Simulating Random Variables

December 1, 2009

We have seen several methods from simulating random variables based on the ability to generated a random variable U that is uniformly distributed on $[0,1]$.

For discrete random variables, the distribution function F_X is constant except for jumps $f_X(x_i)$ at those $x_1 < x_2 < \dots$ that X takes on with positive probability.

In this case, set $x_0 = -\infty$.

$$P\{F_X(x_i) < U \le F_X(x_{i+1})\} = F_X(x_{i+1}) - F_X(x_i) = f_X(x_{i+1})$$

and we can simulate X by assigning the value x_{i+1} whenever $F_X(x_i) < U \le F_X(x_{i+1})$.

1 The Probability Transform

For X a continuous random variable with a density f_x that is positive everywhere in its domain, the distribution function F_X is strictly increasing. In this case F_X has a inverse fuunction. The **probability transform** follows from the fact that

$$U = F_X(X)$$

is uniformly distributed on $[0,1]$. Thus, if we can simulate U, uniformly distributed on $[0,1]$, we can simulated a random variable with distribution F_X via

$$X = F_X^{-1}(U).$$

2 Box-Muller Transform

For the Box-Muller transform, we require two random variables U, V, uniformly distributed on $[0,1]$. Set

$$R = \sqrt{-2\log V} \qquad \text{and} \qquad \Theta = 2\pi U.$$

and

$$Z_1 = R\cos\Theta = \sqrt{-2\log V}\cos(2\pi U), \qquad \text{and} \qquad Z_2 = \sqrt{-2\log V}\sin(2\pi U).$$

Then X and Y are independent standard normal random variables. To obtain two standard normal random variables with correlation ρ, take

$$X = Z_1 \qquad \text{and} \qquad Y = \rho Z_1 + \sqrt{1-\rho^2}Z_2.$$

3 Rejection Sampling

Let Y be a continuous random variable with density f_Y on an interval $[0,1]$ and choose c greater than the maximum value $f_Y(y)$.

Now choose a point $(U,V) \in [0,1] \times [0,c]$ uniformly on the rectangle. If the value is below the graph of f_Y then choose $Y = y$, otherwise reject the value and try again.

Then, consider

$$P\{V \le y | U < f_Y(V)\} = \frac{P\{V \le y, U < f_Y(V)\}}{P\{U < f_Y(V)\}}.$$

and

$$P\{V \le y, U < f_Y(V)\} = \int_0^y \int_0^{f_Y(v)} \frac{1}{c}\, du\, dv = \frac{1}{c}\int_0^y f_Y(v) du\, dv = \frac{1}{c}P\{Y \le y\}.$$

To find $P\{U < f_Y(V)\}$, take $y = 1$ in the previous computation.

$$P\{U < f_Y(V)\} = \frac{1}{c}P\{Y \le 1\} = \frac{1}{c}.$$

Taken together, we find that

$$P\{V \le y | U < f_Y(V)\} = P\{Y \le y\}.$$

Note that the probability of accepting the choice is $1/c$. To improve on this, we can choose a second density f_V so the distribution is easy to compute but looks like f_u. We could them transform both U and V using F_V. Then, the transformed Y is close to uniform and the using the algorithm above, the often accept the transformed value of Y. We then obtain the desired value by applying F_V^{-1}.

Combining all these steps, we have the method **rejection sampling**:

To generate a random variable Y with density f_Y, first set

$$c = \sup_y \frac{f_Y(y)}{f_V)}.$$

Then,

1. Generate two independent variables - U that is uniform on $[0,1]$ and V with density f_V
2. If $U < f_Y(V)/(cf_V(v))$, set $Y = V$. Otherwise, return to step 1.

As before,

$$P\{V \le y | U < f_Y(V)/(cf_V(v))\} = \frac{P\{V \le y, U < f_Y(V)/(cf_V(v))\}}{P\{U < f_Y(V)/(cf_V(V))\}}.$$

The numerator

$$\begin{aligned} P\{V \le y, U < f_Y(V)/(cf_V(v))\} &= \int_0^y \int_0^{f_Y(v)/(cf_V(v))} f_V(v) du\, dv = \int_0^y \frac{f_Y(v)}{cf_V(v)} f_V(v) du\, dv \\ &= \frac{1}{c}\int_0^y f_Y(v) du\, dv = \frac{1}{c}P\{Y \le y\} \end{aligned}$$

The value of the denominator $1/c$ can be found taking $y = \infty$ in the previous computation. Therefore,

$$P\{V \le y | U < f_Y(V)/(cf_V(v))\} = P\{Y \le y\}.$$