In a simulation study, let us say that we are interested in determining θ, a parameter connected with some stochastic model. To estimate θ, the model is simulated to obtain the output datum X which is such that $\theta = E(X)$. The simulation is terminated when M simulation runs have been performed and the estimate of θ is given by $\widehat{\theta}_M = \frac{1}{M}\sum_{i=1}^{M} X_i$. Because this results in an unbiased estimator of θ, it follows that the mean square error is equal to its variance, i.e, $MSE = E(\widehat{\theta}_M - \theta)^2 = Var(\widehat{\theta}_M) = \frac{Var(X)}{M}$.

Now, if we can obtain another unbiased estimator of θ having smaller variance than does θ_M, we would have obtained an improved estimator. There are different methods that one can attempt to use in order to reduce the variance of the simulation estimate. These methods are known as variance reduction techniques. Methods of variance reduction can sometimes bring enormous improvements compared to simple Monte Carlo.

1 Antithetic Variates

Suppose, to obtain a simulation estimate of $\theta = E(X)$, we have generated identically distributed random variables X_1 and X_2 having mean θ. Then

$$Var\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{4}[Var(X_1) + Var(X_2) + 2Cov(X_1, X_2)] = \frac{1}{2}Var(X_1)(1+\rho),$$

where ρ denotes the correlation between X_1 and X_2. Clearly, if X_1 and X_2 were independent (which is the case in the simple Monte Carlo), then this variance equals $Var(X_1)/2$ and it can be reduced further if X_1 and X_2 were negatively correlated. Antithetic variables take advantage of this fact. Observe that the variance could be double of the simple Monte Carlo case, when they are perfectly positively correlated.

Definition: Random variables X, Y on the same probability space are antithetic if they have the same distribution and their covariance is negative.

To see how to arrange for X_1 and X_2 to be negatively correlated, we use the following result (which can be generalized to the case of random vector as well).

Result : If f and g are both monotonic increasing functions or both monotonic decreasing functions then, for any random variable X,

$$Cov(f(X), g(X)) \geq 0.$$

Proof :

Let Y be a random variable that is independent of X with the same distribution. Then,

$$(f(X) - f(Y))(g(X) - g(Y)) \geq 0.$$

So the random variable $(f(X) - f(Y))(g(X) - g(Y)) \geq 0$ must have a nonnegative expected value. Hence,

$$\begin{aligned} 0 &\leq E[(f(X) - f(Y))(g(X) - g(Y))] \\ &= E[f(X)g(X)] - E[f(X)g(Y)] - E[f(Y)g(X)] + E[f(Y)g(Y)]. \end{aligned}$$

Since X and Y are i.i.d the right hand side simplifies to,

$$2E[f(X)g(X)] - 2E[f(X)]E[g(X)] = 2Cov(f(X), g(X)),$$

giving us the required result.

Note that if f is a monotonic increasing function, then so is $-f(1-x)$. Similarly, if f is a monotonic decreasing function, then so is $-f(1-x)$. In either case, applying the above result to a random number U (i.e., U is uniformly distributed on $(0,1)$) gives

$$Cov(f(U), -f(1-U)) \geq 0 \quad \Rightarrow \quad Cov(f(U), f(1-U)) \leq 0.$$

In fact, it can be shown that, if $h(x_1, \ldots, x_n)$ is a monotone function of each of its arguments and $U_1, \ldots, U_n$ are independent random numbers, then $Cov(h(U_1, \ldots, U_n), h(1-U_1, \ldots, 1-U_n)) \leq 0$.

Now, since $1-U$ is also a random number when U is, suppose that $X_1 = h(U_1, \ldots, U_n)$ and $X_2 = h(1-U_1, \ldots, 1-U_n)$. Then X_2 has the same distribution as X_1. In addition, since $1-U$ is negatively correlated with U, we might expect that X_2 might be negatively correlated with X_1. And, this is indeed the case when h is monotone as seen above.

Thus, after having generated $U_1, \ldots, U_n$ to compute X_1, we do better by just using $1-U_1, \ldots, 1-U_n$ to compute X_2. The benefit is not only the reduction is variance of the estimator, but also the saving of the time to generate a second set of random numbers.

Analysis of the Uniform Case :

To understand how the antithetic variate technique works, consider the more general case of approximation, $I = E[f(U)]$ where $U \sim \mathcal{U}[0,1]$ for some function f. The standard Monte Carlo estimate (in this case) is

$$I_{2M} = \frac{1}{2M} \sum_{i=1}^{2M} f(U_i)$$

with i.i.d $U_i \sim \mathcal{U}[0,1]$. The antithetic alternative is

$$\widehat{I}_M = \frac{1}{M} \sum_{i=1}^{M} \frac{f(U_i) + f(1-U_i)}{2}$$

with i.i.d $U_i \sim \mathcal{U}[0,1]$. It can be shown that (on the lines as above),

$$Var\left(\frac{f(U_i) + f(1-U_i)}{2}\right) = \frac{1}{2}\left(Var(f(U_i)) + Cov(f(U_i), f(1-U_i))\right)$$

The success of the antithetic scheme hinges on whether $Var\left(\frac{f(U_i)+f(1-U_i)}{2}\right)$ is smaller than $Var(f(U_i)$. If f is monotonic (as given above), then we have $Cov(f(U_i), f(1-U_i)) \leq 0$. Hence,

$$Var\left(\frac{f(U_i) + f(1-U_i)}{2}\right) \leq \frac{1}{2} Var(f(U_i)).$$