Probability and Statistics

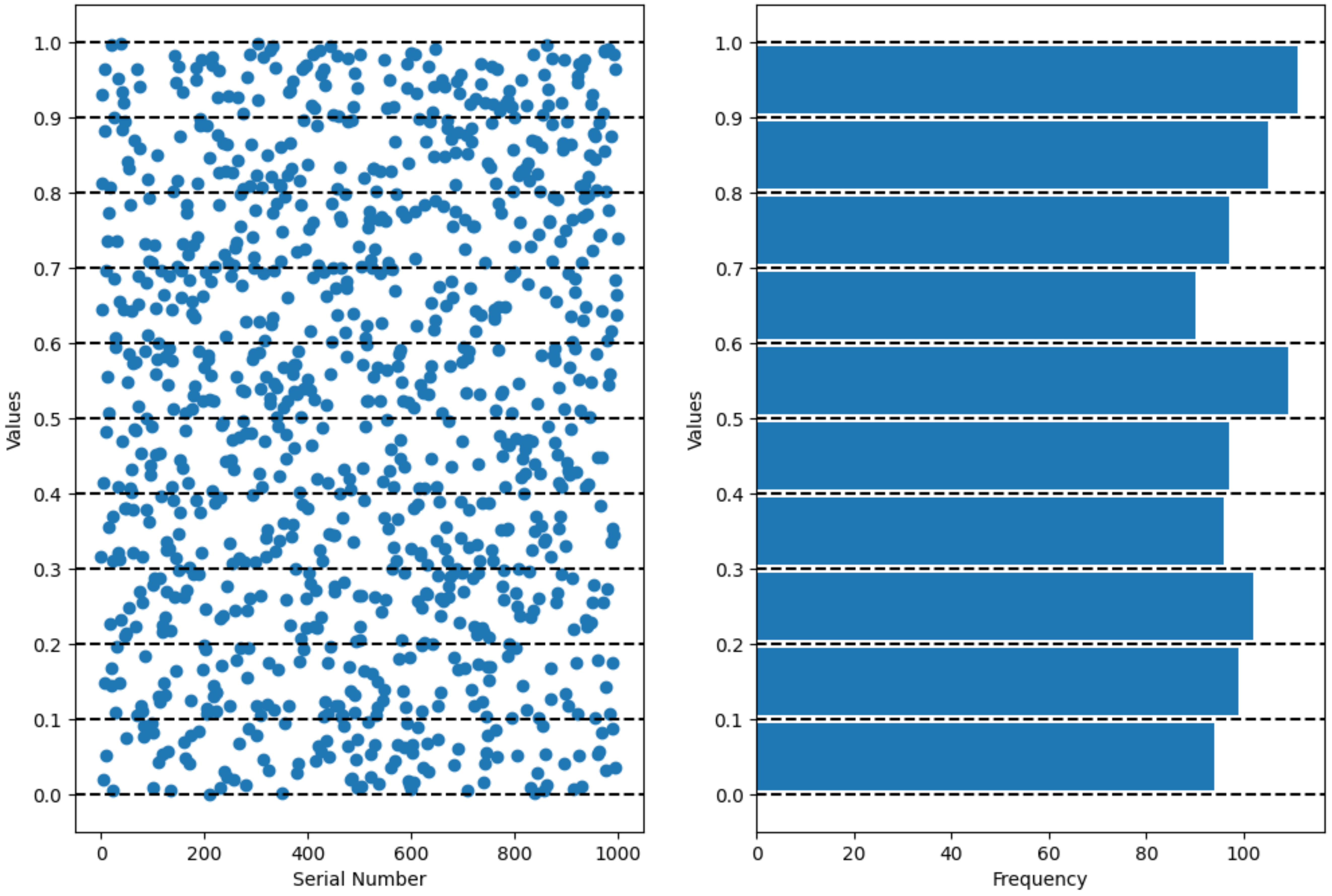
Part-4

MA2103 - 2023

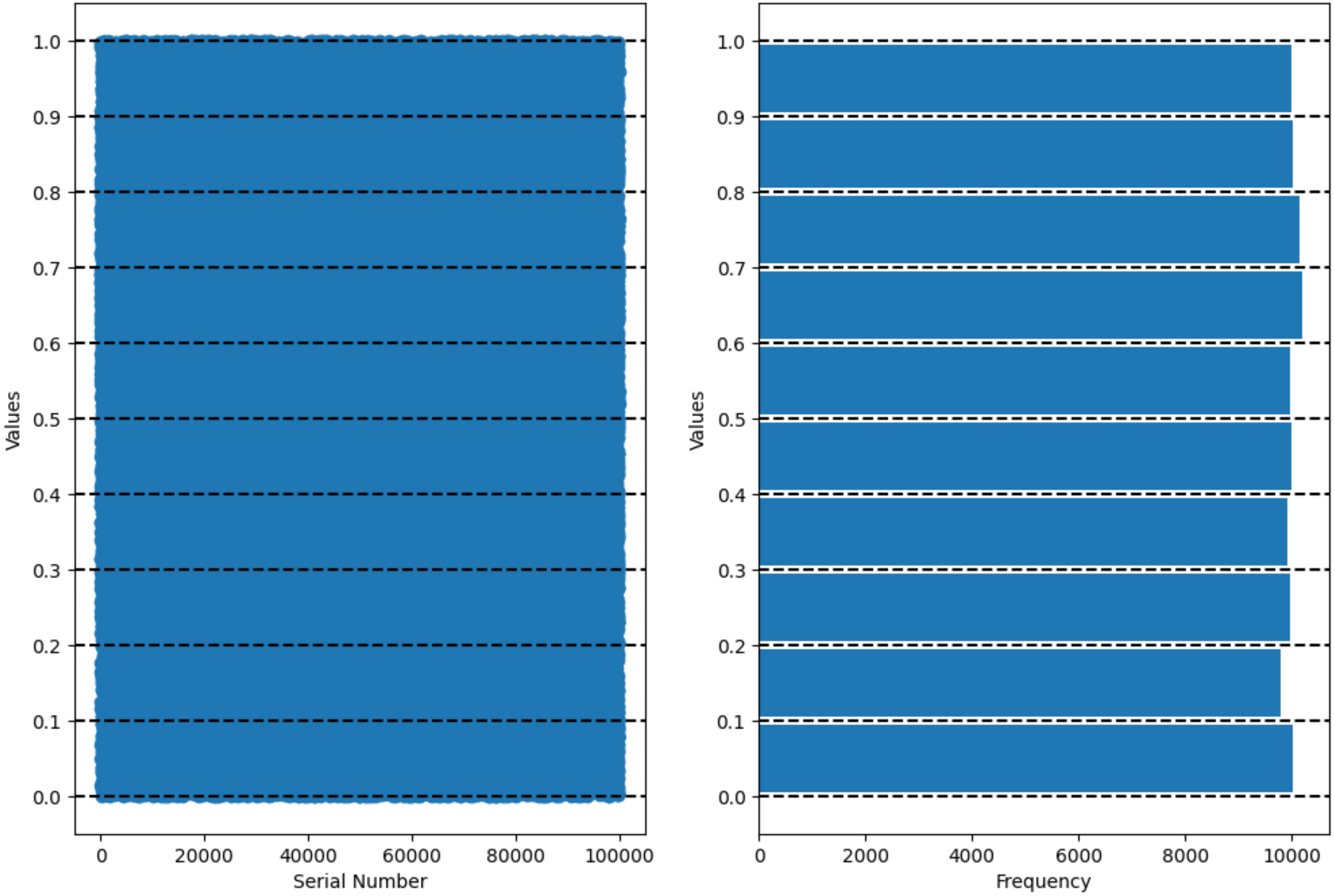
Rajesh Kumble Nayak

Uniform distribution

A random number with uniform distribution in the interval $[a, b]$, has equal probability of finding any number in the interval $[a, b]$



As number of random numbers become more and more, they becomes more or less uniform!



Back to Sampling

Let's look at a function $f(x) = e^{-x^2}$ in the interval $[0,2]$ and sample using two methods!

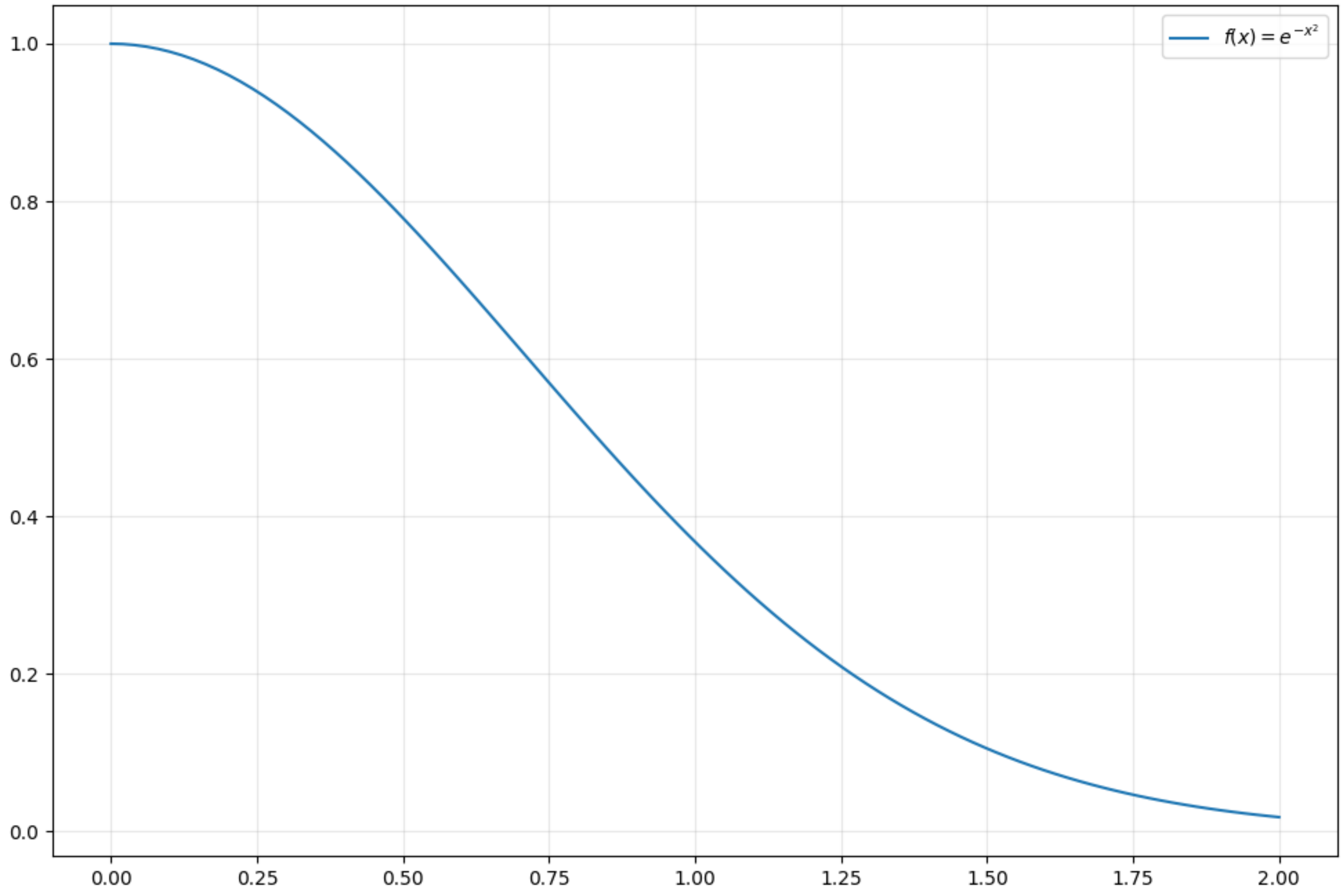
Let's say we want to find the average of this function!

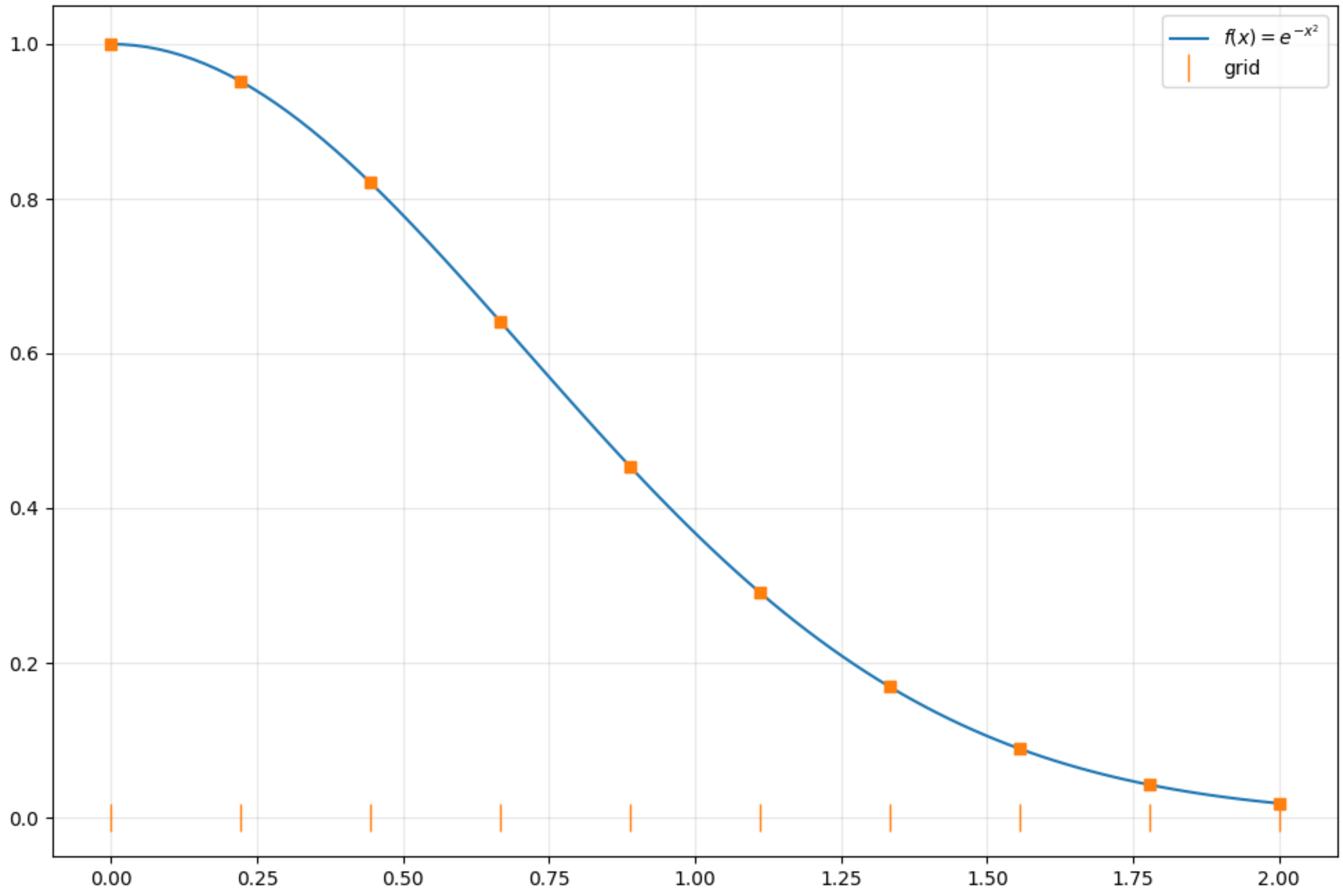
✤Grid based, i.e at uniform interval

✤Random sampling with uniform random numbers in the $[0,2]$

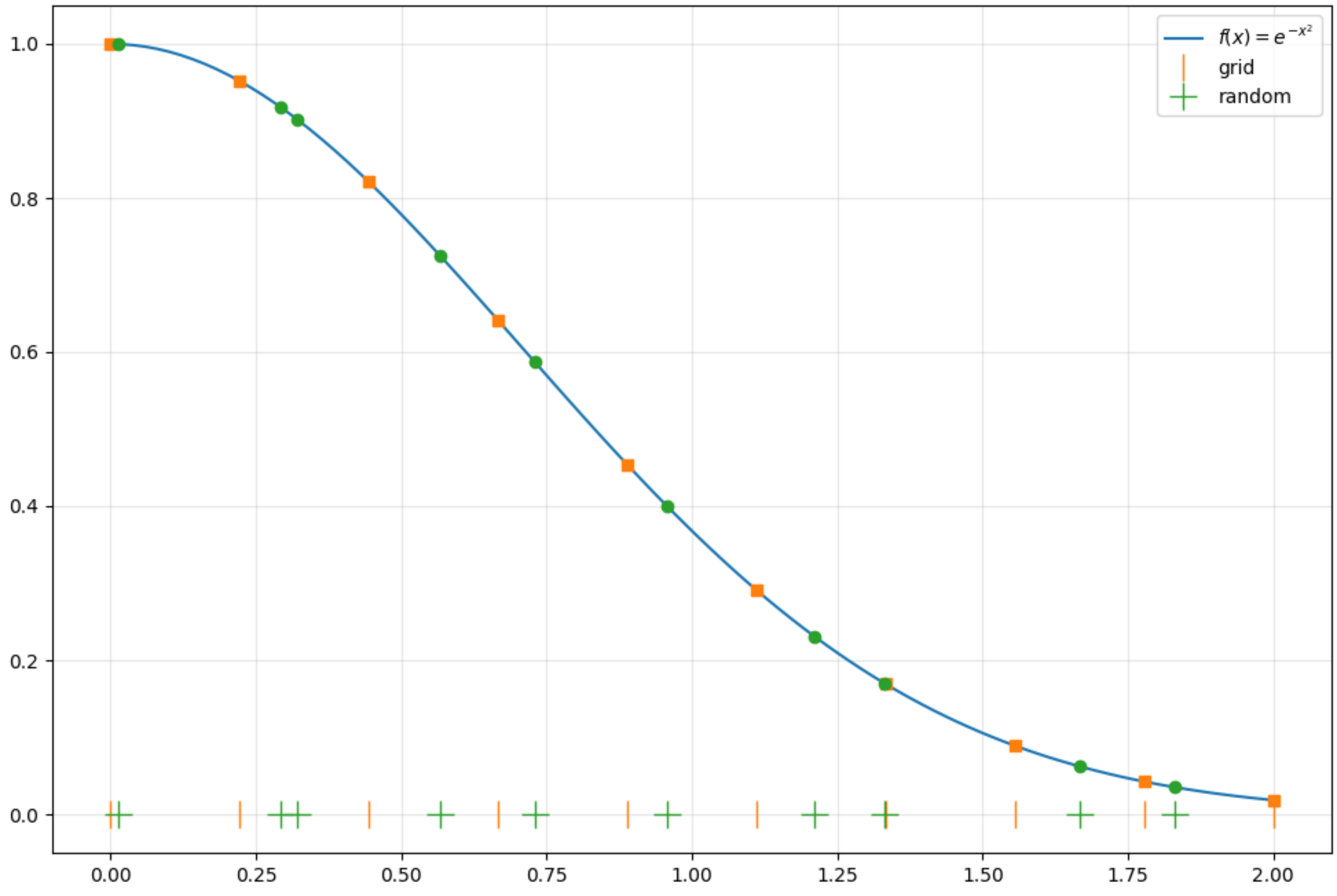
Analytically, we can integrate this function and the average is

$$A = \frac{1}{2}\int_0^2 e^{-x^2}dx = 0.441041$$


$f(x) = e^{-x^2}$
1.0
0.8
0.6
0.4
0.2
0.0
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00



$f(x) = e^{-x^2}$
grid
1.0
0.8
0.6
0.4
0.2
0.0
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00



$f(x) = e^{-x^2}$
grid
random
1.0
0.8
0.6
0.4
0.2
0.0
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00

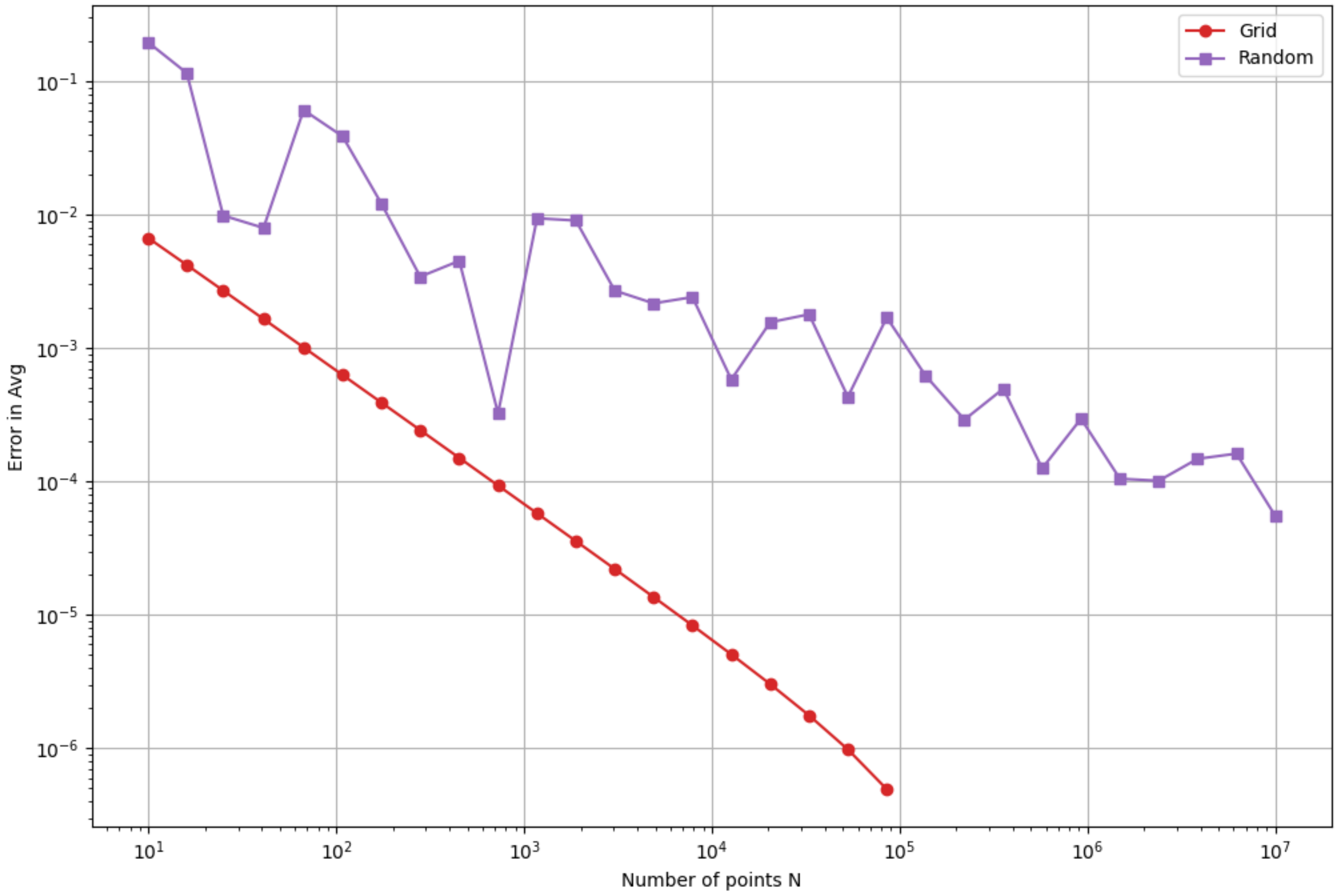
Let's say we want to find the average of this function!

$$A = \frac{1}{2}\int_0^2 e^{-x^2} dx = 0.441041$$

$$A_{grid} = \frac{1}{N}\sum_{i=0}^{N} f(x_n) = 0.4477, \quad x_n = \frac{2n}{N}$$

$$A_{random} = \frac{1}{N}\sum_{i=0}^{N} f(x_n) = 0.48540, \quad x_n \text{uniform random number} \, [0,2]$$

As N increases, the results get more and more accurate



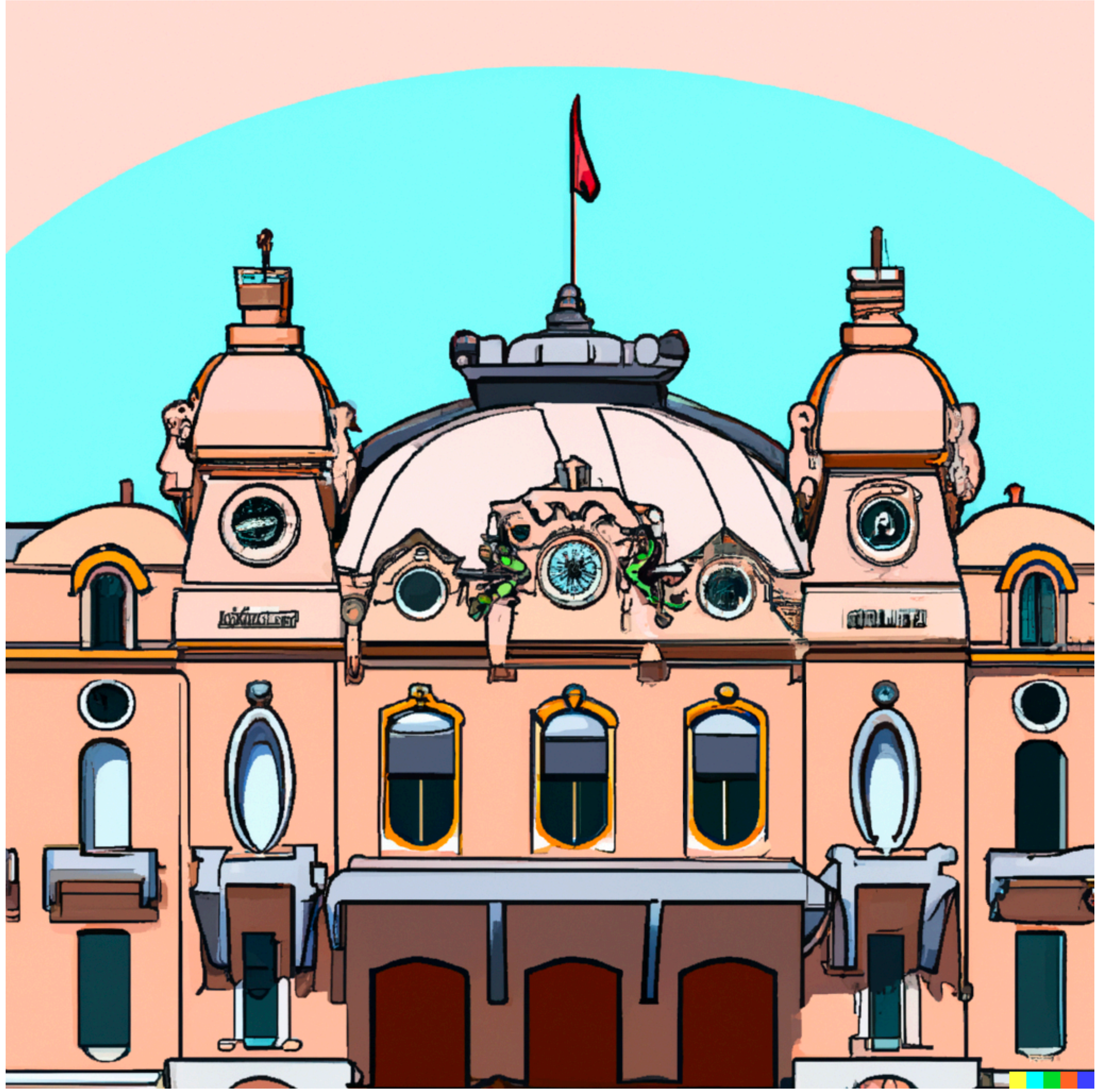
Grid
Random
Error in Avg
Number of points N
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{1}
10^{2}
10^{3}
10^{4}
10^{5}
10^{6}
10^{7}

Summary

★Our results summaries that grid based sampling works very well

★Random sampling with uniform distribution never take over gird based method at least for function of single variable

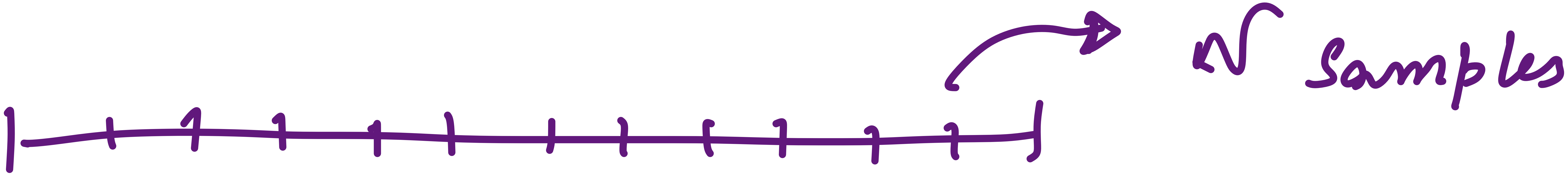
★ Random sampling is called `Monte-Carlo' m



What about Multivariable functions

When we are interested to compute mean, for function of 2 variables, it is like computing area.

Errors can be taken as variance, it is valid for both uniform sampling and random sampling

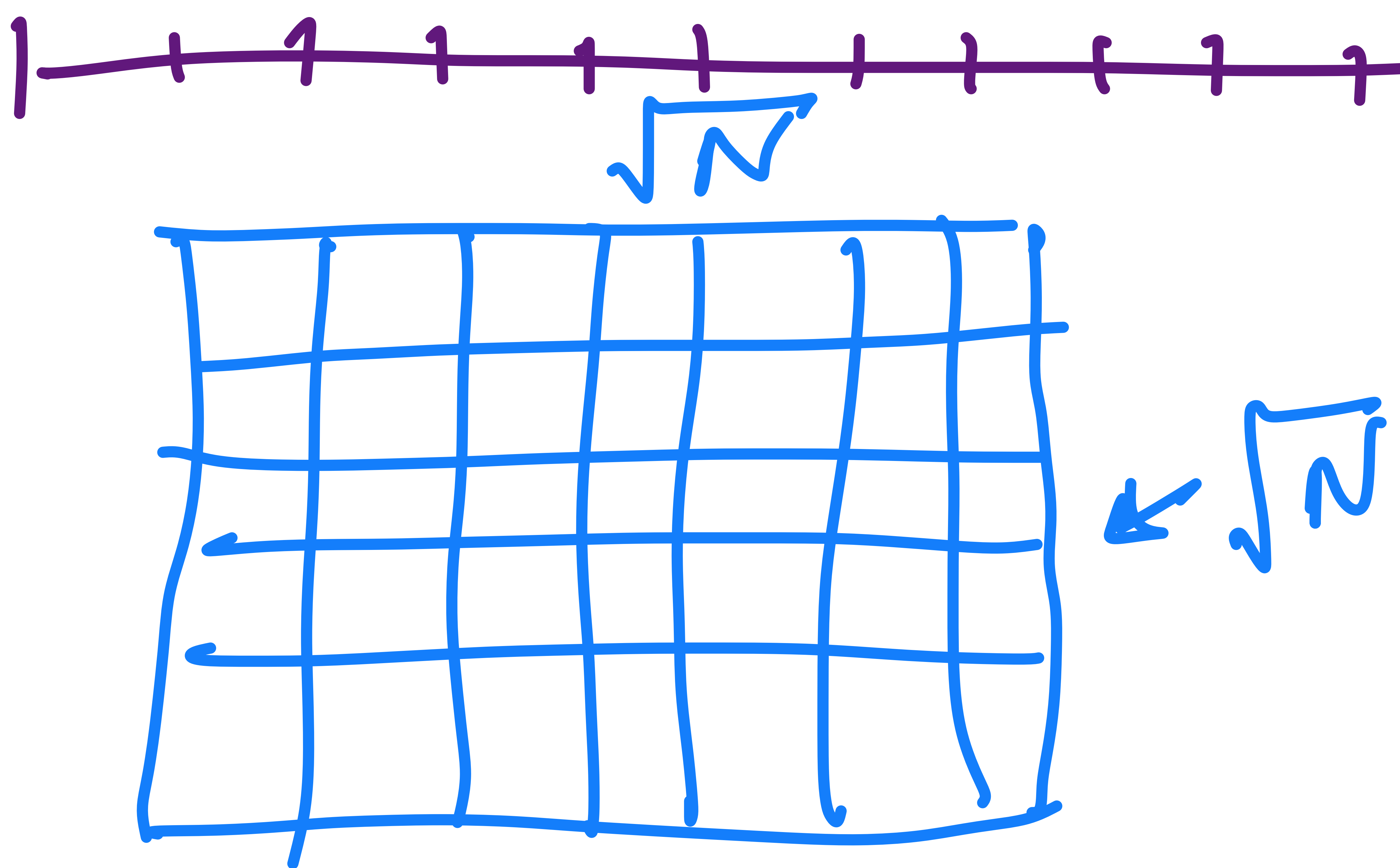


What about Multivariable functions

When we are interested to compute mean, for function of 2 variables, it is like computing area.

Errors can be taken as variance, it is valid for both uniform sampling and random sampling

N samples



effectively we have $\sqrt{N}$ point per dimension

What about Multivariable functions

When we are interested to compute mean, for function of 2 variables, it is like computing area.

Errors can be taken as variance, it is valid for both uniform sampling and random sampling

As number of variable or the dimension of the parameter space increases the Monte-Carlo method become more and more effective.

However, uniform sampling might not always efficient way of sampling. But it gives a good starting point.

Same is true in the Statistical sampling. Often statistical hypothesis has large number of parameters or dimension and Monte-Carlo method is often useful