AMS 597: Statistical Computing

Pei-Fen Kuan (c)

Applied Math and Stats, Stony Brook University

Generating random variables

- Aim: Simulate random variables from specified probability distributions
- A suitable generator of uniform pseudo random numbers is essential. Methods for generating r.v. from other probability distributions all depend on the uniform random number generator
- In this course, we assume that a suitable uniform pseudo random number generator is available

Generating random variables

- We have previously learnt to generate random variables using built in functions in R, e.g. `rnorm`, `rbeta`, etc
- We will learn a few algorithms for generating random variables

Quantile-quantile plot

- However, we first introduce a simple visualization method to compare if two random number generators are equivalent
- One simple way is a graphical display called quantile-quantile plot (Q-Q plot)
- Quantile point q_p for a random variable X is the point such that $F_X(q_p) = P(X \leq q_p) = p$
- The quantile is a function of probability $q_p = F_X^{-1}(p)$
- If q_p^X and q_p^Y are quantile functions of random variables X and Y, Q-Q plot of X and Y is the plot of (q_p^X, q_p^Y) for all p

Quantile-quantile plot

- Exercise: Generate the Q-Q plot of $X \sim \exp(2)$ and $Y \sim \exp(2)$. Since they are identical distribution, we would expect the line $y = x$ as the Q-Q plot.

Inverse Transform Method

- Theorem (Probability Integral Transformation)
  If X is a continuous random variable with cdf $F_X(x)$, then $U = F_X(X) \sim U(0,1)$

Inverse Transform Method

- The method can be applied for generating continuous or discrete random variables.
- Suppose we want to generate $x \sim F_X$
  1. Derive the inverse function $F_X^{-1}(u)$
  2. Write a command or function to compute $F_X^{-1}(u)$
  3. Generate a random u from $U(0,1)$
  4. Define $x = F_X^{-1}(u)$

Inverse Transform Method

- Exercise: Simulate from 10000 X's from Exp(2), i.e., $E(X) = 1/2$ using the inverse transform method

Inverse Transform Method

- Exercise: Generate 1000 random numbers whose probability density function is $f(x) = x$ for $(0, c)$ and 0 otherwise. (Figure out the appropriate c)

Inverse Transform Method

- Theoretically you can generate any random variables from the uniform distribution but the disadvantage of the integral transform method is that it is practical only if the c.d.f. are explicitly available. For instance the c.d.f. of a normal r.v. cannot be expressed explicitly.

Inverse Transform Method

- If $X \sim N(0,1)$, the c.d.f. of X is given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(\frac{z^2}{2}\right) dz.$$

It can be shown that

$$\Phi^{-1}(x) \approx t - \frac{a_0 + a_1 t}{1 + b_1 t + b_2 t^2}$$

for some constants a_i and b_i ($t^2 = -2\log x$), where $a_0 = 2.30753, a_1 = 0.27061, b_1 = 0.99229, b_2 = 0.04481$

Inverse Transform Method

- Exercise: Simulate 10000 $N(0,1)$ using the inverse transform method.

Acceptance-Rejection Method

- Suppose that X and Y are random variables with pdf f and g respectively, and there exists a constant c such that

$$f(t)/g(t) \leq c$$

for all t such that $f(t) > 0$, and that we know how to generate Y. Then the acceptance-rejection method can be used to generate X

Acceptance-Rejection Method

1. Generate Y from $g(y)$
2. Generate U from $U(0,1)$
3. If $U \leq f(Y)/cg(Y)$, then set $X = Y$ (accept); otherwise go back to 1 (reject)

- For the proof, please see page 1-2 of http://www.columbia.edu/~ks20/4703-Sigman/4703-07-Notes-ARM.pdf

Acceptance-Rejection Method

- Exercise: Simulate 10000 r.v's that follow probability density $f(x) = 2x$, $0 < x < 1$ using AR method.

Acceptance-Rejection Method

- Exercise: Simulate 10000 r.v's from Beta(2,2) using AR method. On average, how many iterations needed using the tightest bound?
- How about using $c = 6$?

Acceptance-Rejection Method

- AR method may not be an efficient algorithm
- To use the accept-reject method, the distributions f and g should be somewhat similar to have a sufficiently good algorithm

Transformation Methods

- If $X \sim Ga(a, \lambda)$ and $Y \sim Ga(b, \lambda)$, X and Y are independent, then $X/(X+Y) \sim \text{Beta}(a, b)$
- For the proof, please see https://online.stat.psu.edu/stat414/lesson/23/23.2

Transformation Methods

- Exercise: Simulate 10000 r.v's from Beta(2,2) using transformation method

Transformation Methods

- Box-Muller Method
  If U_1 and U_2 are independent $U(0, 1)$ random variables, then

$$X_1 = \sqrt{-2 \log U_1} \cos(2\pi U_2)$$

$$X_2 = \sqrt{-2 \log U_1} \sin(2\pi U_2)$$

  are independent $N(0, 1)$
- For the proof, please see https://mathworld.wolfram.com/Box-MullerTransformation.html

Transformation Methods

- Exercise: Simulate 10000 $N(0, 1)$ using the Box Muller method.