MONTE CARLO TECHNIQUES

Generating pseudo-random numbers

from **UNIFORM distribution:**

We take the interval to be $[0, 1]$; a simple transformation can change this to any $[a, b]$.

The simplest (and usually sufficient) technique uses `multiplicative congruential generator`, which works as follows:

$$\begin{aligned} X_{n+1} &= aX_n \mod m \\ U_{n+1} &= \frac{X_{n+1}}{m} \end{aligned}$$

where a is the `multiplier`, m the `modulus`, and X_0 the `seed`.

`EXAMPLE:` $a = 22$, $m = 25$, $X_0 = 17$, yields (for X_1, X_2, ...): 24, 3, 16, 2, 19, 18, 21, 14, 8, 1, 22, 9, 23, 6, 7, 4, 13, 11, 17, 24, 3,...(repeating).

The most practical choice of m is 2^ℓ (where ℓ is the number of bits a specific computer uses for integer multiplication).

One can show that, in this case, the longest sequence one can generate is of length $2^{\ell-2}$ ($\ell > 3$), but only when X_0 is odd and $a = 3$ or $5 \bmod 8$ (when $a = 7$ or $9 \bmod 16$, the length is $2^{\ell-3}$, etc.).

Thus, for a 48 bit machine, we get a sequence of length $2^{46} = 7.036\,9 \times 10^{13}$ (a million of these every second would keep you going for over two years - 64 bit machine gives you enough numbers for over half a million years).

But there are other possibilities: $m = 2^{31} - 1$ has been a fairly popular choice (the longest length of the corresponding sequence is $2^{31} - 2$, i.e. goes over all integers but 0, since $2^{31} - 1$ is a prime number).

Also, a slightly more sophisticated generator is of the `linear congruential` type where, after each multiplication, we also add a number, thus

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

where c is usually equal to 1 (I have also seen 11).

Exponential distribution

Solving

$$F(x) = 1 - e^{-x/\beta} = U$$

for x yields

$$X = -\beta \ln(1 - U)$$

Gamma

Just add k independent exponentials.

Normal (standardized)

Since we don't have an explicit expression for $F(z)$, the following `polar method` is used:

Generate X and Y independently from $\mathcal{U}(-1, 1)$. Check if these are inside the unit

circle, i.e. $W \equiv X^2 + Y^2 < 1$. If not, discard and try again. Then

$$\begin{aligned} Z_1 &= X\sqrt{\frac{-2\ln W}{W}} \\ Z_2 &= Y\sqrt{\frac{-2\ln W}{W}} \end{aligned}$$

are independent from $\mathcal{N}(0,1)$.

`Proof:` Obviously, $f(x,y) = \frac{1}{\pi}$. Also, the inverse transformation, namely

$$\begin{aligned} x &= \frac{z_1}{\sqrt{z_1^2+z_2^2}} \cdot \exp\left(-\frac{z_1^2+z_2^2}{4}\right) \\ y &= \frac{z_2}{\sqrt{z_1^2+z_2^2}} \cdot \exp\left(-\frac{z_1^2+z_2^2}{4}\right) \end{aligned}$$

has a Jacobian with the absolute value of

$$\frac{1}{2}\exp\left(-\frac{z_1^2+z_2^2}{2}\right)$$

Chi-square

Add k independent Z_i^2.

Bernoulli

If $U < p$ return 1, otherwise return 0.

Binomial

Add n independent Bernoulli RVs.

Geometric

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

where brackets imply truncation to an integer.

`Proof:` Suppose Y is exponential with the mean of β. Then, $[Y]$ has the following probability function: $f(i) = \exp(-\frac{i}{\beta}) - \exp(-\frac{i+1}{\beta}) \equiv p \cdot q^i$ with $p = 1 - \exp(-\frac{1}{\beta})$. Solving this for β, we get $-\frac{1}{\ln(1-p)}$.

Poisson

Count how many independent exponentially distributed CVs (with $\beta = 1$) can you fit into a interval of length λ.

In general (hypergeometric, etc.), we would compare U to the corresponding $F(i)$.

Selecting a RIS (without replacement)

or, equivalently, select k integers out of $1..n$, thus:

Set $j = k$, then, for i from 1 to n, if $U < \frac{j}{n-i+1}$ select i and decrease j by 1.

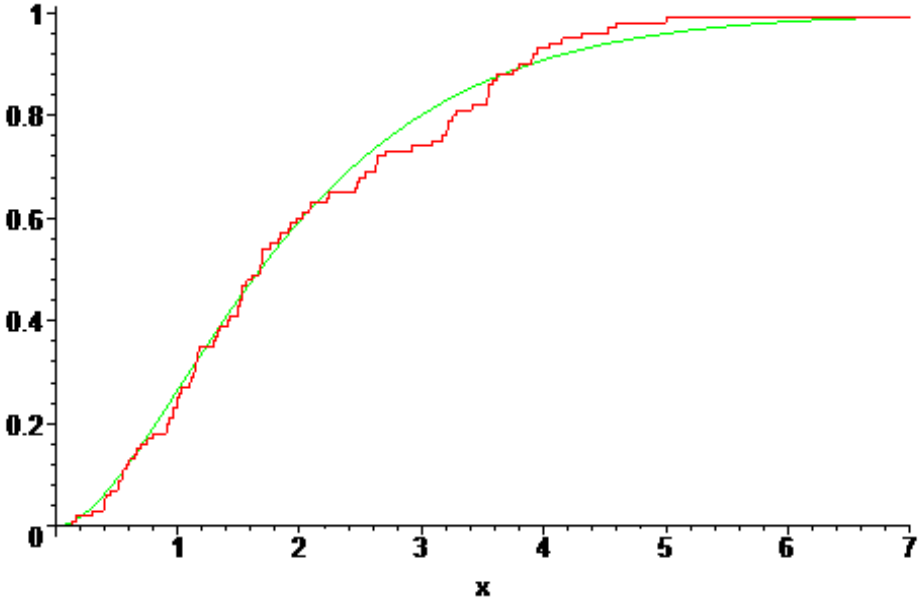
Estimating distribution function $F(x)$ and pdf $f(x)$

For $F(x)$, we use

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i < x)$$

where I (the `indicator function`) equals 1 when the condition is met, 0 otherwise.

`Example` (gamma$_{2,1}$):



To estimate $f(x)$ is a bit more difficult, we use

$$\frac{1}{nh}\sum_{i=1}^{n}\phi(\tfrac{x-X_i}{h})$$

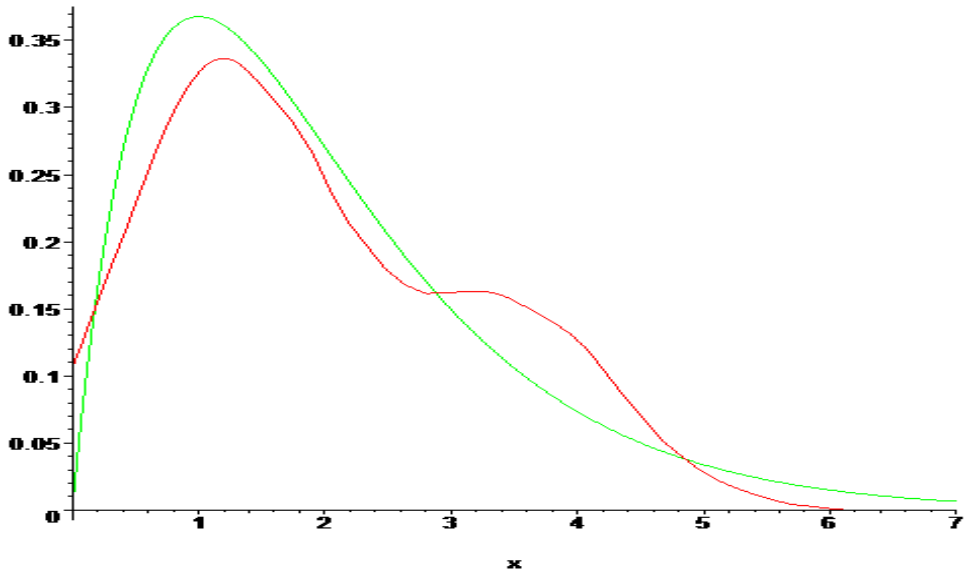
where

$$\phi(u)=\frac{3}{4}\max(1-u^2,0)$$

and

$$h\approx\frac{2\times\text{st.dev.}}{n^{0.2}}$$

Example:



Sampling a multivariate distribution (Quantum-Chemistry application)

Suppose we have a function of several variables, say $\psi(\mathbf{R}) \geq 0$. Then,

$$f(\mathbf{R}) = \frac{\psi(\mathbf{R})}{\int \cdots \int \psi(\mathbf{R}) d\mathbf{R}}$$

is clearly a multivariate pdf. To generate a sample of configurations from this distribution, we first realize that $f(\mathbf{R})$ meets the following PDE

$$-\frac{1}{2}\boldsymbol{\nabla}^2 f + \boldsymbol{\nabla} \bullet \mathbf{F}\, f = 0$$

where

$$\mathbf{F} \equiv \frac{\boldsymbol{\nabla}\psi(\mathbf{R})}{2\psi(\mathbf{R})}$$

Proof:

$\boldsymbol{\nabla} \bullet \mathbf{F}\, f = \frac{\boldsymbol{\nabla} \bullet \boldsymbol{\nabla}\psi(\mathbf{R})}{2c} = \frac{\boldsymbol{\nabla}^2\psi(\mathbf{R})}{2c}$ and $\boldsymbol{\nabla}^2 f = \frac{\boldsymbol{\nabla}^2\psi(\mathbf{R})}{c}$.

We will now try to solve

$$-\frac{1}{2}\boldsymbol{\nabla}^2 f + \boldsymbol{\nabla} \bullet \mathbf{F}\, f = -\frac{\partial f}{\partial t}$$

instead (and take its stationary solution).

There is a Green's-function solution to

$$-\frac{1}{2}\boldsymbol{\nabla}^2 f = -\frac{\partial f}{\partial t}$$

namely

$$(2\pi t)^{-d/2} \exp\left[-\frac{(\mathbf{R}-\mathbf{R}_0)^2}{2t}\right]$$

(d is the dimension of $\mathbf{R}$) and to

$$\boldsymbol{\nabla} \bullet \mathbf{F}\ f = -\frac{\partial f}{\partial t}$$

namely

$$\delta[\mathbf{R}-\tilde{\mathbf{R}}(t)] \text{ where } \frac{d\tilde{\mathbf{R}}(t)}{dt} = \mathbf{F}[\tilde{\mathbf{R}}(t)] \text{ and } \tilde{\mathbf{R}}(0) = \mathbf{R}_0$$

Each of these has a nice statistical and 'physical' interpretation, called `diffusion` and `drift` respectively.

The Green's function to the original equation does not have a simple analytical form, but for small t (called `time step`) it can be approximated by applying drift and diffusion consecutively (this is called an `iteration`). Using a specific (small) time step, we first generate an arbitrary `ensemble` of configurations which we then keep on advancing (by repeated drift/diffusion) until a stationary situation is reached (we can tell by monitoring one or more averages). We then continue averaging, until a good estimate (for each quantity of interest) is obtained. Note that consecutive iteration averages are not independent (since we are dealing with a complicated time-series process).

There are two ways of removing the time-step error:

Use several time steps (e.g. $0.01, 0.02, 0.03$), then extrapolate all your averages to $t = 0$.

Use **Metropolis sampling**, which removes the time-step error (even when t is relatively large).

This is done by computing, at the end of each iteration, the following quantity (one for every configuration)

$$T \equiv \frac{\psi(\mathbf{R})}{\psi(\mathbf{R}_0)} \cdot \exp\left(\frac{[(\mathbf{R}-\mathbf{R}_0-t\mathbf{F}(\mathbf{R}_0)]^2 - [(\mathbf{R}_0-\mathbf{R}-t\mathbf{F}(\mathbf{R})]^2}{2t}\right)$$

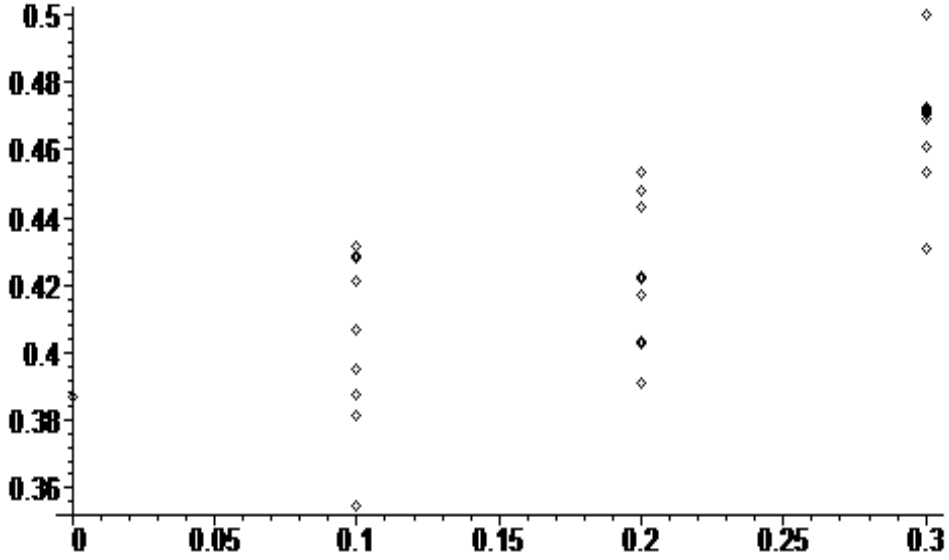
where $\mathbf{R}_0$ represents the starting location, $\mathbf{R}$ is the new location (after drift and diffusion). The actual 'move' is then accepted with probability T (a uniform random number needs to be generated, to make this decision). If a move is rejected, we simply return to $\mathbf{R}_0$ (stay put). One should monitor 'staleness' of each configuration. Note that T may occasionally be bigger than 1 (in which case we automatically accept the move).

`Example` (one-dimensional):

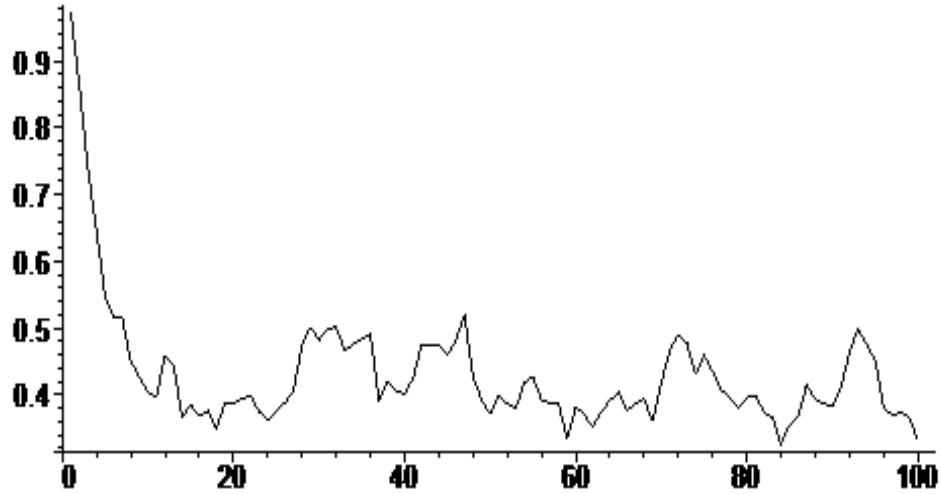
$$\psi = \exp(-R^2/2)\sqrt{\ln(2+R^2)}$$

Using 200 configurations and 100 iterations at $t = 0.1$, 0.2 and 0.3, we get, for the

distribution's variance:



This is how the individual results looked like, for $t = 0.1$:



With Metropolis, we get at $t = 0.3$