STAT 3006: Statistical Computing
Lecture 5*

5 February

4 Generating Random Deviates

4.1 The Inverse Method

Definition 4.1. For a non-decreasing function F on $\mathbb{R}$, the generalized inverse of F, written as F^-, is the function defined by

$$F^-(u) = \inf\{x : F(x) \geq u\}.$$

In the figure 1, the curve is a distribution function F. Select a value u_1 between 0 and 1. There is a unique value t_1, such that $F(t_1) = u_1$, so we denote t_1 by $F^-(u_1)$. Select a value u_2, and there exists an interval $[t_3, t_2]$ satisfying $F(t) = u_2$ for each $t \in [t_3, t_2]$. We denote t_3 by $F^-(u_2)$ because of t_3 is the minimum of all points satisfying $F(t) \geq u_2$. Select values u_3 or u_4. Although there is no point t such that $F(t) = u_3$ or u_4, the minimum of $\{t : F(t) \geq u_3\} = [t_3, \infty]$ or $\{t : F(t) \geq u_4\} = [t_3, \infty]$ does exit, which is t_3. Therefore, we also denote t_3 by $F^-(u_3)$ or $F^-(u_4)$. $F^-(u_2) = F^-(u_3) = F^-(u_4)$.

Proposition 4.1. *If $U \sim Unif[0, 1]$, then the random variable $F^-(U)$ has the distribution F.*

Proof. $\forall u \in [0, 1]$ and $\forall x \in$ the domain of F, we always have

$$F(F^-(u)) \geq u \tag{4.1}$$
$$F^-(F(x)) \leq x. \tag{4.2}$$

Equation (4.1) holds based on the right continuity of the distribution function F. Equation (4.2) holds because of the definition of the generalized inverse of F.

Subsequently, $\forall x$, we investigate the two sets $\{u : F^-(u) \leq x\}$ and $\{u : u \leq F(x)\}$. When $F^-(u) \leq x$, $F(F^-(u)) \leq F(x)$ as F is non-decreasing. According to Equation (4.1), we have

*If you have any question about the note, please send an email to xyluo@link.cuhk.edu.hk

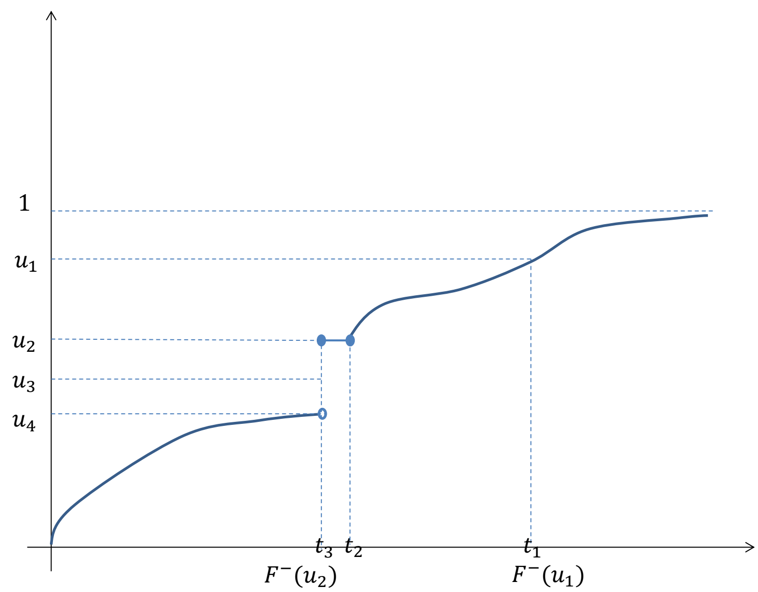

Figure 1: Figure description for the generalized inverse function $F^-(x)$.

$u \leq F(x)$. When $u \leq F(x)$, notice that F^- is also non-decreasing, so $F^-(u) \leq F^-(F(x))$. According to Equation (4.2), we have $F^-(u) \leq x$. Therefore, $\{u : F^-(u) \leq x\} = \{u : u \leq F(x)\}$.

It follows that $P(F^-(U) \leq x) = P(U \leq F(x)) = F(x)$. It indicates that $F^-(U)$ has the distribution F.

Algorithm: Inverse method for drawing samples from F.

Input: the distribution function F, total sample number N.
Initialize: $n \leftarrow 0$.
Repeat
  generate a uniform random number U_n from $[0, 1]$; $X_n \leftarrow F^-(U_n)$; $n \leftarrow n + 1$;
Until $n == N$.
Output:$\{X_1, \ldots, X_N\}$ are N samples from distribution F.

Example: sampling from the exponential distribution $f(x) = e^{-x}$.
$F(x) = 1 - e^{-x}$, $F^-(u) = -log(1 - u)$. Sample $U \sim Unif[0, 1]$ and let X be $-log(1 - U)$. Notice $1 - U$ also follows $Unif[0, 1]$, so $X = -log(U)$ also follows $f(x) = e^{-x}$.

4.2 The Accept-Reject Method

We are interested in sampling from a probability density function f. $f(x)$ can be written as $\int_0^{f(x)} du = \int_0^{\infty} I(u \leq f(x))du$. From this perspective, we interpret $f(x)$ as the marginal distribution of $Unif\{(x, u) : 0 < u < f(x)\}$. The U is called *auxiliary variable*.

Theorem 4.2. *(Fundamental Theorem of Simulation) Simulating* $X \sim f(x) \Leftrightarrow$ *simulating* $(X, U) \sim Unif\{(x, u) : 0 < u < f(x)\}$.

Proof. $F_X(x) = P(X \leq x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(0 \leq u \leq f(t), t \leq x)dudt = \int_{-\infty}^{x}\int_{0}^{f(t)} dudt = \int_{-\infty}^{x} f(t)dt$. Therefore, $f_X(x) = F'_X(x) = f(x)$.

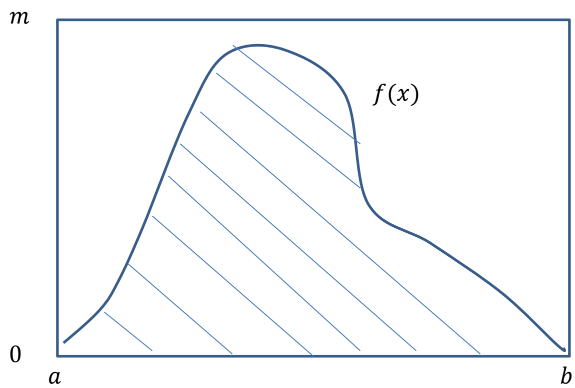

Figure 2: f's upper bound in a special case.

Proposition 4.3. *When the region* $\{(x, u) : u \leq f(x)\}$ *is bounded by a rectangle* $[a, b] \times [0, m]$ *(see Figure 2), simulating uniform samples*(X, U) *from* $\{(x, u) : u \leq f(x)\} \Leftrightarrow$ *simulating uniform samples from* $[a, b] \times [0, m]$ *but only accepting samples in the* $\{(x, u) : u \leq f(x)\}$.

Proof. The theorem is equivalent to $X \sim f(x) \Leftrightarrow Y \sim Unif[a, b]$, $U \sim Unif[0, m]$ and accept Y as X if $U \leq f(Y)$.

$$\begin{aligned} P(X \leq x) &= P(Y \leq x | U \leq f(Y)) \\ &= \frac{P(Y \leq x, U \leq f(Y))}{P(U \leq f(Y))} \\ &= \frac{\int_a^x \frac{1}{b-a} \int_0^{f(y)} \frac{1}{m} dudy}{\int_a^b \frac{1}{b-a} \int_0^{f(y)} \frac{1}{m} dudy} \\ &= \frac{\int_a^x f(y)dy}{1} \\ &= \int_a^x f(y)dy. \end{aligned}$$

Furthermore, we can calculate that the acceptance rate is $P(U \leq f(Y)) = \frac{1}{m(b-a)}$.

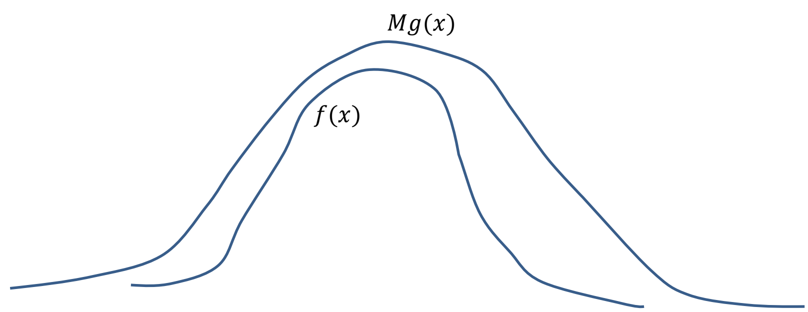

Figure 3: f's upper bound in the general case.

Proposition 4.4. *When the region* $\{(x,u) : u \leq f(x)\}$ *is bounded by the region* $\{(x,u) : u \leq M \cdot g(x)\}$ *(see Figure 3), where* $f(x) \leq Mg(x) \forall x \in dom(f)$*, simulating uniform samples*(X,U) *from* $\{(x,u) : u \leq f(x)\} \Leftrightarrow$ *simulating uniform samples from* $\{(x,u) : u \leq M \cdot g(x)\}$ *but only accepting samples in the* $\{(x,u) : u \leq f(x)\}$*.*

Remark 1. : Usually, drawing samples from g is much easier than drawing samples from f, so we "reshape" samples from g to "construct" samples from f.

Proof. The theorem is equivalent to $X \sim f(x) \Leftrightarrow Y \sim g(y)$, $U|Y = y \sim Unif[0, Mg(y)]$ and accept Y as X if $U \leq f(Y)$.

$$\begin{aligned}
P(X \leq x) &= P(Y \leq x | U \leq f(Y)) \\
&= \frac{P(Y \leq x, U \leq f(Y))}{P(U \leq f(Y))} \\
&= \frac{\int_{-\infty}^{x} \int_0^{f(y)} \frac{1}{Mg(y)} du g(y) dy}{\int_{-\infty}^{\infty} \int_0^{f(y)} \frac{1}{Mg(y)} du g(y) dy} \\
&= \int_{-\infty}^{x} f(y) dy.
\end{aligned}$$

Moreover, the acceptance rate is $P(U \leq f(Y)) = \frac{1}{M}$.

Algorithm: Accept-Reject method for drawing samples from f.

Input: target pdf f, proposal pdf g, constant M ($f(x) \leq Mg(x)$), total sample number N.
Initialize: $n \leftarrow 0$.
repeat
  generate a uniform random number U_n from $[0, 1]$;
  generate a random number Y_n from $g(y)$;
  if $U_n \leq \frac{f(Y_n)}{Mg(Y_n)}$
   $X_n \leftarrow Y_n$;
   $n \leftarrow n + 1$;
  end if
until $n == N$.
Output:$\{X_1, \ldots, X_N\}$ are N samples from distribution f.

4.3 Sequential Sampling

Proposition 4.5. *sampling* $(x_1, x_2, \ldots, x_k)$ *from* $f(x_1, x_2, \ldots, x_k)$ *is equivalent to first sampling* $x_1 \sim f(x_1)$*, second sampling* $x_2 \sim f(x_2|x_1)$*, ..., and finally sampling* $x_k \sim f(x_k|x_1, \ldots, x_{k-1})$.

Example: draw a sample $(X_1, X_2, \ldots, X_k)$ *from a multinomial distribution* $Multi(n; p_1, p_2, \ldots, p_k)$*, where* X_j *denotes the count of type* j $(1 \leq j \leq k)$, $\sum_{j=1}^k X_j = n$, $\sum_{j=1}^k p_j = 1$.

$X_1 \sim Binomial(n, p_1)$, $X_2|X_1 = x_1 \sim Binomial(n - x_1, \frac{p_2}{1-p_1})$, ..., $X_{k-1}|X_1 = x_1, \ldots, X_{k-2} = x_{k-2} \sim Binomial(n \sum_{i=1}^{k-2} x_i, \frac{p_{k-1}}{1-\sum_{i=1}^{k-2} p_i})$*, and* $X_k = n - \sum_{i=1}^{k-1} X_i$.