Chapter 3 - Monte Carlo Methods

Monte Carlo Methods.

Prof. Alex Alvarez, Ali Raisolsadat

School of Mathematical and Computational Sciences
University of Prince Edward Island

Monte Carlo Estimation Error Rate (fixed n)

Goal: We introduce the problem $\theta = E(g(X))$ where $X \sim f(x)$. Assume that $Var(g(X)) < \infty$ (variance exists and it is finite).

Simulation: $X_1, X_2, \ldots, X_n$ generated random numbers/vectors from probability distribution $f(x)$, i.e. $X_k \overset{d}{=} X$.

Estimator: $\hat{\theta}_n = \frac{1}{n} \sum_{j=1}^{n} g(X_j)$

Error in the estimation: $\epsilon_n = \hat{\theta}_n - \theta$

$$E(\epsilon_n) = E(\hat{\theta}_n - \theta) = E(\hat{\theta}_n) - \theta = 0$$

The estimator is unbiased: there is not systematic error.

Mean Square Error (MSE) and Standard Error(SE)

$$
\begin{aligned}
MSE(\hat{\theta}_n) &= E(\hat{\theta}_n - \theta)^2 \\
&= Var(\hat{\theta}_n) + (\text{bias}(\hat{\theta}_n))^2 \\
&= Var(\hat{\theta}_n) = \frac{1}{n^2} Var\left(\sum_{j=1}^{n} g(X_j)\right) = \frac{1}{n} Var(g(X))
\end{aligned}
$$

and

$$
RMSE(\hat{\theta}_n) = \sqrt{MSE(\hat{\theta}_n)} = \frac{1}{\sqrt{n}}\sigma(g(X))
$$

Remarks on Monte Carlo Integration

- The estimator

$$\hat{\theta}_n = \frac{1}{n} \sum_{j=1}^{n} g(X_j)$$

  is unbiased, i.e. $\mathbb{E}[\hat{\theta}_n] = \theta$. By the **Law of Large Numbers (LLN)**, the error converges to zero as $n \to \infty$:

$$\hat{\theta}_n \xrightarrow{a.s.} \theta$$

- The **speed of convergence** of Monte Carlo integration is

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

  regardless of the dimension of the integral.

- The convergence rate of Monte Carlo integration does not depend on the dimension d of the problem.

- In contrast, the convergence rate of conventional deterministic methods (e.g. trapezoidal or Simpson's rule) is

$$O(n^{-2/d}),$$

  which deteriorates rapidly as d increases.

Monte Carlo and the Curse of Dimensionality

- **Theoretical Perspective**: The convergence rate of Monte Carlo integration,

$$\text{RMSE}(\hat{\theta}_n) = O(n^{-1/2}),$$

  is independent of the dimension of the problem.

- **Practical Perspective:** Although the rate does not worsen with dimension, the variance of the estimator may increase rapidly:

$$Var(g(X)) \text{ may grow with } d,$$

  making the estimator less efficient in high dimensions.

- **Comparison**:

Monte Carlo:	$O(n^{-1/2})$	(dimension-free)
Trapezoid / Quadrature:	$O(n^{-2/d})$	(slower as $d \uparrow$)

- **Note**: Monte Carlo does not suffer from the curse of dimensionality in its convergence rate, but it faces a practical curse due to increasing variance and sampling difficulty in high dimensions.

Choice of Sample Size

By using that the Mean Square Error of the Monte Carlo estimator is given by

$$MSE(\hat{\theta}_n) = \frac{1}{n} Var(g(X))$$

If we want $MSE \leq \epsilon^2$ (or equivalently that the $SE \leq \epsilon$) then we must perform a number of simulations n such that:

$$n \geq \frac{Var(g(x))}{\epsilon^2}$$

Choice of Sample Size Example

Example: Assume that $Var(g(x)) = 5$. What should be the minimum sample size that we use for the basic Monte Carlo estimator assuming that we want to achieve a standard error $SE \leq 0.1$.

$$n \geq \frac{Var(g(x))}{\epsilon^2} = \frac{5}{(0.1)^2} = 500$$

.

Notice that if we want $SE \leq 0.05$ (meaning that we reduce the standard error by a half) we will need to use a sample size given by

$$n \geq \frac{Var(g(x))}{\epsilon^2} = \frac{5}{(0.05)^2} = 2000$$

or four times the number of simulations.

Improved error bounds

Assume that $Var(g(X)) = \sigma^2$ is known. We can get better error bounds by using the Central Limit Theorem. We know that for large values of n we have approximately:

$$\frac{\sqrt{n}}{\sigma}\left(\hat{\theta}_n - \theta\right) \sim N(0,1)$$

Then, if we want $P\left(|\hat{\theta}_n - \theta| \leq \epsilon\right) \geq 1 - \alpha$ (a confidence interval) then we only need to use a number of simulations n such that

$$n \geq \frac{Z^2_{1-\alpha/2}\sigma^2}{\epsilon^2}$$

where Z_α is the α-level percentile of the standard normal distribution.

In particular, if $\alpha = 0.05$ we get $n \geq \frac{1.96^2\sigma^2}{\epsilon^2}$

Remarks on Sample Size

- A critical assumption is that $Var(g(X)) < \infty$.
  If $Var(g(X)) = \infty$ the law of large numbers still applies but not the central limit theorem. Convergence is slower in that case, and more sophisticated results are needed to stablish error bounds.
- The number of simulations n needed to achieve a specified error is proportional to the variance of $g(X)$. For large values of the variance we will need a large amount of simulations.
- There are more elaborate variants of the Monte Carlo method that try to deal with this issue. Many of these methods are referred to as methods of variance reduction. We will study some of these methods during the next few classes.

Unknown Variance

What if the variance is not known?

In most cases, the variance is not known. There may be different alternatives to deal with this issue.

One alternative is to use an upper bound for the variance (if available) in place of the variance.

Another alternative is to get an estimate for the variance σ^2 from a relatively small generated sample, and then use the estimate in place of the variance.

Homework

Consider a random variable U uniformly distributed on $[0, 1]$

a) Find the exact value of $\theta = E(U^3)$
b) Find $Var(U^3)$
c) Using the result in part b), determine the number of simulations n that are needed in order to achieve an absolute error of less than 0.005 in the Monte Carlo estimation of θ with confidence level given by $\alpha = 0.01$.
d) Write a computer program implementing the Monte Carlo method with the required number of simulations according to part c) and verify that we get the pre-determined accuracy.