The Inverse Transform Method

Question: How can we use a random number generator that samples from a uniform distribution on $[0,1]$ to sample from another distribution?

Recall that a uniform random variable U on $[0,1]$ has cumulative distribution function

$$F_U(x) = P(U \leq x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x > 1. \end{cases}$$

1 Discrete Case

Let X be a discrete random variable with $p_i = P(X = x_i)$, $i = 1, \ldots, n$. If U is a uniform random variable and $0 \leq a \leq b \leq 1$ then

$$P(a \leq U \leq b) = P(U \leq b) - P(U \leq a) = F_U(b) - F_U(a) = b - a.$$

Hence for every n

$$P(p_1 + \ldots + p_{n-1} \leq U \leq p_1 + \ldots + p_n) = p_n.$$

Now let $Y = \Phi(U)$ be a function of the random variable U defined by

$$Y = \Phi(U) = \begin{cases} x_1 & \text{if } U \leq p_1 \\ x_2 & \text{if } p_1 \leq U \leq p_1 + p_2 \\ \vdots & \vdots \qquad \texttt{modified to <} \\ x_n & \text{if } p_1 + \ldots + p_{n-1} \leq U \leq p_1 + \ldots + p_n \end{cases}$$

Then Y has the same distribution as X. Therefore, if $u_1, \ldots, u_k$ is a sample from a uniform distribution then $\Phi(u_1), \ldots, \Phi(u_k)$ is a sample from the distribution of X. Notice that this definition can be extended to discrete random variables with an infinite number of states (e.g. a Poisson distribution to model the number of customers arriving in a fixed time interval). The function Φ is then defined by

$$\Phi(U) = \begin{cases} x_1 & \text{if } U \leq p_1 \\ x_2 & \text{if } p_1 \leq U \leq p_1 + p_2 \\ \vdots & \vdots \qquad \textbf{modified to <} \\ x_i & \text{if } p_1 + \ldots + p_{i-1} \leq U \leq p_1 + \ldots + p_i \\ \vdots & \vdots \end{cases}$$

2 The continuous case

Now suppose X is a random variable with a cumulative distribtion function $F_X(x) = P(X \leq x)$. Suppose that F_X is strictly increasing in the following sense: If $x \geq y$ and $0 < F_X(x)$ and $F_X(y) < 1$ then $F_X(x) < F_X(y)$ [1]. That is the case for the normal, the exponential, and most other continuous distributions; it is not the case for discrete distribution - that's why we had the extra case above. If F_X is strictly increasing in the sense explained, then for each $0 < u < 1$ the equation $F_X(x) = u$ has a unique solution, call it $x = F_X^{-1}(u)$. Moreover, the function F_X^{-1} is strictly increasing on the interval $(0, 1)$ where it is defined. As in the discrete case we are interested in finding a function Φ such that $Y = \Phi(U)$ has the same distribution as X, where U is a uniform random variable. Notice first

$$P(U \leq F_X(x)) = F_X(x)$$

for each x such that $F_X(x) \in [0, 1]$. Since F_X^{-1} is strictly increasing the event $U \leq F_X(x)$ occurs if and only if the event $F_X^{-1}(U) \leq F_X^{-1}(F_X(x))$ occurs. By the definition of F_X^{-1}, $F_X^{-1}(F_X(x))$ is the solution of the equation $F_X(x) = u$, where $u = F_X(x)$. This solution is obviously x itself. Hence the event $U \leq F_X(x)$ occurs if and only if $F_X^{-1}(U) \leq x$ and we conclude that

$$F_X(x) = P(U \leq F_X(x)) = P(F_X^{-1}(U) \leq x).$$

[1]Notice that $F_X(x) \leq F_X(y)$ is satisfied by EVERY cumulative distribution function. If X has a continuous density function f then the cumulative distribution function is strictly increasing in our sense if the support of f (i.e. the set of reals on which f is positive) has no holes. Here, I define a hole to be an interval of the type $[a, b]$ with $a < b$.

In other words, the random variable $Y = F_X^{-1}(U)$ has the same distribution as X. Therefore, if $u_1, \ldots, u_k$ are sampled from a uniform distribution and x_i is the solution of $F_X(x) = u_i$ then $x_1, \ldots, x_k$ are sampled from a distribution with cumulative distribution function F_X.