Random Variables Generation

Revised version of the slides based on the book
Discrete-Event Simulation: a first course
L.L. Leemis & S.K. Park

Section(s) 6.1, 6.2, 7.1, 7.2

 0-13-142917-5

Introduction

- Monte Carlo Simulators differ from Trace Driven simulators because of the use of Random Number Generators to represent the variability that affects the behaviors of real systems.
- Uniformly distributed random variables are the most elementary representations that we can use in Monte Carlo simulation, but are not enough to capture the complexity of real systems.
- We must thus devise methods for generating instances (*variates*) of arbitrary random variables
- Properly using uniform random numbers, it is possible to obtain this result.
- In the sequel we will first recall some basic properties of *Discrete* and *Continuous* random variables and then we will discuss several methods to obtain their variates

Basic Probability Concepts

- *Empirical Probability*, derives from performing an experiment many times n and counting the number of occurrences n_a of an event $\mathcal{A}$
  - The *relative frequency* of occurrence of event $\mathcal{A}$ is n_a/n
  - The *frequency theory of probability* asserts that the relative frequency converges as $n \to \infty$

$$\Pr(\mathcal{A}) = \lim_{n \to \infty} \frac{n_a}{n}$$

- *Axiomatic Probability* is a formal, set-theoretic approach
  - Mathematically construct the sample space and calculate the number of events $\mathcal{A}$

- The two are complementary!

Example

- Roll two dice and observe the up faces

$$\begin{array}{cccccc}
(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6)
\end{array}$$

- If the two up faces are summed, an integer-valued random variable, say X, is defined with possible values 2 through 12

$$\begin{array}{lccccccccccc}
\text{sum, } x: & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\
\Pr(X = x): & \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{6}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} & \frac{2}{36} & \frac{1}{36}
\end{array}$$

- $\Pr(X = 7)$ could be estimated by replicating the experiment many times and calculating the relative frequency of occurrence of 7's

Discrete Random Variables

- A random variable X is *discrete* if and only if its set of possible values $\mathcal{X}$ is finite or, at most, countably infinite
- A discrete random variable X is uniquely determined by
  - Its set of possible values $\mathcal{X}$
  - Its *probability distribution function* (pdf):
    A real-valued function $f(\cdot)$ defined for each $x \in \mathcal{X}$ as the probability that X has the value x

    $$f(x) = Pr(X = x)$$

    By definition,

    $$\sum_{x} f(x) = 1$$

Examples

- **Example 1** X is *Equilikely* (Discrete Uniform dist.) (a, b)
  $|\mathcal{X}| = b - a + 1$ and each possible value is equally likely

$$f(x) = \frac{1}{b - a + 1} \qquad x = a, a + 1, \ldots, b$$

- **Example 2** Roll two fair dice
  If X is the sum of the two up faces, $\mathcal{X} = \{x | x = 2, 3, \ldots, 12\}$
  From example 2.3.1,

$$f(x) = \frac{6 - |7 - x|}{36} \qquad x = 2, 3, \ldots, 12$$

or alternatively

$$f(x) = \begin{cases} \dfrac{x-1}{36} & 2 \le x \le= 7 \\ \dfrac{13-x}{36} & 8 \le x \le= 12 \end{cases}$$

- **Example 3** Toss a coin *until the first* tail occurs
- Assume the coin has p as its probability of a head
- If X is the number of heads, $\mathcal{X} = \{x | x = 0, 1, 2, ...\}$ and the pdf is

$$f(x) = p^x(1-p) \qquad x = 0, 1, 2, ...$$

- X is *Geometric*(p) and the set of possible values is *infinite*
- Verify that $\sum_x f(x) = 1$:

$$\sum_x f(x) = \sum_{x=0}^{\infty} p^x(1-p) = (1-p)(1+p+p^2+p^3+p^4+\cdots) = 1$$

Cumulative Distribution Function

- The *cumulative distribution function(cdf)* of X is the real-valued function $F(\cdot)$ for each $x \in \mathcal{X}$ as

$$F(x) = Pr(X \leq x) = \sum_{t \leq x} f(t)$$

- If X is *Equilikely*(a, b) then the cdf is

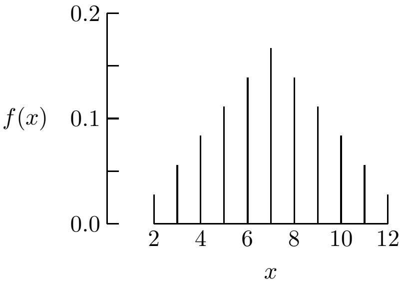
$$F(x) = \sum_{t=a}^{x} 1/(b - a + 1) = (x-a+1)/(b-a+1) \qquad x = a, a+1, \ldots, b$$

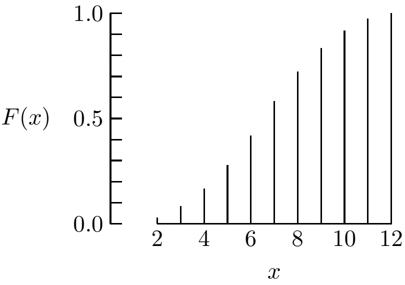
- If X is *Geometric*(p) then the cdf is

$$F(x) = \sum_{t=0}^{x} p^t(1-p) = (1-p)(1+p+\cdots+p^x) = 1-p^{x+1} \qquad x = 0, 1, 2, \ldots$$

Example

- No simple equation for $F(\cdot)$ for sum of two dice
- $|\mathcal{X}|$ is small enough to tabulate the cdf





Relationship Between cdfs and pdfs

- A cdf can be generated from its corresponding pdf by recursion

  For example, $\mathcal{X} = \{x | x = a, a+1, ..., b\}$

$$\begin{aligned} F(a) &= f(a) \\ F(x) &= F(x-1) + f(x) \qquad x = a+1, a+2, ..., b \end{aligned}$$

- A pdf can be generated from its corresponding cdf by subtraction

$$\begin{aligned} f(a) &= F(a) \\ f(x) &= F(x) - F(x-1) \qquad x = a+1, a+2, ..., b \end{aligned}$$

- A discrete random variable can be defined by specifying *either* its pdf or its cdf

Other cdf Properties

- A cdf is strictly monotone increasing:
  if $x_1 < x_2$, then $F(x_1) < F(x_2)$
- The cdf values are bounded between 0.0 and 1.0
- Monotonicity of $F(\cdot)$ is the basis to generate discrete random variates as we will see soon

Mean and Standard Deviation

- The *mean* μ of the discrete random variable X is

$$\mu = \sum_x x f(x)$$

- The variance of X is

$$\sigma^2 = \sum_x (x-\mu)^2 f(x) \qquad \text{or} \qquad \sigma^2 = \left(\sum_x x^2 f(x)\right) - \mu^2$$

- The corresponding *standard deviation* σ is

$$\sigma = \sqrt{\sum_x (x-\mu)^2 f(x)} \qquad \text{or} \qquad \sigma = \sqrt{\left(\sum_x x^2 f(x)\right) - \mu^2}$$

Examples

If X is *Equilikely*(a, b) then the mean and standard deviation are

- $$\begin{aligned}\mu &= \sum_{x=a}^{b} \frac{x}{b-a+1} = \frac{1}{b-a+1}\left[\sum_{x=1}^{b} x - \sum_{x=1}^{a-1} x\right] \\ &= \frac{1}{b-a+1}\left[\frac{b(b+1)}{2} - \frac{(a-1)a}{2}\right] = \frac{b^2-b-a^2+a}{2(b-a+1)} = \frac{a+b}{2}\end{aligned}$$

- $$\begin{aligned}\sigma^2 &= \sum_{x=a}^{b} \frac{(x-\mu)^2}{b-a+1} = \sum_{x=a}^{b} \frac{x^2}{b-a+1} - \mu^2 = \frac{1}{b-a+1}\left[\sum_{x=1}^{b} x^2 - \sum_{x=1}^{a-1} x^2\right] - \mu^2 \\ &= \frac{1}{b-a+1}\left[\frac{b(b+1)(2b+1)}{6} - \frac{(a-1)a(2a-1)}{6}\right] - \mu^2 \\ &= \frac{(b-a+1)^2-1}{12} \quad (^*)\end{aligned}$$

When X is *Equilikely*$(1, 6)$, $\mu = 3.5$ and $\sigma = \sqrt{\frac{35}{12}} \cong 1.708$

(*)recall that the sum of the squares of the first n natural numbers is

$$S_n^2 = \frac{n(n+1)(2n+1)}{6}$$

Roll two fair dice: average

If X is the sum of two dice then

$$
\begin{aligned}
\mu &= \sum_{x=2}^{12} x f(x) = \sum_{x=2}^{12} x \ \frac{6-|7-x|}{36} \\
&= \sum_{x=2}^{6} [x + (14-x)] \ \frac{6-7+x}{36} + 7 \ \frac{6-|7+7|}{36} \\
&= \sum_{x=2}^{6} 14 \ \frac{x-1}{36} + \frac{7}{6} = \frac{7}{18} \sum_{x=2}^{6} (x-1) + \frac{7}{6} \\
&= \frac{7}{18} \sum_{k=0}^{4} (2+k-1) + \frac{7}{6} = \frac{7}{18} \sum_{k=0}^{4} (1+k) + \frac{7}{6} \\
&= \frac{7}{18} \left[5 + \frac{4*(4+1)}{2} \right] + \frac{7}{6} = \frac{7}{18} 10 + \frac{21}{18} = \\
&= \frac{70}{18} + \frac{21}{18} = \frac{7*18}{18} = 7
\end{aligned}
$$

Roll two fair dice: alternative derivations

$$\begin{aligned}
\mu &= \sum_{x=2}^{12} x f(x) = \sum_{x=2}^{7} x \, \frac{x-1}{36} + \sum_{x=8}^{12} x \, \frac{13-x}{36} \\
&= \sum_{x=2}^{6} x \, \frac{x-1}{36} + 7\frac{6}{36} + \sum_{y=6}^{2} (14-y) \, \frac{y-1}{36}
\end{aligned}$$

where $y = 14 - x$. With this substitution it is easy to see that

$$\begin{aligned}
\mu &= \sum_{x=2}^{6} [x + (14-x)] \, \frac{x-1}{36} + \frac{7}{6} &= \sum_{x=2}^{6} 14 \, \frac{x-1}{36} + \frac{7}{6} \\
&= \frac{7}{18} \sum_{x=2}^{6} (x-1) + \frac{7}{6} &= \frac{7}{18} \sum_{k=0}^{4} (2+k-1) + \frac{7}{6} \\
&= \frac{7}{18} \sum_{k=0}^{4} (1+k) + \frac{7}{6} &= \frac{7}{18} \left[5 + \frac{4*(4+1)}{2} \right] + \frac{7}{6} \\
&= \frac{7}{18} 10 + \frac{21}{18} = \frac{70}{18} + \frac{21}{18} &= \frac{7*18}{18} = 7
\end{aligned}$$

Roll two fair dice: variance

If X is the sum of two dice then

$$\sigma^2 = \sum_{x=2}^{12} (x - \mu)^2 f(x) = 35/6 = 5.81\overline{3}$$

Standard Deviation

$$\sigma = \sqrt{35/6} \cong 2.415$$

Another Example

- If X is *Geometric*(p) then the mean and standard deviation are

$$
\begin{aligned}
\mu &= \sum_{x=0}^{\infty} x f(x) = \sum_{x=1}^{\infty} x p^x (1-p) = \cdots = \frac{p}{1-p} \\
\sigma^2 &= \left(\sum_{x=0}^{\infty} x^2 f(x) \right) - \mu^2 = \left(\sum_{x=1}^{\infty} x^2 p^x (1-p) \right) - \frac{p^2}{(1-p)^2} \\
&\vdots \\
\sigma^2 &= \frac{p}{(1-p)^2} \\
\sigma &= \frac{\sqrt{p}}{(1-p)}
\end{aligned}
$$

Expected Value

- The mean of a random variable is also known as the *expected value*
- The expected value of the discrete random variable X is

$$E[X] = \sum_x x f(x) = \mu$$

- Expected value refers to the *expected average* of a large sample $x_1, x_2, \ldots, x_n$ corresponding to X: $\bar{x} \to E[X] = \mu$ as $n \to \infty$.
- The *most likely* value x (with largest $f(x)$) is the *mode*, which can be different from the expected value

Example

- Toss a fair coin until the first tail appears
  - $Pr\{T\} = 1/2$
  - $Pr\{HT\} = 1/4$
  - $Pr\{HHT\} = 1/8$
  - ...
- The most likely number of heads is 0
- The expected number of heads is 1
- 0 occurs with probability 1/2 and
  1 occurs with probability 1/4

  The most likely value is twice as likely as the expected value

- For some random variables, the mean and mode may be the same

  For the sum of two dice, the most likely value and expected value are both 7

More on Expectation

- Define function $h(\cdot)$ for all possible values of X
  $h(\cdot) : \mathcal{X} \rightarrow \mathcal{Y}$
- $Y = h(X)$ is a *new* random variable, with possible values $\mathcal{Y}$
- The expected value of Y is

$$E[Y] = E[h(X)] = \sum_{x} h(x)f(x)$$

  Note: in general, this is *not* equal to $h(E[X])$

Examples

- If $y = (x-\mu)^2$ with $\mu = E[X]$,

$$\begin{aligned} E[Y] &= E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x) \\ &= \sum_x x^2 f(x) - \sum_x 2x\mu f(x) + \sum_x \mu^2 f(x) \\ &= E[X^2] - 2\mu E[X] + \mu^2 = E[X^2] - E[X]^2 = \sigma^2 \end{aligned}$$

- If $y = x^2 - \mu^2$,

$$E[Y] = E[X^2 - \mu^2] = \sum_x (x^2 - \mu^2) f(x) = \left(\sum_x x^2 f(x)\right) - \mu^2 = \sigma^2$$

- $E[X^2] \geq E[X]^2$ with equality if and only if X is not really random

Examples (cont)

- If $Y = aX + b$ for constants a and b,

$$E[Y] = E[aX+b] = \sum_x (ax+b)f(x) = a\left(\sum_x xf(x)\right)+b = aE[X]+b$$

- Suppose
  - X is the number of heads before the first tail
  - Win \$2 for every head and let Y be the amount you win
- The possible values Y you win are defined by

$$y = h(x) = 2x \qquad x = 0, 1, 2, \ldots$$

- Your *expected winnings* are

$$E[Y] = E[2X] = 2E[X] = 2$$

Continuous Random Variables

- A random variable X is *continuous* if and only if its set of possible values $\mathcal{X}$ is a *continuum*
- A continuous random variable X is uniquely determined by
  - Its set of possible values $\mathcal{X}$
  - Its *probability density function* (pdf):
    A real-valued function $f(\cdot)$ defined for each $x \in \mathcal{X}$

    $$\int_a^b f(x)dx = \Pr(a \leq X \leq b)$$

    By definition,

    $$\int_{\mathcal{X}} f(x)dx = 1$$

  - In practice it is common to assume that $\mathcal{X}$ is an open interval (a, b) where a may be $-\infty$ and b may be $+\infty$

Example

- X is $Uniform(a, b)$

  $\mathcal{X} = (a, b)$ and all values in this interval are equally likely

$$f(x) = \frac{1}{b - a} \qquad a < x < b$$

- In the continuous case,
  - $\Pr(X = x) = 0$ for *any* $x \in \mathcal{X}$
  - If $[a, b] \subseteq \mathcal{X}$,

$$\begin{aligned} \int_a^b f(x)dx &= \Pr(a \leq X \leq b) = \Pr(a < X \leq b) \\ &= \Pr(a \leq X < b) = \Pr(a < X < b) \end{aligned}$$

Cumulative Distribution Function

- The *cumulative distribution function(cdf)* of the continuous random variable X is the real-valued function $F(\cdot)$ for each $x \in \mathcal{X}$ as

$$F(x) = \Pr(X \leq x) = \int_{t \leq x} f(t)dt$$

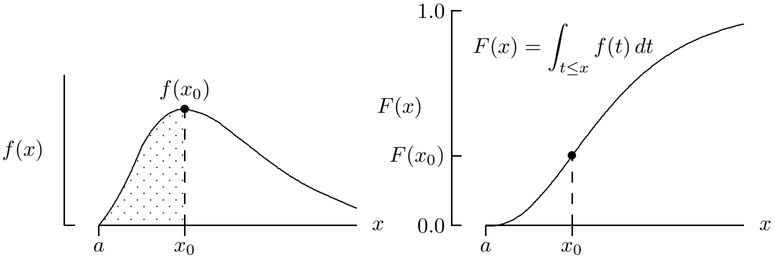
- **Example:** If X is $\mathit{Uniform}(a, b)$, the cdf is

$$F(x) = \int_{t=a}^{x} \frac{1}{(b-a)} dt = \frac{x-a}{b-a} \qquad a < x < b$$

- In special case where U is $\mathit{Uniform}(0, 1)$, the cdf is

$$F(u) = \Pr(U \leq u) = u \qquad 0 \leq u \leq 1$$

Relationship between pdfs and cdfs



- Shaded area in pdf graph equals $F(x_0)$

More on cdfs

- The cdf is strictly monotone increasing:
  if $x_1 < x_2$, then $F(x_1) < F(x_2)$
- The cdf is bounded between 0.0 and 1.0
- The cdf can be obtained from the pdf by integration
  The pdf can be obtained from the cdf by differentiation as

$$f(x) = \frac{d}{dx}F(x) \qquad x \in \mathcal{X}$$

- A continuous random variable model can be specified by $\mathcal{X}$ and either the pdf or the cdf

Example: *Exponential*(η)

- $X = -\eta \ln(1 - U)$ where U is *Uniform*$(0, 1)$
- The cdf of X is

$$
\begin{aligned}
F(x) = \Pr(X \leq x) &= \Pr(-\eta \ln(1 - U) \leq x) \\
&= \Pr(1 - U \geq \exp(-x/\eta)) \\
&= \Pr(U \leq 1 - \exp(-x/\eta)) \\
&= 1 - \exp(-x/\eta)
\end{aligned}
$$

- The pdf of X is

$$
f(x) = \frac{d}{dx} F(x) = \frac{d}{dx}(1 - \exp(-x/\eta)) = \frac{1}{\eta} \exp(-x/\eta) \qquad x > 0
$$

- Often the Exponential distribution is defined in terms of a parameter $\lambda = 1/\eta$ which is called the *rate* of the distribution.

Mean and Standard Deviation

- The *mean* μ of the continuous random variable X is

$$\mu = \int_x xf(x)dx$$

- The corresponding *variance* σ^2 is

$$\sigma^2 = \int_x (x-\mu)^2 f(x)dx \qquad \text{or} \qquad \sigma^2 = \left(\int_x x^2 f(x)dx\right) - \mu^2$$

- The *standard deviation* σ is

$$\sigma = \sqrt{\int_x (x-\mu)^2 f(x)dx} \qquad \text{or} \qquad \sigma = \sqrt{\left(\int_x x^2 f(x)dx\right) - \mu^2}$$

Examples: Uniform distribution

Uniform(0,1)

- $$f(x) = 1$$
- $$\mu = \int_{x=0}^{1} x\, dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}$$
- σ^2 $$= \int_{x=0}^{1} (x-\mu)^2 dx = \frac{x^3}{3}\Big|_0^1 - \mu^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

Uniform(a,b)

- $$f(x) = \frac{1}{b-a}$$
- $$\mu = \int_{x=a}^{b} \frac{x}{b-a} dx = \frac{x^2}{2(b-a)}\Big|_a^b = \frac{a+b}{2}$$
- $$\sigma^2 = \int_{x=a}^{b} \frac{(x-\mu)^2}{b-a} dx = \frac{x^3}{3(b-a)}\Big|_a^b - \mu^2 = \frac{b^3-a^3}{3(b-a)} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$

Examples: Exponential

- If X is *Exponential*(η),

$$\mu = \int_x x f(x) dx = \int_0^\infty \frac{x}{\eta} \exp(-x/\eta) dx = \eta \int_0^\infty t \exp(-t) dt = \cdots = \eta$$

$$\sigma^2 = \left(\int_0^\infty \frac{x^2}{\mu} \exp(-x/\eta) dx \right) - \eta^2 = \cdots = \eta^2$$

Expected Value

- The mean of a continuous random variable is also known as the *expected value*
- The expected value of the continuous random variable X is

$$\mu = E[X] = \int_x x f(x) dx$$

- The variance is the expected value of $(X - \mu)^2$

$$\sigma^2 = E[(X - \mu)^2] = \int_x (x - \mu)^2 f(x) dx$$

- In general, if $Y = g(X)$, the expected value of Y is

$$E[Y] = E[g(X)] = \int_x g(x) f(x) dx$$

Example

- If continuous random variable $Y = aX + b$ for constants a and b,

$$E[Y] = E[aX + b] = aE[X] + b$$

$$VAR[Y] = VAR[aX + b] = a^2 VAR[X]$$

Scaling and Shifting

- Suppose X is a random variable with mean μ and standard deviation σ
- Define random variable $X' = aX + b$ for constants a, b
- The mean μ' and standard deviation σ' of X' are

$$\mu' = E[X'] = E[aX + b] = aE[X] + b = a\mu + b$$

$$(\sigma')^2 = E[(X' - \mu')^2] = E[(aX - a\mu)^2] = a^2 E[(X - \mu)^2] = a^2 \sigma^2$$

Therefore,

$$\mu' = a\mu + b \qquad \text{and} \qquad \sigma' = |a|\sigma$$

Example

- Suppose Z is a random variable with mean 0 and standard deviation 1
- Construct a new random variable X with *specified* mean μ and standard deviation σ
- Define $X = \sigma Z + \mu$
- $E[X] = \sigma E[Z] + \mu = \mu$
- $E[(X - \mu)^2] = E[\sigma^2 Z^2] = \sigma^2 E[Z^2] = \sigma^2$

Coefficient of variation

- A measure of the dispersion of the values of a random variable X with respect to their mean is the *Coefficient of Variation* defined in the following way

$$CV[X] = \frac{\sigma}{\mu}$$

- An alternative definition is

$$CV^2[X] = \frac{VAR[X]}{(E[X])^2} = \frac{\sigma^2}{\mu^2}$$

- Random variables with Coefficients of Variation larger than 1 are considered to be affected by large *variability*
- Exponentially distributed random variables have $CV = CV^2 = 1$

Generating Random Variates

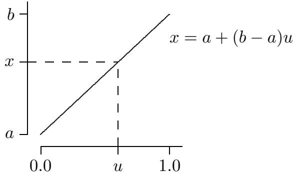
Random Variates

- A *Random Variate* is an algorithmically generated realization of a random variable
- $u = \texttt{Random}()$ generates a *Uniform*$(0, 1)$ random variate
- The generation of *Uniform* random variates (Random Numbers for short) is the basis for generating variates (instances) of arbitrary Discrete and/or Continuous random variables.

Preliminary to the discussion of the general methods for the generation of variates of arbitrary distributions is the introduction of the simple transformations which allow to generate instances of *Uniform*(a, b) and *Equilikely*(a, b)

Uniform Random Variates

We can generate a *Uniform*(a, b) variate from a *Uniform*$(0, 1)$ variate by simply applying a **Scaling** and **Shifting** transformation
Scaling coefficient $\alpha = (b - a)$; Shifting coefficient $\beta = a$



Generating a Uniform Random Variate

```
double Uniform(double a, double b)     /* use a < b */ {
    return (a + (b - a) * Random());
}
```

	Average μ	Variance σ^2
Uniform (0,1)	$1/2$	$1/12$
Uniform (a,b)	$(a + b)/2$	$(b - a)^2/12$

Equilikely Random Variates

- *Uniform*$(0, 1)$ random variates can also be used to generate an *Equilikely*(a, b) random variate

$$
\begin{aligned}
0 < u < 1 \quad & \iff \quad 0 < (b - a + 1)u < b - a + 1 \\
& \iff \quad 0 \leq \lfloor (b - a + 1)u \rfloor \leq b - a \\
& \iff \quad a \leq a + \lfloor (b - a + 1)u \rfloor \leq b \\
& \iff \quad a \leq x \leq b
\end{aligned}
$$

- Specifically, $x = a + \lfloor (b - a + 1)\, u \rfloor$

Generating an Equilikely Random Variate

```
long Equilikely(long a, long b)      /* use a < b */ {
    return (a + (long) ((b - a + 1) * Random()));
}
```

Examples

- **Example 1** To generate a random variate x that simulates rolling two fair dice and summing the resulting up faces, use

```
x = Equilikely(1, 6) + Equilikely(1, 6);
```

  Note that this is *not* equivalent to

```
x = Equilikely(2, 12);
```

- **Example 2** To select an element x at random from the array $a[0]$, $a[1]$, ..., $a[n-1]$ use

```
i = Equilikely(0, n - 1);
x = a[i];
```

Methods for Generation of Instances of Arbitrary Random Variables

The generation of instances of *non-uniform* random variables can be performed using three different methods

- Inverse Transformation Method
- Acceptance–Rejection Method
- Composition Method

The Inverse Transformation method relies on the explicit knowledge of the Cumulative Distribution Function of the random variable. The Acceptance–Rejection and Composition methods assume that the Distribution Functions (in case of discrete random variables) or the Density Functions (in case of continuous random variables) are known.

Inverse Distribution Function

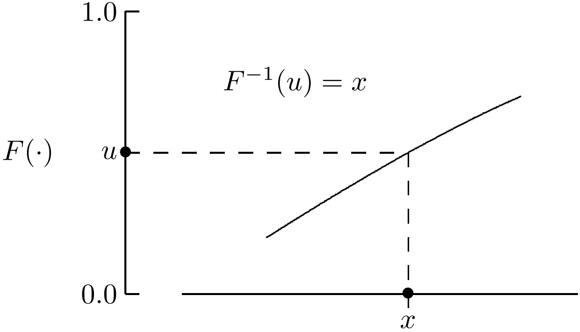
- The *inverse distribution function (idf)* of X is the function $F^{-1} : (0,1) \to \mathcal{X}$ for all $u \in (0,1)$ as

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

  where $x \in \mathcal{X}$ is the unique possible value for $F(x) = u$
- There is a one-to-one correspondence between possible values $x \in \mathcal{X}$ and cdf values $u = F(x) \in (0,1)$
  - Assumes the cdf is strictly monotone increasing
  - True if $f(x) > 0$ for all $x \in \mathcal{X}$

Continuous Random Variable idfs

- The idf for a continuous random variable is a true inverse



- Can sometimes determine the idf in "closed form" by solving $F(x) = u$ for x

Idf for Uniform

- If X is *Uniform*(a, b), $u = F(x) = (x - a)/(b - a)$ for $a < x < b$

$$x = F^{-1}(u) = a + (b - a)u \qquad 0 < u < 1$$

Idf for Exponential

- If X is *Exponential*(η), $u = F(x) = 1 - \exp(-x/\eta)$ for $x > 0$

$$x = F^{-1}(u) = -\eta \ln(1 - u) \qquad 0 < u < 1$$

Idf for Quadratic

- If X is a continuous variable with possible value $0 < x < b$ and pdf $f(x) = 2x/b^2$, the cdf is $u = F(x) = (x/b)^2$

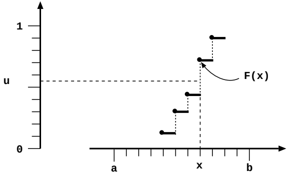
$$x = F^{-1}(u) = b\sqrt{u} \qquad 0 < u < 1$$

Discrete Random Variable idfs

- In the case of discrete random variables, the *inverse distribution function (idf)* of X is the function $F^* : (0,1) \to \mathcal{X}$ for all $u \in (0,1)$ as

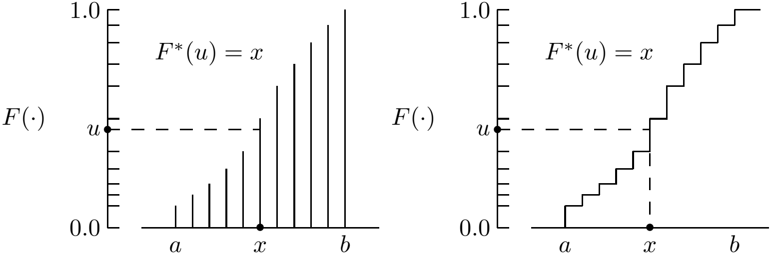
$$F^*(u) = \min_x \{x : u < F(x)\}$$

  $F(\cdot)$ is the cdf of X
- That is, if $F^*(u) = x$, x is the smallest possible value of X for which $F(x)$ is greater than u



Discrete Random Variabe: idf identification

Two common ways of plotting a cdf with $\mathcal{X} = \{a, a+1, \ldots, b\}$:



Theorem

Let $\mathcal{X} = \{a, a+1, \ldots, b\}$ where b may be ∞ and $F(\cdot)$ be the cdf of X For any $u \in (0,1)$,

- *if $u < F(a)$, $F^*(u) = a$*
- *else $F^*(u) = x$ where $x \in \mathcal{X}$ is the unique possible value of X for which $F(x-1) \le u < F(x)$*

Example: Idf for Equilikely

If X is *Equilikely*(a, b),

$$F(x) = \frac{x - a + 1}{b - a + 1} \quad x = a, a + 1, \ldots, b$$

- For $0 < u < F(a)$, $\qquad F^*(u) = a$
- For $F(a) \le u < 1$,

$$\begin{aligned} F(x-1) \le u < F(x) \quad &\iff \quad \frac{(x-1) - a + 1}{b - a + 1} \le u < \frac{x - a + 1}{b - a + 1} \\ &\iff \quad x \le a + (b - a + 1)u < x + 1 \end{aligned}$$

- Therefore, for all $u \in (0, 1)$

$$F^*(u) = a + \lfloor (b - a + 1)u \rfloor$$

Example: Idf for Geometric

If X is *Geometric*(p),

$$f(x) = (1-p)p^x \quad x = 0, 1, 2, \ldots$$
$$F(x) = 1 - p^{x+1} \quad x = 0, 1, 2, \ldots$$

- For $0 < u < F(0)$, $\qquad F^*(u) = 0$
- For $F(0) \leq u < 1$,

$$\begin{aligned}
F(x-1) \leq u < F(x) &\iff 1 - p^x \leq u < 1 - p^{x+1} \\
&\iff -p^x \leq u - 1 < -p^{x+1} \\
&\quad\vdots \\
&\iff x \leq \frac{\ln(1-u)}{\ln(p)} < x + 1
\end{aligned}$$

- For all $u \in (0, 1)$

$$F^*(u) = \left\lfloor \frac{\ln(1-u)}{\ln(p)} \right\rfloor$$

Idf for Bernoulli

- The discrete random *variable* X with possible values $\mathcal{X} = \{0, 1\}$ is said to be *Bernoulli (p)* if
  - $X = 1$ with probability p and $X = 0$ with probability $1 - p$
- The pdf: $f(x) = p^x(1-p)^{1-x}$ for $x \in \mathcal{X}$
- The cdf: $F(x) = (1-p)^{1-x}$ for $x \in \mathcal{X}$
- Let $F(x) = u$, then $x = 0$ iff $0 < u < 1 - p$
- From this last observation follows that the idf is:

$$F^*(u) = \begin{cases} 0 & 0 < u < 1 - p \\ 1 & 1 - p \leq u < 1 \end{cases}$$

Random Variate Generation By Inversion

- X is a random variable with idf $F^{-1}(\cdot)$
- Continuous random variable U is *Uniform*$(0, 1)$
- Z is the random variable defined by $Z = F^{-1}(U)$

Theorem

Z and X are identically distributed

This theorem applies to Continuous and Discrete random variables by making the proper use of $F^{-1}(\cdot)$ or $F^{*}(\cdot)$ and states that they can be generated with *one* call to `Random()`, provided that the idf is known

Algorithm

If X is a random variable with idf $F^{-1}(\cdot)$, a random variate x can be generated as

```
u = Random();
return F^{-1}(u);
```

Notice: replace $F^{-1}(\cdot)$ with $F^{*}(\cdot)$ in case of discrete random variables

Inversion Examples

Generating a *Uniform*(a, b) Random Variate

```
u = Random();
return a + (b - a) * u;
```

Generating an *Exponential*(η) Random Variate

```
u = Random();
return −η * log(1 - u);
```

More Inversion Examples

Generating an *Equilikely*(a, b) Random Variate

```
u = Random();
return a + (long) (u * (b - a + 1));
```

Generating a *Geometric*(p) Random Variate

```
u = Random()
return (long) (log(1.0 - u) / log(p));
```

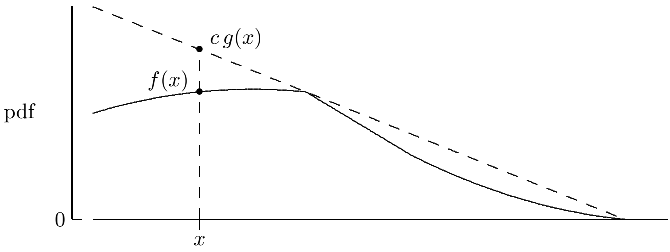
Generating a *Bernoulli*(p) Random Variate

```
u = Random();
if (u < 1-p)
    return 0;
else
    return 1;
```

Acceptance–Rejection

- Acceptance–Rejection is used to generate random variates
  - Usually *continuous* random variates (focus of this lecture)
  - For distributions whose idf cannot be computed efficiently

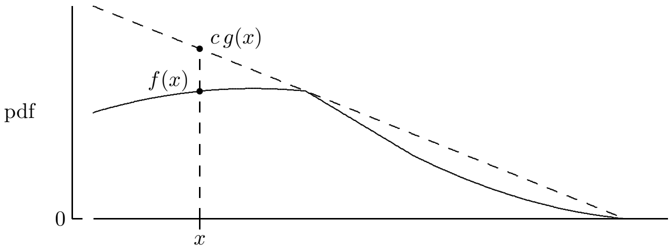
Definition



Acceptance–rejection for generating random variates

- Let X be a continuous r.v. with possible values $\mathcal{X}$, pdf $f(\cdot)$
- Choose *majorizing* pdf $g(\cdot)$, constant $c > 1$ such that
  - $f(x) \leq c\,g(x), \quad$ for all $x \in \mathcal{X}$
  - idf $G^{-1}(\cdot)$ associated with $g(\cdot)$ can be evaluated *efficiently*

Algorithm



Algorithm to generate random variate with distribution $f(\cdot)$

```
do {
    u = Random();
    x = G^{-1}(u);      // x has distribution g(·)
    v = Random();
} while (c * g(x) * v > f(x)); // rejection criterion
return x;
```

Algorithm

Notice the difference between the following two algorithms with and without majorizing function

Algorithm using a majorizing distribution (algorithm from previous slide)

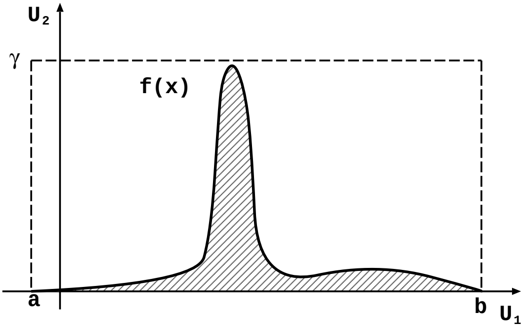
```
do {
    u = Random();
    x = G^{-1}(u);        // x has distribution g(·)
    v = Random();
} while (c * g(x) * v > f(x)); // rejection criterion
return x;
```

Algorithm to generate random variate with distribution $f(\cdot)$ having a support interval (a, b) and maximal height M

```
do {
    u = Random();
    x = a + u * (b − a);        // x is Unif(a, b)
    v = Random();
} while (M * v > f(x)); // rejection criterion
return x;
```

Efficiency

- Efficiency depends on how well $c\,g(x)$ approximates $f(x)$



Composition

- Composition method is used to generate random variates
  - Usually *continuous* random variates (focus of this lecture)
  - For distributions whose idf cannot be computed efficiently
  - Distributions can be represented as weighted sums of elementary distributions whose idf can be computed efficiently

Algorithm

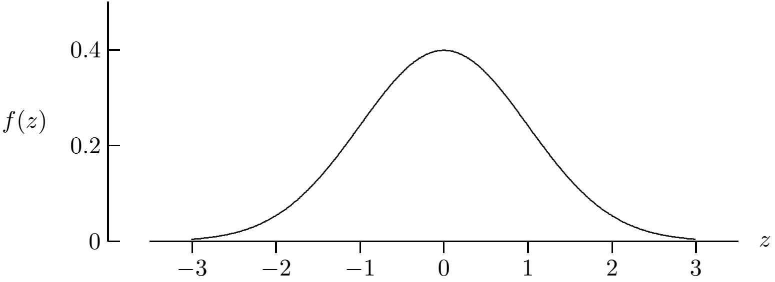
Algorithm to generate random variate with distribution $f(\cdot)$

(*Initialization*)
$A_1 := \alpha_1$;
for $i := 2$ **to** n **do**
    $A_i := A_{i-1} + \alpha_i$;
(*Select the component distribution*)
"Generate $Y = \textit{Uniform}\,(0,1)$";
$r := 0$
repeat
    $r := r + 1$;
until $(Y \leq A_r)$;
"Generate an instance V_r of a random variable with distribution $g_r(v)$";
(*Return the obtained instance*)
$X := V_r$

- **Standard Normal Random Variable**

  Z is *Normal*$(0, 1)$ if and only if the set of all possible values is $\mathcal{Z} = (-\infty, \infty)$ and the pdf is

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp(-z^2/2) \qquad -\infty < z < \infty$$



Standard Normal Random Variable

- If Z is *Normal*$(0, 1)$, Z is "standardized"

  The mean is

  $$\mu = \int_{-\infty}^{\infty} zf(z)dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z \exp(-z^2/2)dz = \cdots = 0$$

  The variance is

  $$\sigma^2 = \int_{-\infty}^{\infty} (z-\mu)^2 f(z)dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 \exp(-z^2/2)dz = \cdots = 1$$

- The cdf is

  $$F(z) = \int_{-\infty}^{z} f(t)dt \qquad -\infty < z < \infty$$

Normal Random Variable

- The continuous random variable X is *Normal*(μ, σ) iff

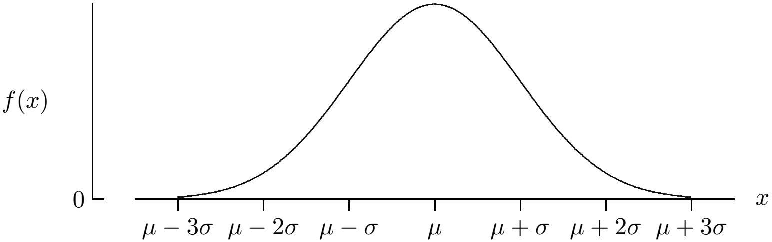
$$X = \sigma Z + \mu$$

  where $\sigma > 0$ and Z is *Normal*$(0, 1)$
- The mean of X is μ and the standard deviation is σ
- *Normal*(μ, σ) is constructed from *Normal*$(0, 1)$
  - by "shifting" the mean from 0 to μ via the addition of μ
  - by "scaling" the standard deviation from 1 to σ via multiplication by σ

pdf of Normal Random Variable

The pdf of *Normal*(μ, σ) is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-(x-\mu)^2/2\sigma^2)$$



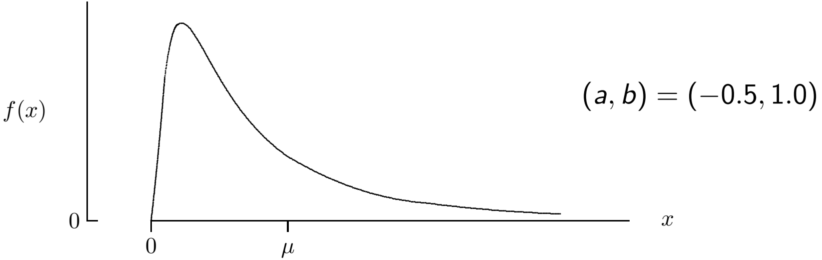
Lognormal Random Variable

- The continuous random variable X is *Lognormal*(a, b) if and only if
$$X = \exp(a + bZ)$$
where Z is *Normal*$(0, 1)$ and $b > 0$
- *Lognormal*(a, b) is also based on transforming *Normal*$(0, 1)$
    - The transformation is non-linear

pdf of Lognormal Random Variable

- The pdf of *Lognormal*(a, b) is

$$f(x) = \frac{1}{bx\sqrt{2\pi}} \exp(-(\ln(x) - a)^2/2b^2) \qquad x > 0$$



- $\mu = \exp(a + b^2/2)$ Above, $\mu = 1.0$
- $\sigma = \exp(a + b^2/2)\sqrt{\exp(b^2) - 1}$ Above, $\sigma \simeq 1.31$

Erlang Random Variable

- The continuous random variable X is *Erlang*(n, b) iff

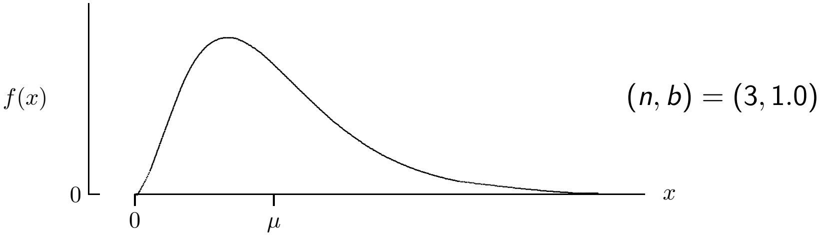
$$X = X_1 + X_2 + \cdots + X_n$$

where $X_1, X_2, \cdots, X_n$ are *iid Exponential*(b) random variables

pdf of Erlang Random Variable

- The pdf of *Erlang*(n, b) is

$$f(x) = \frac{1}{b(n-1)!}(x/b)^{n-1}\exp(-x/b) \qquad x > 0$$



- For $(n, b) = (3, 1.0)$, $\mu = 3.0$ and $\sigma \simeq 1.732$

Hyper-Exponential Distribution

- The continuous random variable X is *Hyper-Exponential*(k) iff its pdf is a composition of elementary exponential distributions

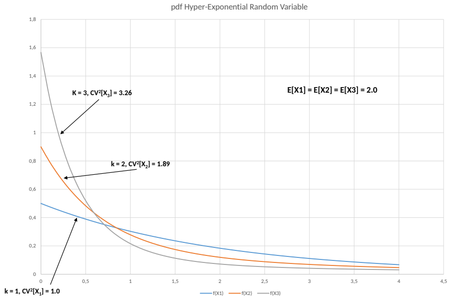
$$f_X(x) \;=\; \sum_{i=1}^{k} \alpha_i f_{X_i}(x) \;=\; \sum_{i=1}^{k} \alpha_i \frac{1}{\eta_i} e^{-\frac{x}{\eta_i}} \;=\; \sum_{i=1}^{k} \alpha_i \lambda_i e^{-\lambda_i x}$$

where

$$\sum_{i=1}^{k} \alpha_i \;=\; 1$$

Hyper-Exponential Distribution

- The definition of the distribution of an Hyper-Exponential random variable requires the specification of the weights and of the rates of the individual exponential components
- The pdf of an Hyper-Exponential distribution has the following shape



- $f(X_1) => \alpha_1 = 1; \lambda_1 = 0,5;$
- $f(X_2) => \alpha_1 = \alpha_2 = 1/2; \lambda_1 = 0,3; \lambda_2 = 1.5;$
- $f(X_3) => \alpha_1 = \alpha_2 = \alpha_3 = 1/3; \lambda_1 = 0.2; \lambda_2 = 1.5; \lambda_3 = 3;$

Hyper-Exponential Distribution

- Expected value and Variance of this random variable can be easily computed from the definition of its pdf
  - For what concerns the Expected value we easily obtain

$$\begin{aligned} E[X] &= \int_0^\infty x f_X(x)dx = \int_0^\infty x \sum_{i=1}^k \alpha_i f_{X_i}(x)dx \\ &= \sum_{i=1}^k \left(\alpha_i \int_0^\infty x f_{X_i}(x)dx \right) = \sum_{i=1}^k \alpha_i \, \eta_i = \sum_{i=1}^k \alpha_i \, \frac{1}{\lambda_i} \end{aligned}$$

  - For what concerns the Variance, we must first compute the Second Moment

$$E[X^2] = \int_0^\infty x^2 f_X(x)dx = 2\sum_{i=1}^k \alpha_i \, \eta_i^2 = 2\sum_{i=1}^k \alpha_i \, \frac{1}{\lambda_i^2}$$

  - an then obtain $VAR[X] = E[X^2] - (E[X])^2$

Normal Distribution

- If $U_1, U_2, \ldots, U_{12}$ is an *iid* sequence of *Uniform*$(0, 1)$,

$$Z = U_1 + U_2 + \ldots + U_{12} - 6$$

  is approximately *Normal*$(0, 1)$
    - The mean is 0.0 and the standard deviation is 1.0
    - Possible values are $-6.0 < z < 6.0$
    - Justification is provided by the central limit theorem
    - This algorithm is: portable, robust, relatively efficient and clear

Normal and Lognormal Random Variates

- Random variates corresponding to *Normal*(μ, σ) and *Lognormal*(a, b) can be generated by using a *Normal*$(0, 1)$ random variate generator

Example: Generating a *Normal*(μ, σ) Random Variate

```
z = Normal(0.0, 1.0);
return μ + σ * z;
```

Example: Generating a *Lognormal*(a, b) Random Variate

```
z = Normal(0.0, 1.0);
return exp(a + b * z);
```

- Both algorithms are essentially ideal

Alternative Random Variate Generation Algorithms

- **Erlang Random Variates**

  An *Erlang*(n, b) random variate can be generated by summing n *Exponential*(b) random variates

Generating an *Erlang*(n, b) Random Variate

```
x = 0.0;
for (i = 0; i < n; i++)
    x += Exponential(b);
return x;
```

- The algorithm is: portable, exact, robust, and clear
- The algorithm is **not** efficient (it is $\mathcal{O}(n)$)

Modified Algorithms for Erlang Random Variates

- To increase computational efficiency, use

Generating an *Erlang*(n, b) Random Variate

```
t = 1.0;
for (i = 0; i < n; i++)
  t *= (1.0 - Random());
return -b * log(t);
```

- This algorithm requires only one `log()` evaluation, rather than n
- Can further improve efficiency by using `t *= Random();`
- The algorithm remains $\mathcal{O}(n)$, so is not efficient if n is large

Generation of Hyper-Exponential Random Variates

- Use the Composition algorithm
- The generation of instances of the hyper-exponential distribution, we can use the previous algorithm randomly selecting the exponential distribution to be used for each variate

Generation of a random variate with hyper-exponential distribution

$$f_X(x) = \sum_{i=1}^{k} \alpha_i \frac{1}{\eta_i} e^{-\frac{x}{\eta_i}} = \sum_{i=1}^{k} \alpha_i \lambda_i e^{-\lambda_i x}$$

(*Initialization*)
$A_1 := \alpha_1$;
for $i := 2$ **to** n **do**
    $A_i := A_{i-1} + \alpha_i$;
(*Select the component distribution*)
"Generate $Y = \mathit{Uniform}\ (0, 1)$";
$j := 0$
repeat
    $r := j + 1$;
until $(Y \leq A_j)$;
"Generate an instance V_j of an *exponential*(η_j)";
(*Return the obtained instance*)
$X := V_j$

Library rvgs

- Includes 6 discrete random variate generators
  - `long Bernoulli(double` p`)`
  - `long Binomial(long` n`, double` p`)`
  - `long Equilikely(long` a`, long` b`)`
  - `long Geometric(double` p`)`
  - `long Pascal(long` n`, double` p`)`
  - `long Poisson(double` μ`)`

- and 7 continuous random variate generators
  - `double Chisquare(long` n`)`
  - `double Erlang(long` n`, double` b`)`
  - `double Exponential(double` μ`)`
  - `double Lognormal(double` a`, double` b`)`
  - `double Normal(double` μ`, double` σ`)`
  - `double Student(long` n`)`
  - `double Uniform(double` a`, double` b`)`