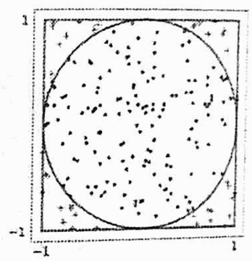
13.5 Monte Carlo simulation

Monte Carlo (MC) simulation is a computational method that involves random numbers. The computation of π through random numbers, as described below, is basically a MC simulation where random numbers are drawn from a uniform distribution. That means, we essentially consider some random sample points of the entire space. This is an example of random *sampling*.

13.5.1 Value of pi



Laplace's method (1886) of calculating π is based on random numbers. Consider a circle of radius 1 unit inscribed in a square of length 2 units. The centre of the circle coincides with the centre of the square. Take one side of the square along x- axis and the adjacent side along y- axis in [-1, 1].

Now, imagine that we throw darts towards a square dart board randomly. The ratio of hits inside the circle to the total trials should be the same as the ratio of the areas of the circle and the square. This approximation is better as we consider larger number of throws. The ratio of the areas of the inscribed circle to that of square is $\pi/4$. Therefore, the ratio of hits, multiplied by 4, will be an estimate of *pi*.

Each point on the square can be represented by $(x, \ y)$ coordinates. The dart, hitting at a random point anywhere inside the square, is given by two random values of x and y in [-1, 1]. The condition of the dart hitting inside the circle: $x^2 + y^2 \leq 1$. We take uniform random numbers in the range.

13.5.2 Monte Carlo integration

From the mean value statistics, we can write,

$$\bar{f} \approx \frac{1}{b-a}\int_a^b f(x)dx$$

$$(b-a)\bar{f} \approx \int_a^b f(x)dx$$

$$(b-a)\frac{1}{N}\sum_{i=1}^{N} f(x_i) \approx \int_a^b f(x)dx$$

The last expression tells us that we can approximate the integral by calculating the mean value of the function over some discrete points and then multiply it by the range of the integral. By law of large numbers, the mean value on left tends to the actual value of integral when $N \to \infty$.

The idea of MC integration is to generate many values at random points (taken from some distribution) within the range of integration and then to compute the mean value. We can think that `(b - a)/N` is the mean width of the subdivisions which is multiplied by the heights to find the total area of the rectangular strips under the curve.

Geometric interpretation:

In MC integration, we essentially pick up some random points (just like that in estimating π) to target the area under a curve in order to measure it. The task is to find out the area under the curve in the domain of integration. To begin with, we draw a rectangular box which will contain the curve as shown below. In this case, the side of the box along the x-axis will be given by the two limits $[0, \pi]$. The side along the y-axis will be given by the minimum and maximum values of the function. In practice, we take the height something greater than the maximum value of the function so as not to miss any point falling in the box.

Imagine, we throw random darts (random points inside the box) and count how many of them hit below the curve. Consider, A = area of the box, I = area under the curve (Integral), n = number of dots under the curve, N = Number of trials. So, we can write, $\frac{I}{A} = \frac{n}{N}$, which means, $I = A * n/N$, as $N \to \infty$.

Algorithm:

1. Calculate the area of the rectangle around the curve:
   width = x- range of integration, height = 1.2 * max of y- value (1.2 factor is arbitrary.)
2. Call a pair of random numbers: x, y, in the domain given by the sides of the box.
3. Check if a random point (x, y) falls under the curve: $y <' f(x)$. Then
   Count points (n) falling under the curve out of total trials (N).
4. Compute the Integral: $I = A * n/N$