Generating discrete Random variable

Arabin Kumar Dey

Assistant Professor

Department of Mathematics
Indian Institute of Technology Guwahati

Talk at Indian Institute of Technology Guwahati

6-th january, 2012

The Inverse transform method

- Consider a r.v.X with pmf $P(X = x_j) = p_j\ ,\ \ j = 0, 1, \cdots, \sum_j p_j = 1$

- To generate a value from this distribution, first we generate a random number U and set

$$X = \begin{cases} x_0, & \text{if } \ U < p_0 \\ x_1, & \text{if } \ p_0 \leq U < p_0 + p_1 \\ \vdots & \\ x_j & \textit{if } \ \sum_{i=0}^{j-1} p_i \leq U < \sum_{i=0}^{j} p_i \\ \vdots & \end{cases}$$

- This is called the inverse transform method.
- Proof to be outlined in class.

$$
\begin{aligned}
P(X = x_j) &= P\left(\sum_{i=0}^{j-1} p_i \leq U < \sum_{i=0}^{j} p_i\right) \\
&= P\left(U < \sum_{i=0}^{j} p_i\right) - P\left(U < \sum_{i=0}^{j-1} p_i\right) \\
&= \sum_{i=0}^{j} p_i - \sum_{i=0}^{j-1} p_i \\
&= p_j
\end{aligned}
$$

Some remarks

- The method can be written algorithmically as generate a random number U
  if $U < p_0$, set $X = x_0$ and STOP
  if $U < p_0 + p_1$, set $X = x_1$ and STOP
  if $U < p_0 + p_1 + p_2$, set $X = x_2$ and STOP
  $\vdots$
- If the $X_i' s$ are ordered like $X_0 < X_1 < X_2 < \cdots$ so that the cdf $F(X_k) = \sum_{i=0}^{k} p_i$ and that
  X equals to X_j if $F(X_{j-1}) \leq U < F(X_j)$.
  Therefore, after generating U, we determine the value of X by looking for the interval.$[F(x_{j-1}), F(X_j)]$ in which it lies (or equivalently finding the inverse of U).

Illustrative example

Suppose we want to simulate from the discrete distribution with $P(X = 1) = 0.20$, $P(X = 2) = 0.25$, $P(X = 3) = 0.40$, $P(X = 4) = 0.15$

We do the following:

generate a random number U

if $U < 0.20$, set $X = 1$ and STOP

if $U < 0.45$, set $X = 2$ and STOP

if $U < 0.85$, set $X = 3$ and STOP

otherwise set X = 4.

It is suggested the following could be more efficient:

generate a random number U

if $U < 0.40$, set $X = 3$ and STOP

if $U < 0.65$, set $X = 2$ and STOP

if $U < 0.85$, set $X = 1$ and STOP

otherwise set X = 4.

Generating a discrete uniform random variable

- In the discrete uniform distribution, we have equal probabilities: $P(X = j) = \frac{1}{n}$, for $j = 1, 2, \ldots, n$.
- To simulate from this distribution, we generate a random number U and then set
  $X = j \quad \text{if } \frac{j-1}{n} \leq U < \frac{j}{n}$.
- This condition is equivalent to if $j - 1 \leq nU < j$, that is
  X=Int(nU)+1 ,
  where Int(X) is the greatest integer part of X.