1 Monte Carlo Simulation

1.1 Monte Carlo Algorithm

For the risk-neutral process

$$\frac{dS}{S} = rdt + \sigma dW \tag{1}$$

the price of a European option is

$$V(S_0, 0) = e^{-rT}\mathbb{E}(V(S_T, T))$$

where e^{-rT} is the discounting factor. We can use the following algorithm to estimate the price $V(S_0, 0)$:

The Monte Carlo (MC) simulation:

1. For $k = 1, 2, \cdots, N\{$

   (a) Choose a seed $SEED(k)$ *(It is important to select a different seed for each cycle);*

   (b) Integrate the SDE $dS = rSdt + \sigma SdW$ over the interval $[0, T]$ to produce a trajectory, in turn, $S^{(k)}(T)$, for a given S_0.

   (c) Evaluate the payoff

   $$P_k \equiv V\left(S^{(k)}(T), T\right)$$

   which is a random variable

   $\}$

   (After the completion of this N Monte Carlo cycles, we have obtained a population of $\{P_k\}$, which are statistically independent.)

2. Estimating the price:

   (a) Estimate the risk-neural expectation:

   $$\epsilon \equiv \frac{1}{N}\sum_{k=1}^{N} P_k$$

   (b) To estimate the discounted price:

   $$\hat{V} \equiv e^{-rT}\epsilon$$

   which is an approximate price for $V(S_0, 0)$.

Simulation Efficiency:

1. Note that

$$\bar{X} \equiv \frac{1}{N} \sum_{k=1}^{N} X_k,$$

$$\sigma_{\bar{X}}^2 = \frac{1}{N} \sigma_X^2$$

Using $\sigma_{\bar{X}}$ as our estimate of error, we have

$$\text{error}^2 \sim \frac{1}{N} \sigma_X^2$$

Therefore

$$\text{the computational work } \propto N \propto \frac{\sigma_X^2}{\text{error}^2}$$

if we assume that the major work is in the Monte Carlo cycle. We note that doubling accuracy needs quadrupling the work.

To obtain relative accurate results, we need the number, N, of realizations to be sufficiently large, e.g., $N \sim 10^4$ to 10^5.

2. Obviously, for the geometric Brownian motion (1), there is no competition against the Black-Scholes formula. MC is used for models for which no simple Black-Scholes-like formula exists. For example, r, σ follows its own SDE or coupled SDE:

(a) Example 1:

$$dr = -\alpha (r - R)\, dt + \sigma r^{\beta} dW, \quad \alpha > 0$$

for some general β.

(b) Example 2:

$$\begin{cases} dS = \sigma S dW^{(1)} \\ d\sigma = -(\sigma - \zeta)\, dt + \nu_0 \sigma dW^{(2)} \\ d\zeta = -\gamma (\zeta - \sigma)\, dt \end{cases}$$

with $\mathbb{E}\left(dW^{(1)} dW^{(2)}\right) = \rho dt$. HW: To price a European call with the following data:

$$\nu_0 = 0.25,\ \gamma = 15,\ S_0 = 10, K = 15, \sigma_0 = 0.25, \zeta_0 = 0.2, \rho = 0.9$$

3. Now let us obtain some simple formula to quantify simulation efficiency. For an estimator of variance σ^2, the computation work is

$$\text{total work} = \tau N$$

where N is the number of MC cycles and τ is the computational work per cycle. Therefore,

$$\text{total work} \propto \tau \frac{\sigma_X^2}{\text{error}^2}.$$

Suppose we have two different estimators,

$$\begin{aligned} \text{Estimator 1: variance} &= \sigma_1^2, \quad \text{work per cycle } = \tau_1 \\ \text{Estimator 2: variance} &= \sigma_2^2, \quad \text{work per cycle } = \tau_2 \end{aligned}$$

To quantify the efficiency gained using Estimator 2 against Estimator 1, we can use **the acceleration factor** in work for the same error:

$$\mathbf{a} \equiv \frac{\tau_1 \frac{\sigma_1^2}{\text{error}^2}}{\tau_2 \frac{\sigma_2^2}{\text{error}^2}} = \frac{\tau_1 \sigma_1^2}{\tau_2 \sigma_2^2}$$

Defining the ratio of computation work per cycle:

$$\eta \equiv \frac{\tau_1}{\tau_2}$$

thus the acceleration factor in work for the same accuracy is

$$\mathbf{a} = \eta \frac{\sigma_1^2}{\sigma_2^2}$$

1.2 Variance Reduction

A major issue in the MC simulation is to increase simulation efficiency. Note that a straightforward implementation of MC may be useless for practical problems because of its cost. To speed up an MC algorithm, the main idea is to reduce the variance. Commonly used approaches include:

1. Antithetic variates;
2. Control variates;
3. Importance sampling.

There are other methods we may not be able to cover due to time, such as stratification, moment matching, and low-discrepancy sequences.

1.2.1 Antithetic Variates

The Method: First, we use an example to illustrate the idea.

Example: Evaluation of a European call on the underlying

$$dS = rSdt + \sigma SdW$$

At $t = T$, the solution of this process is

$$S(T) = S_0 e^{\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z}, \qquad Z \sim \mathcal{N}(0,1)$$

the payoff is

$$V(T) = (S(T) - K)_+$$

For each realization, $Z^{(k)}, k = 1, 2, \cdots N$, we have

$$S_+^{(k)}(T) \equiv S_0 e^{\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z^{(k)}}$$

For the same $Z^{(k)}$, we produce another S by changing the sign of Z :

$$S_-^{(k)}(T) \equiv S_0 e^{\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}\left(-Z^{(k)}\right)}$$

Therefore, the payoffs corresponding these two stock prices are

$$\begin{aligned} V_+^{(k)}(T) &= \left(S_+^{(k)}(T) - K\right)_+ \\ V_-^{(k)}(T) &= \left(S_-^{(k)}(T) - K\right)_+ \end{aligned}$$

respectively, where $V_-^{(k)}(T)$ is the antithetic variate. Finally, we use the following estimator for the European call:

$$\begin{aligned} V_{AT}^{(k)}(T) &\equiv \frac{1}{2}\left(V_+^{(k)}(T) + V_-^{(k)}(T)\right), \\ \bar{V}(0) &\equiv \frac{1}{N}\sum_{k=1}^{N} V_{AT}^{(k)}(T), \end{aligned}$$

which uses the antithetic variates. Note that $V_+^{(k)}(T), V_-^{(k)}(T)$ are statistically *dependent* for a fixed k, but $V_{AT}^{(k)}(T), k = 1, 2, \cdots, N$ are **statistically independent**.

Note that it is straightforward to implement the antithetic variates in MC cycles.

Efficiency of Antithetic Variates:

1. In the MC cycle, approximately, there is a doubling of work of a straightforward MC due to the computation of antithetic variates, therefore,

$$\eta_{AT} \simeq \frac{1}{2}.$$

2. Since $V_{AT} = \frac{1}{2}(V_+ + V_-)$, the variance for the antithetic estimator is

$$\mathrm{Var}V_{AT} = \frac{1}{4}\left[\mathrm{Var}V_+ + 2\mathrm{Cov}(V_+, V_-) + \mathrm{Var}V_-\right]$$

Note that

$$\mathrm{Var}V_+ = \mathrm{Var}V_-$$

$$\therefore \mathrm{Var}V_{AT} = \frac{1}{2}\left(\mathrm{Var}V_+ + \mathrm{Cov}(V_+, V_-)\right)$$

Therefore, the acceleration factor is

$$\begin{aligned}
\mathbf{a} &= \eta_{AT}\frac{\sigma_1^2}{\sigma_2^2} \\
&= \frac{1}{2}\frac{\mathrm{Var}V_+}{\mathrm{Var}V_{AT}} \\
&= \frac{\mathrm{Var}V_+}{\mathrm{Var}V_+ + \mathrm{Cov}(V_+, V_-)}
\end{aligned}$$

$$\therefore \mathbf{a} = \frac{1}{1 + \rho_{V_+V_-}}$$

in which

$$\rho_{V_+V_-} = \frac{\mathrm{Cov}(V_+, V_-)}{\sqrt{\mathrm{Var}V_+\mathrm{Var}V_-}} = \frac{\mathrm{Cov}(V_+, V_-)}{\mathrm{Var}V_+}$$

If the antithetic variates indeed make the estimator more efficient, then

$$\mathbf{a} > 1$$

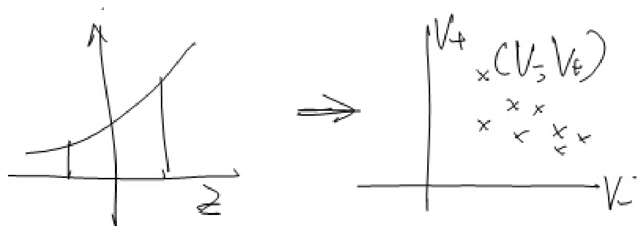
i.e.,

$$\rho_{V_+V_-} < 0$$

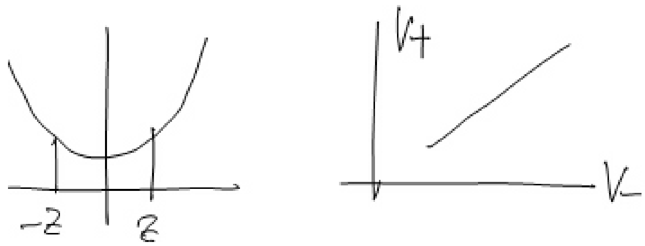
that is, V_+, V_- must be anti-correlated in order that there is a reduction of computation work.

(a) If the payoff function is linear with respect to Z, then $\rho_{V_+V_-} = -1$, thus, $\mathbf{a} \to \infty$, i.e., we need only one MC cycle to obtain the final result.

(b) If the payoff function is a monotonic function of Z, then $\rho_{V_+V_-} < 0$. Hence, increase of efficiency by using the antithetic variates.



(c) If the payoff function is almost symmetric with respect to Z, then $\rho_{V_+V_-} > 0$, $\mathbf{a} < 1$. Hence, reduction of efficiency.



Therefore, an incorrect implementation of antithetic variates may reduce the computational efficiency.

1.2.2 Control Variates

The approach of control variates is a better way for variance reduction than antithetic variates. Sometimes, the speed-up can be significant, e.g., 10 to 10^2 folds.

The idea Instead of using the original variates, we try to find a control instrument and consider the following estimator

$$\bar{V}^c = \bar{V} + \beta \left(V^a - \bar{V}^a \right) \tag{2}$$

where $\bar{V}$ is the uncontrolled estimator of the value, in which we are interested originally, $\bar{V}^a$ is the estimator for the control variate instrument, V^a is the precise value of the control variate instrument, and $\bar{V}^c$ is the so-called controlled estimator. Here, V^a is assumed to be *known precisely* (either by analytic formula or some other numerical computation). The aim of constructing this new estimator is to achieve

$$\mathbb{E}\bar{V}^c = \mathbb{E}\bar{V}$$

while

$$\mathrm{Var}\bar{V}^c \ll \mathrm{Var}V$$

Therefore, β is a parameter to be optimized in order to achieve this reduction.

Now we analyze a bit about the construction (2): Writing out the estimators, Eq. (2) is

$$\bar{V}^c = \frac{1}{N}\sum_{k=1}^{N} V_k + \beta\left(V^a - \frac{1}{N}\sum_{k=1}^{N} V_k^a\right)$$

and the associated variance is

$$\begin{aligned}
\mathrm{Var}\bar{V}^c &= \frac{1}{N}\sum_{k=1}^{N}\mathrm{Var}V + \frac{\beta^2}{N}\sum_{k=1}^{N}\mathrm{Var}V^a - \frac{2\beta}{N^2}\sum_{k=1}^{N}\sum_{k'=1}^{N}\mathrm{Cov}\left(V, V^a\right) \\
&= \mathrm{Var}V - 2\beta\mathrm{Cov}\left(V, V^a\right) + \beta^2\mathrm{Var}V^a
\end{aligned}$$

Therefore,

$$\mathrm{Var}\bar{V}^c = \sigma_V^2 - 2\beta\rho_{VV^a}\sigma_V\sigma_{V^a} + \beta^2\sigma_{V^a}^2$$

where $\sigma_V^2 \equiv \mathrm{Var}V, \sigma_{V^a}^2 \equiv \mathrm{Var}V^a, \rho_{VV^a} \equiv \mathrm{Cov}(V, V^a)/\sigma_V\sigma_{V^a}$. Now we select a value of β to minimize the variance of the controlled estimator, i.e.,

$$\begin{aligned}
\frac{\partial}{\partial\beta}\,\mathrm{Var}\bar{V}^c &= 0 \\
\textit{i.e.,}\quad -2\rho_{VV^a}\sigma_V\sigma_{V^a} + 2\beta\sigma_{V^a}^2 &= 0
\end{aligned}$$

$$\begin{aligned}
\therefore\quad \beta_{op} &= \frac{\rho_{VV^a}\sigma_V\sigma_{V^a}}{\sigma_{V^a}^2} \\
&= \frac{\rho_{VV^a}\sigma_V}{\sigma_{V^a}}
\end{aligned}$$

which yields the minimum variance of $\bar{V}^c$,

$$\begin{aligned}
\left(\mathrm{Var}\bar{V}^c\right)_{\min} &= \frac{4\sigma_{V^a}^2\sigma_V^2 - 4\rho_{VV^a}^2\sigma_V^2\sigma_{V^a}^2}{4\sigma_{V^a}^2} \\
&= \sigma_V^2\left(1 - \rho_{VV^a}^2\right)
\end{aligned}$$

Efficiency of Control Variates: The acceleration factor for the control variates thus is

$$\begin{aligned}
\mathbf{a} &= \eta_{CV}\frac{\mathrm{Var}V}{\left(\mathrm{Var}\bar{V}^c\right)_{\min}} = \eta_{CV}\frac{\sigma_V^2}{\sigma_V^2\left(1 - \rho_{VV^a}^2\right)} \\
&= \frac{\eta_{CV}}{1 - \rho_{VV^a}^2}
\end{aligned}$$

If the control variate instrument is designed such that $\rho_{VV^a}^2 \approx 0.95$, then $\mathbf{a} = 10$ if we have $\eta_{CV} = 1/2$ — that is, ten times faster than the uncontrolled estimator. Of course, if we can find a controlled variate V^a that is perfectly correlated with the variate V we intended to compute, i.e., $\rho_{VV^a}^2 = 1$, then $\mathbf{a} = \infty$, i.e., we need only one MC cycle to obtain the value of the instrument. This can be understood as follows, the value between V and V^a is merely a rescaling due the prefect correlation ($\rho^2_{VV^a} = 1$) (note that the value of V^a is known precisely). In general, if the control variates speed up the computation, this requires $\mathbf{a} > 1$, i.e.,

$$\frac{\eta_{CV}}{1-\rho^2_{VV^a}} > 1$$

or

$$\rho^2_{VV^a} > 1 - \eta_{CV}.$$

If the computational work of the control instrument is about the same as that of the originally uncontrolled, then

$$\eta_{CV} \approx \frac{1}{2}$$

therefore,

$$\rho > \rho_{\min} = \sqrt{1-\eta_{CV}} = \sqrt{1-\frac{1}{2}} \approx 0.707 \tag{3}$$

which constrains what controlled variates can be used in order to achieve efficiency improvement. Note that the constraint (Eq. (3)) can be also expressed as

$$\mathrm{Cov}\left(V, V^a\right) > \frac{1}{2}\left(\mathrm{Var}V\mathrm{Var}V^a\right)^{\frac{1}{2}} \qquad \left(\because \rho^2 > \rho^2_{\min} = \frac{1}{2}\right).$$

Note that

1. The control process does not have to be a true financial instrument;
2. The criteria for the control variates are
   (a) Any function of the same underlying process with a **known** expectation (even if known up to high accuracy through numerical computation);
   (b) This function is highly correlated with the instrument we are interested in.

Examples of control variates:

1. Asian options: There are two kinds of averages:
   (a) Arithmetic average, i.e.,

$$\hat{S} \equiv \frac{1}{T}\int_0^T S\left(t\right) dt$$

   which is the limit of the sum

$$\hat{S} \equiv \frac{1}{T}\sum_{i=1}^{n} S\left(t_i\right)$$

(b) Geometric average, i.e.,

$$\hat{S}_g \equiv \exp\left(\frac{1}{T}\int_0^T \log S(t)\,dt\right)$$

which is the limit of

$$\begin{aligned}\left(\prod_{i=1}^{n} S(t_i)\right)^{\frac{1}{n}} &= \exp\left[\frac{1}{n}\log\prod_i S(t_i)\right] \\ &= \exp\left[\frac{1}{n}\sum_i \log S(t_i)\right]\end{aligned}$$

For example, the payoff for a geometric averaged Asian call is

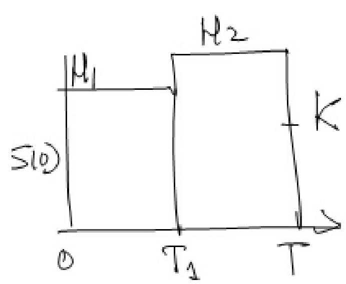
$$V_g(T) = \left(\hat{S}_g - K\right)_+$$

whereas the payoff of the arithmetic averaged Asian call is

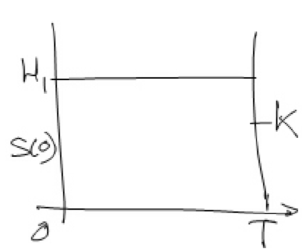
$$V(T) = \left(\hat{S} - K\right)_+$$

It turns out that there is a closed form for the geometric averaged Asian option. So it can be used as a control for the corresponding arithmetic averaged Asian option. We can use the MC simulation to estimate the price of an arithmetic averaged Asian option with the geometric averaged Asian option as the control variate instrument for variance reduction.

2. Barrier options: To price the following barrier:



we can use this following barrier as a control (using the PDE method to get V^a precisely, say):



1.2.3 Importance Sampling

As we know, pricing European options is merely to compute the expectation

$$V(0) = \mathbb{E}[F(S_T)]$$

where $F(S_T)$ is the discounted payoff function. Suppose the random variable S_T has the pdf $p(S_T)$, then

$$V(0) = \int F(x)\, p(x)\, dx \tag{4}$$

where $x = S_T$. The Monte Carlo simulation uses the estimator

$$V(0) = \frac{1}{N}\sum_{k=1}^{N} F(S_k)$$

where S_k has the density $p(S_k)$. Suppose $p(x)$ is known, if we sample x according to $p(x)$, we may waste a lot of work in the area where $F(x)$ is extremely small. The method of importance sampling address this issue to improve efficiency.

Basic Idea: Choose another pdf $\pi(x)$, then Eq. (4) can be written as

$$V(0) = \int F(x)\,\frac{p(x)}{\pi(x)}\pi(x)\, dx \tag{5}$$

then the MC estimator becomes

$$V(0) = \frac{1}{N}\sum_{k=1}^{N} F\left(\hat{S}_k\right)\frac{p\left(\hat{S}_k\right)}{\pi\left(\hat{S}_k\right)}$$

where $\hat{S}_k$ has the pdf $\pi\left(\hat{S}_k\right)$. The pdf $p(x)$ is called the *nominal* pdf and $\pi(x)$ is the so-called *importance* pdf. By carefully choosing $\pi(x)$, we can achieve

$$\mathrm{Var}\left(F\left(\hat{S}_k\right)\frac{p\left(\hat{S}_k\right)}{\pi\left(\hat{S}_k\right)}\right) \ll \mathrm{Var}\left(F\left(S_k\right)\right)$$

where S_k has the density $p(S_k)$ and $\hat{S}_k$ has the pdf $\pi\left(\hat{S}_k\right)$. Note that, if the evaluation of $\frac{p\left(\hat{S}_k\right)}{\pi\left(\hat{S}_k\right)}$ requires no significant additional cost, then we achieve reduction of variance, thus, increasing efficiency. The ratio $\frac{p\left(\hat{S}_k\right)}{\pi\left(\hat{S}_k\right)}$ is called the likelihood ratio.

An Extreme Case — "Optimal Importance pdf": We have the following optimal importance pdf, if $F(S) > 0\ \forall S$,

$$\pi_{op}(x) = \frac{F(x)\,p(x)}{V(0)}$$

Why? Let's evaluate the variance

$$\begin{aligned}
&\mathrm{Var}\left[F\left(\hat{S}_k\right)\frac{p\left(\hat{S}_k\right)}{\pi\left(\hat{S}_k\right)}\right] \qquad (6)\\
&= \mathbb{E}\left[\left(F\left(\hat{S}_k\right)\frac{p\left(\hat{S}_k\right)}{\pi\left(\hat{S}_k\right)}\right)^2\right] - \underbrace{\mathbb{E}^2\left[F\left(\hat{S}_k\right)\frac{p\left(\hat{S}_k\right)}{\pi\left(\hat{S}_k\right)}\right]}_{=[V(0)]^2}\\
&= \int\left(F(x)\frac{p(x)}{\pi(x)}\right)^2\pi(x)\,dx - V^2(0)\\
&= \int\frac{[F(x)\,p(x)]^2}{\pi(x)}dx - V^2(0)
\end{aligned}$$

For the optimal $\pi_{op} = \frac{F(x)p(x)}{V(0)}$, the variance (6) becomes

$$\begin{aligned}
\mathrm{Var}\left[F\left(\hat{S}_k\right)\frac{p\left(\hat{S}_k\right)}{\pi_{op}\left(\hat{S}_k\right)}\right] &= \int\frac{[F(x)\,p(x)]^2}{\frac{F(x)p(x)}{V(0)}}dx - V^2(0)\\
&= V(0)\underbrace{\int F(x)\,p(x)\,dx}_{=V(0)} - V^2(0)\\
&= V^2(0) - V^2(0)\\
&= 0
\end{aligned}$$

We have completely removed variance by using this optimal importance pdf! However, unfortunately, this optimal $\pi_{op}(x)$ is useless for estimating the value $V(0)$ since it requires the value of $V(0)$ itself in the formula. Note that the effect under this importance pdf π_{op} for the integrand in Eq. (5) is

$$F(x)\frac{p(x)}{\pi(x)} \overset{\pi=\pi_{op}}{=} V(0)$$

i.e., under $\pi_{op}(x)$, the random variable, $F(x)\frac{p(x)}{\pi(x)}$, to be averaged is constant. In general, this is the aim of the importance pdf — to make the new $F(x)\frac{p(x)}{\pi(x)}$ in Eq. (5) as flat as possible, thus achieving a reduction of variance.

Example: Question: How to choose a process such that this reduction becomes possible? Let us consider a European call option on

$$\frac{dS}{S} = rdt + \sigma dW \tag{7}$$

which is the risk-neutral process. The value is given by

$$V(0) = \mathbb{E}\left[e^{-rT}(S(T) - K)_+\right]$$

for which our MC estimator would be

$$V(0) = \frac{1}{N}\sum_{k=1}^{N} e^{-rT}\left(S^{(k)}(T) - K\right)_+$$

by constructing N trajectories of the stock price movement using N MC cycles. Denoting the discounted payoff $e^{-rT}(S-K)_+$ by $F(S)$, the estimator becomes

$$V(0) = \frac{1}{N}\sum_{k=1}^{N} F\left(S^{(k)}(T)\right).$$

Intuitively, if we can have more paths end in the money, our estimator would become more efficient. This is the idea we will examine below. The risk-neural process of the underlying is

$$\frac{dS}{S} = rdt + \sigma dW.$$

We want to increase the drift by c to make more trajectories terminate in the money. Note that

$$\begin{aligned}\frac{dS}{S} &= rdt + \sigma dW \\ &= (r+c)\,dt + \sigma dW - cdt \\ &= (r+c)\,dt + \sigma\left(dW - \frac{c}{\sigma}dt\right)\end{aligned}$$

Let

$$d\hat{W} \equiv dW - \frac{c}{\sigma}dt \tag{8}$$

then

$$\frac{dS}{S} = (r+c)\,dt + \sigma d\hat{W}. \tag{9}$$

It can be shown that $\hat{W}(t)$ is still a Wiener process under a different measure. We can use the Monte Carlo simulation to produce trajectories for this new process (9). The question is how to choose the value of c. We will examine the consequence of this drift-enhanced process for the MC efficiency.

Note that the stock price at maturity is

$$S(T) = S_0 e^{\left(r - \frac{1}{2}\sigma^2\right)T + \sigma W(T)}$$

under the risk-neutral process (7) with

$$\text{the pdf of } W(T) = \frac{1}{\sqrt{2\pi T}} e^{-\frac{W(T)^2}{2T}}$$

thus the price of the option is

$$\begin{aligned} V(0) &= \mathbb{E}_W \left[F\left(S_0 e^{\left(r-\frac{1}{2}\sigma^2\right)T+\sigma W(T)} \right) \right] \\ &= \int F\left(S_0 e^{\left(r-\frac{1}{2}\sigma^2\right)T+\sigma W(T)} \right) \frac{1}{\sqrt{2\pi T}} e^{-\frac{W(T)^2}{2T}} dW(T) \end{aligned} \tag{10}$$

which evaluates the expectation using the usual pdf as in Eq. (4).

Next changing $W(T)$ to

$$\hat{W}(T) = W(T) - \frac{c}{\sigma}T, \qquad (\because \text{ Eq. (8)})$$

Eq. (10) becomes

$$\begin{aligned} V(0) &= \int F\left(S_0 e^{\left(r-\frac{1}{2}\sigma^2\right)T+\sigma\left[\hat{W}(T)+\frac{c}{\sigma}T\right]} \right) \frac{1}{\sqrt{2\pi T}} e^{-\frac{\left[\hat{W}(T)+\frac{c}{\sigma}T\right]^2}{2T}} d\hat{W}(T) \\ &= \int \frac{F\left(S_0 e^{\left(r+c-\frac{1}{2}\sigma^2\right)T+\sigma\hat{W}(T)} \right)}{\exp\left[\frac{c^2}{2\sigma^2}T + \frac{c}{\sigma}\hat{W}(T)\right]} \frac{1}{\sqrt{2\pi T}} e^{-\frac{\hat{W}(T)^2}{2T}} d\hat{W}(T) \end{aligned}$$

Note that $\exp\left[\frac{c^2}{2\sigma^2}T + \frac{c}{\sigma}\hat{W}(T)\right]$ in the denominator is a random variable. Denote this by

$$\zeta(T) \equiv \frac{c^2}{2\sigma^2}T + \frac{c}{\sigma}\hat{W}(T),$$

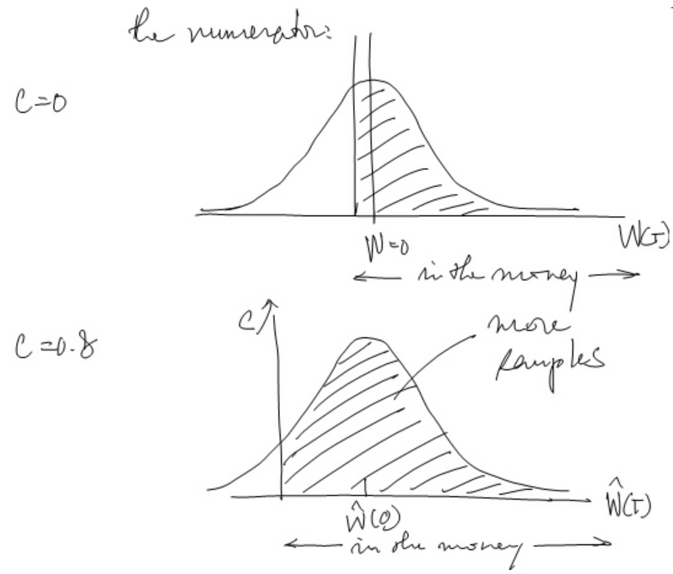
we have

$$V(0) = \mathbb{E}_{\hat{W}} \left[\frac{e^{-rT}\left(S_0 e^{\left(r+c-\frac{1}{2}\sigma^2\right)T+\sigma\hat{W}(T)} - K \right)_+}{\zeta(T)} \right]$$

which has the form of Eq. (5).

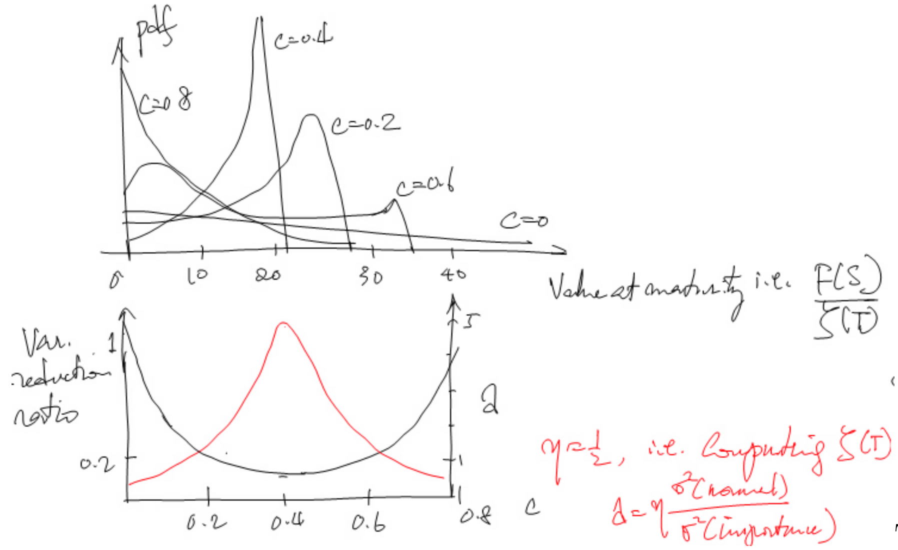
Note that there are two effects by the enhancement of the drift from r to $r+c$:

1. Because of the positive value of c, we have more samplings *in the money* for the numerator:



$S(0) = 100, K = 100, r = 0.1, \sigma = 0.3, T = 0.75.$

2. Because of $\zeta(T)$, a large value of c leads to a suppression of the integrand, thus, leading to a possible increase of variance:



The red curve is the acceleration factor for different values of c

By plotting the distribution of $\frac{F(S(T))}{\zeta(T)}$, we can see that the maximal variance reduction is achieved for moderately large values of c, e.g., $c \approx 0.4$. For too large or too small values of c, the variance can be large. The ratio of work η is about $\frac{1}{2}$ since an extra computational work is need to obtain $\zeta(T)$ under the drift-enhanced process. It turns out that the maximum acceleration factor one can achieve is about a factor of 5 for $c \sim 0.4$.