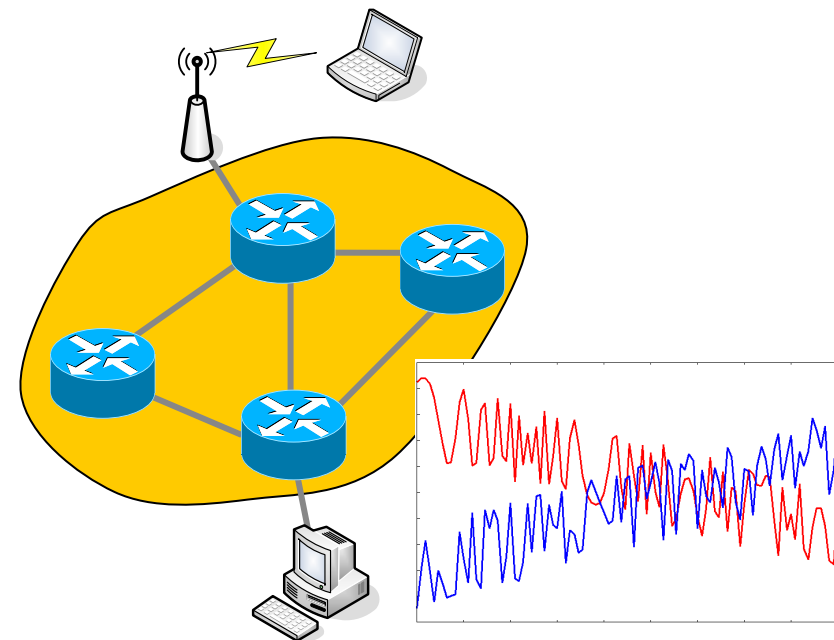


Chapter 7

Random-Variate Generation



Contents

- Inverse-transform Technique
- Acceptance-Rejection Technique
- Special Properties

Purpose & Overview

- Develop understanding of generating samples from a specified distribution as input to a simulation model.
- Illustrate some widely-used techniques for generating random variates:
 - Inverse-transform technique
 - Acceptance-rejection technique
 - Special properties

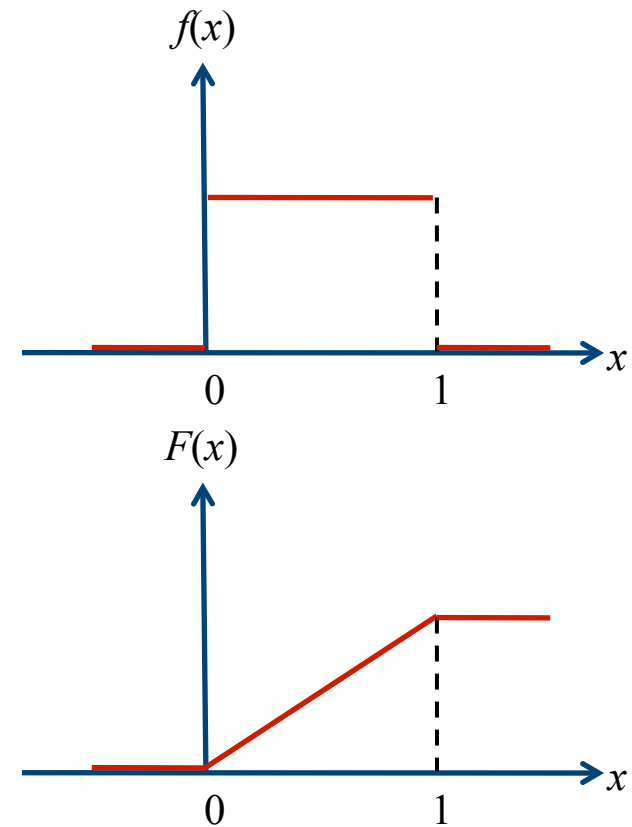
Preparation

- It is assumed that a source of uniform $[0,1]$ random numbers exists.
 - Linear Congruential Method (LCM)
- Random numbers R, R_1, R_2, \dots with
 - PDF

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

- CDF

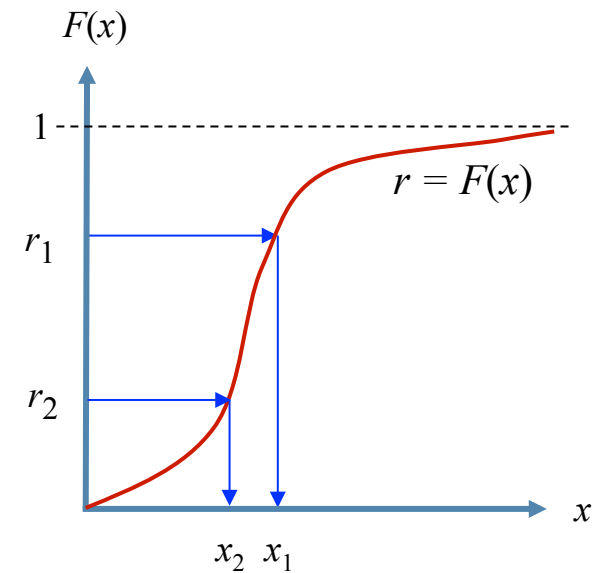
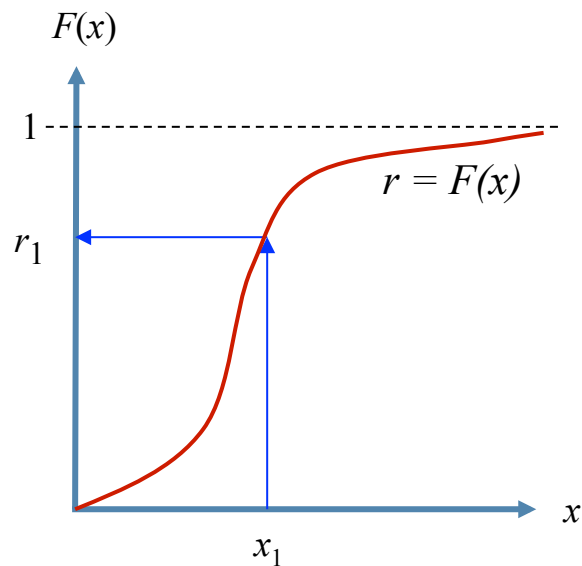
$$F_R(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$



Inverse-transform Technique

Inverse-transform Technique

- The concept:
 - For CDF function: $r = F(x)$
 - Generate r from uniform $(0,1)$, a.k.a $U(0,1)$
 - Find x , $x = F^{-1}(r)$



Inverse-transform Technique

- The inverse-transform technique can be used in principle for any distribution.
- Most useful when the CDF $F(x)$ has an **inverse** $F^{-1}(x)$ which is easy to compute.
- Required steps
 1. Compute the CDF of the desired random variable X
 2. Set $F(X) = R$ on the range of X
 3. Solve the equation $F(X) = R$ for X in terms of R
 4. Generate uniform random numbers R_1, R_2, R_3, \dots and compute the desired random variate by $X_i = F^{-1}(R_i)$

Inverse-transform Technique: Example

- Exponential Distribution

- PDF

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

- CDF

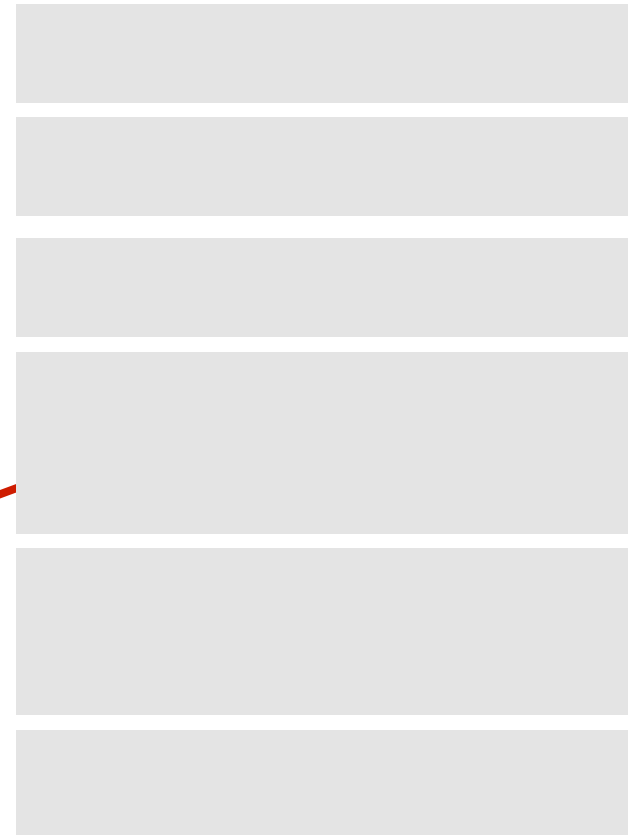
$$F(x) = 1 - e^{-\lambda x}$$

- Simplification

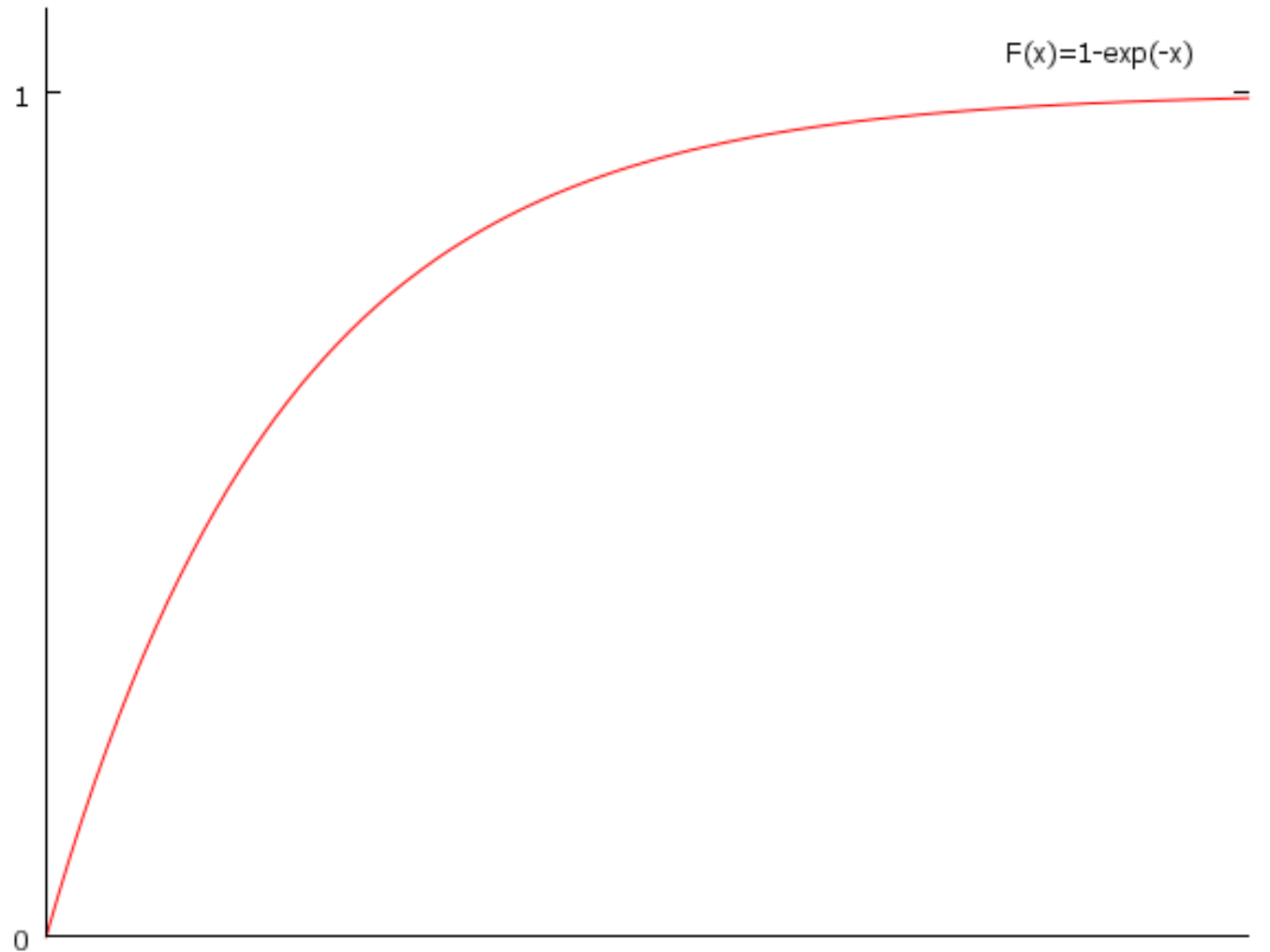
$$X = -\frac{\ln(R)}{\lambda}$$

- Since R and $(1-R)$ are uniformly distributed on $[0,1]$

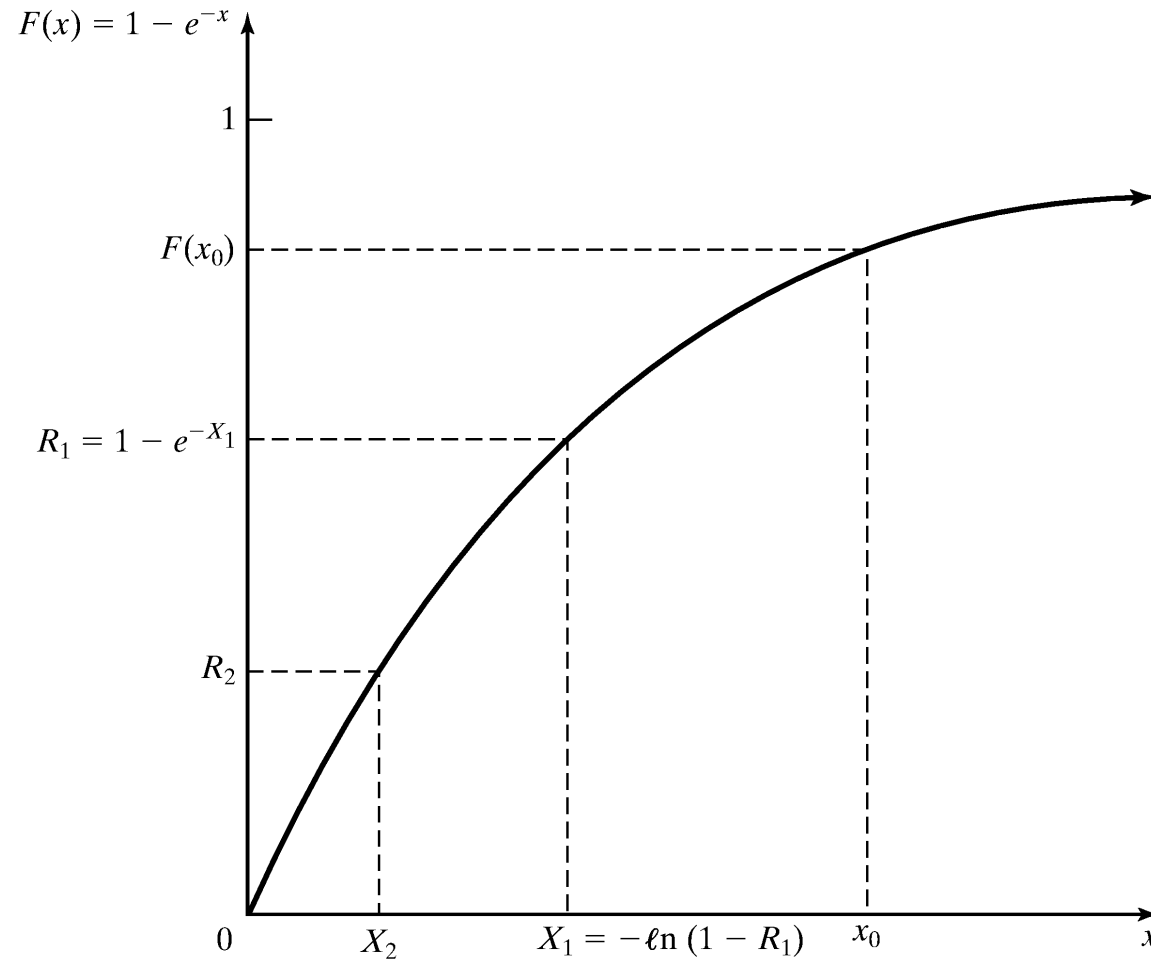
- To generate $X_1, X_2, X_3 \dots$



Inverse-transform Technique: Example



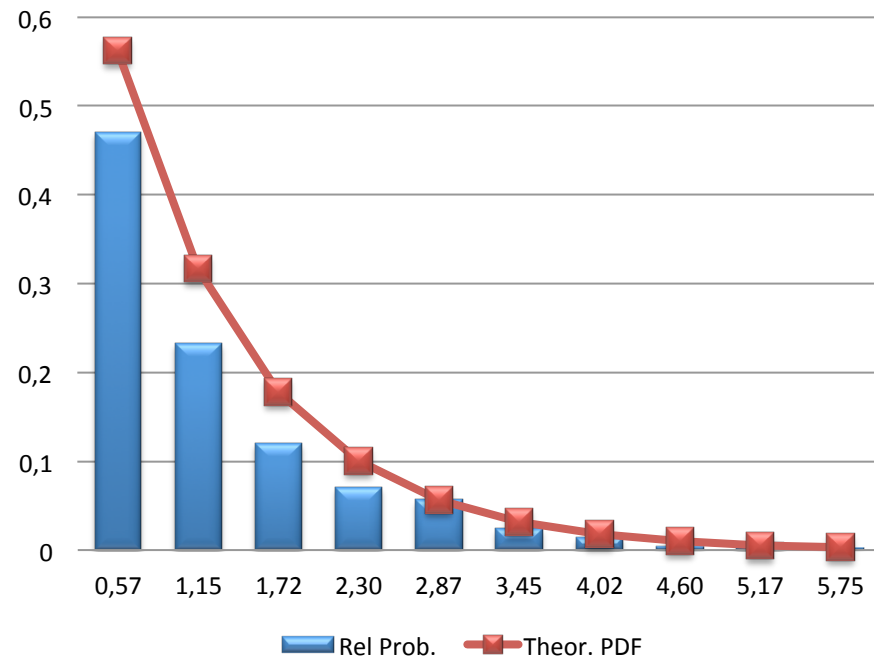
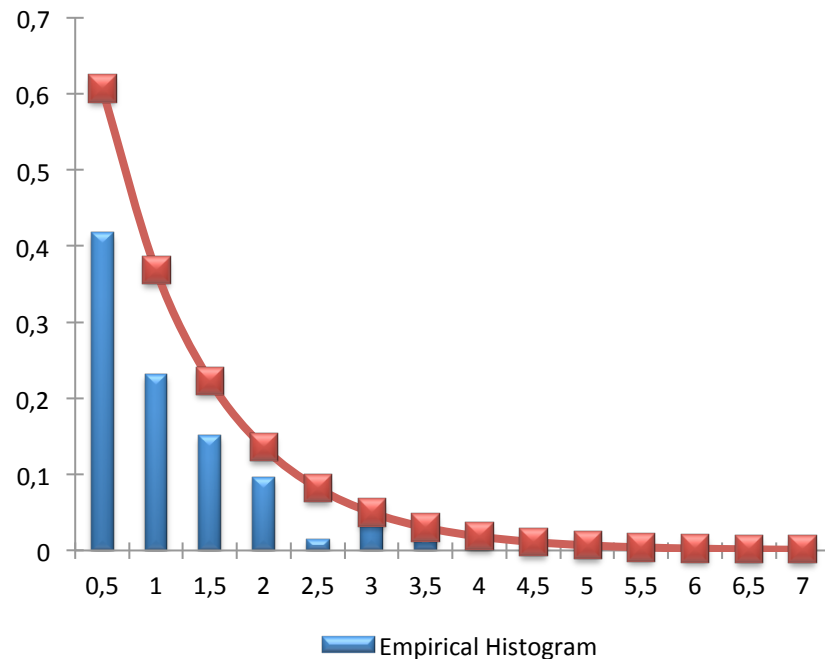
Inverse-transform Technique: Example



Inverse-transform technique for $\exp(\lambda = 1)$

Inverse-transform Technique: Example

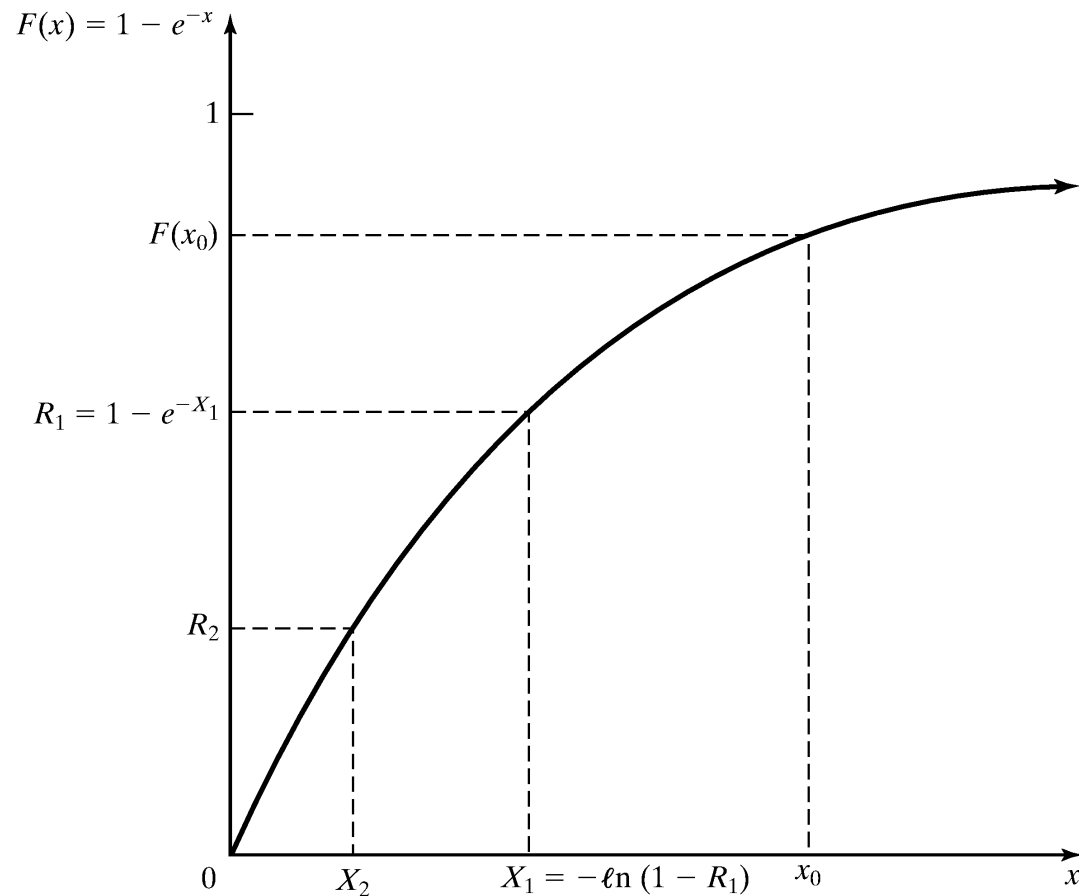
- Example:
 - Generate 200 or 500 variates X_i with distribution $\exp(\lambda=1)$
 - Generate 200 or 500 R_s with $U(0,1)$, the histogram of X_s becomes:



Inverse-transform Technique

- Check: Does the random variable X_1 have the desired distribution?

$$P(X_1 \leq x_0) = P(R_1 \leq F(x_0)) = F(x_0)$$



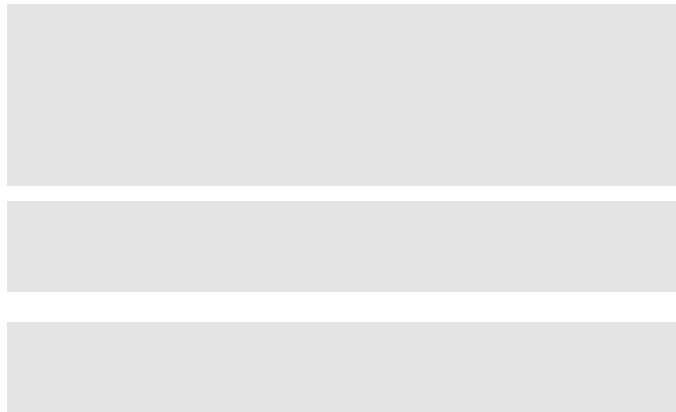
Inverse-transform Technique: Other Distributions

- Examples of other distributions for which inverse CDF works are:
 - Uniform distribution
 - Weibull distribution
 - Triangular distribution

Inverse-transform Technique: Uniform Distribution

- Random variable X uniformly distributed over $[a, b]$

$$F(X) = R$$



Inverse-transform Technique: Weibull Distribution

- The Weibull Distribution is described by

- PDF

$$f(x) = \frac{\beta}{\alpha^\beta} x^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^\beta}$$

- CDF

$$F(X) = 1 - e^{-\left(\frac{x}{\alpha}\right)^\beta}$$

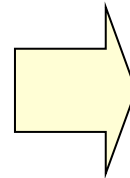
- The variate is



Inverse-transform Technique: Triangular Distribution

- The CDF of a Triangular Distribution with endpoints $(0, 2)$ is given by

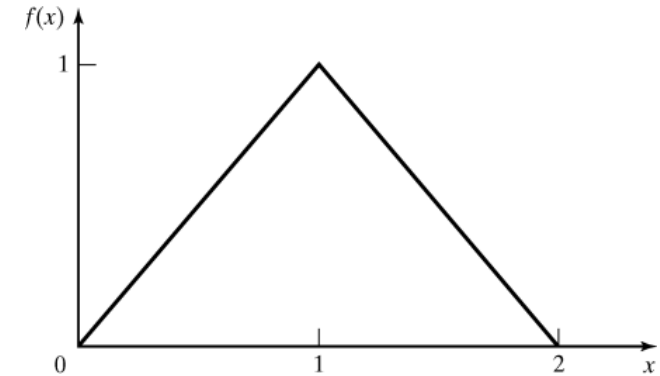
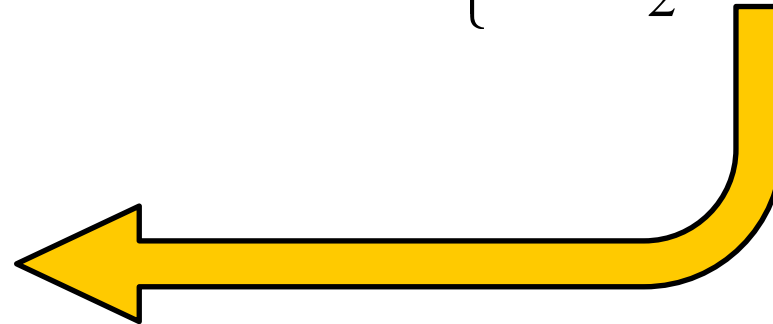
$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x^2}{2} & 0 < x \leq 1 \\ 1 - \frac{(2-x)^2}{2} & 1 < x \leq 2 \\ 1 & x > 2 \end{cases}$$



$$R(X) = \begin{cases} \frac{X^2}{2} & 0 \leq X \leq 1 \\ 1 - \frac{(2-X)^2}{2} & 1 \leq X \leq 2 \end{cases}$$

- X is generated by

$$X = \begin{cases} \sqrt{2R} & 0 \leq R \leq \frac{1}{2} \\ 2 - \sqrt{2(1-R)} & \frac{1}{2} < R \leq 1 \end{cases}$$



Inverse-transform Technique: Empirical Continuous Distributions

- When theoretical distributions are not applicable
- To collect empirical data:
 - Resample the observed data
 - Interpolate between observed data points to fill in the gaps

Inverse-transform Technique: Empirical Continuous Distributions

- For a small sample set (size n):
 - Arrange the data from smallest to largest

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$$

- Set $x_{(0)}=0$
- Assign the probability $1/n$ to each interval $x_{(i-1)} \leq x \leq x_{(i)}$ $i = 1, 2, \dots, n$
- The slope of each line segment is defined as

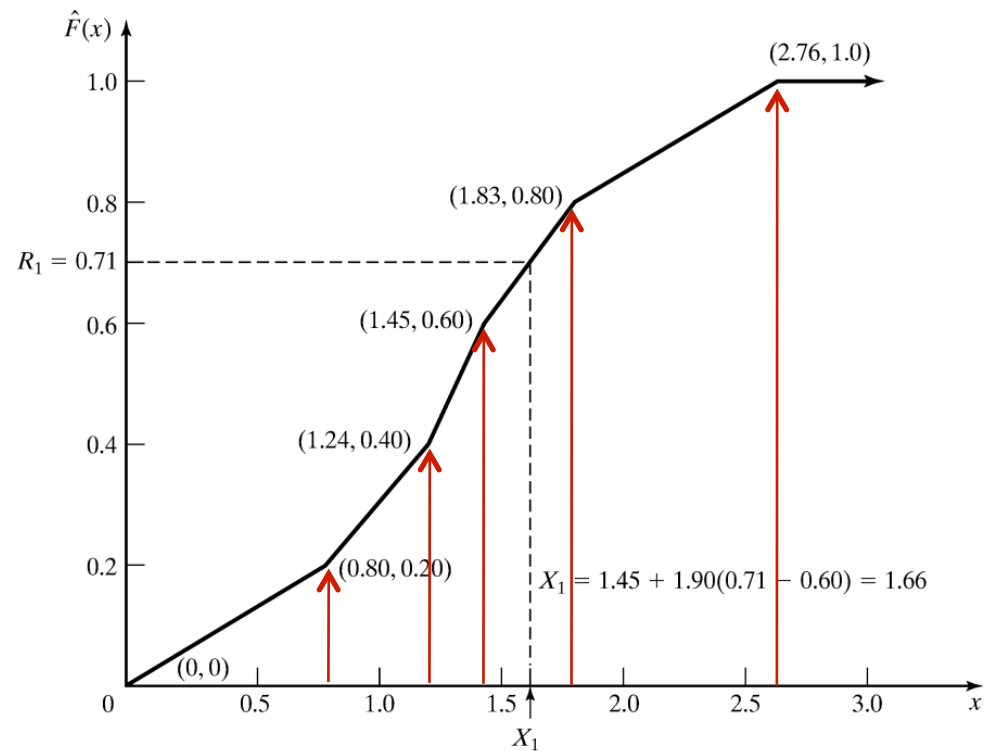
$$a_i = \frac{\frac{x_{(i)} - x_{(i-1)}}{n} - \frac{x_{(i-1)} - x_{(i-2)}}{n}}{\frac{i}{n} - \frac{(i-1)}{n}} = \frac{x_{(i)} - x_{(i-1)}}{1}$$

- The inverse CDF is given by

$$X = \hat{F}^{-1}(R) = x_{(i-1)} + a_i \left(R - \frac{(i-1)}{n} \right) \quad \text{when} \quad \frac{(i-1)}{n} < R \leq \frac{i}{n}$$

Inverse-transform Technique: Empirical Continuous Distributions

i	Interval	PDF	CDF	Slope a_i
1	$0.0 < x \leq 0.8$	0.2	0.2	4.00
2	$0.8 < x \leq 1.24$	0.2	0.4	2.20
3	$1.24 < x \leq 1.45$	0.2	0.6	1.05
4	$1.45 < x \leq 1.83$	0.2	0.8	1.90
5	$1.83 < x \leq 2.76$	0.2	1.0	4.65



$$R_1 = 0.71$$

$$\begin{aligned} X_1 &= x_{(4-1)} + a_4(R_1 - (4-1)/n) \\ &= 1.45 + 1.90(0.71 - 0.6) \\ &= 1.66 \end{aligned}$$

Inverse-transform Technique: Empirical Continuous Distributions

- What happens for large samples of data?
 - Several hundreds or tens of thousand
- First summarize the data into a frequency distribution with smaller number of intervals
- Afterwards, fit continuous empirical CDF to the frequency distribution
- Slight modifications
 - Slope

$$a_i = \frac{x_{(i)} - x_{(i-1)}}{c_i - c_{i-1}}$$

c_i cumulative probability of the first i intervals

- The inverse CDF is given by

$$X = \hat{F}^{-1}(R) = x_{(i-1)} + a_i(R - c_{i-1}) \quad \text{when } c_{i-1} < R \leq c_i$$

Inverse-transform Technique: Empirical Continuous Distributions

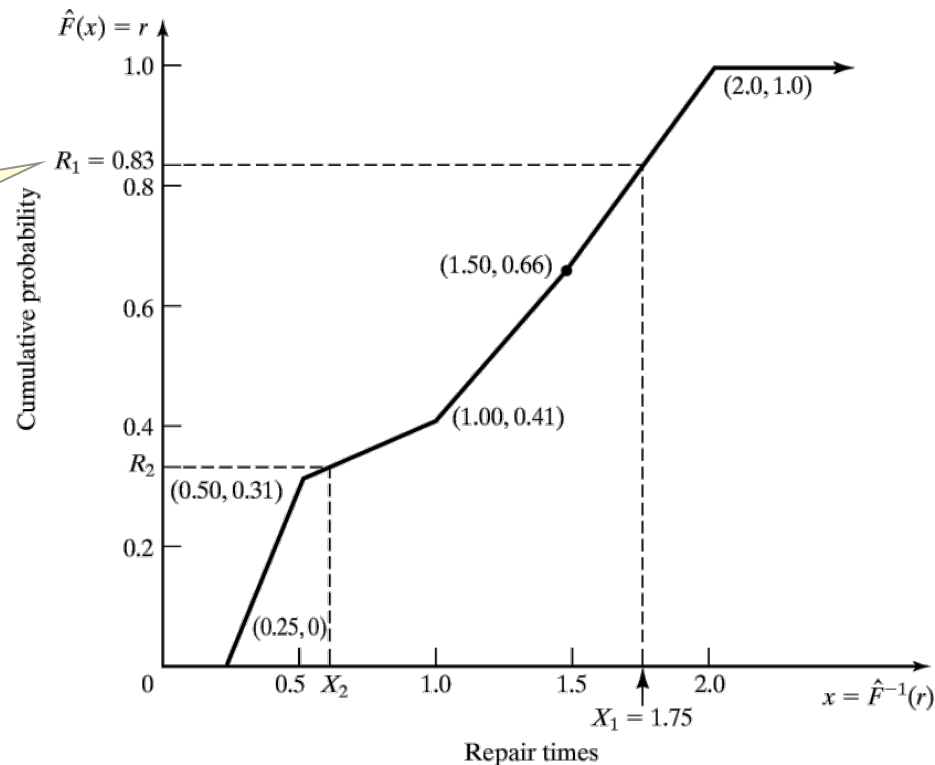
- Example: Suppose the data collected for 100 repair times are:

Interval (Hours)	Relative Frequency	Cumulative Frequency, c_i	Slope, a_i
$0.25 \leq x \leq 0.5$	0.31	0.31	0.81
$0.5 \leq x \leq 1.0$	0.10	0.41	5.00
$1.0 \leq x \leq 1.5$	0.25	0.66	2.00
$1.5 \leq x \leq 2.0$	0.34	1.00	1.47

Consider $R_1 = 0.83$:

$$c_3 = 0.66 < R_1 < c_4 = 1.00$$

$$\begin{aligned} X_1 &= x_{(4-1)} + a_4(R_1 - c_{(4-1)}) \\ &= 1.5 + 1.47(0.83 - 0.66) \\ &= 1.75 \end{aligned}$$



Inverse-transform Technique: Empirical Continuous Distributions

- Problems with empirical distributions
 - The data in the previous example is restricted:
Range $0.25 \leq X \leq 2.0$
 - The underlying distribution might have a wider range
 - Thus, try to find a theoretical distribution
- Hints for building empirical distributions based on frequency tables
 - It is recommended to use relatively short intervals
 - Number of bins increases
 - This will result in a more accurate estimate

Inverse-transform Technique: Continuous Distributions

- A number of continuous distributions do not have a closed form expression for their CDF, e.g.

- Normal $F(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt$
- Gamma
- Beta

- The presented method does not work for these distributions
- Solution
 - Approximate the CDF or numerically integrate the CDF
- Problem
 - Computationally slow

Inverse-transform Technique: Discrete Distribution

- All discrete distributions can be generated via inverse-transform technique
- Methods:
 - numerically
 - table-lookup procedure
 - algebraically
 - a formula
- Examples of application
 - Empirical
 - Discrete uniform
 - Geometric

Inverse-transform Technique: Discrete Distribution

- Example: Suppose the number of shipments, x , on the loading dock of a company is either 0, 1, or 2
 - Data - Probability distribution:

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

- The inverse-transform technique as table-lookup procedure

$$F(x_{i-1}) = r_{i-1} < R \leq r_i = F(x_i)$$

- Set $X = x_i$

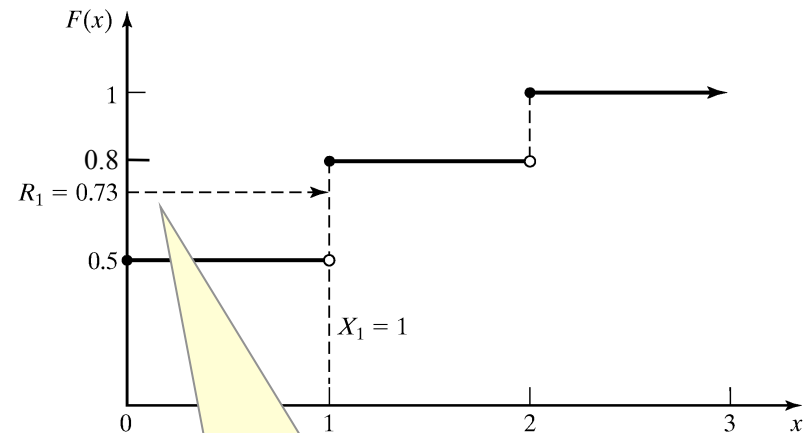
Inverse-transform Technique: Discrete Distribution

Method - Given R , the generation scheme becomes:

$$x = \begin{cases} 0, & R \leq 0.5 \\ 1, & 0.5 < R \leq 0.8 \\ 2, & 0.8 < R \leq 1.0 \end{cases}$$

Table for generating the discrete variate X

i	Input r_i	Output x_i
1	0.5	0
2	0.8	1
3	1.0	2



Consider $R_1 = 0.73$:
 $F(x_{i-1}) < R \leq F(x_i)$
 $F(x_0) < 0.73 \leq F(x_1)$
 Hence, $X_1 = 1$

Acceptance-Rejection Technique

Acceptance-Rejection Technique

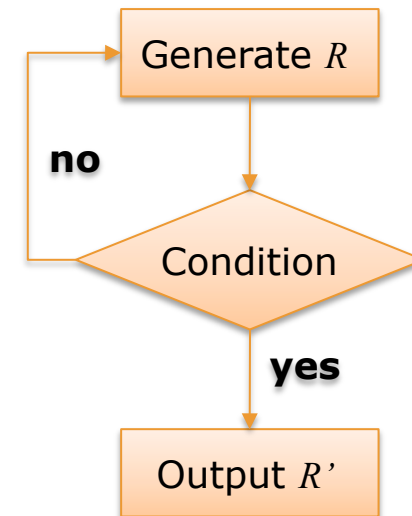
- Useful particularly when inverse CDF does not exist in closed form
 - Thinning
- Illustration: To generate random variates: $X \sim U(1/4, 1)$

Procedure:

Step 1. Generate $R \sim U(0, 1)$

Step 2. If $R \geq 1/4$, accept $X=R$

Step 3. If $R < 1/4$, reject R , return to Step 1



- R does not have the desired distribution, but R conditioned (R') on the event $\{R \geq 1/4\}$ does.
- Efficiency: Depends heavily on the ability to minimize the number of rejections.

Acceptance-Rejection Technique: Poisson Distribution

- Probability mass function of a Poisson Distribution

$$P(N = n) = \frac{\alpha^n}{n!} e^{-\alpha}$$

- Exactly n arrivals during one time unit

$$A_1 + A_2 + \cdots + A_n \leq 1 < A_1 + A_2 + \cdots + A_n + A_{n+1}$$

- Since interarrival times are exponentially distributed we can set

$$A_i = \frac{-\ln(R_i)}{\alpha}$$

- Well known, we derived this generator in the beginning of the class

Acceptance-Rejection Technique: Poisson Distribution

- Substitute the sum by

$$\sum_{i=1}^n \frac{-\ln(R_i)}{\alpha} \leq 1 < \sum_{i=1}^{n+1} \frac{-\ln(R_i)}{\alpha}$$

- Simplify by

- multiply by $-\alpha$, which reverses the inequality sign
- sum of logs is the log of a product

$$\sum_{i=1}^n \ln(R_i) \geq -\alpha > \sum_{i=1}^{n+1} \ln(R_i)$$

$$\ln \prod_{i=1}^n R_i \geq -\alpha > \ln \prod_{i=1}^{n+1} R_i$$

- Simplify by $e^{\ln(x)} = x$

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

Acceptance-Rejection Technique: Poisson Distribution

- Procedure of generating a Poisson random variate N is as follows
 1. Set $n=0, P=1$
 2. Generate a random number R_{n+1} , and replace P by $P \times R_{n+1}$
 3. If $P < \exp(-\alpha)$, then accept $N=n$
 - Otherwise, reject the current n , increase n by one, and return to step 2.

Acceptance-Rejection Technique: Poisson Distribution

- Example: Generate three Poisson variates with mean $\alpha=0.2$
 - $\exp(-0.2) = 0.8187$
- Variate 1
 - Step 1: Set $n = 0, P = 1$
 - Step 2: $R1 = 0.4357, P = 1 \times 0.4357$
 - Step 3: Since $P = 0.4357 < \exp(-0.2)$, **accept $N = 0$**
- Variate 2
 - Step 1: Set $n = 0, P = 1$
 - Step 2: $R1 = 0.4146, P = 1 \times 0.4146$
 - Step 3: Since $P = 0.4146 < \exp(-0.2)$, **accept $N = 0$**
- Variate 3
 - Step 1: Set $n = 0, P = 1$
 - Step 2: $R1 = 0.8353, P = 1 \times 0.8353$
 - Step 3: Since $P = 0.8353 > \exp(-0.2)$, reject $n = 0$ and return to Step 2 with $n = 1$
 - Step 2: $R2 = 0.9952, P = 0.8353 \times 0.9952 = 0.8313$
 - Step 3: Since $P = 0.8313 > \exp(-0.2)$, reject $n = 1$ and return to Step 2 with $n = 2$
 - Step 2: $R3 = 0.8004, P = 0.8313 \times 0.8004 = 0.6654$
 - Step 3: Since $P = 0.6654 < \exp(-0.2)$, **accept $N = 2$**

Acceptance-Rejection Technique: Poisson Distribution

- It took five random numbers to generate three Poisson variates
- In long run, the generation of Poisson variates requires some overhead!

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

Special Properties

Special Properties

- Based on features of particular family of probability distributions
- For example:
 - Direct Transformation for normal and lognormal distributions
 - Convolution

Direct Transformation

- Approach for $N(0,1)$
 - PDF

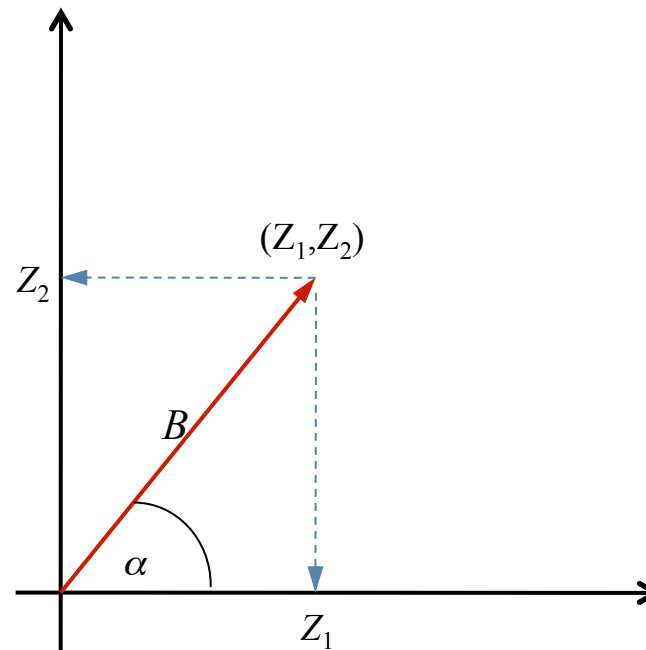
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

- CDF, No closed form available

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

Direct Transformation

- Approach for $N(0,1)$
 - Consider two standard normal random variables, Z_1 and Z_2 , plotted as a point in the plane:
 - In polar coordinates:
 - $Z_1 = B \cos(\alpha)$
 - $Z_2 = B \sin(\alpha)$



Direct Transformation

- Chi-square distribution
 - Given k independent $N(0, 1)$ random variables X_1, X_2, \dots, X_k , then the sum is according to the Chi-square distribution

- PDF

$$\chi_k^2 = \sum_{i=1}^k X_i^2$$

$$f(x, k) = \frac{1}{\Gamma\left(\frac{k}{2}\right) 2^{\frac{k}{2}}} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$$

Direct Transformation

- The following relationships are known
 - $B^2 = Z_1^2 + Z_2^2 \sim \chi^2$ distribution with 2 degrees of freedom = $\exp(\lambda = 1/2)$.
 - Hence:

$$B = \sqrt{-2 \ln R}$$

- The radius B and angle α are mutually independent.

$$\begin{aligned} Z_1 &= \sqrt{-2 \ln R_1} \cos(2\pi R_2) \\ Z_2 &= \sqrt{-2 \ln R_1} \sin(2\pi R_2) \end{aligned}$$

Direct Transformation

- Approach for $N(\mu, \sigma^2)$:
 - Generate $Z_i \sim N(0,1)$

$$X_i = \mu + \sigma Z_i$$

- Approach for Lognormal(μ, σ^2):
 - Generate $X \sim N(\mu, \sigma^2)$

$$Y_i = e^{X_i}$$

Direct Transformation: Example

- Let $R_1 = 0.1758$ and $R_2 = 0.1489$
- Two standard normal random variates are generated as follows:

$$Z_1 = \sqrt{-2\ln(0.1758)} \cos(2\pi \cdot 0.1489) = 1.11$$

$$Z_2 = \sqrt{-2\ln(0.1758)} \sin(2\pi \cdot 0.1489) = 1.50$$

- To obtain normal variates X_i with mean $\mu = 10$ and variance $\sigma^2 = 4$

$$X_1 = 10 + 2 \cdot 1.11 = 12.22$$

$$X_2 = 10 + 2 \cdot 1.50 = 13.00$$

Convolution

- Convolution
 - The sum of independent random variables
- Can be applied to obtain
 - Erlang variates
 - Binomial variates

Convolution

- Erlang Distribution
 - Erlang random variable X with parameters (k, θ) can be depicted as the sum of k independent exponential random variables $X_i, i = 1, \dots, k$ each having mean $1/(k \theta)$



Summary

- Principles of random-variate generation via
 - Inverse-transform technique
 - Acceptance-rejection technique
 - Special properties
- Important for generating continuous and discrete distributions