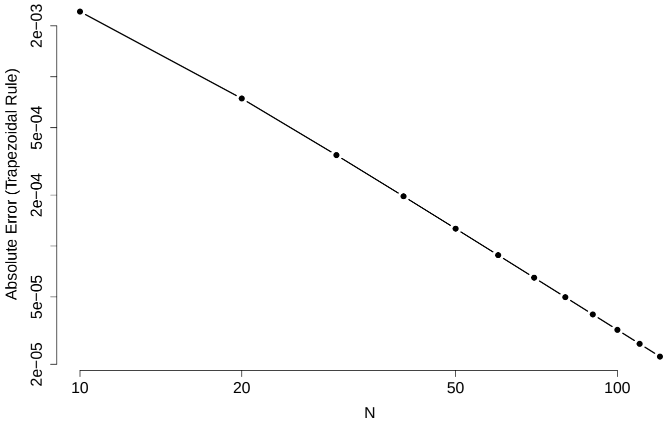
Monte Carlo Integration 2

Mihir Arjunwadkar

Centre for Modeling and Simulation
Savitribai Phule Pune University

Why use Monte Carlo integration at all?

Apply Trapezoidal Rule to $I = \frac{1}{\sqrt{2\pi}} \int_3^\infty e^{-x^2/2} dx$:



- Trapezoidal Rule: Absolute error of $\sim 1 \times 10^{-5}$ with only about 100 points.
- Gaussian Quadrature: far better approximation in far fewer # of points!

Then why use Monte Carlo integration at all?

Error behaviour: Numerical integration in 1 dimension

Consider approximating integral of a function f of a *single* variable x:

$$I = \int_a^b g(x)dx$$

Error in the numerical approximation:

$$\epsilon \propto \delta^k,$$

where the grid spacing

$$\delta = (b - a)/n$$

and k (*order* of the integration method) is 2 for trapezoidal, 3 or 4 for Simpson, etc.

Because

$$\delta \propto n^{-1},$$

we have

$$\epsilon \propto n^{-k}.$$

Error behaviour: Numerical integration in d dimensions

Consider approximating integral of a function f of d variables $x_1, \ldots, x_d$:

$$I = \int \ldots \int g(x_1, \ldots, x_d) dx_1, \ldots, dx_d$$

Error in the numerical approximation (trapezoidal, Simpson, etc.)

$$\epsilon \propto \delta^k.$$

Now, the grid size along any dimension is

$$\delta^d \propto n^{-1}, \text{ that is, } \delta \propto n^{-1/d}.$$

Hence

$$\epsilon \propto n^{-k/d}.$$

Larger the d, slower the convergence! Curse of Dimensionality

Monte Carlo integration in d dimensions

Multi-dimensional integral

$$\begin{aligned} I &= \int \dots \int g(x_1, \dots, x_d) dx_1, \dots, dx_d \\ &= \int \dots \int h(x_1, \dots, x_d) f(x_1, \dots, x_d) dx_1, \dots, dx_d \end{aligned}$$

with

$$\begin{aligned} f(x_1, \dots, x_d) &\geq 0 \\ \int \dots \int f(x_1, \dots, x_d) dx_1, \dots, dx_d &= 1 \\ h(x_1, \dots, x_d) f(x_1, \dots, x_d) &= g(x_1, \dots, x_d) \end{aligned}$$

for each $(x_1, \dots, x_d)$.

Algorithm

1. Generate $(X_1^{(1)}, X_2^{(1)}, \ldots, X_d^{(1)}), \ldots, (X_1^{(n)}, X_2^{(n)}, \ldots, X_d^{(n)}) \sim f$
2. Estimator I as $\widehat{I}_n = \frac{1}{n} \sum_{i=1}^{n} h(X_1^{(i)}, \ldots, X_d^{(i)})$
3. $\widehat{\text{Var}}\left(\widehat{I}_n\right) = \frac{1}{n} \times \frac{1}{n-1} \sum_{i=1}^{n} \left(h(X_1^{(i)}, \ldots, X_d^{(i)}) - \widehat{I}_n \right)^2$
4. Etc.

Error behaviour: Monte Carlo integration in d dimensions

Because we *estimate* the value of I
the error in the estimate is expected to be

$$O(n^{-1/2})$$

independent of the # of dimensions!

Error behaviour: Monte Carlo integration in d dimensions

For any k (i.e., order of the numerical integration method), for sufficiently large number d of dimensions, we will have

$$k/d < 1/2.$$

This means that, beyond this value of d, the error in numerical approximation will be more than that in the Monte Carlo estimate.

more pointers here

Monte Carlo integration is therefore particularly useful when dealing with high-dimensional integrals. High-dimensional integrals often occur in statistical physics, Bayesian inference, etc.

Toy example

Estimating / approximating
volume of d-dimensional unit hypersphere

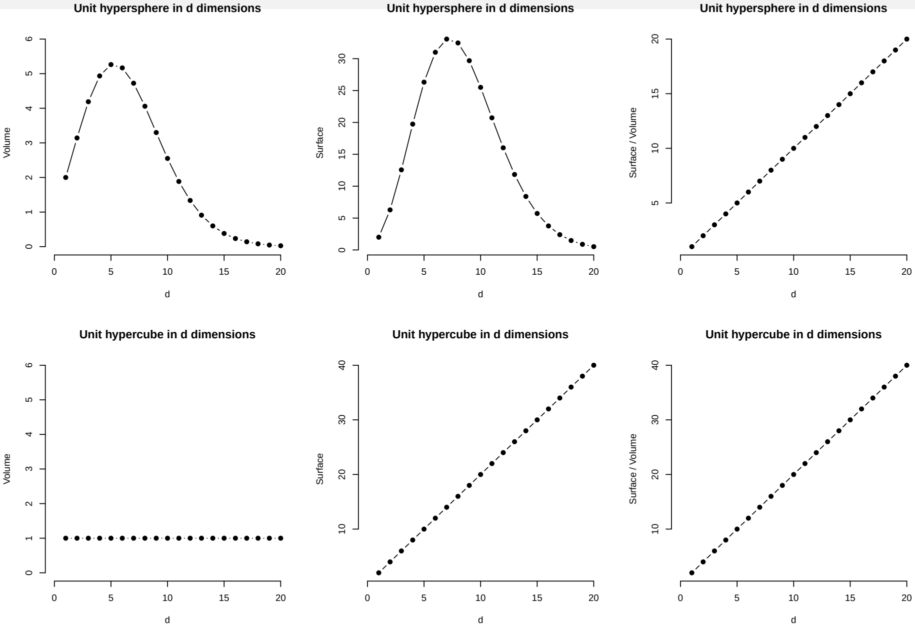
Statutory Warning

High-dimensional spaces and objects
can be
strange and non-intuitive

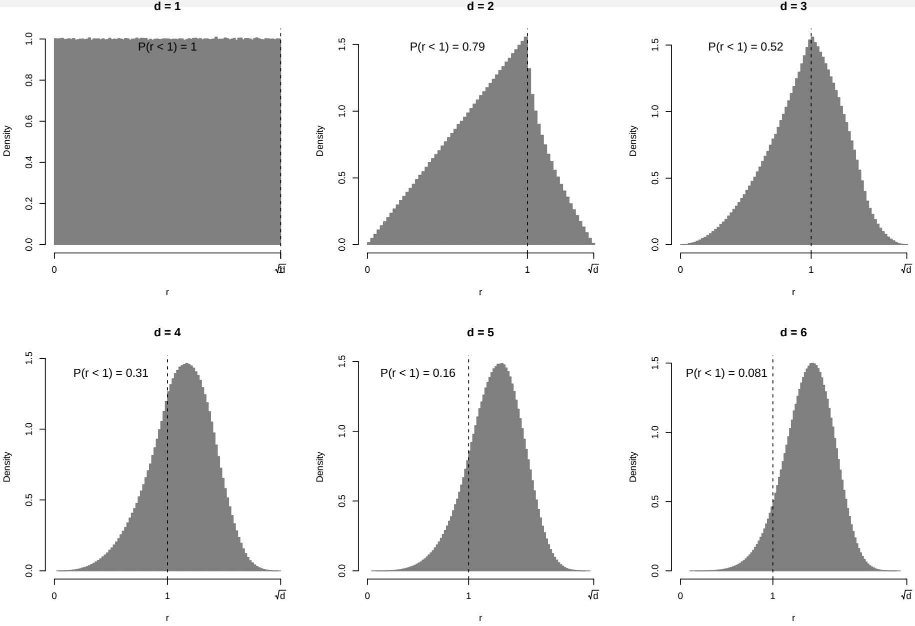
Hypershperes and hypercubes in d dimensions

	Volume	Surface
Hypercube	L^d	$2dL^{d-1}$
Hypersphere	$\frac{\pi^{\frac{d}{2}}}{\Gamma\left(1+\frac{d}{2}\right)} R^d$	$\frac{\pi^{\frac{d}{2}}}{\Gamma\left(1+\frac{d}{2}\right)} dR^{d-1}$

Hypershperes and hypercubes in *d* dimensions



Distance distribution inside unit hypercubes



20-dimensional watermelons

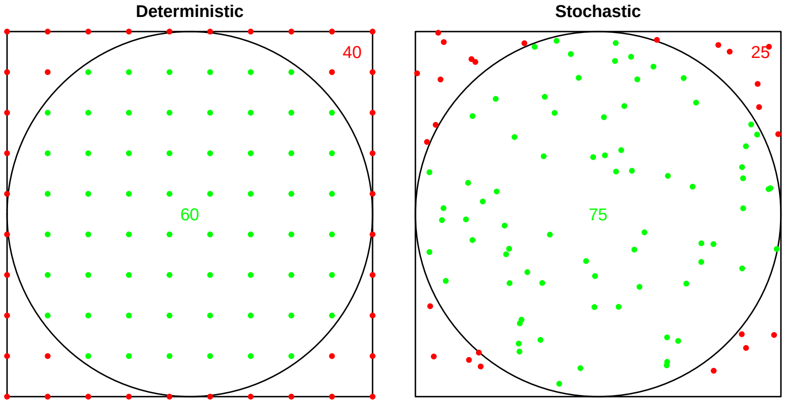
5.14. A property of the n-dimensional volume. It consists in the fact that for very large n the "*volume of an n-dimensional figure is concentrated near its surface.*" For example, the volume of the spherical ring between spheres of radius 1 and $1-\epsilon$ equals $b_n[1-(1-\epsilon)^n]$, which, for fixed arbitrarily small ϵ, but increasing n approaches b_n. A 20-dimensional watermelon with a radius of 20 cm. and a skin with a thickness of 1 cm. is nearly two-thirds skin:

$$1-\left((1-\frac{1}{20}\right)^{20}\approx 1-e^{-1}.$$

p.124
Kostrikin & Manin
Linear Algebra & Geometry
Gordon & Breach Science Publishers (1989?)

Courtesy: Prof. Anil Gangal

Estimating volume: deterministic & stochastic dartboards



$$\text{Area approximation or estimate} = 4 \times \frac{\#\text{ of points inside } \bigcirc}{\#\text{ of points inside } \square}.$$

Volume of d-dimensional unit hypersphere: MCI

Formally,

$$\begin{aligned} V(d) &= 2^d \int_0^1 \dots \int_0^1 h(x_1, \dots, x_d) dx_1 \dots dx_d \\ &= 2^d \int \dots \int h(x_1, \dots, x_d) f(x_1, \dots, x_d) dx_1 \dots dx_d, \end{aligned}$$

where

$$\begin{aligned} h(x_1, \dots, x_d) &= \begin{cases} 1 & \sum_{i=1}^{d} x_i^2 \leq 1 \\ 0 & \text{otherwise} \end{cases} \\ f(x_1, \dots, x_d) &= \begin{cases} 1 & 0 \leq x_1, \dots, x_d \leq 1 \\ 0 & \text{otherwise} \end{cases}. \end{aligned}$$

$f(x_1, \dots, x_d)$ :: uniform PDF over d-dimensional unit hypercube

Algorithm

1. Generate $(X_1^{(1)}, X_2^{(1)}, \ldots, X_d^{(1)}), \ldots, (X_1^{(n)}, X_2^{(n)}, \ldots, X_d^{(n)}) \sim f$
2. Estimator $V(d)$ as $\widehat{V}_n(d) = \frac{1}{n} \sum_{i=1}^{n} h(X_1^{(i)}, \ldots, X_d^{(i)})$
3. $\widehat{\text{Var}}\left(\widehat{V}_n(d)\right) = \frac{1}{n} \times \frac{1}{n-1} \sum_{i=1}^{n} \left(h(X_1^{(i)}, \ldots, X_d^{(i)}) - \widehat{V}_n(d)\right)^2$
4. Etc.

Effectively, step 2 yields

$$\widehat{V}_n(d) = 2^d \times \frac{\#\text{ of points inside } \bigcirc}{\#\text{ of points inside } \square}.$$

To do ...

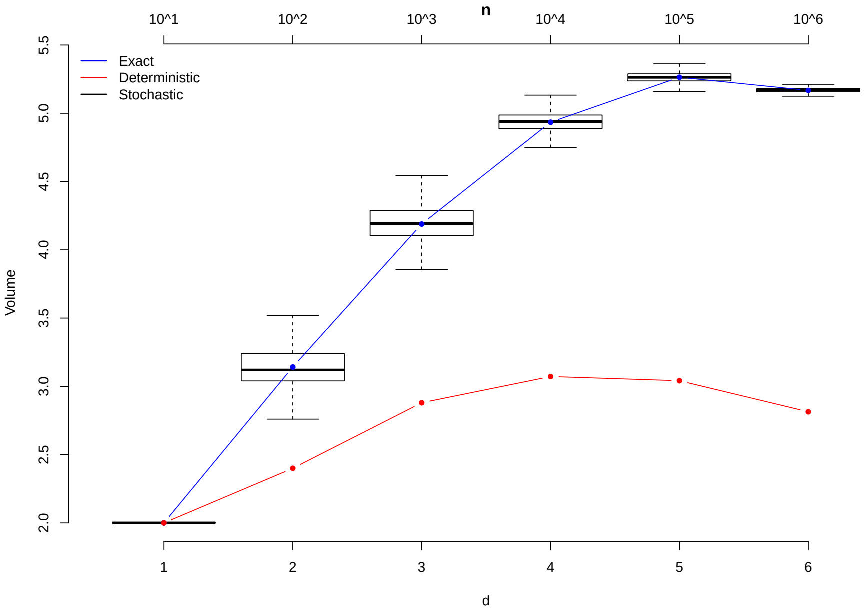
Let us now

- apply this MCI estimator to the deterministic and stochastic grids;
- compare results the exact volume volume of a d-dimensional unit hypersphere:

$$V(d) = \frac{\pi^{\frac{d}{2}}}{\Gamma\left(1 + \frac{d}{2}\right)};$$

- compare stochastic and deterministic results for the same n and d.

Deterministic vs. stochastic dartboard comparison



Hypersphere volume estimation using MCI: A surprise

$\widehat{\mathrm{SE}}\left(\widehat{V}_n(d)\right) \approx d^2$ :: not constant :: not clear why