

Step 1. Set $n = 0, P = 1$.

Step 2. $R_1 = 0.8353, P = 1 \cdot R_1 = 0.8353$.

Step 3. Since $P \geq e^{-\alpha}$, reject $n = 0$ and return to Step 2 with $n = 1$.

Step 2. $R_2 = 0.9952, P = R_1R_2 = 0.8313$.

Step 3. Since $P \geq e^{-\alpha}$, reject $n = 1$ and return to Step 2 with $n = 2$.

Step 2. $R_3 = 0.8004, P = R_1R_2R_3 = 0.6654$.

Step 3. Since $P < e^{-\alpha}$, accept $N = 2$.

n	R_{n+1}	P	*Accept/Reject*	*Result*
0	0.4357	0.4357	$P < e^{-\alpha}$ (accept)	$N = 0$
0	0.4146	0.4146	$P < e^{-\alpha}$ (accept)	$N = 0$
0	0.8353	0.8353	$P \geq e^{-\alpha}$ (reject)	
1	0.9952	0.8313	$P \geq e^{-\alpha}$ (reject)	
2	0.8004	0.6654	$P < e^{-\alpha}$ (accept)	$N = 2$

Poisson variates that we have generated

What if λ is high? (1)

- As we noticed, higher λ imposes more cost to the acceptance/rejection technique
  - Specially when $\lambda \geq 15$
  - An alternate approach is to use an **approximation** technique based on the standard normal distribution
- Set: $Z = \frac{N-\lambda}{\sqrt{\lambda}}$
  - Where Z is approximately a normally distributed variable with mean 0 and variance 1
- Generate standard normal variate Z (pair wise):

$$Z_1 = (-2\ln R_1)^{1/2}\cos(2\pi R_2)$$
$$Z_2 = (-2\ln R_1)^{1/2}\sin(2\pi R_2)$$

Generate 2 random numbers, and get 2 standard normal variates

What if λ is high? (2)

- Z_1, and Z_2 are standard normal variates, which are totally independent
- But your Poisson variates have not been yet generated
  - To do so, set N equal to the following:

$$N = \lceil \lambda + \sqrt{\lambda} Z - 0.5 \rceil$$

  - For both Z_1, and Z_2
  - This equation has been obtained based on the conducted transformation in the previous slide
- Note: If $\lambda + \sqrt{\lambda} Z - 0.5 < 0$, put N=0
- Not forget that this approximation technique is not an acceptance/rejection technique!

Non-Stationary Poisson Process (NSPP)

- Recall: a NSPP is a Poisson arrival process with an arrival rate that varies with time
  - Our goal here is to use a special case of acceptance/rejection technique known as **thinning** to produce exponentially distributed interarrivals with rate $\lambda(t), 0 \leq t \leq T$
- Idea behind thinning:
  - 1st step: Generate a **stationary** Poisson arrival process at the fastest rate, $\lambda^* = \text{Max}\ \lambda(t)$, assign t=0, and i=1
    - i indicates the ith event (or customer)
  - 2nd step: Generate an exponential variate E with rate λ^*, and put t=t+E
  - 3rd step: Generate $R \sim U[0,1]$, if $R \leq \lambda(t)/\ \lambda^*$ → assign $\tau_i = t$, i=i+1
    - τ_i is the arrival time of the ith NSPP event
  - 4th step: Return to step 2 (whether accepted or not)
- Where does thinning come with low efficiency?

Example

SHARIF UNIVERSITY OF TECHNOLOGY

- For the following NSPP, generate a random variate using the thinning technique

	t (min)	Mean Time Between Arrivals (min)	Arrival Rate $\lambda(t)$ (#/min)
1^{st} 60 min	0	15	1/15
1^{st} 60 min	60	12	1/12
	120	7	1/7
4^{th} 60 min	180	5	1/5
4^{th} 60 min	240	8	1/8
	300	10	1/10
	360	15	1/15
	420	20	1/20
	480	20	1/20

Step 1: $\lambda^* = \max \lambda(t) = 1/5$, $t = 0$ and $i = 1$

Step 2: For random number R = 0.2130

$E = -5\ln(0.213) = 13.13$

$t = t+E = 13.13$

Within the 1^{st} 60 min

Step 3: Generate R = 0.8830

$\lambda(13.13)/\lambda^*=(1/15)/(1/5)=1/3$

Since R>1/3, do not accept the arrival time

Step 2: For random number R = 0.5530

$E = -5\ln(0.553) = 2.96$

$t = t+E = 13.13 + 2.96 = 16.09$

Step 3: Generate R = 0.0240

$\lambda(16.09)/\lambda^*=(1/15)/(1/5)=1/3$

Since R<1/3, $T_1 = t = 16.09$

and $i = i + 1 = 2$

Variate Generation for Special Cases

- Among other distributions, there are special techniques used for variate generation, for instance:
  - Normal, and log-normal distributions
    - You can find out more in the textbook
  - The standard normal distribution, which we have already discussed
  - Beta distribution (which could be obtained from gamma distribution)
  - Erlang, and binomial distributions
    - **Convolution technique** is used: In this technique, to produce a random variate following an intended distribution, two or more random variables from another distribution are added up
      - For the Erlang distribution:

$$X = \sum_{i=1}^{k} X_i \longrightarrow X = \sum_{i=1}^{k} -\frac{1}{k\theta} \ln R_i = -\frac{1}{k\theta} \ln \left(\prod_{i=1}^{k} R_i \right)$$

A sequence of R_is is used for a single X

Summary

- Principles of random-variate generation via:
  - Inverse-transform technique
    - The CDF of the target distribution must be reversible
      - Unless, we need another method
  - Acceptance-rejection technique
    - Discussed for Poisson and NSPP
  - Special cases
    - Convolution technique for Erlang variates
- These techniques are important for generating variates with continuous or discrete distributions to obtain appropriate inputs to be fed into our simulator