Sampling from continuous distributions

Inverse Transform Method:

Let the random variable X have a continuous and increasing distribution function F. Denote the inverse of F by F^{-1} . Then X can be generated as follows:

- Generate U from U(0, 1);
- Return $X = F^{-1}(U)$.

If F is not continuous or increasing, then we have to use the *generalized* inverse function

$$F^{-1}(u) = \min\{x : F(x) \ge u\}.$$



Examples:

- X = a + (b a)U is uniform on (a, b);
- $X = -\ln(U)/\lambda$ is exponential with parameter λ ;
- $X=(-\ln(U))^{1/a}/\lambda$ is Weibull, parameters a and $\lambda.$

Unfortunately, for many distribution functions we do not have an easy-touse (closed-form) expression for the inverse of F.



Composition method:

This method applies when the distribution function F can be expressed as a mixture of other distribution functions F_1, F_2, \ldots ,

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

where

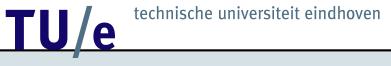
$$p_i \ge 0, \qquad \sum_{i=1}^{\infty} p_i = 1$$

The algorithm is as follows:

• First generate an index *I* such that

$$P(I=i) = p_i, \qquad i = 1, 2, \dots$$

• Generate a random variable X with distribution function F_I .



Examples:

• Hyper-exponential distribution:

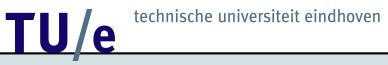
$$F(x) = p_1 F_1(x) + p_2 F_2(x) + \dots + p_k F_k(x), \qquad x \ge 0,$$

where $F_i(x)$ is the exponential distribution with parameter μ_i , $i = 1, \ldots, k$.

• Double-exponential (or Laplace) distribution:

$$f(x) = \begin{cases} \frac{1}{2}e^x, & x < 0; \\ \\ \frac{1}{2}e^{-x}, & x \ge 0, \end{cases}$$

where f denotes the density of F.



Convolution method:

In some case X can be expressed as a sum of independent random variables Y_1,\ldots,Y_n , so

 $X = Y_1 + Y_2 + \dots + Y_n.$

where the Y_i 's can be generated more easily than X.

Algorithm:

- Generate independent Y_1, \ldots, Y_n , each with distribution function G;
- Return $X = Y_1 + \cdots + Y_n$.



Example:

If X is Erlang distributed with parameters n and μ , then X can be expressed as a sum of n independent exponentials Y_i , each with mean $1/\mu$.

Algorithm:

- Generate n exponentials Y_1, \ldots, Y_n , each with mean μ ;
- Set $X = Y_1 + \dots + Y_n$.

More efficient algorithm:

- Generate n uniform (0, 1) random variables U_1, \ldots, U_n ;
- Set $X = -\ln(U_1U_2\cdots U_n)/\mu$.

Acceptance-Rejection method:

Denote the density of X by f. This method requires a function g that majorizes f,

$$g(x) \ge f(x)$$

for all x. Now g will not be a density, since

$$c = \int_{-\infty}^{\infty} g(x) dx \ge 1.$$

Assume that $c < \infty$. Then h(x) = g(x)/c is a density.

Algorithm:

- ı. Generate Y having density h;
- 2. Generate U from $U(0,1), \, {\rm independent} \, {\rm of} \, Y;$
- 3. If $U \leq f(Y)/g(Y)$, then set X = Y; else go back to step 1.

The random variable X generated by the above algorithm has density f.

/ department of mathematics and computing science



Validity of the Acceptance-Rejection method:

Note

$$P(X \le x) = P(Y \le x | Y \text{accepted}).$$

Now,

$$P(Y \le x, Y \text{accepted}) = \int_{-\infty}^{x} \frac{f(y)}{g(y)} h(y) dy = \frac{1}{c} \int_{-\infty}^{x} f(y) dy,$$

and thus, letting $x \to \infty$ gives

$$P(\text{Yaccepted}) = \frac{1}{c}.$$

Hence,

$$P(X \le x) = \frac{P(Y \le x, Y \text{accepted})}{P(Y \text{accepted})} = \int_{-\infty}^{x} f(y) dy.$$

/ department of mathematics and computing science



Note that the number of iterations is geometrically distributed with mean *c*.

How to choose g?

- \bullet Try to choose g such that the random variable Y can be generated rapidly;
- The probability of rejection in step 3 should be small; so try to bring c close to 1, which mean that g should be close to f.

TU/e technische universiteit eindhoven

Example:

The Beta(4,3) distribution has density

$$f(x) = 60x^3(1-x)^2, \qquad 0 \le x \le 1.$$

The maximal value of f occurs at x=0.6, where f(0.6)=2.0736. Thus, if we define

$$g(x) = 2.0736, \qquad 0 \le x \le 1,$$

then g majorizes f.

Algorithm:

I. Generate Y and U from U(0, 1);

2. If

$$U \le \frac{60Y^3(1-Y)^2}{2.0736},$$

then set X = Y; else reject Y and return to step 1.

Generating Normal random variables

Methods:

- Acceptance-Rejection method
- Box-Muller method

Acceptance-Rejection method:

If X is N(0,1), then the density of $\left|X\right|$ is given by

$$f(x) = \frac{2}{\sqrt{2\pi}}e^{-x^2/2}, \qquad x > 0.$$

Now the function

$$g(x) = \sqrt{2e/\pi}e^{-x}$$

majorizes f. This leads to the following algorithm:

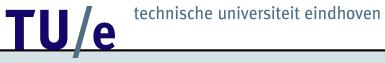
- ı. Generate an exponential Y with mean ı;
- 2. Generate U from U(0, 1), independent of Y;

3. If

$$U \le e^{-(Y-1)^2/2},$$

then accept Y; else reject Y and return to step 1.

4. Return X = Y or X = -Y, both with probability 1/2.



Box-Muller method:

If U_1 and U_2 are independent U(0,1) random variables, then

$$X_{1} = \sqrt{-2 \ln U_{1}} \cos(2\pi U_{2})$$

$$X_{2} = \sqrt{-2 \ln U_{1}} \sin(2\pi U_{2})$$

are independent standard normal random variables.