Random Object Generation[1]

Starting point: uniform random variable U on $[0,1]$. Good method is a trade secrete. One general idea is to use

$$X_{n+1} = aX_n + c \ (\mathrm{mod}\ m)$$

for LARGE positive integer m and some positive integers a, c; then $U_n = X_n/m,\ n = 1, \ldots, N$ will be (approximately) iid uniform on $[0,1]$ for some finite value of N.

Direct generation using the *inverse probability integral transform*: if F_X is the cdf of X and

$$F_X^{\leftarrow}(x) = \inf\{t > 0 : F(t) \geq x\}$$

is the *left continuous* inverse of F_X, the random variable $F_X^{\leftarrow}(U)$ has the same distribution as X. The proof is by direct computation: using the definition of $F_X^{\leftarrow}$, show that $\mathbb{P}\big(F_X^{\leftarrow}(U) \leq x\big) = \mathbb{P}\big(U \leq F_X(x)\big) = F_X(x)$. Note that if F_X is continuous and strictly increasing, then $F_X^{\leftarrow} = F_X^{-1}$, the inverse function.

Examples

(1) Exponential random variable X with mean θ: $F_X(x) = 1 - e^{-x/\theta}$, $F_X^{\leftarrow}(x) = F_X^{-1}(x) = -\theta \ln(1-x)$, and so [using that U and $1-U$ have the same distribution] the random variable $-\theta \ln U$ an exponential random variable with mean θ.

(2) The standard Cauchy random variable ξ: the cdf is $F_\xi(x) = (1/2) + (1/\pi)\tan^{-1}(x)$, so $\tan\big(\pi U - (\pi/2)\big)$ has the standard Cauchy distribution.

(3) The Box-Muller transformation. If U and V are iid is uniform on $(0,1)$, then

$$X = \sqrt{-2\ln(U)}\cos(2\pi V), \quad Y = \sqrt{-2\ln(U)}\sin(2\pi V)$$

are iid standard normal. This is named after British statistician George Edward Pelham Box (1919 – 2013) and American computer scientist Mervin Edgar Muller (1928–2018). When a *very* large number of Gaussian random variables is necessary, this method might not work because of all the computer power necessary to evaluate the log and trig functions.

(4) If X and Y are iid standard normal, then X/Y has standard Cauchy distribution and $X^2 + Y^2$ has χ^2_2 distribution.

Further developments

- A discrete random variable X with $\mathbb{P}(X = a_k) = p_k > 0$, $k = 1, \ldots, M$: generate a random variable U that is uniform on $(0,1)$ and then set $X = a_1$ if $U \leq p_1$, $X = a_2$ if $p_1 < U \leq p_1 + p_2$, etc.
- Geometric random variable $G_0(p)$: generate a random variable U that is uniform on $(0,1)$ and then

$$G_0(p) = \left\lfloor \frac{\ln U}{\ln(1-p)} \right\rfloor.$$

- Poisson random variable with mean λ:

$$\mathcal{P}(\lambda) = \min\left(n : \prod_{k=1}^{n} U_k < e^{-\lambda}\right) - 1 = \max\left(n : \prod_{k=1}^{n} U_k \geq e^{-\lambda}\right); \quad U_k \text{ iid uniform on } [0,1].$$

Indeed, we have $\mathbb{P}(\mathcal{P}(\lambda) > n) = \mathbb{P}(\mathrm{Gamma}(n+1, \lambda) \leq 1)$, $\mathrm{Gamma}(n+1) = \sum_{k=1}^{n+1} V_k$ for iid exponential V_k with mean $1/\lambda$, $\lambda V_k = -\ln U_k$.

- Uniform on $\mathbb{S}^n$, the n-dimensional sphere $\{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = 1\}$: generate $X_1, \ldots, X_n$ iid standard normal and compute $R = \sqrt{X_1^2 + \cdots + X_n^2}$, then $U = (X_1/R, \ldots, X_n/R)$ will be uniform on $\mathbb{S}^n$.
- Rejection methods: for example to generate a uniform in the unit disk, generate X, Y that are iid in $[-1,1]$, and then keep the result if $X^2 + Y^2 \leq 1$.

[1]Sergey Lototsky, USC