Random numbers

Clearly a number that is produced on the computer in a deterministic way cannot be truly random. So a valid question is what do we mean by a “random number” and how can we test it?

We consider pseudo-random numbers that share some properties with random numbers but obviously are reproducible on the computers and therefore are not truly random.

A useful definition of true random numbers is lack of correlations. If we consider the product of two random numbers \( r_1 \) and \( r_2 \), -- \( r_1 \cdot r_2 \) and we average over possible values of \( r_1 \) and \( r_2 \), we should have \( \langle r_1 \cdot r_2 \rangle = \langle r_1 \rangle \langle r_2 \rangle \).

So a test for a random number generator would be 
\[
\prod_{i=1}^{N} r_i = \prod_{i=1}^{N} \langle r_i \rangle
\]

Essentially all the existing random number generators fail eventually on this kind of test. The common generators are cyclic in nature. There is an \( L \) -- large integer such that \( r_{i+L} = r_i \), hence the number of random numbers that can be generated is finite.

Widely used random number generators are based on the following simple (and fast) operations:

\[
I_{k+1} = \alpha I_k + \beta \pmod{m}
\]

The integers \( I_k \) are between zero and m-1. Dividing by m provides a floating point between 0 and 1. If all is well the sequence of the integers is uniformly distributed at the interval [0,1]

Example: \( \alpha = 2,147,437,301 \) \( \beta = 453,816,981 \) \( m=2^{32} \)

- Using random numbers suggests a procedure to estimate \( \pi \)

To improve the quality and the “randomness” of numbers generated by the above procedure it is useful to have a long vector of random numbers and to shuffle them (randomly)

In MATLAB

Rand(n,m) provide an nxm matrix of random numbers.

The above procedure provides random numbers generated from a uniform distribution. Can we generate random numbers from other probability distribution (e.g. normal)?
A general procedure for doing it is based on the probability function. Let \( p(x)dx \) be the probability of finding \( x \) and \( x + dx \). Suppose that we want to generate a series of points \( x \) and then compute a function of these points \( y(x) \). What will be the distribution of the \( y \)-s? It will be connected to the probability function of the \( x \)-s.

\[
p(x)dx = p(y)|dy|
\]

\[
p(y) = p(x) \left| \frac{dx}{dy} \right|
\]

Example: suppose \( y(x) = -\log_e(x) \)

\[
p(y)dy = \left| \frac{dx}{dy} \right| = e^{-y}dy
\]

Another example: Gaussian

We want

\[
p(y)dy = \frac{1}{\sqrt{2\pi}} \exp\left[-y^2/2\right]dy
\]

select

\[
y_1 = \sqrt{-2\log(x_1)} \cos(2\pi x_2)
\]

\[
y_2 = \sqrt{-2\log(x_1)} \sin(2\pi x_2)
\]

\[
x_1 = \exp\left[-(y_1^2 + y_2^2)/2\right]
\]

\[
x_2 = \frac{1}{2\pi} \arctan\left( \frac{y_2}{y_1} \right)
\]

\[
\begin{vmatrix}
\frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\
\frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2}
\end{vmatrix} = -\left[ \frac{1}{\sqrt{2\pi}} \exp\left(-y_1^2/2\right) \right] \left[ \frac{1}{\sqrt{2\pi}} \exp\left(-y_2^2/2\right) \right]
\]

Note that there is one-to-one correspondence between \( x \) and \( y \).