Lecture 8: The Monte Carlo method for estimating expectations

Professor Ilias Bilionis

The uncertainty propagation problem

PREDICTIVE
SCIENCE LABORATORY

The uncertainty propagation problem

- You are given a function $g(x)$ representing a physical model.

  x: input , $g(x)$: output

- The inputs of the model are uncertain.

- You represent this uncertainty with a random variable:

$$X \sim p(x)$$

- You would like to quantify your uncertainty about the model output:

$$\boxed{Y = g(X)}$$

PREDICTIVE SCIENCE LABORATORY

The uncertainty propagation problem

- We would like to estimate the expected value of the output:

$$\mathbb{E}[Y] = \mathbb{E}[g(X)] = \int g(x)\, p(x)\, dx$$

The uncertainty propagation problem

- We would like to estimate the variance of the output:

$$\mathbb{V}[Y] = \int (g(x) - \mathbb{E}[Y])^2 p(x) dx$$

$$= \mathbb{E}[[g(X)]^2] - (\mathbb{E}[g(X)])^2$$

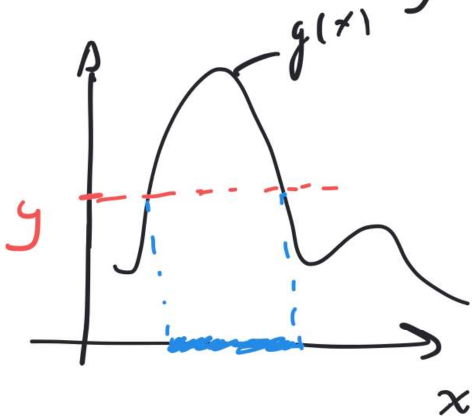
$$\mathbb{E}[[g(X)]^2] = \int g^2(x) p(x) dx$$

The uncertainty propagation problem

- Or maybe the probability that the output exceeds a threshold:

$$p(Y \geq y) = \int 1_{[y,\infty)}(g(x))\,p(x)\,dx = \mathbb{E}\left[1_{[y,\infty)}(g(X))\right]$$

$$= \mathbb{E}\left[1_{[y,\infty)}(Y)\right]$$



$$1_{[y,\infty)}(g(x)) = \begin{cases} 1, & \text{if } g(x) \geq y \\ 0, & \text{otherwise} \end{cases}$$

↘ The indicator function of the set $[y,\infty)$

PREDICTIVE SCIENCE LABORATORY

The uncertainty propagation problem

- Notice that all these statistics are essentially expectations of functions of X.

- We must learn how to do such integrals!

PREDICTIVE
SCIENCE LABORATORY

Lecture 8: The Monte Carlo method for estimating expectations

Professor Ilias Bilionis

The curse of dimensionality

PREDICTIVE
SCIENCE LABORATORY

The curse of dimensionality

$\underline{X} = (X_1, \dots, X_d)$, $X_i \sim U([0,1])$ independent

$$p(\underline{x}) = \prod_{i=1}^{d} p(x_i) = \prod_{i=1}^{d} 1_{[0,1]}(x_i) = 1_{[0,1]^d}(\underline{x})$$

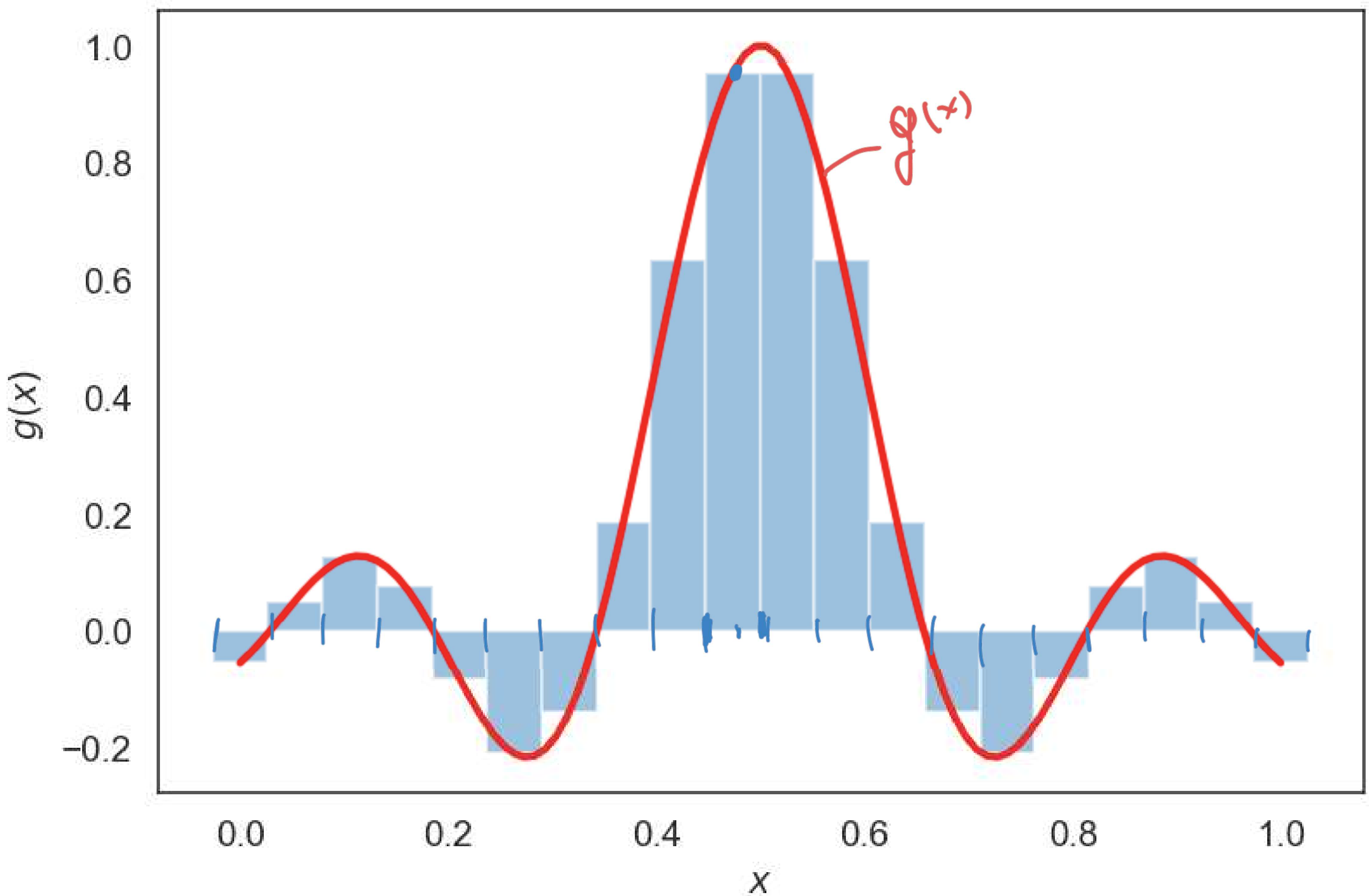
- Take the d-dimensional uniform: $X \sim U([0,1]^d)$.

- Take a function $g(x)$.

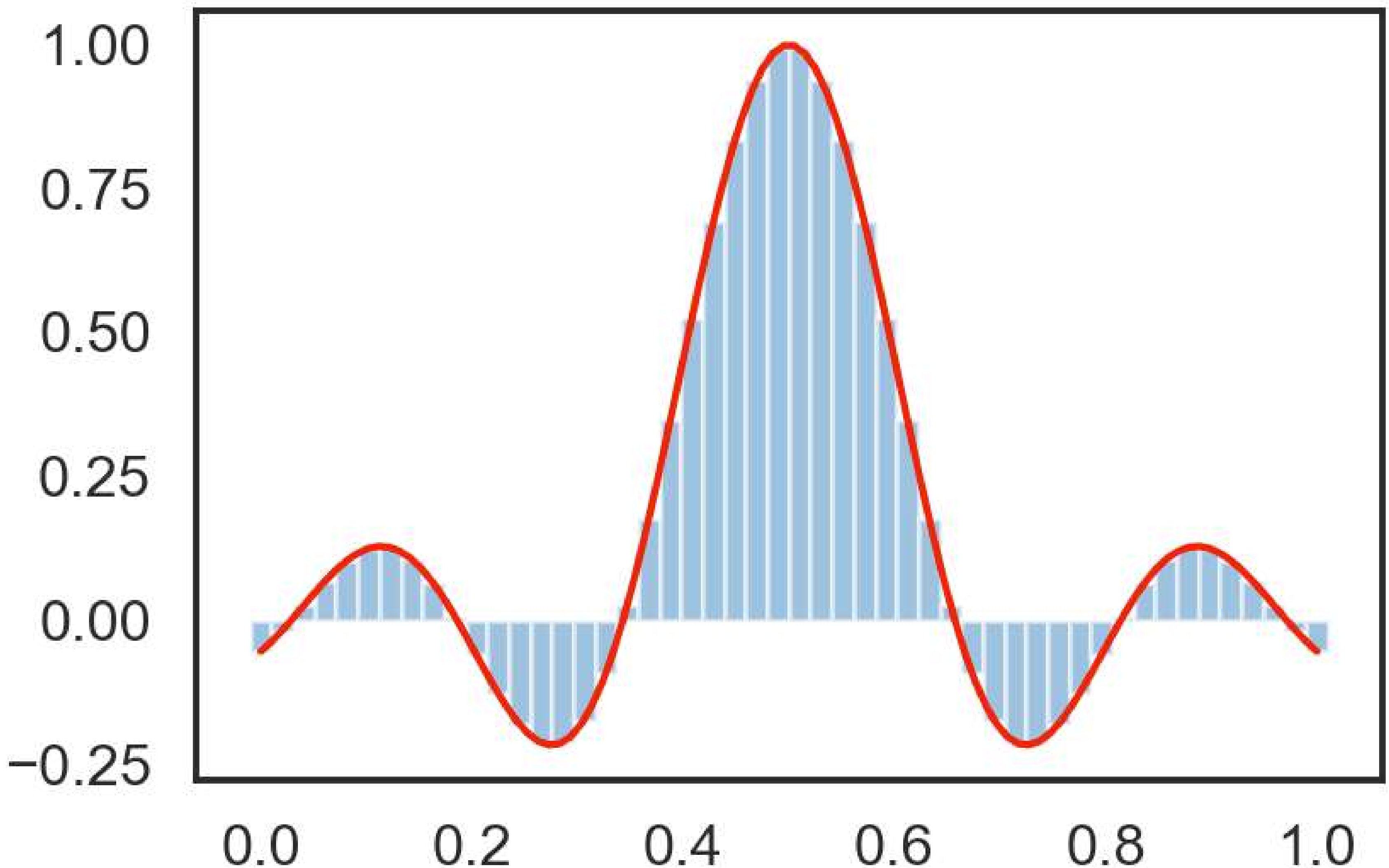
- We would like to estimate:

$$\boxed{\mathbb{E}[g(X)] = \int g(x)\underline{p(x)}dx}$$

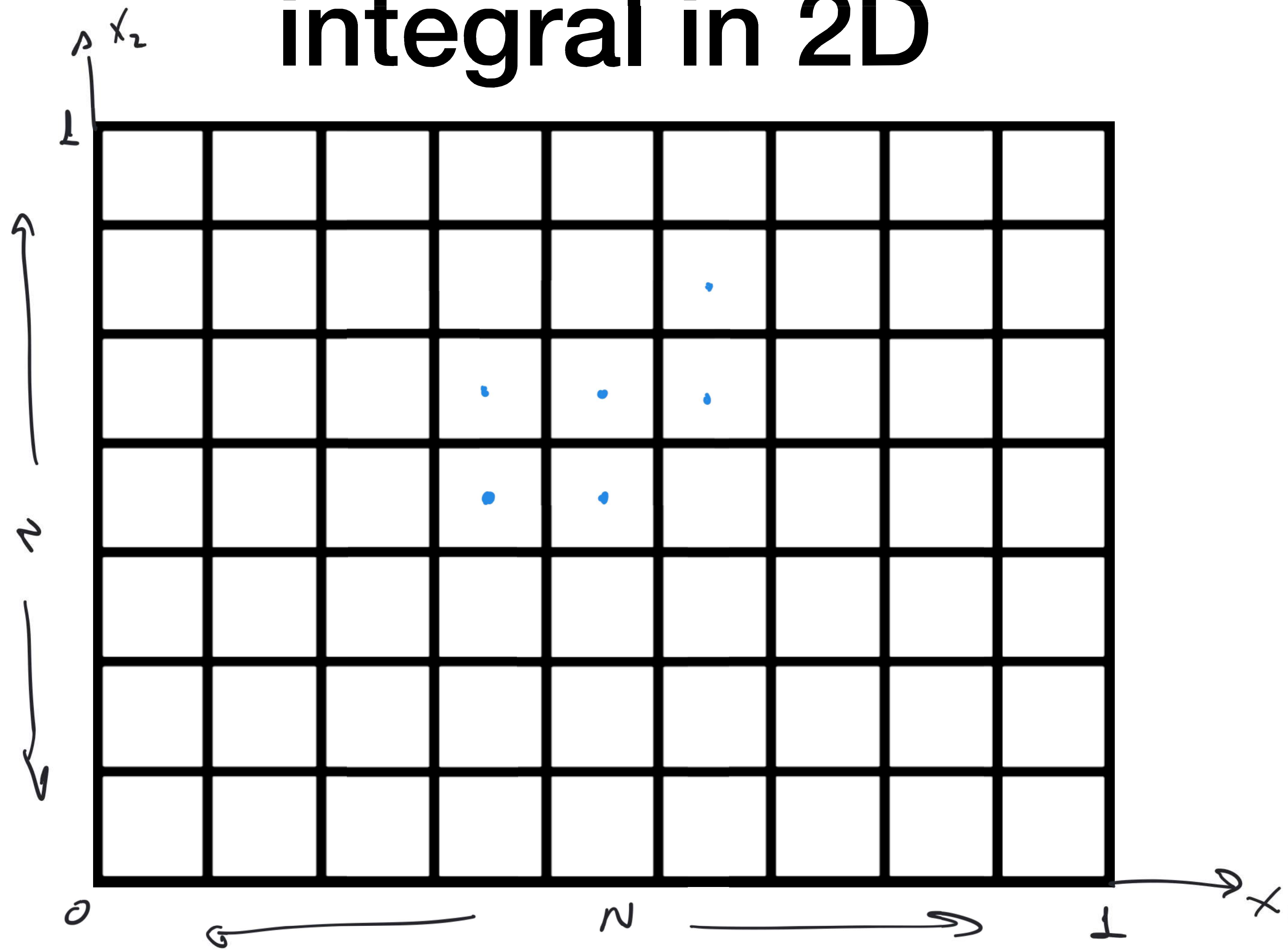
Example: Evaluating integral in 1D



Example: Evaluating integral in 1D



Example: Evaluating integral in 2D



PREDICTIVE SCIENCE LABORATORY

The curse of dimensionality

- Use n equidistant points per dimension.
- You will have n^d boxes each with volume n^{-d}.
- You can evaluate the integral by:

$$\boxed{\mathbb{E}[g(X)] \approx n^{-d} \sum_{j=1}^{n^d} g(x_{c,j})}$$

The Curse of dimensionality

n^d

- Assume it takes a millisecond to evaluate the function. $g(x)$
- Take $n = 10$ points per dimension.
- d=2, needs 0.1 seconds.
- d=3, needs 1 second.
- d=5, needs 100 seconds.
- d=6, needs, 1000 seconds or 16 minutes.
- d=10, needs 115 days…
- d=20, needs 3.17 billion years

Lecture 8: The Monte Carlo method for estimating expectations

Professor Ilias Bilionis

The law of large numbers

PREDICTIVE
SCIENCE LABORATORY

The strong law of large numbers

- Take an infinite series of independent random variables $X_1, X_2, \ldots$ with the same distribution (it doesn't matter what distribution).

- The *sample average*: $\frac{X_1 + \ldots + X_N}{N} \longrightarrow \mu$

$$\bar{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i \rightarrow \mu \text{ a.s.}$$

almost surely (measure theory)

where $\boxed{\mu = \mathbb{E}[X_i]}$ (as $N \rightarrow \infty$).

The Monte Carlo method for estimating integrals

- Take a random variable $X \sim p(x)$ and some function $g(x)$.
- We want to estimate the expectation:

$$\boxed{I = \mathbb{E}[g(X)] = \int g(x)p(x)dx}$$

- Make independent identical copies of X: $X_1, X_2, \dots \sim p(x)$
- Consider the also the independent iden. dist.:

$Y_1 = g(X_1), Y_2 = g(X_2), \dots$

- By the strong law of large numbers:

$$I_N = \frac{Y_1 + \dots + Y_N}{N} \longrightarrow \mathbb{E}[Y_i] = \mathbb{E}[g(X_i)] = I \text{ a.s.}$$

PREDICTIVE SCIENCE LABORATORY

Example: 1D expectation

(This is Example 3.4 of Robert & Casella (2004))

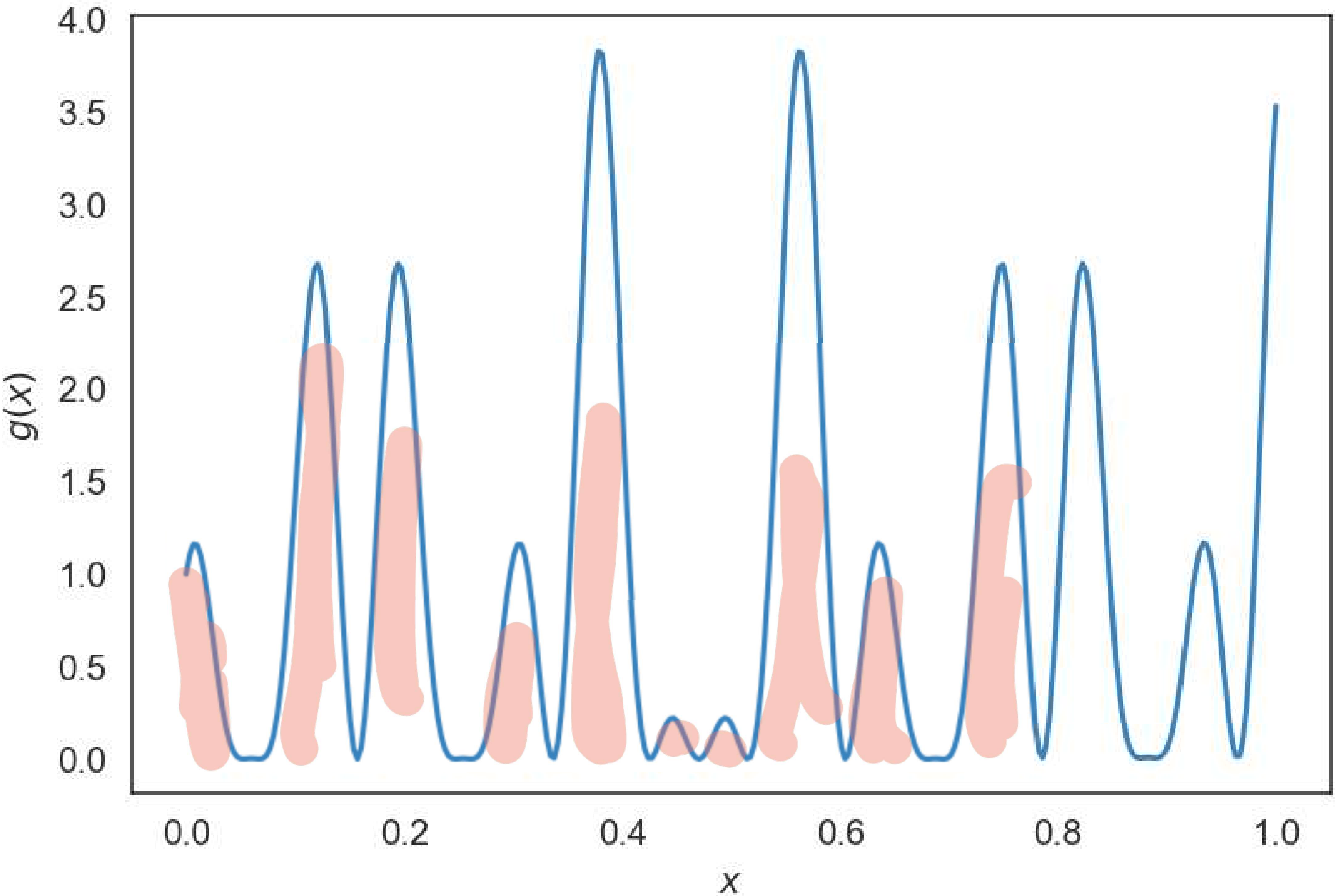
$$X \sim \mathcal{U}([0, 1])$$

$$g(x) = \left(\cos(50x) + \sin(20x)\right)^2$$

The correct value for the integral is:

$$\mathbb{E}[g(X)] = 0.965$$

Example: 1D expectation

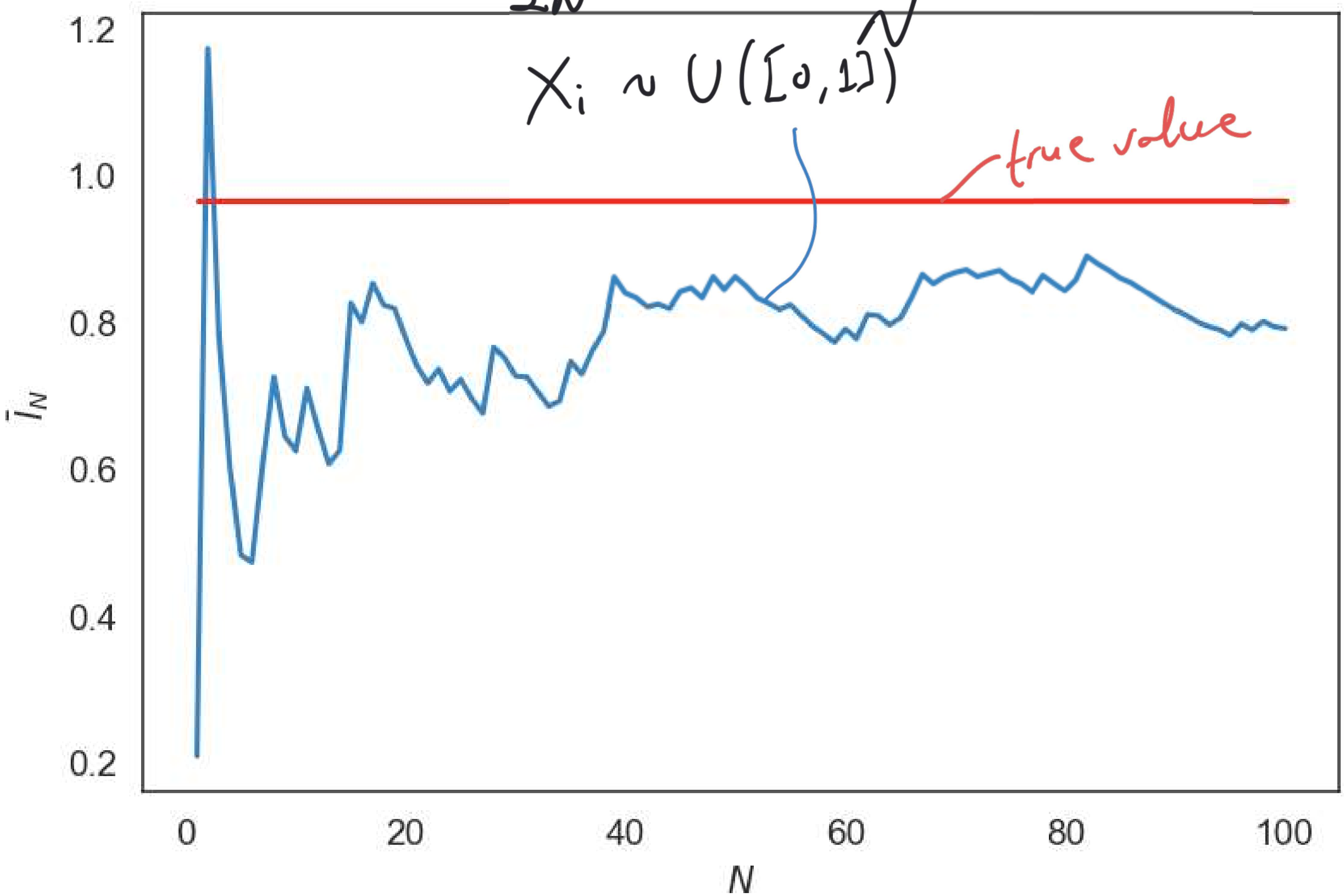


PREDICT
SCIENCE LABORATORY

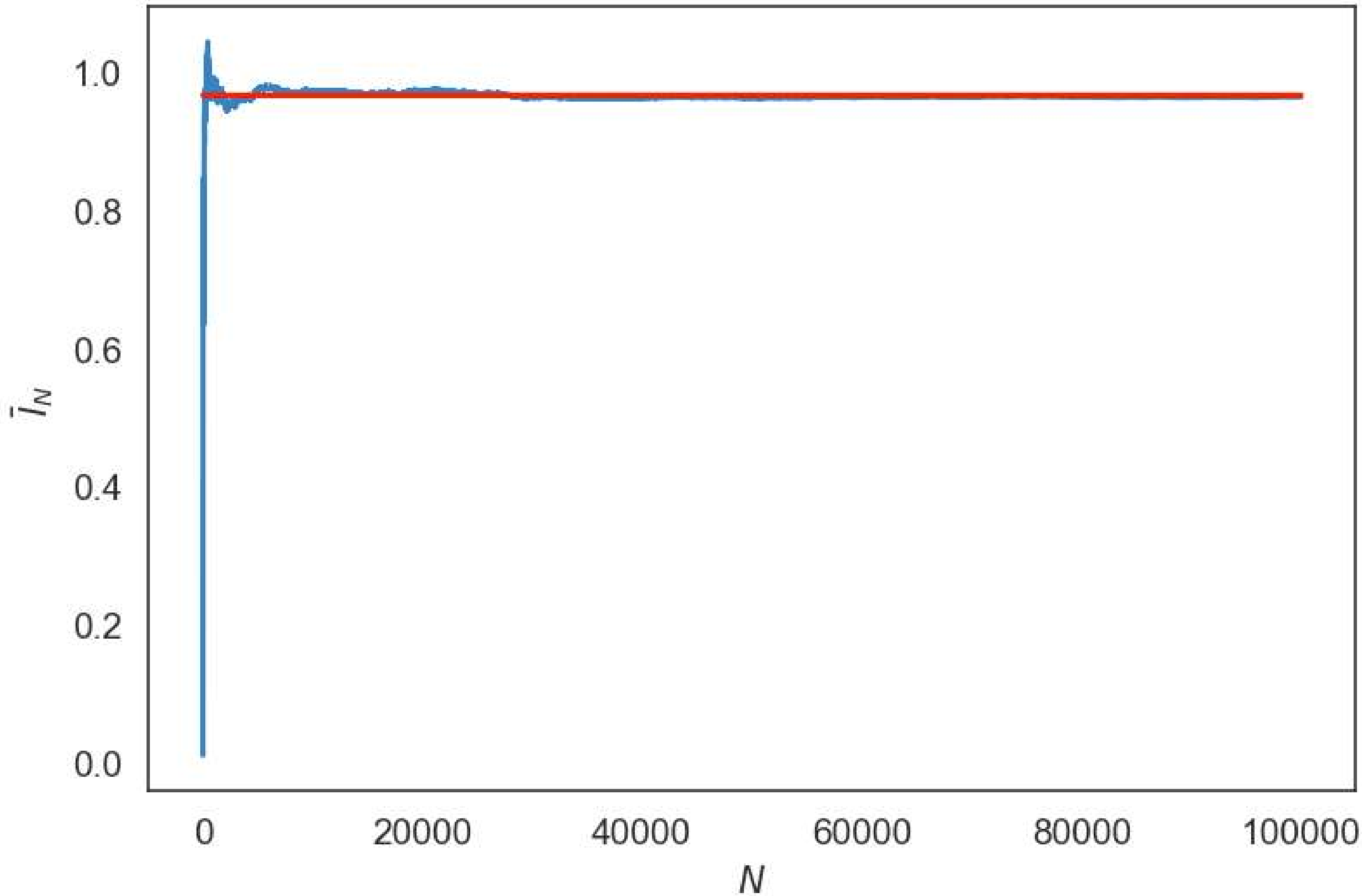
Example: 1D expectation

$$\bar{I}_N = \frac{g(X_1) + \dots + g(X_N)}{N}$$

$$X_i \sim U([0,1])$$



Example: 1D expectation



PREDICTIVE SCIENCE LABORATORY

Lecture 8: The Monte Carlo method for estimating expectations

Professor Ilias Bilionis

Estimating the variance

PREDICTIVE
SCIENCE LABORATORY

Estimating the variance

- Take a random variable $X \sim p(x)$ and some function $g(x)$.
- We would like to estimate the variance:

$$V = \mathbb{V}[g(X)] = \mathbb{E}\left[\left(g(X) - \mathbb{E}[g(X)]\right)^2\right] = \mathbb{E}\left[\left(g(X) - I\right)^2\right]$$

- Note that:

$$V = \mathbb{V}[g(X)] = \mathbb{E}[g^2(X)] - I^2$$

PREDICTIVE
SCIENCE LABORATORY

Estimating the variance

- Take $X_1, X_2, \ldots$ independent identical copies of X.
- Estimate the mean using a sample average:

$$\boxed{\bar{I}_N = \frac{1}{N}\sum_{i=1}^{N} g(X_i)}$$

$\mathbb{E}[g^2(X)]$

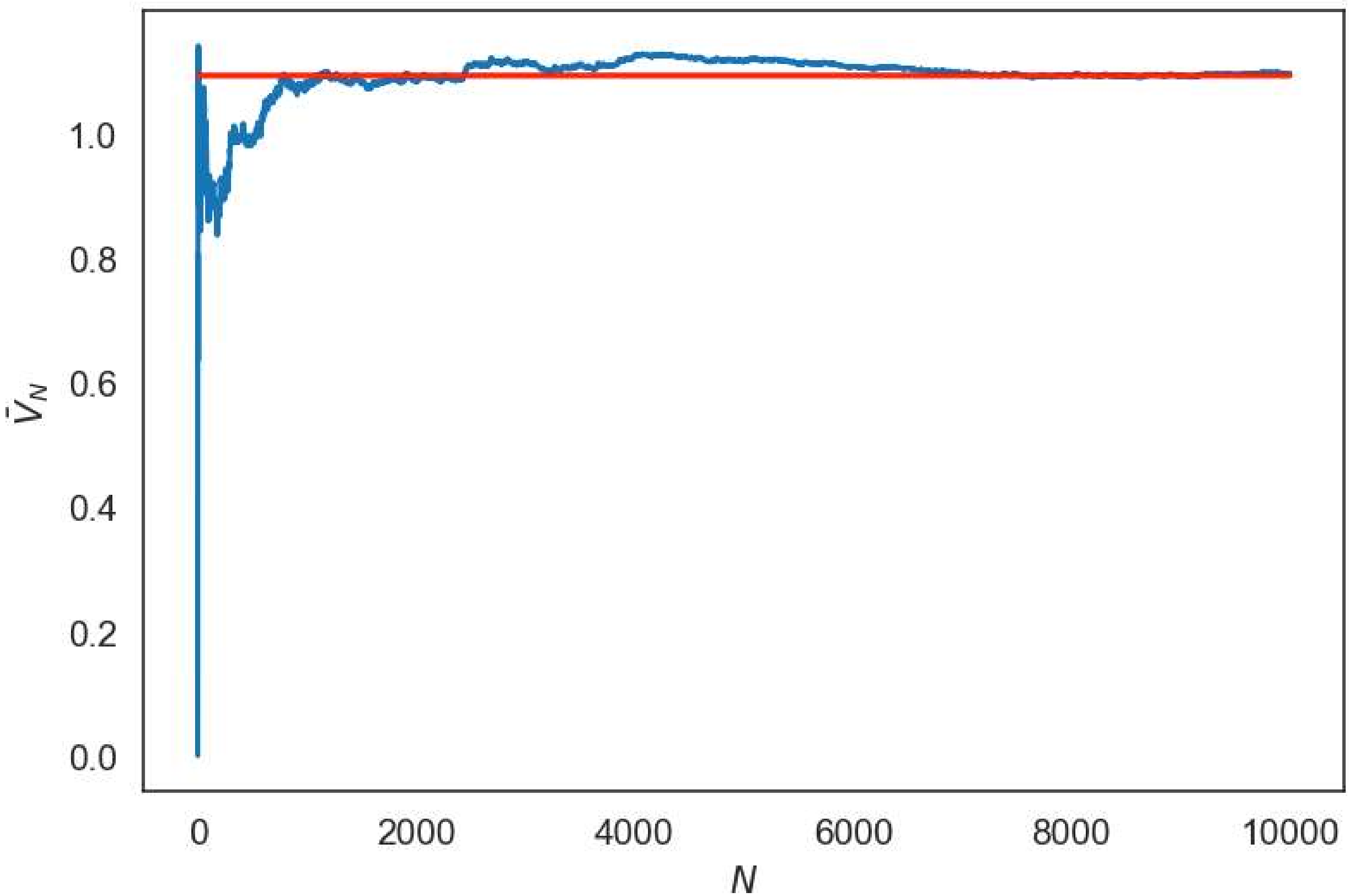
estimate

- Estimate the variance by:

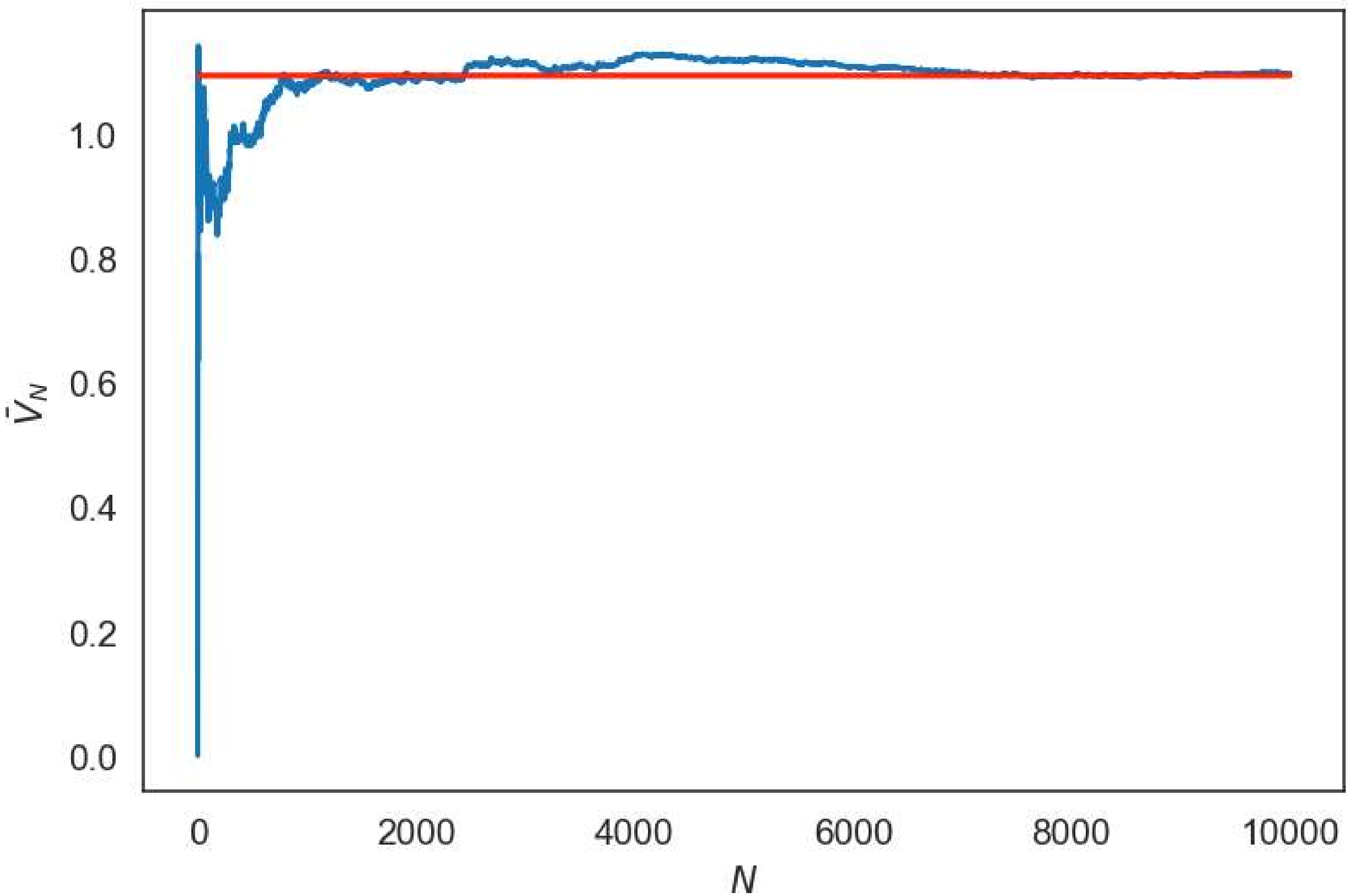
$$\boxed{\bar{V}_N = \frac{1}{N}\sum_{i=1}^{N} g^2(X_i) - \bar{I}_N^2}$$

PREDICTIVE
SCIENCE LABORATORY

Example: 1D variance



Example: 1D variance



PREDICTIVE SCIENCE LABORATORY