

1 MONTE CARLO METHODS

1.1 INTRODUCTION

Monte Carlo methods are computational algorithms that represent a milestone for financial application since it is a go-to way to obtain numerical results by relying on random sampling.

Two of the numerical problems in statistical inference and, consequently, in finance, are optimization and integration. When we approach any sort of statistical inference, we are led to take into account numerical solutions.

The main issue with numerical integration methods is that they are not always able to spot the support of a function to be integrated. Another problem that we face in numerical integration is that it can't easily deal with multidimensional integrals: by splitting the problem in pieces we increase the cost of computation.

A solution is to use simulation methods that can't be wrong since they exploit information included in the probability density associated with the integrals and are faster with respect to numerical integration methods.

1.2 MONTE CARLO INTEGRATION

By approaching Monte Carlo integration, the main problem is to solve the integral:

$$\mathbb{E}_f[h(x)] = \int_{\chi} h(x)f(x)dx, \quad (1.1)$$

where χ represents the support of the density f .

By using Monte Carlo we want to approximate this function by generate a sample (X_1, \dots, X_n) from f with the empirical average:

$$\bar{h}_n = \frac{1}{n} \sum_{j=1}^n h(x_j). \quad (1.2)$$

By the Strong Law of Large Numbers, \bar{h}_n converge to $\mathbb{E}_f[h(x)]$.

The asymptotic variance of the approximation is

$$var(\bar{h}_n) = \frac{1}{n} \int_{\chi} (h(x) - \mathbb{E}_f[h(x)])^2 f(x) d(x), \quad (1.3)$$

which can be estimated from the sample

$$v_n = \frac{1}{n^2} \sum_{j=1}^n [h(x_j) - \bar{h}_n]^2 \quad (1.4)$$

and, due to the Central Limit Theorem, for large n ,

$$\frac{\bar{h}_n - \mathbb{E}_f[h(x)]}{\sqrt{v_n}} \sim N(0, 1). \quad (1.5)$$

1.3 IMPROVE THE QUALITY

There are two main approach in order to improve the quality of the estimation: the first is to increase n : it improves the estimate, but the rate of decrease is about $O(1/\sqrt{n})$; it means that the rate of improvement is getting slower as we keep adding samples.

The second is to decrease σ : in order to achieve this, we can apply various variance reduction techniques: Importance Sampling, Antithetic Variance, Control Variance, Conditional Monte Carlo. In the following subsection we show how to use variance reduction techniques. [1]

1.3.1 IMPORTANCE SAMPLING

This method is based on importance functions which are fictitious distribution, instead of the original ones. Evaluating 1.1 by drawing from f is not optimal since the use of other distribution can improve the variance of the estimator.

It is possible to represent 1.1 in terms of an alternative density $g > 0$ when $h \times f \neq 0$ in the following way:

$$\begin{aligned}\mathbb{E}_f[h(x)] &= \int_{\mathcal{X}} h(x) \frac{f(x)}{g(x)} g(x) dx \\ &= \mathbb{E}_g \left[\frac{h(X)f(X)}{g(X)} \right]\end{aligned}\tag{1.6}$$

This importance sampling fundamental identity justifies the use of the estimator

$$\frac{1}{n} \sum_{j=1}^n \frac{f(X_j)}{g(X_j)} h(X_j) \rightarrow \mathbb{E}_f[h(X)]\tag{1.7}$$

Importance sampling is important since we have to deal with a small amount of restriction on the choice of g ; we are able to chose either from easy-to-simulate or from efficient in the approximation distribution.

On one hand, importance sampling is highly versatile but this property could produce poor outcomes. It is important to consider the variance of the estimator when analysing the fairness of the function g .

While ?? converges to 1.1, the variance of the estimator is finite only when:

$$\mathbb{E}_g \left[h^2(X) \frac{f^2(X)}{g^2(X)} \right] = \mathbb{E}_f \left[h^2(X) \frac{f(X)}{g(X)} \right] = \int_{\mathcal{X}} h^2(X) \frac{f^2(X)}{g(X)} dx < \infty\tag{1.8}$$

is finite. This condition does not forbid functions with lighter tail than f but rather it states that those function could led to infinite variance estimator.

1.3.2 ANTITHETIC VARIATES

The basic concept of this methodology is that by including correlation we can obtain higher efficiency.

Given two samples (X_1, \dots, X_n) and (Y_1, \dots, Y_n) from f used for the estimation of

$$I = \int_R h(x)f(x)dx, \quad (1.9)$$

the estimator

$$\frac{1}{2n} \sum_{i=1}^n [h(X_i) + h(Y_i)] \quad (1.10)$$

is more efficient than the estimator based on the iid sample of size $2n$ if the variables $h(X_i)$ and $h(Y_i)$ are negatively correlated. In this framework, the Y_i 's are the antithetic variables.

It is not always possible to create negative correlation for an arbitrary transform h , even when $\text{corr}(X_i, Y_i) < 0$. A solution, proposed by Rubinstein (1981) [2], is to use uniform variables U_i to generate the X_i 's as $X_i = F^-(U_i)$ and the Y_i 's as $Y_i = F^-(1 - U_i)$.

1.3.3 CONTROL VARIATES

It happens, sometimes, that we know the mean of some functions h_0 under f . The control variates technique reduce the variance by exploiting this information. If δ_1 is an estimator of I and δ_3 is an unbiased estimator of $\mathbb{E}_f[h_0(X)]$, consider the weighted estimator

$$\delta_2 = \delta_1 + \beta(\delta_3 - \mathbb{E}_f[h_0(X)]) \quad (1.11)$$

The estimators δ_1 and δ_2 have the same mean and

$$\text{var}(\delta_2) = \text{var}(\delta_1) + \beta^2 \text{var}(\delta_3) + 2\beta \text{cov}(\delta_1, \delta_3) \quad (1.12)$$

For the optimal choice,

$$\beta^* = -\frac{\text{cov}(\delta_1, \delta_3)}{\text{var}(\delta_3)} \quad (1.13)$$

getting,

$$\frac{\text{var}(\delta_2)}{\text{var}(\delta_1)} = (1 - \rho_{13}^2) \quad (1.14)$$

where ρ_{13}^2 is the correlation coefficient between δ_1 and δ_3 ; by doing this, the control variate brings to a decreased variance for δ_2

1.3.4 CONDITIONING

To estimate $\theta = \mathbb{E}[X]$ is sometimes useful to condition with respect to another variable Y :

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] \quad (1.15)$$

since $Var(X) \geq Var(\mathbb{E}[X|Y])$ as shown by the conditional variance formula:

$$Var(X) = \mathbb{E}[Var(X|Y)] + Var(\mathbb{E}[X|Y]) \quad (1.16)$$

we know that $\mathbb{E}[X|Y]$ is an unbiased estimator for θ which helps to reduce variance and if $\mathbb{E}[X|Y]$ can be correctly computed the additional noise coming from simulating X given Y can be eliminated.

1.4 MONTE CARLO FOR OPTION PRICING

In terms of theory, Monte Carlo valuation relies on risk neutral valuation. Here, the price of an option is its discounted expected value.

First, we need to generate a large number of random price paths for the underlying via simulation: in order to get it, we draw from a prior distribution and then we compute the value of the option for each path.

These payoffs are then averaged and discounted: the result is the value of the option.

This approach, although, is relatively straightforward; it allows for increases in complexity: an option on equity may be modelled with one source of uncertainty which is the price of the underlying stock.

Here the price of the underlying instrument S_t , is usually modelled such that it follows a geometric Brownian motion with constant drift μ , and volatility σ .

Thus:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \quad (1.17)$$

where dW_t , is found by drawing from a Gaussian. This approach brings to the first covered model: Black-Scholes Model. Then, we can use Monte Carlo method to price more complex derivatives via simulation.

REFERENCES

- [1] Christian P. Robert & George Casella: Introducing Monte Carlo Methods with R, Springer, 2010
- [2] Reuven Y. Rubinstein & Dirk P. Kroese: Simulation and the Monte Carlo Method, Wiley, New York, 1981