Introduction to Simulation of Random Variables

(gustavoc@msu.edu)

Monte Carlo (MC) Methods play a central role in statistical analysis, including power and type-I error rate analysis and in both Likelihood and Bayesian inference. The basic idea underlying MC methods is very simple: **we use samples** from the distribution of random variables (or functions thereof) to evaluates features of that distribution (e.g., expected value, variance, quantiles) are very difficult or cannot be derived analytically.

The first step in a MC study is then to sample random variables. In this note I review a few basic algorithms for simulating random variables and provide examples using R.

1. Random numbers & seeds

Computers do not generate random numbers. Everything in an algorithm is deterministic. However, there are well-established algorithms that will generate sequences of numbers that resemble IID sequences from uniform random variables. A random number generator implements an algorithm that, starting from a seed (an integer) generates a deterministic sequence of numbers that resemble IID samples from uniform distributions. Given the seed the sequence is deterministic but we regarded it as random. Try in R the following code

```
set.seed(1923)
runif(5)
runif(4)
set.seed(1923)
runif(5)
runif(4)
```

1. Transformations of random variables

Suppose that $Y = g(X)$ where X is a random variable with pdf $f_X(x)$ and $g(X)$ is a one-to-one monotonic map, then

$$f_Y(y) = f_X(X(y)\,)\left|\frac{dX(y)}{dy}\right| \qquad [1]$$

where $X(y) = g^{-1}(y)$.

Here $Y = g(X)$ is a transformation and $X(Y) = g^{-1}(Y)$ is the inverse transformation.

Example: sampling an exponential random variable.

The pdf of an exponential random variable is given by

$$f_Y(y) = \lambda e^{-\lambda y} \quad \lambda, y > 0$$

Result: if $X \sim U(0,1)$ and $Y = -\frac{\log(X)}{\lambda}$ then $Y \sim Exponential(\lambda)$.

Proof:

(1) here $Y = g(X) = -\frac{\log(X)}{\lambda}$; therefore $X = g^{-1}(Y) = e^{-\lambda Y}$.
(2) If $X \sim U(0,1)$=> $f_X(x\,) = 1(0 \leq x \leq 1)$, then using [1]
(3) $f_Y(y) = 1(g^{-1}(Y) \in [0,1]) \left|\frac{dX(y)}{dy}\right| = 1(y > 0)e^{-\lambda y}|-\lambda| = \lambda e^{-\lambda y} \quad y > 0$

Generalization: The above results generalizes to random vectors, in which case $|\frac{dX(y)}{dy}|$ needs to be replaced with the absolute value of the Jacobian of the transformation (i.e., the absolute value of the determinant of the matrix of first derivatives of the inverse-transformation) evaluated at y. Note: this result holds for one-to-one maps. It would not hold, for instance for $g(X) = X^2$.

Example:

```
lambda=5;n=1e5
x=runif(n)
y=-log(x)/lambda
hist(y,n)
mean(y); 1/lambda
var(y); 1/(lambda^2)
y2=rexp(n=n,rate=lambda)
par(mfrow=c(1,2))
hist(y,50); hist(y2,50)
```

[See map of distributions from Casella and Berger in our website]

2. Sampling Bernoulli Random Variables

A Bernoulli random variable (X) with success probability θ can be sampled as follows: sample `U~uniform(n=1,min=0,max=1)`, set X to be 0 if U<p, 1 otherwise, that is `X=ifelse(U<θ,0,1)`.

Why does it work? $p(X = 0) = p(U < \theta) = \int_0^\theta 1 d\theta = \theta$, that is the cdf of the uniform is the 45 degree line in the unit square, that is `p==qunif(p)` for all values of p in [0,1].

4. The inverse probability method

The previous method illustrates how uniform RVs can be used to sample discrete RVs. But what about continuous RVs?

Any continuous RV can be generated from a uniform RV using the inverse probability method. Recall that the cdf of a RV is given by $p(X < q) = \Phi(q) = \int_{-\infty}^{q} p(x)dx$. The cdf maps from values of the random variable to probabilities. The inverse cdf (or quantile function) maps from probabilities to quantiles, that is $\Phi^{-1}(\theta) = q : \Phi(q) = \theta$. In R the cdf of a random variable has prefix 'p' (e.g., pnorm) and the inverse cdf has prefix 'q' (e.g., qnorm).

The inverse probability method generate a random draw for an arbitrary RV, with known inverse-cdf, by: (i) drawing a uniform [0,1] RV (say U~runif(1)) and then evaluating the inverse cdf of X at U. The following example illustrates this:

```
mu=5
SD=4
N=1e5

u=runif(n)
Y1=qnorm(u, sd=SD,mean=mu)
Y2=rnorm(n,sd=SD,mean=mu)
# Compare the empirical quantiles of Y1 and Y2…
```

Why does the inverse probability method works? It works because the cdf(X) follows a uniform distribution no matter what the distribution of X is. Here are a few examples:

```
N=1e5

X=rexp(n=N,rate=4)
hist(pexp(X,rate=4),50)

X=rnorm(n=N,mean=12,sd=5)
hist(pnorm(X,mean=12,sd=5),50)
```