

Random Variate Generation

Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

These slides are available on-line at:

<http://www.cse.wustl.edu/~jain/cse567-06/>



1. Inverse transformation
2. Rejection
3. Composition
4. Convolution
5. Characterization

Random-Variate Generation

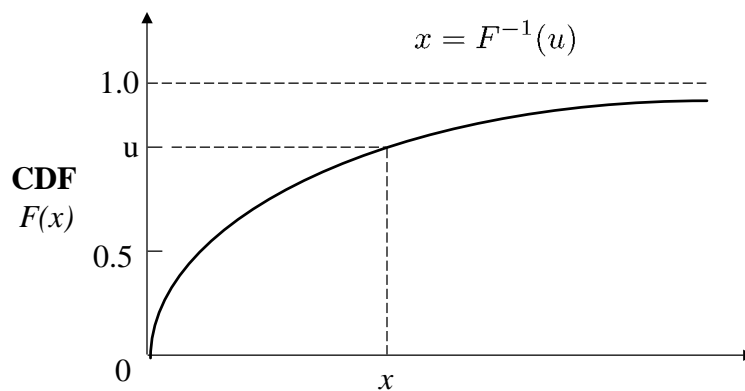
- General Techniques
- Only a few techniques may apply to a particular distribution
- Look up the distribution in Chapter 29

Inverse Transformation

- Used when F^{-1} can be determined either analytically or empirically.

$$u = F(x) \sim U(0, 1)$$

$$x = F^{-1}(u)$$



Proof

Let $y = g(x)$, so that $x = g^{-1}(y)$.

$$\begin{aligned}F_Y(y) &= P(Y \leq y) = P(x \leq g^{-1}(y)) \\ &= F_X(g^{-1}(y))\end{aligned}$$

If $g(x) = F(x)$, or $y = F(x)$

$$F(y) = F(F^{-1}(y)) = y$$

And:

$$f(y) = dF/dy = 1$$

That is, y is uniformly distributed between 0 and 1.

Example 28.1

- For exponential variates:

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

The CDF $F(x) = 1 - e^{-\lambda x} = u$ or, $x = -\frac{1}{\lambda} \ln(1 - u)$

- If u is $U(0,1)$, $1-u$ is also $U(0,1)$
- Thus, exponential variables can be generated by:

$$x = -\frac{1}{\lambda} \ln(u)$$

Example 28.2

- The packet sizes (trimodal) probabilities:

Size	Probability
64 Bytes	0.7
128 Bytes	0.1
512 Bytes	0.2

- The CDF for this distribution is:

$$F(x) = \begin{cases} 0.0 & 0 \leq x < 64 \\ 0.7 & 64 \leq x < 128 \\ 0.8 & 128 \leq x < 512 \\ 1.0 & 512 \leq x \end{cases}$$

Example 28.2 (Cont)

- The inverse function is:

$$F^{-1}(u) = \begin{cases} 64 & 0 < u \leq 0.7 \\ 128 & 0.7 < u \leq 0.8 \\ 512 & 0.8 < u \leq 1 \end{cases}$$

Generate $u \sim U(0, 1)$

$u \leq 0.7 \Rightarrow \text{Size} = 64$

$0.7 < u \leq 0.8 \Rightarrow \text{size} = 128$

$0.8 < u \Rightarrow \text{size} = 512$

- Note: CDF is *continuous from the right*
 - \Rightarrow the value on the right of the discontinuity is used
 - \Rightarrow The inverse function is continuous from the left
 - $\Rightarrow u=0.7 \Rightarrow x=64$

Applications of the Inverse-Transformation Technique

Distribution	CDF $F(x)$	Inverse
Exponential	$1 - e^{-x/a}$	$-a \ln(u)$
Extreme value	$1 - e^{-e^{\frac{x-a}{b}}}$	$a + b \ln \ln u$
Geometric	$1 - (1 - p)^x$	$\left\lceil \frac{\ln(u)}{\ln(1-p)} \right\rceil$
Logistic	$1 - \frac{1}{1 + e^{\frac{x-\mu}{b}}}$	$\mu - b \ln\left(\frac{1}{u} - 1\right)$
Pareto	$1 - x^{-a}$	$1/u^{1/a}$
Weibull	$1 - e^{-(x/a)^b}$	$a(\ln u)^{1/b}$

Washington University in St. Louis

CSE567M

©2006 Raj Jain

28-9

Rejection

- Can be used if a pdf $g(x)$ exists such that $c g(x)$ majorizes the pdf $f(x) \Rightarrow c g(x) \geq f(x) \forall x$
- Steps:
 1. Generate x with pdf $g(x)$.
 2. Generate y uniform on $[0, cg(x)]$.
 3. If $y \leq f(x)$, then output x and return.
Otherwise, repeat from step 1.
 \Rightarrow Continue *rejecting* the random variates x and y until $y \geq f(x)$
- Efficiency = how closely $c g(x)$ envelopes $f(x)$
Large area between $c g(x)$ and $f(x) \Rightarrow$ Large percentage of (x, y) generated in steps 1 and 2 are rejected
- If generation of $g(x)$ is complex, this method may not be efficient.

Washington University in St. Louis

CSE567M

©2006 Raj Jain

28-10

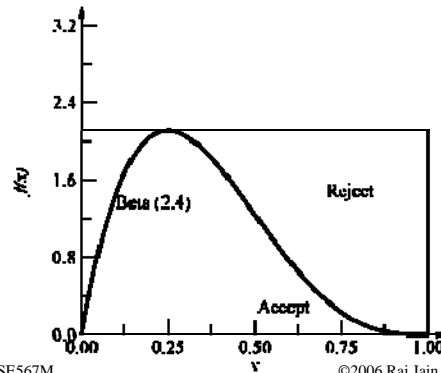
Example 28.2

- Beta(2.4) density function:
 $f(x) = 20x(1-x)^3 \quad 0 \leq x \leq 1$
 $c=2.11$ and $g(x) = 1 \quad 0 \leq x \leq 1$

- Bounded inside a rectangle of height 2.11

⇒ Steps:

- Generate x uniform on $[0, 1]$.
- Generate y uniform on $[0, 2.11]$.
- If $y \leq 20x(1-x)^3$, then output x and return. Otherwise repeat from step 1.



Washington University in St. Louis

CSE567M

©2006 Raj Jain

28-11

Composition

- Can be used if CDF $F(x) =$ Weighted sum of n other CDFs.

$$F(x) = \sum_{i=1}^n p_i F_i(x)$$

- Here, $p_i \geq 0$, $\sum_{i=1}^n p_i = 1$, and F_i 's are distribution functions.

- n CDFs are composed together to form the desired CDF
Hence, the name of the technique.

- The desired CDF is decomposed into several other CDFs
⇒ Also called **decomposition**.

- Can also be used if the pdf $f(x)$ is a weighted sum of n other pdfs:

$$f(x) = \sum_{i=1}^n p_i f_i(x)$$

Washington University in St. Louis

CSE567M

©2006 Raj Jain

28-12

Steps:

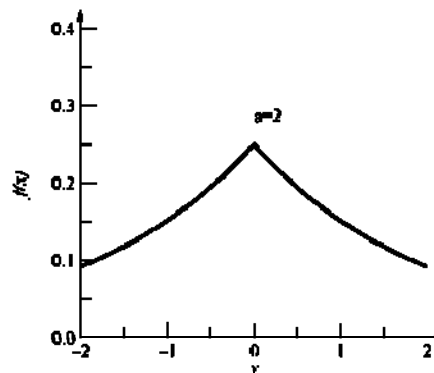
- Generate a random integer I such that:

$$P(I = i) = p_i$$

- This can easily be done using the inverse-transformation method.
- Generate x with the i th pdf $f_i(x)$ and return.

Example 28.4

- pdf: $f(x) = \frac{1}{2a} e^{-|x|/a}$
- Composition of two exponential pdf's
- Generate
 - $u_1 \sim U(0, 1)$
 - $u_2 \sim U(0, 1)$
- If $u_1 < 0.5$, return; otherwise return $x = a \ln u_2$.
- Inverse transformation better for Laplace



Convolution

- ❑ Sum of n variables: $x = y_1 + y_2 + \dots + y_n$
- ❑ Generate n random variate y_i 's and sum
- ❑ For sums of two variables, pdf of $x =$ convolution of pdfs of y_1 and y_2 . Hence the name
- ❑ Although no convolution in generation
- ❑ If pdf or CDF = Sum \Rightarrow Composition
- ❑ Variable $x =$ Sum \Rightarrow Convolution

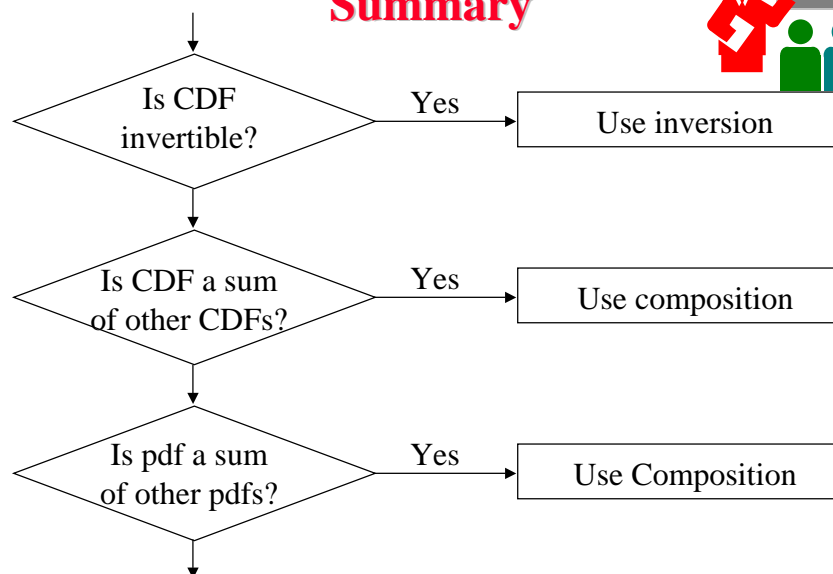
Convolution: Examples

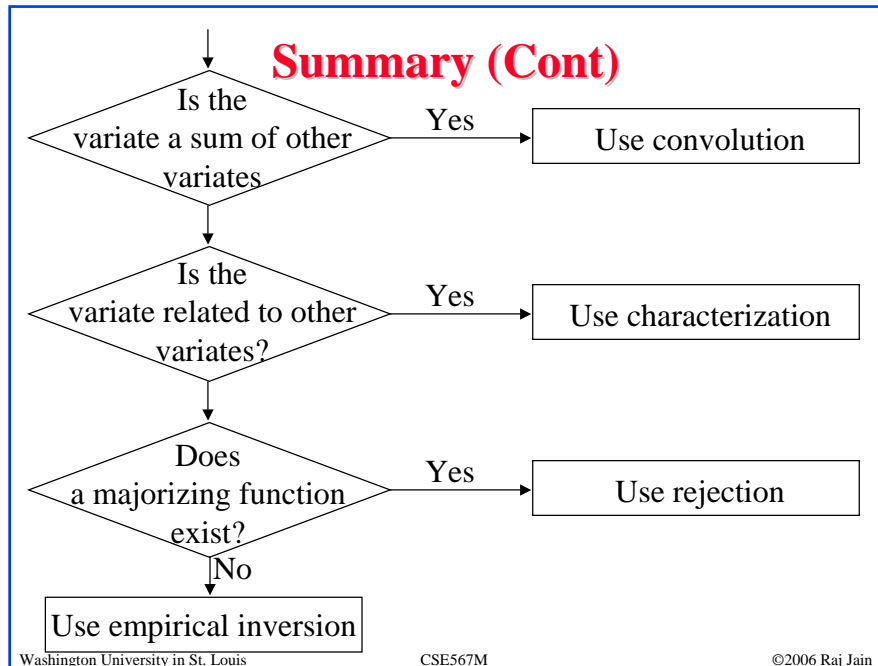
- ❑ Erlang- $k = \sum_{i=1}^k$ Exponential _{i}
- ❑ Binomial(n, p) = $\sum_{i=1}^n$ Bernoulli(p)
 \Rightarrow Generated n $U(0,1)$,
return the number of RNs less than p
- ❑ $\chi^2(v) = \sum_{i=1}^v N(0,1)^2$
- ❑ $\Gamma(a, b_1) + \Gamma(a, b_2) = \Gamma(a, b_1 + b_2)$
 \Rightarrow Non-integer value of $b =$ integer + fraction
- ❑ $\sum_{i=1}^n$ Any = Normal $\Rightarrow \sum U(0,1) =$ Normal
- ❑ $\sum_{i=1}^m$ Geometric = Pascal
- ❑ $\sum_{i=1}^2$ Uniform = Triangular

Characterization

- Use special characteristics of distributions \Rightarrow **characterization**
- Exponential inter-arrival times \Rightarrow Poisson number of arrivals
 \Rightarrow Continuously generate exponential variates until their sum exceeds T and return the number of variates generated as the Poisson variate.
- The a^{th} smallest number in a sequence of $a+b+1$ $U(0,1)$ uniform variates has a $\beta(a, b)$ distribution.
- The ratio of two unit normal variates is a Cauchy(0, 1) variate.
- A chi-square variate with even degrees of freedom $\chi^2(\nu)$ is the same as a gamma variate $\gamma(2, \nu/2)$.
- If x_1 and x_2 are two gamma variates $\gamma(a, b)$ and $\gamma(a, c)$, respectively, the ratio $x_1/(x_1+x_2)$ is a beta variate $\beta(b, c)$.
- If x is a unit normal variate, $e^{\mu+\sigma x}$ is a lognormal(μ, σ) variate.

Summary





28-19

Exercise 28.1

- A random variate has the following triangular density:

$$f(x) = \min(x, 2 - x) \quad 0 \leq x \leq 2$$
- Develop algorithms to generate this variate using each of the following methods:
 - a. Inverse-transformation
 - b. Rejection
 - c. Composition
 - d. Convolution

Washington University in St. Louis CSE567M ©2006 Raj Jain

28-20

Homework

- A random variate has the following triangular density:

$$f(x) = \min(x, 8 - x) \quad 0 \leq x \leq 8$$

- Develop algorithms to generate this variate using each of the following methods:
 - a. Inverse-transformation
 - b. Rejection
 - c. Composition
 - d. Convolution