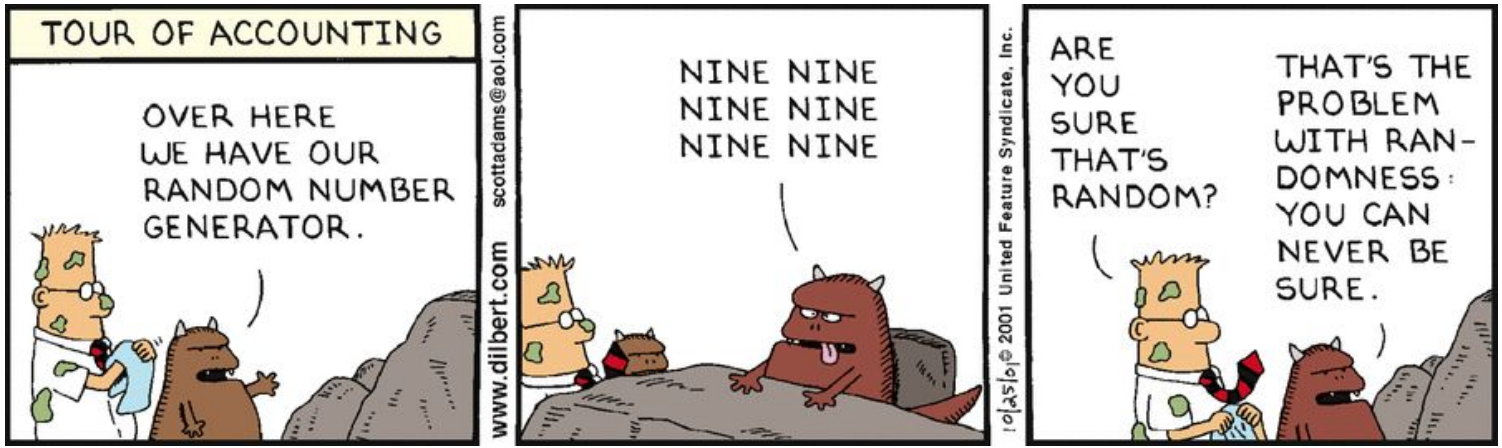
Monte Carlo Integration

Matt Palmer - PHYS 113



What are Monte Carlo methods?

- **A broad class of computational algorithms relying on *random sampling* to achieve results.**
- We can use randomness to solve principle based problems.
- Complex uses in the physical sciences, but also simpler uses including:
  - Integration
  - Simulation and optimization
  - Inverse problems

Visualization of a *Rayleigh-Taylor instability problem*— instability of an interface between two fluids of different densities occurring when the lighter fluid is pushing the heavier fluid.

What are Monte Carlo methods?

General Monte Carlo Process:

1. Define a domain.
2. Generate random inputs into the domain.
3. Perform a deterministic* computation on those inputs
4. Aggregate data.

Discussed in G&T problems 3.36, 3.60 & in B&B 28.9

Notes:

1. Non-uniformly distributed inputs lead to poor approximations
2. Large numbers of inputs are necessary for accurate results.

*given a particular input, will always produce the same output.

Monte Carlo Integration - mathematics

Consider a multidimensional indefinite integral (though it doesn't have to be)

$$I = \int_\Omega f(\overline{\mathbf{x}})\, d\overline{\mathbf{x}} \quad \text{with volume} \quad V = \int_\Omega d\overline{\mathbf{x}}$$

The simple Monte Carlo approach is to sample points uniformly in Ω. Thus, with N uniform samples:

$$\overline{\mathbf{x}}_1, \cdots, \overline{\mathbf{x}}_N \in \Omega \quad \text{we approximate} \quad I \approx Q_N \equiv V \frac{1}{N} \sum_{i=1}^{N} f(\overline{\mathbf{x}}_i) = V\langle f \rangle$$

Which holds thanks to the law of large numbers,

$$\lim_{N\to\infty} Q_N = I.$$

Monte Carlo Integration - an example; G&T 3.60

	x_i, y_i		x_i, y_i
1	0.984, 0.246	6	0.637, 0.581
2	0.860, 0.132	7	0.779, 0.218
3	0.316, 0.028	8	0.276, 0.238
4	0.523, 0.542	9	0.081, 0.484
5	0.349, 0.623	10	0.289, 0.032

Table 3.6: A sequence of ten random pairs of numbers (see Problem 3.60).

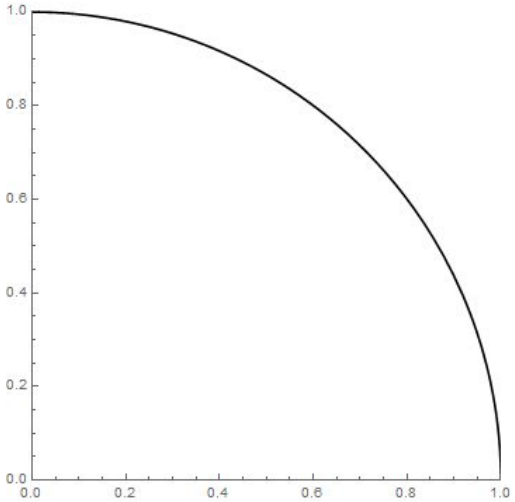
Problem 3.60. Monte Carlo integration

Consider the ten pairs of numbers (x_i, y_i) given in Table 3.6. The numbers are all in the range $0 < x_i, y_i \leq 1$. Imagine that these numbers were generated by counting the clicks generated by a Geiger counter of radioactive decays, and hence they can be considered to be a part of a sequence of random numbers. Use this sequence to estimate the magnitude of the integral

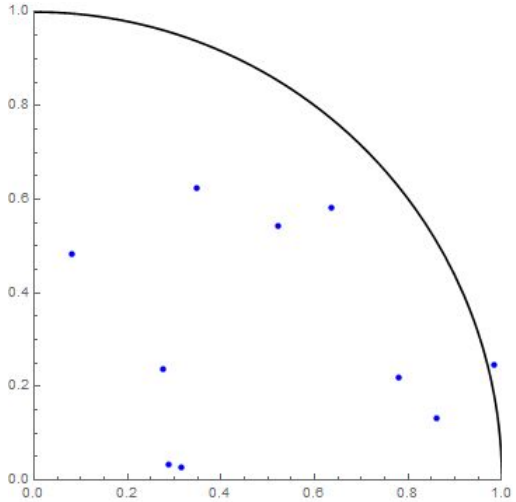
$$F = \int_0^1 dx \sqrt{(1 - x^2)}. \tag{3.194}$$

If you have been successful in estimating the integral in this way, you have found a simple version of a general method known as *Monte Carlo integration*.[14]

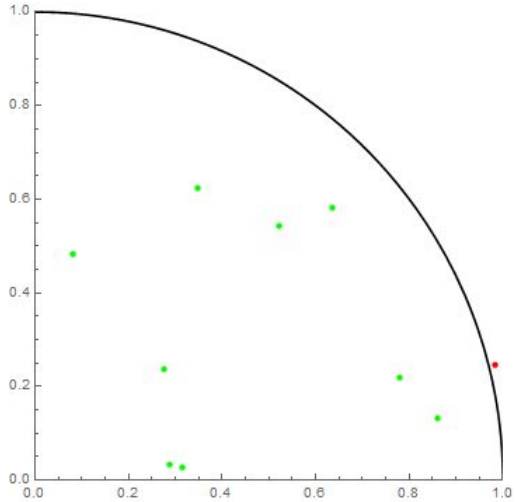
$$F = \int_0^1 dx \sqrt{(1 - x^2)}$$



Step 1: plot a function in some domain (here it's [0,1],[0,1])



Step 2: plot pairs of random numbers.



Step 3: count how many are underneath our function.

Step 4: (total area of domain) x (fraction of points in function) = Monte Carlo Integral! In this case, our estimate is F=.9 (bad)

Monte Carlo Integration - an example; G&T 3.60

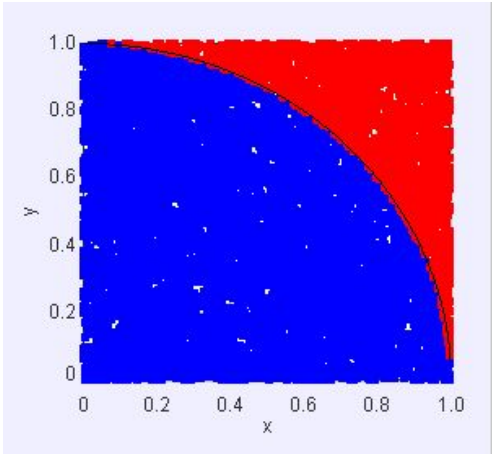
(a) Show analytically that the integral in (3.194) is equal to $\pi/4$.

(b) Use Program `MonteCarloEstimation` to estimate the integral (3.194) by Monte Carlo integration. Determine the error (the magnitude of the deviation from the exact answer) for trials of n pairs of points equal to $n = 10^4$, 10^6, and 10^8. Does the error decrease with increasing n on the average?

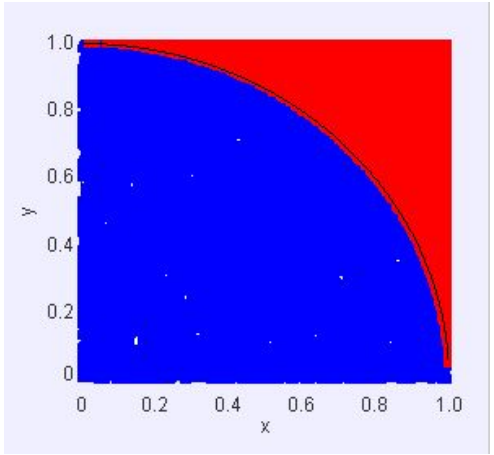
(c) Estimate the integral using $n = 1000$. Repeat for a total of ten trials using a different random number seed each time. The easiest way to do so is to press the Reset button and then press the Calculate button. The default is for the program to choose a new seed each time based on the clock. Is the magnitude of the variation of your values of the same order as the error between the average value and the exact value? For a large number of trials, the error is estimated from the standard error of the mean, which approximately equals the standard deviation divided by the square root of the number of trials. □

Monte Carlo Integration - an example; G&T 3.60

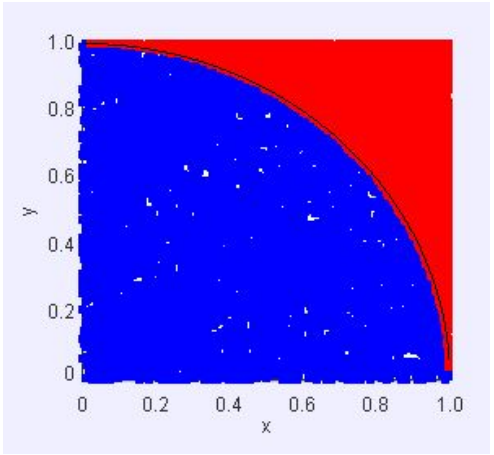
A) → To the board!
B)



$N = 10^4$, error = .0046



$N = 10^6$, error = .0003



$N = 10^8$, error = 0 (!)

Yes! As *n* increases, the error decreases across the board! Note: the third estimation took a non-trivial amount of time.

Monte Carlo Integration - an example; G&T 3.60

C)

											Mean
Value	.785	.776	.787	.773	.788	.790	.795	.780	.790	.781	.7845
Error	.0004	.0047	-.0016	.0124	-.0026	-.0046	-.0090	.0054	-.0046	.0044	.0009

No! The error for the average value is one order of magnitude smaller on average.

Monte Carlo Integration

An extension of G&T 3.60… → Mathematica!

Next Level MC integration— Importance Sampling

If we consider our integral, we can increase the accuracy of our integration by placing more points where the integrand is large.

- Define a weight function *w(x)* that tells us which regions are significant.

Thus, we redefine $I = \int_V d\mathbf{x} \frac{f(\mathbf{x})}{w(\mathbf{x})} w(\mathbf{x}).$ Now, $\overline{I} = \frac{1}{N} \sum_{i=1}^{N} \frac{f(\mathbf{x}_i)}{w(\mathbf{x}_i)}$

For example, if our integrand has a large peak around $x=x_0$, we might use a Gaussian weight: $w(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_0)^2}{2\sigma^2}}$

Thanks to appropriately placed sampling points, this Monte Carlo integration will require less computational power for better estimates.

Sources:

http://www.chem.utoronto.ca/~jmschofi/simulation/partintegration.pdf

https://en.wikipedia.org/wiki/Monte_Carlo_method

http://mathworld.wolfram.com/MonteCarloIntegration.html

https://www.scratchapixel.com/lessons/mathematics-physics-for-computer-graphics/monte-carlo-methods-in-practice/monte-carlo-integration