Systems Simulation
Chapter 8: Random-Variate Generation

Fatih Cavdur
fatihcavdur@uludag.edu.tr

May 15, 2014

Introduction

- This chapter deals with procedures for sampling from a variety of widely-used continuous and discrete distributions.
- The purpose of the chapter is to explain and illustrate some widely-used techniques for generating random variates.
- The techniques mentioned here are the inverse-transform technique, the acceptance-rejection technique.

Introduction-cont.

Assumption

- We assume that we have $U[0,1]$ RVs $R_1, R_2, \ldots$ where

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

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

Inverse-Transform (IT) Technique

- The IT technique can be used to sample from the exponential, the uniform, the Weibull, the triangular distributions and from empirical distributions.
- It is also the underlying principle for sampling from a wide variety of discrete distributions.
- We will explain it for the exponential distribution and then apply to the others.
- Computationally, it is the most straightforward technique, but not always the most efficient.

IT for the Exponential Distribution

The PDF and CDF of the exponential RV X are

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$F(X) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

λ can be interpreted as the mean number of occurrences per time unit, and the mean of X is $E(X_i) = 1/\lambda$.
The IT can be utilized in principle for any distribution, but it is most useful when the inverse of the CDF $F(X)$, F^{-1} is easily computed.

IT for the Exponential Distribution-cont.

Step (1) Compute the CDF of the RV X.

For the ED, it is $F(X) = 1 - e^{-\lambda x}, x \geq 0$.

Step (2) Set $F(X) = R$ on the range of X.

For the ED, it is $1 - e^{-\lambda X} = R$ on the range $x \geq 0$.

Step (3) Solve the equation $F(X) = R$ for X in terms of R.

$$\begin{aligned} 1 - e^{-\lambda X} &= R \\ e^{-\lambda X} &= 1 - R \\ -\lambda X &= \ln(1 - R) \\ X &= -\frac{1}{\lambda} \ln(1 - R) \rightarrow X = F^{-1}(R) \end{aligned}$$

IT for the Exponential Distribution-cont.

Step (4) Generate RNs $R_1, R_2, \ldots$ and compute the RVs using $X_i = F^{-1}(R_i)$.

For the ED, it is

$$X = F^{-1}(R) = -\frac{1}{\lambda}\ln(1-R) \Rightarrow X_i = -\frac{1}{\lambda}\ln(1-R_i)$$

Since both R_i and $1 - R_i$ are uniform, we can write

$$X_i = -\frac{1}{\lambda}\ln(R_i)$$

Exponential Distribution-cont.

Example

Table : Generation of ED-RVs with Mean 1

i	1	2	3	4	5
R_i	0.1306	0.0422	0.6597	0.7965	0.7696
X_i	0.1400	0.0431	1.0780	1.5920	1.4680

IT for the Exponential Distribution-cont.

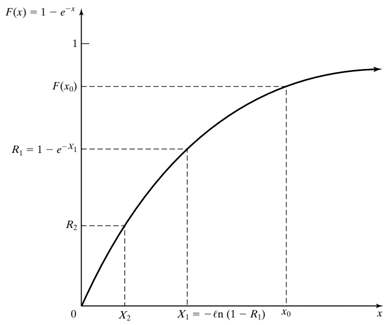

Figure : Graphical View of the IT Technique

Building on Exponential Distribution

$$Y = 2\sum_{i=1}^{v} X_i \sim \text{ C-S } (2v)$$

$$Y = \beta\sum_{i=1}^{\alpha} X_i \sim \text{ gamma } (\alpha, \beta)$$

$$Y = \frac{\sum_{i=1}^{a} X_i}{\sum_{i=1}^{a+b} X_i} \sim \text{ beta } (a, b)$$

Uniform Distribution

Step (1) The CDF is given by

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

$$F(X) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & x > b \end{cases}$$

Step (2) Set $F(X) = (X - a)/(b - a) = R$

Step (3) Solve for X in terms of R to obtain $X = a + (b - a)R$

Weibull Distribution

Step (1) The CDF is given, when $v = 0$, by

$$f(x) = \begin{cases} \frac{\beta}{\alpha^\beta} x^{\beta-1} e^{-(x/\alpha)^\beta}, & x \ge 0 \\ 0, & \text{otherwise} \end{cases}$$

$$F(X) = 1 - e^{-(x/\alpha)^\beta}, x \ge 0$$

Step (2) Set $F(X) = 1 - e^{-(x/\alpha)^\beta} = R$

Step (3) Solve for X in terms of R to obtain
$X = \alpha[-\ln(1 - R)]^{1/\beta}$

IT Example for the Triangular Distribution

Triangular Distribution (with end points (0, 2) and mode 1

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

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

IT Example for the Triangular Distribution-cont.

Triangular Distribution (with end points (0, 2) and mode 1)

For $0 \leq X \leq 1$,

$$R = \frac{X^2}{2}$$

and for $1 \leq X \leq 2$,

$$R = 1 - \frac{(2-x)^2}{2}$$

Thus,

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

IT Example for Empirical Continuous Distributions

We have the following data: 2.76, 1.83, 1.80, 1.45, 1.24 The data are arranged from smallest to largest. The smallest possible value is assumed to be 0, so we define $x_{(0)} = 0$. Each interval has equal probability of $1/n = 1/5$. The slope of the ith line segment is

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

The inverse CDF, when $(i-1)/n < R < i/n$, is given by

$$X = \widehat{F}^{-1}(R) = x_{(i-1)} + a_i \left(R - \frac{i-1}{n} \right)$$

IT Example for Empirical Continuous Distributions-cont.

For example, for $R_1 = 0.71$, we have

$$X_1 = x_{(4-1)} + a_4 \left(R_1 - \frac{4-1}{n} \right) = 1.45 + 1.90(0.71 - 0.60) = 1.66$$

i	Interval $x_{(i-1)} < x < x_{(i)}$	Probability $1/n$	C.P. i/n	Slope a_i
1	$0.00 < x \leq 0.80$	0.2	0.2	4.00
2	$0.80 < x \leq 1.24$	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

IT Example for Empirical Continuous Distributions-cont.

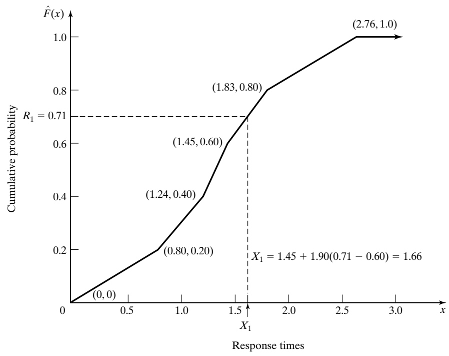

Figure : CDF for the Example

Continuous Distributions without Closed-Form Inverse

- Some distributions do not have a closed form expressions for their CDF or its inverse, such as normal, gamma and beta distributions.
- If we are willing to approximate the inverse CDF, or numerically integrate, we can use the IT method for RV generation.
- A simple approximation, for instance, to the inverse CDF of the normal distribution is proposed by Schmeiser (1979).

$$X = F^{-1}(R) \approx \frac{R^{0.135} - (1-R)^{0.135}}{0.1975}$$

Normal Approximation

$$\Phi(x) \approx 1 - \phi(x)[b_1 t + b_2 t^2 + b_3 t^4 + b_4 t^4 + b_5 t^5], \; x > 0$$

where

$$t = (1 + px)^{-1}$$

and

$$p = 0.2316419, b_1 = 0.31938, b_2 = -0.35656$$

$$b_3 = 1.78148, b_4 = -1.82125, b_5 = 1.33027$$

Discrete Distributions

An Empirical Discrete Distribution Example

The PMF and CDF are given as follows:

$$\begin{aligned} p(0) &= P(X = 0) = 0.50 \\ p(1) &= P(X = 1) = 0.30 \\ p(2) &= P(X = 2) = 0.20 \end{aligned}$$

$$F(X) = \begin{cases} 0.0, & x \le 0 \\ 0.5, & 0 \le x < 1 \\ 0.8, & 1 \le x < 2 \\ 1.0, & x \ge 2 \end{cases}$$

Discrete Distributions

An Empirical Discrete Distribution Example

For generating discrete RVs, the IT technique becomes a table-lookup procedure in this example. For $R = R_1$, if

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

then, set $X_1 = x_i$. We have the following generation scheme here:

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

Discrete Distributions

An Empirical Discrete Distribution Example

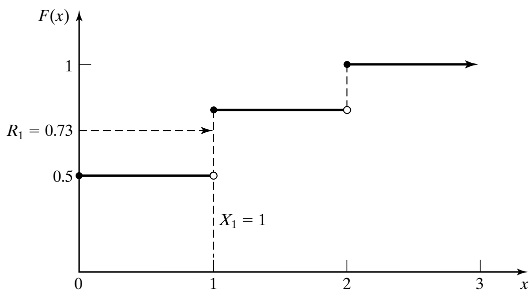

Figure : CDF for the Example

Discrete Distributions

Discrete Uniform Distribution Example

The PMF and CDF are given as

$$p(x) = \frac{1}{k}, x = 1, 2, \ldots, k$$

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{k}, & 1 \leq x < 2 \\ \frac{2}{k}, & 2 \leq x < 3 \\ \vdots, & \vdots \\ \frac{k-1}{k}, & k-1 \leq x < k \\ 1, & k \leq x \end{cases}$$

Discrete Distributions

Discrete Uniform Distribution Example

Using $F(x_{i-1}) = r_{i-1} < R \leq r_i = F(X_i)$, we have the following.

$$r_{i-1} = \frac{i-1}{k} < R \leq r_i = \frac{i}{k}$$

Solving it for i

$$i - 1 < Rk < i \Rightarrow Rk \leq i < Rk + 1$$

From the above inequality, we obtain

$$X = \lceil Rk \rceil$$

Acceptance-Rejection Technique

Uniform Distribution

Step (1) Generate a RN R.

Step (2) -

(a) If $R \geq 1/4$, accept $X = R$, then, go to Step 3.

(b) If $R < 1/4$, reject R, then, go to Step 1.

-

Step (3) If another RV needed, go to Step 1, otherwise stop.

Acceptance-Rejection Technique

Poisson Distribution

The PMF of a Poisson RV is

$$p(n) = P(N = n) = \frac{e^{-\alpha}\alpha^n}{n!}, n = 0, 1, 2, \ldots$$

We can write

$$N = n \Leftrightarrow A_1 + A_2 + \ldots + A_n \leq 1 < A_1 + A_2 + \ldots + A_n + A_{n+1}$$

Now, we let $A_i = (-1/\alpha) \ln R_i$ (the IT method for the ED)

Acceptance-Rejection Technique

Poisson Distribution

Using the inequality in the previous slide, we obtain

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

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

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

Acceptance-Rejection Technique

Poisson RV Generation Procedure

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

Step (2) Generate a RN R_{n+1}, and replace P by PR_{n+1}.

Step (3) If $P < e^{-\alpha}$, then, accept $N = n$; otherwise reject n, increase n by once, and go to step 2.

Acceptance-Rejection Technique

Non-Stationary Poisson RV Generation Procedure

Step (1) Let $\lambda^* = \max_{0 \leq t \leq T} \lambda(t)$ be the max of the arrival rate function and set $t = 0$ and $i = 1$.

Step (2) Generate E from the exponential distribution with rate λ^* and let $t = t + E$ (arrival time of the stationary Poisson process)

Step (3) Generate RN R. If $R \leq \lambda(t)/\lambda^*$, then, let $\tau_i = t$ and $i = i + 1$.

Step (4) Go to step 2.

Acceptance-Rejection Technique

Gamma RV Generation Procedure

Step (1) Compute $a = 1/(2/\beta - 1)^{1/2}, b = \beta - \ln 4$

Step (2) Generate R_1 and R_2. Set $V = R_1/(1 - R_1)$

Step (3) Compute $X = \beta V^a$.

Step (4) -

(a) If $X > b + (\beta a + 1)\ln(V) - \ln(R_1^2 R_2)$, reject X and return to step 2.

(b) If $X \leq b + (\beta a + 1)\ln(V) - \ln(R_1^2 R_2)$, accept X.

-

Step (5) X has mean and variance both equal to β. If it is desired to have mean $1/\theta$ and variance $1/\beta\theta^2$, replace X by $X/(\beta\theta)$.

Special Properties

Direct Transformation for the Normal and Lognormal Distributions

Consider two standard normal RVs, Z_1 and Z_2, plotted as a point in the plane as shown in the figure on the next slide, and

$$\begin{aligned} Z_1 &= B\cos\theta \\ Z_2 &= B\sin\theta \end{aligned}$$

It is known that $B^2 = Z_1^2 + Z_2^2$ has a chi-square distribution with 2 degrees of freedom, which is equivalent to an ED with mean 2. So, we can write, $B = (-2\ln R)^{1/2}$, and hence,

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

Special Properties

Direct Transformation for the Normal and Lognormal Distributions

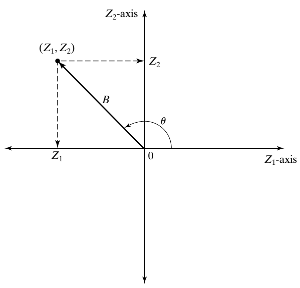

Figure : Polar Representation of a Pair of Std. Nor. Variables

Special Properties

Convolution Method-Erlang Distribution

An Erlang RV X with parameters (k, θ) can be shown to be the sum of k independent exponential RVs, $X_i, i = 1, \ldots, k$ each with mean $1/k\theta$. The convolution approach is to generate $X_1, \ldots, X_k$, then, sum them to get X. Therefore,

$$\begin{aligned} X &= \sum_{i=1}^{k} -\frac{1}{k\theta} \ln R_i \\ &= -\frac{1}{k\theta} \ln \left(\prod_{i=1}^{k} R_i \right) \end{aligned}$$

Special Properties

More Special Properties-Beta Distribution

Assume that $X_1 \sim G(\beta_1, \theta_1 = 1/\beta_1)$ and $X_2 \sim G(\beta_2, \theta_2 = 1/\beta_2)$, and X_1 and X_2 are independent. Then,

$$Y = \frac{X_1}{X_1 + X_2}$$

has a beta distribution with β_1 and β_2 on the interval $(0, 1)$. If we want Y to be defined on (a, b), then,

$$Y = a + (b - a) \left(\frac{X_1}{X_1 + X_2} \right)$$

Summary

- Reading HW: Chapter 8.
- Chapter 8 Exercises.