Simulating an IID Sequence from an Arbitrary Distribution

Brian Hunt and David Levermore
University of Maryland
AMSC/MATH 420, Spring 2013

Pseudo-Random Number Generators

- Many algorithms have been developed to simulate random IID sequences (though nothing a computer does is truly random).
- These algorithms usually generate numbers that are uniformly distributed, typically between 0 and 1.
- Most mathematical/statistical software systems have a built-in routine to convert the uniformly distributed pseudorandom numbers into normally distributed numbers with mean 0 and variance 1.
- In MATLAB, the commands `rand` and `randn` generate pseudoranom numbers that are uniformly and normally distributed, respectively.

Arbitrary Probability Distributions

- Suppose you want do simulate an IID sequence of real numbers with a different (but still continuous) distribution, whose probability density function ("pdf") is $q(X)$.
- The simplest algorithm to do so involves the corresponding cumulative distribution function ("cdf")

$$Q(X) = \int_{-\infty}^{X} q(x)dx.$$

- Recall that the meaning of the pdf is that $\int_a^b q(x)dx$ is the probability that the random variable X is between a and b. Also, Q is nondecreasing with $Q(-\infty) = 0$ and $Q(\infty) = 1$.

Key Observation

- Since $Q(x)$ is the probability that a randomly chosen X is $\leq x$, the values of $Q(X)$ are uniformly distributed between 0 and 1. Here's why:
- The probability that X is between a and b is $Q(b) - Q(a)$, so the probability that $Q(X)$ is between $Q(a)$ and $Q(b)$ is also $Q(b) - Q(a)$.
- Since $Q(a)$ and $Q(b)$ can take on all values between 0 and 1, $Q(X)$ is a random variable whose probability of being between c and d is $d - c$ for all $0 < c < d < 1$. This is the standard uniform random variable.

The Punch Line (and Algorithm)

- Therefore, if Y is a random variable that is uniformly distributed between 0 and 1, then $Q^{-1}(Y)$ is distributed according to the desired distribution with pdf q.
- Algorithm: Use `rand` or the corresponding pseudorandom number generator for your software system to simulate a uniform IID sequence of real numbers, and apply Q^{-1} to each of these numbers to simulate an IID sequence with pdf q.
- For some pdfs q, one can find a formula for Q^{-1}. Otherwise, one needs to compute enough values of Q in order to reasonably approximate Q^{-1}.

Example: Exponential Distribution

- The exponential distribution with mean β has pdf $q(x) = \exp(-x/\beta)/\beta$ for $x \geq 0$ and $q(x) = 0$ for $x < 0$.
- The corresponding cdf is $Q(x) = 1 - \exp(-x/\beta)$ for $x \geq 0$ and $Q(x) = 0$ for $x < 0$.
- Then $Q^{-1}(y) = -\beta \log(1 - y)$ for $0 < y < 1$.
- If Y is uniformly distributed then so is $1 - Y$; thus, the logarithms of uniformly distributed numbers are exponentially distributed.