A Brief Introduction into the Monte-Carlo Method and Generation of the Standard Normal Random Deviate.

Kenneth C. Syracuse, Ph.D., CQE
Chief Scientist
Balan Biomedical, Inc.

A Balan Biomedical, Inc. WHITE paper

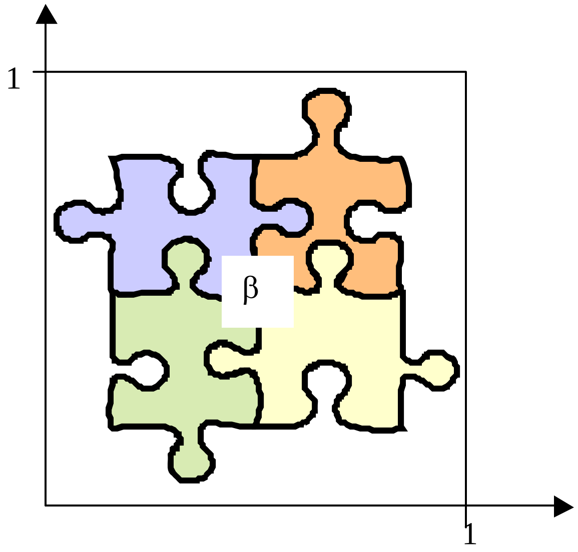
February, 2009

Monte Carlo Simulation:

Although the name is synonymous with the gambling mecca of the world, the concept of the monte-carlo simulation is not a method to improve your chances of beating the house. A monte-carlo simulation is a method of solving mathematical problems by simulation of the associated random variables.

The origin of the name “monte-carlo” is simply that the method derives its utility through the use of random numbers. One of the simplest mechanical devices for generating random numbers is a roulette wheel.

An example will be used to illustrate the basic concept:



Suppose the desired computation is the area of the shaded object, say β. Let’s allow that β is inscribed within the unit area depicted on the axes. Generate N random values each having equal probability of occurring within the unit area. Count the number of them falling within the zone circumscribed by β, say n. Clearly the ratio n/N provides an estimate of the area of β relative to the unit area. This “hit or miss” method can be expanded to multidimensional problems where standard mathematical procedures are not effective.

There are two advantages to using the Monte-Carlo method. The first is that it is relatively simple to implement. In the example above, each “trial” involves the generation of a random deviate in a space and comparing the result to some condition. Secondly, the error in calculation is roughly equivalent to $(\delta/N)^{1/2}$, where δ is some constant and N represents the number of trials. Note that increasing N decreases the error.

The key to using the monte-carlo method is in understanding that the method can be used to simulate any process, which is influenced by random factors. In addition, the method enables the artificial construction of a probabilistic model, which can then be used to solve the problem. It is in this mode that the author will develop the necessary computational algorithms.

Simulating random variables.

A random deviate is a variable whose value in any given circumstance, is unknown. However, what is known are the values it can assume and the probabilities associated with those values. Therefore, the result of a single trial cannot be predicted with any great precision, however, the results can be predicted with high reliability with a large number of these trials. Thus, to define a random deviate, we must indicate its’ value space and the associated probabilities.

Because the author’s primary interest is with continuous random variables, the description of the discrete case will be omitted.

Continuous Random Variables

A continuous random variable *g* is defined by specifying an interval containing all its possible values and a function *f(x)* which is called the probability density of the random variable (ref. Appendix E). The physical meaning of *f* is as follows: Let *(a, b)* be an arbitrary interval contained in (α, β). Consider only those *(a, b)* pairs for which $(\alpha \le a < b \le \beta)$. The probability that X falls in the interval *(a, b)* is equal to

$$P\{a < x < b\} = \int_a^b f(x)dx$$

When X is defined over the unit interval (0, 1) and having appropriate probability density (i.e., $\int_0^1 f(x)dx = 1$), then X is said to be uniformly distributed (0, 1) when the probability that X lies in the interval *(a, b)* is $\int_a^b f(x)dx = b - a$. That is, the probability is equal to the length of the sub-interval. Straightforward calculation yields the mean:

$$E(x) = \int_0^1 x f(x) dx = \left.\frac{x^2}{2}\right|_0^1 = {}^1\!/_2, \qquad \text{and}$$

$$Var(x) = E(x^2) - \{E(x)\}^2 = \int_0^1 x^2 f(x) dx - {}^1\!/_4 = {}^1\!/_{12}.$$

That is to say that the random variable X is *uniformly* distributed over the unit interval if and only if its probability density function is given by:

$$f(x) = \frac{1}{(b - a)}, \qquad \textit{for all } a < b.$$

This event is denoted by $X \sim U(0,1)$.

Theorem (without proof): Let X be a continuous random variable with distribution function $F(.)$. Then let the random variable $Y = F(x)$. Let the distribution function of Y be given by $G(y) = P(Y \leq y)$. Then Y is distributed as $U(0,1)$. The result follows from Theorem 17, in Appendix D. And so to generate random deviates from a continuous distribution it is sufficient to generate random deviates from a U(0,1) distribution and use the inverse of the distribution function associated with the desired random deviate.

Generating the $N(\mu, \sigma^2)$ *random deviate:*

Consider a normal random variable ξ, with expected value equal to and standard deviation $\sigma = 1$. Let η be a random variable with the same distribution, but independent of ξ. Then the probability density of a random point with Cartesian points (ξ, η) in the *(x,y)* plane is equal to the product of their marginal one-dimensional densities. Thus,

$$p(x, y) = 1/\sqrt{2\pi}\, e^{\frac{-x^2}{2}} * 1/\sqrt{2\pi}\, e^{\frac{-y^2}{2}} = 1/2\pi\, e^{\frac{-(x^2+y^2)}{2}}$$

Next, introduce the typical polar coordinate mapping

$$x = r\cos\varphi \text{ and } y = r\sin\varphi .$$

Now let ρ and θ denote the polar coordinates of the (ξ, η) pair:

$$\xi = \rho\cos\theta \text{ and } \eta = \rho\sin\theta .$$

The joint density becomes:

$$p'(r, \varphi) = p(x, y)\left|\frac{\partial(x, y)}{\partial(r, \varphi)}\right| = \frac{r}{2\pi} e^{-r^2/2}$$

The individual densities of ρ and θ by be found by integration:

$$p_1(r) = \int_0^{2\pi} p'(r, \varphi)d\varphi = \frac{r}{2} e^{-r^2/2} \text{, and}$$

$$p_2(\varphi) = \int_0^{\infty} p'(r, \varphi)dr = \frac{1}{2\pi} .$$

Since the product $p_1(r) * p_2(\varphi) = p'(r, \varphi)$, ρ and θ are independent. Now

$$F_1(r) = \int_0^r p_1(t)dt = 1 - e^{-r^2/2}, \qquad 0 < r < \infty \text{, and}$$

$$F_2(\varphi) = \int_0^{\varphi} p_2(t)dt = \varphi/2\pi, \qquad 0 < \varphi < 2\pi .$$

Next, let $\gamma_1 = e^{-\rho^2/2}$, and $\gamma_2 = \theta/2\pi$, then

$$F_1(\rho) = 1 - \gamma_1 \text{, and } F_2(\theta) = \gamma_2 .$$

Solving for ρ and φ we have:

$$\rho = (-2\ln\gamma_1)^{1/2} \text{, and } \varphi = 2\pi\gamma_2 ,$$

and finally,

$$\xi = (-2 \ln \gamma_1)^{\frac{1}{2}} \cos 2\pi\gamma_2, \text{ and } \eta = (-2 \ln \gamma_1)^{\frac{1}{2}} \sin 2\pi\gamma_2.$$

This allows a tractable method for the computation of two normal random deviates with 0 mean and a variance of 1, the standard normal random deviate. A simple linear transformation performed on the *N(0,1)* random deviate will yield a normally distributed random deviate with the desired parameters.