Monte Carlo Methods in Finance

Arthur Li

April 6, 2025

Contents

1 Introductory Topics

1.1 Introduction to Monte Carlo Methods

Consider the problem of computing an expectation

$$I = \mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x)\, p(x)\, dx,$$

where X is a random variable with probability density function (pdf) $p(x)$.
Approximate I by drawing independent samples $X_1, X_2, \ldots, X_n$ from $p(x)$ and computing the sample average:

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^{n} f(X_i).$$

In financial engineering the price of a derivative security is often given by an expected value; hence, pricing reduces to computing such an expectation—even when the integral is high-dimensional.

Example 1.1.1. [European Call Option]
A European call option gives the holder the right to buy the underlying stock at a fixed strike K at maturity T. Its payoff is:

$$(S_T - K)^+ = \max\{S_T - K, 0\}.$$

Under risk-neutral pricing, the option price at time $t = 0$ is

$$C = e^{-rT}\, \mathbb{E}\left(S_T - K)^+\right].$$

In the Black–Scholes model the terminal stock price is modeled as

$$S_T = S_0 \exp\left\{\left(r - \frac{1}{2}\sigma^2\right) T + \sigma\sqrt{T}\, Z\right\}, \quad Z \sim N(0,1).$$

Algorithm 1: Monte Carlo Algorithm for Option Pricing

---

for $i = 1$ to n **do**
  Generate Z_i from $N(0,1)$.
  Set $S_T = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) T + \sigma\sqrt{T} Z_i\right)$
  Set $\alpha_i = e^{-rT}(S_T - K)^+$
end for
Set $\hat{\alpha}_n = (\alpha_1 + \cdots + \alpha_n)/n$

Example 1.1.2. [Asian Option Pricing]
The payoffs depend on average of values of underlying asset at intermediate dates

$$S_{t_{j+1}} = S_{t_j} \exp\left(\left(r - \frac{1}{2}\sigma^2\right)(t_{j+1} - t_j) + \sigma\sqrt{t_{j+1} - t_j} Z_{j+1}\right)$$
$$\overline{S} = \frac{1}{m} \sum_{i=1}^{m} S_{t_i}, \quad 0 = t_0 < t_1 < \cdots < t_m = T$$

Value of option at t_0 is expected discounted payoff

$$\alpha = \mathbb{E}[e^{-rT}(\overline{S} - K)^+]$$

where expectation is taken with respect to distribution of $\overline{S}$.
Compute $\alpha = \mathbb{E}[f(X)]$, where

$$X = (Z_1, \ldots, Z_m)' \sim N(0, I_m)$$
$$f(X) = e^{-rT}(\overline{S}(Z_1, \ldots, Z_m) - K)^+$$

Algorithm 2: Monte Carlo Algorithm for Asian Option Pricing

for $i = 1$ to n **do**
  for $j = 1$ to m **do**
    Generate Z_{ij} from $N(0,1)$.
    Set $S_{t_j} = S_{t_{j-1}} \exp\left(\left(r - \frac{1}{2}\sigma^2\right)(t_j - t_{j-1}) + \sigma\sqrt{t_j - t_{j-1}} Z_{ij}\right)$
  end for
  Set $\overline{S} = (S_{t_1} + \cdots + S_{t_m})/m$
  Set $\alpha_i = e^{-rT}(\overline{S} - K)^+$
end for
Set $\hat{\alpha}_n = (\alpha_1 + \cdots + \alpha_n)/n$

Theorem 1.1.3. *Law of Large Numbers (Weak)*
Let $X_1, X_2, \ldots$ be i.i.d. random variables with finite expectation $E[X_1] = \mu$. Then,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu \quad \text{as } n \to \infty,$$

where $\xrightarrow{p}$ denotes convergence in probability.

Theorem 1.1.4. *Law of Large Numbers (Strong)*
Let $X_1, X_2, \ldots$ be i.i.d. random variables with finite expectation $E[X_1] = \mu$. Then,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{a.s.} \mu \quad \text{as } n \to \infty.$$

That is, the sample average converges almost surely to the true mean.

Theorem 1.1.5. *Central Limit Theorem*
Let $X_1, X_2, \ldots$ be i.i.d. random variables with finite mean $E[X_i] = \mu$ and finite variance $\mathrm{Var}(X_i) = \sigma^2 > 0$. Define

$$S_n = \sum_{i=1}^{n} X_i.$$

Then, as $n \to \infty$, the normalized sum converges in distribution to a standard normal random variable:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1).$$

Remark 1.1.6. The accuracy of sample average $\hat{\alpha}_n$ compared to true value α:

$$\alpha = \mathbb{E}[f(X)], \quad \hat{\alpha}_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

Converge of $\hat{\alpha}_n$ to α as $n \to +\infty$ is guaranteed by law of large numbers and central limit theorem.

$$\mathbb{E}[\hat{\alpha}_n] = \alpha = \mathbb{E}[f(x)]$$

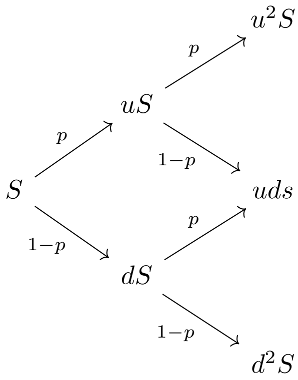
$\hat{\alpha}_n$ is unbiased. Also,

$$\mathrm{Var}[\hat{\alpha}_n] = \frac{\sigma_f^2}{n}, \quad \sigma_f^2 = \mathrm{Var}[f(x)].$$

Then standard deviation of $\hat{\alpha}_n = \frac{\sigma_f}{\sqrt{n}}$, the standard error.
Note $O(n^{-1/2})$ converge rate holds for all dimensions.

1.2 Random Variables

Example 1.2.1. [Binomial Pricing Model]
Consider the following model for stock price. Denote by S_i the price at the i-th time step. If the current price is S, then at the next time step the price either moves up to uS with probability p or moves down to dS with probability $1-p$, where $0<d<u$. Given $S_0 = x$, find the distribution of S_n.



[Solution] Suppose that among the first n time steps, there are k steps at which the stock price moves up. Then there are $(n-k)$ steps at which the stock price moves down, and

$$S_n = u^k d^{n-k} S_0 = u^k d^{n-k} x.$$

Since the number of time steps at which the stock price moves up is a binomial random variable with parameters p and n, the distribution of S_n is given by

$$P\left(S_n = u^k d^{n-k} x\right) = \binom{n}{k} p^k \left(1-p\right)^{n-k}, \quad k = 0, 1, \ldots, n$$

Example 1.2.2. [Call Option Expected Value]
The evaluation of call options often involves the calculation of expected values such as $\mathbb{E}\left[(S-K)^+\right]$, where S is the price of the underlying stock, and K is the strike price. Assume S is lognormally distributed with parameters μ and σ^2. Compute this expected value.
[Solution] Since $X = \log S \sim N(\mu, \sigma^2)$, it follows that

$$\begin{aligned}
\mathbb{E}\left[(S-K)^+\right] &= \int_{\log K}^{\infty} \left(e^x - K\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \\
&= \int_{\theta}^{\infty} \left(e^{\mu+\sigma z} - K\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\, dz, \quad z = \frac{x-\mu}{\sigma}, \theta = \frac{\log K - \mu}{\sigma} \\
&= I - II \\
I &= e^{\mu+\frac{1}{2}\sigma^2} \int_{\theta}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-\sigma)^2}\, dz \\
&= e^{\mu+\frac{1}{2}\sigma^2} \Phi(\sigma - \theta) \\
II &= K\Phi(-\theta)
\end{aligned}$$

where Φ is the c.d.f. of $N(0,1)$.
Hence $\mathbb{E}[(S-K)^+] = e^{\mu+\frac{1}{2}\sigma^2}\Phi(\sigma-\theta) - K\Phi(-\theta)$.

Example 1.2.3. [Credit Risk Model]
Consider a credit risk model where losses are due to the default of obligors on contractual payments. Suppose there are m obligors and the ith obligor defaults if and only if $X_i \geq x_i$ for some random variable X_i and given level x_i.The random variable X_i is assumed to take the form

$$X_i = \rho_i Z + \sqrt{1-\rho_i^2}\epsilon_i$$

where $Z, \epsilon_1, \ldots, \epsilon_m$ are independent standard normal r.v. and ρ_i are constants satisfying $-1 < \rho_i < 1$. Compute the probability that none of the obligors defaults:

$$P\{X_1 < x_1,\ X_2 < x_2,\ \ldots,\ X_m < x_m\}$$

[Solution] Given $Z = z$, we have

$$P\{X_1 < x_1,\ \ldots,\ X_m < x_m \mid Z = z\} = \prod_{i=1}^{m} P\{X_i < x_i \mid Z = z\}.$$

Note that

$$X_i < x_i \iff \rho_i Z + \sqrt{1-\rho_i^2}\epsilon_i < x_i \iff \epsilon_i < \frac{x_i - \rho_i Z}{\sqrt{1-\rho_i^2}}, \quad i = 1, 2, \ldots, m.$$

Hence, we have

$$\begin{aligned} P\{X_1 < x_1,\ \ldots,\ X_m < x_m \mid Z = z\} &= \prod_{i=1}^{m} P\left\{\epsilon_i < \frac{x_i - \rho_i z}{\sqrt{1-\rho_i^2}}\right\}, \text{ where } \ (\epsilon_i \sim N(0,1)) \\ &= \prod_{i=1}^{m} \Phi\left(\tfrac{x_i - \rho_i z}{\sqrt{1-\rho_i^2}}\right), \end{aligned}$$

It follows by law of total probability that

$$\begin{aligned} P\{X_1 < x_1,\ \ldots,\ X_m < x_m \mid Z = z\} &= \int_{\mathbb{R}} P\left\{X_1 < x_1,\ \ldots,\ X_m < x_m \middle| \ Z = z\right\} f(z)\, dz \\ &= \int_{\mathbb{R}} \prod_{i=1}^{m} \Phi\left(\tfrac{x_i - \rho_i z}{\sqrt{1-\rho_i^2}}\right)\ f(z)\, dz, \end{aligned}$$

Example 1.2.4. [Jump Diffusion Model]
Consider the following jump diffusion model for the stock price S:

$$S = e^Y, \quad Y = X_1 + \sum_{i=1}^{X_2} Z_i,$$

where $X_1 \sim N(\mu, \sigma^2)$, $X_2 \sim \text{Poisson}(\lambda)$, $Z_i \sim N(0, \nu^2)$, and $X_1, X_2, \{Z_i\}$ are independent.
The evaluation of call options on the stock involves expected values such as

$$\mathbb{E}[(S-K)^+]$$

Compute this expected value. *[Solution]* Note that

$$X_2 \sim \text{Poisson}(\lambda), \quad \Omega = \{0, 1, 2, \ldots\}, \quad P\{X_2 = n\} = e^{-\lambda}\frac{\lambda^n}{n!}, \quad n = 0, 1, 2, \ldots$$

Recall that if $S = e^Y$ and $Y \sim N(\mu, \sigma^2)$, then $S \sim \text{Lognormal}(\mu, \sigma^2)$. Also,

$$\begin{aligned} \mathbb{E}[(S-K)^+] &= e^{\mu + \frac{1}{2}\sigma^2}\Phi(\sigma - \theta) - K\Phi(-\theta) \\ S &= e^Y, \quad Y = X_1 + \sum_{i=1}^{X_2} Z_i \end{aligned}$$

Given $X_2 = n$, and $Y = X_1 + Z_1 + \cdots + Z_n$, then

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E}[X_1] + \sum_{i=1}^{n} \mathbb{E}[Z_i] = \mu \\ \text{Var}(Y) &= \text{Var}(X_1) + \sum_{i=1}^{n} \text{Var}(Z_i) = \sigma^2 + n\nu^2 S \quad = e^Y, \quad y \sim N(\mu, \sigma^2 + n\cdot\nu^2), \quad S \sim \log N(\mu, \sigma^2 + n \cdot \nu^2) \end{aligned}$$

By tower property of conditional expectations,

$$\begin{aligned} \mathbb{E}[(S-K)^+] &= \mathbb{E}[\mathbb{E}[(S-K)^+ \mid X_2]] \\ &= \sum_{n=0}^{\infty} v_n \mathbb{P}(X_2 = n) = e^{-\lambda}\sum_{n=0}^{\infty} v_n \frac{\lambda^n}{n!} \end{aligned}$$

1.3 Brownian Motion

Definition 1.3.1. *Brownian Motion*
A stochastic process $W = \{W_t : t \geq 0\}$ (i.e. a collection of random variables indexed by time t) is called a *standard Brownian motion* if the following conditions hold:

1. Every sample path of the process is continuous.
2. $W_0 = 0$.
3. The process has independent increments; that is, for any sequence $0 = t_0 < t_1 < \cdots < t_n$, the increments
$$W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$$
are independent random variables.
4. For any $s \geq 0$ and $t \geq 0$, the increment $W_{s+t} - W_s$ is normally distributed with mean 0 and variance t.

Definition 1.3.2. *Brownian Motion with Drift*
For constants μ and $\sigma > 0$, we call a process $X(t)$ a Brownian motion with drift μ and diffusion coefficient σ^2 (abbreviated $X \sim \mathrm{BM}(\mu, \sigma^2)$) if
$$\frac{X(t) - \mu t}{\sigma}$$
is a standard Brownian motion.
We may construct $X(t)$ from a standard Brownian motion W by setting
$$X(t) = \mu t + \sigma W(t)$$
It follows that $X(t) \sim \mathcal{N}(\mu t, \sigma^2 t)$

Remark 1.3.3. *Joint Distribution of Brownian Motion*
Joint distribution of $\big(W(t_1), \dots, W(t_n)\big) \in \mathbb{R}^n$:

1. $\big(W(t_1), \dots, W(t_n)\big)$ is a linear transformation of the vector of increments
$$(W(t_1), W(t_2) - W(t_1), W(t_3) - W(t_2), \dots, W(t_n) - W(t_{n-1}))$$
which are independent and normally distributed.
Thus, $(W(t_1), \dots, W(t_n))$ has a multivariate normal distribution.
2. Since $\mathbb{E}[W(t)] = 0$ for all t, the mean vector is $\mathbf{0}$.
3. For covariance matrix C, first consider any $0 < s < t < T$. Then
$$\begin{aligned} \mathrm{Cov}[W(s), W(t)] &= \mathrm{Cov}[W(s), W(s) + (W(t) - W(s))] \\ &= \mathrm{Cov}[W(s), W(s)] + \mathrm{Cov}[W(s), (W(t) - W(s))] \\ &= s + 0 = s \end{aligned}$$
Hence, $C_{ij} = \min(t_i, t_j)$

Remark 1.3.4. *Properties of Standard Brownian Motion*
Suppose $W = \{W_t : t \geq 0\}$ is a standard Brownian motion. Then:

1. **Symmetry:** The process $-W = \{-W_t : t \geq 0\}$ is a standard Brownian motion.
2. Fix an arbitrary $S > 0$ and define $B_t = W_{t+S} - W_S$ for $t \geq 0$.
Then $B = \{B_t : t \geq 0\}$ is a standard Brownian motion.

Remark 1.3.5. *Random Walk Construction*
In the simulation of Brownian motions, focus on simulating values $(W(t_1), \dots, W(t_n))$ or $(X(t_1), \dots, X(t_n))$ at a fixed set of points $0 < t_1 < \cdots < t_n$.
Let $Z_1, \dots, Z_n$ be independent standard normal random variables. For a standard Brownian motion, set $W(0) = 0$. Then
$$W(t_{i+1}) = W(t_i) + \sqrt{t_{i+1} - t_i} Z_{i+1}, \quad i = 0, \dots, n-1.$$
For $X \sim \mathrm{BM}(\mu, \sigma^2)$ with constants μ and σ and given $X(0)$, set
$$X(t_{i+1}) = X(t_i) + \mu(t_{i+1} - t_i) + \sigma\sqrt{t_{i+1} - t_i} Z_{i+1}, \quad i = 0, \dots, n-1.$$

Remark 1.3.6. *Running Maximum of Brownian Motion*
Define the *running maximum* of a standard Brownian motion W by time t as

$$M_t = \max_{0 \le s \le t} W_s$$

Distribution (pdf) of M_t is

$$f(x) = \frac{2}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right), \quad x \ge 0$$

The joint distribution (pdf) of (W_t, M_t) is

$$f(x,y) = \begin{cases} \dfrac{2\,(2y-x)}{\sqrt{2\pi\, t^3}} \exp\left(-\frac{(2y-x)^2}{2t}\right), & \text{if } x \le y \text{ and } y \ge 0, \\ 0, & \text{otherwise.} \end{cases}$$

Remark 1.3.7. *Multi-Dimensional Brownian Motion*
Let $B = \{B_t : t \ge 0\}$ be a stochastic process where B_t is a d-dimensional random vector for each t. Let $\Sigma = [\Sigma_{ij}]$ be a $d \times d$ symmetric, positive-definite matrix.
The process B is a d-dimensional Brownian motion with covariance Σ if:

1. Every sample path of B is continuous.
2. The process has independent increments; i.e., for any sequence $0 = t_0 < t_1 < \cdots < t_n$, the increments

   $$B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$$

   are independent random vectors.
3. For any $s \ge 0$ and $t \ge 0$, the increment $B_{s+t} - B_s$ is a jointly normal r.v. with mean 0 and covariance $t\,\Sigma$.
4. If $\Sigma = I_d$, the process B is called a d-dimensional *standard* Brownian motion.

1.4 Derivatives and Black-Scholes Prices

For financial derivatives, assume the price of the underlying asset is given by a geometric Brownian motion:

$$S_t = S_0 \exp\left(\left(\mu - \tfrac{1}{2}\sigma^2\right)t + \sigma W_t\right), \quad W_t \sim \mathcal{N}(0, t)$$

where W_t is a standard Brownian motion, S_0 is the initial asset price, and μ and $\sigma > 0$ are the drift and volatility respectively.

For every fixed $t > 0$, S_t is lognormally distributed:

$$S_t \sim \mathrm{LogN}\left(\log S_0 + \left(\mu - \tfrac{1}{2}\sigma^2\right)t,\ \sigma^2 t\right).$$

In evaluating derivatives, use the risk-neutral dynamics and set $\mu = r$, where r is the risk-free interest rate.

Definition 1.4.1. *Value of Financial Derivative*
The value of a financial derivative is given by the expected discounted payoff:

$$v = \mathbb{E}[e^{-rT}X],$$

where T is the maturity of the derivative and X is the payoff of the derivative at maturity.

Example 1.4.2. [Pricing European Call Option]
Price a call option with strike price K and maturity T. Its payoff at time T is

$$(S_T - K)^+ = \begin{cases} S_T - K, & \text{if } S_T > K, \\ 0, & \text{otherwise.} \end{cases}$$

[Solution] The price of this call option at time $t = 0$ is

$$v = \mathbb{E}[e^{-rT}(S_T - K)^+] = e^{-rT}\mathbb{E}[(S_T - K)^+]$$

Recall that when $S_T \sim \log N(\mu, \sigma^2)$, the price of option is given by

$$\mathbb{E}(S_T - K)^+] = e^{\mu + \frac{1}{2}\sigma^2}\Phi(\sigma - \theta) - K\Phi(-\theta)$$

where $\theta = \frac{\log K - \mu}{\sigma}$. Note that $\mu \to \log S_0 + \left(r - \frac{1}{2}\sigma^2\right)T$, and $\sigma \to \sigma\sqrt{T}$.
In this example, S_T is log-normally distributed as

$$\log N\left(\log S_0 + \left(r - \frac{1}{2}\sigma^2\right)T, \sigma^2 T\right)$$

It follows from previous result that

$$\begin{aligned} v &= S_0\Phi(\sigma\sqrt{T} - \theta) - Ke^{-rT}\Phi(-\theta) \\ \theta &= \frac{1}{\sigma\sqrt{T}}\log\left(\frac{K}{S_0}\right) + \left(\frac{\sigma}{2} - \frac{r}{\sigma}\right)\sqrt{T} \end{aligned}$$

Example 1.4.3. [Pricing Binary Call Option]
A binary call option with maturity T pays 1 if the stock price S_T at time T is at or above a certain level K, and pays nothing otherwise. Price this option.
[Solution] The payoff of the option at time T can be written as

$$X = 1_{\{S_T \geq K\}} = \begin{cases} 1, & S_T \geq K, \\ 0, & \text{otherwise.} \end{cases}$$

The price of the option at time $t = 0$ is

$$v = \mathbb{E}[e^{-rT}X] = e^{-rT}\mathbb{P}(S_T \geq K)$$

Since $\log S_T$ is normally distributed with mean $\log S_0 + (r - \sigma^2/2)T$ and variance $\sigma^2 T$, it follows that

$$\begin{aligned} v &= e^{-rT}\mathbb{P}(\log S_T \geq \log K) \\ &= e^{-rT}\Phi\left(\frac{\log(K/S_0) - \left(r - \frac{1}{2}\sigma^2\right)\, T}{\sigma\,\sqrt{T}}\right) \end{aligned}$$

Example 1.4.4. [Pricing Asian Call Option]
Consider an Asian call option whose payoff at maturity T depends on the average of the underlying stock price

$$X = (\,\overline{S} - K)^+$$

where $\overline{S}$ is the *geometric* average of the stock price

$$\overline{S}_G = \left(\prod_{i=1}^m S_{t_i}\right)^{1/m} = (S_{t_1}\,S_{t_2}\,\cdots\,S_{t_m})^{1/m}$$

for a given set of monitoring dates $0 \leq t_1 < \cdots < t_m \leq T$. Calculate the price of this option.
[Solution] To compute $v = \mathbb{E}[e^{-rT}(\overline{S}_g - K)^+]$, where

$$\begin{aligned} S_{t_i} &= S_0 \exp\left(\left(\mu - \tfrac{1}{2}\sigma^2\right)t_i + \sigma W_{t_i}\right), \quad i = 1, \dots, m \\ \log \overline{S}_g &= \frac{1}{m}\sum_{i=1}^m \log S_t \\ &= \frac{1}{m}\sum_{i=1}^m \left[\log S_0 + \left(\mu - \frac{1}{2}\sigma^2\right)t_i + \sigma dW_{t_i}\right] \\ &= \log S_0 + \frac{1}{m}\sum_{i=1}^m \left[\left(\mu - \frac{1}{2}\sigma^2\right)t_i + \sigma W_{t_i}\right] \end{aligned}$$

Note $\log \overline{S}_g$ is a linear combination of m r.v., hence $\log \overline{S}_g$ is normally distributed.

$$\begin{aligned} \overline{\mu} = \mathbb{E}\left[\log(\overline{S}_g)\right] &= \log S_0 + \frac{1}{m}\sum_{i=1}^m \left(\mu - \tfrac{1}{2}\sigma^2\right)t_i \\ &= \log S_0 + \frac{1}{m}\left(\mu - \tfrac{1}{2}\sigma^2\right)\sum_{i=1}^m t_i \\ \overline{\sigma}^2 = \mathrm{Var}\left(\log(\overline{S}_g)\right) &= \mathrm{Var}\left(\frac{\sigma}{m}\sum_{i=1}^m W_{t_i}\right) \\ &= \frac{\sigma^2}{m^2}\mathrm{Var}\left(\sum_{i=1}^m W_{t_i}\right) \\ &= \frac{\sigma^2}{m^2}\left[\sum_{i=1}^m \mathrm{Var}[W_{t_i}] + 2\sum_{j>i}\mathrm{Cov}(W_{t_i}, W_{t_j})\right] \\ &= \frac{\sigma^2}{m^2}\left[\sum_{i=1}^m t_i + 2\sum_{i=1}^m\sum_{j=i+1}^m t_i\right], \quad \text{where } \mathrm{Cov}(W_s, W_t) = \min\{s, t\} \\ &= \frac{\sigma^2}{m^2}\sum_{i=1}^m [1 + 2(m-i)]\,t_i \end{aligned}$$

Hence we have

$$\log \overline{S}_g \sim N(\overline{\mu}, \overline{\sigma}^2), \quad \overline{S}_g \sim \log N(\overline{\mu}, \overline{\sigma}^2)$$

Thus the price of Asian option is

$$v = \mathbb{E}[e^{-rT}X] = e^{-rT}\left(e^{\overline{\mu} + \frac{1}{2}\overline{\sigma}^2}\Phi(\overline{\sigma} - \theta) - K\Phi(-\theta)\right), \quad \theta = (\log K - \overline{\mu})/\overline{\sigma}$$

Theorem 1.4.5. *Path Dependent Function for Brownian Motion with Drift*
Given a constant θ, let $B = \{B_t : t \geq 0\}$ be a Brownian motion with drift θ, i.e. $B_t = W_t + \theta t$ for $t \geq 0$, where W_t is a standard Brownian motion. Then for any $T > 0$ and any path-dependent function h, we have

$$\mathbb{E}\left[h\left(B_{[0,T]}\right)\right] = \mathbb{E}\left[\exp\left(\theta W_T - \tfrac{1}{2}\theta^2 T\right) h\left(W_{[0,T]}\right)\right]$$

Example 1.4.6. [Pricing Lookback Call Option]
A lookback call option with strike price K and maturity T is a path-dependent option, whose payoff is

$$X = \left(\max_{0 \leq t \leq T} S_t - K\right)^+$$

Assuming $K > S_0$, calculate value of the option $v = \mathbb{E}[e^{-rT}X]$.
[Solution] Given

$$\begin{aligned}
v &= \mathbb{E}[e^{-rT}X], \quad X = \left(\max_{0 \leq t \leq T} S_t - K\right)^+ \\
S_t &= S_0 \exp\left(\left(\sigma - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right) \\
&= S_0 \exp\left(\sigma\left[W_t + \left(\frac{r}{\sigma} - \frac{\sigma}{2}\right) t\right]\right) \\
&= S_0 \exp\left(\sigma\left[W_t + \theta t\right]\right) \\
&= S_0 \exp\left(\sigma B_t\right) \\
X &= \left(\max_{0 \leq t \leq T} S_t - K\right)^+ \\
&= \max_{0 \leq t \leq T}\left(S_0 \exp\left(\sigma B_t\right) - K\right)^+ \\
&= \left(S_0 \exp\left(\sigma \max_{0 \leq t \leq T} B_t\right) - K\right)^+ \\
&= h[B_{[0,T]}] \\
\mathbb{E}[X] &= \mathbb{E}[h[B_{[0,T]}]] \\
&= \mathbb{E}\left[\exp\left(\theta W_T - \frac{1}{2}\theta^2 T\right) \cdot h[W_{[0,T]}]\right] \\
&= \mathbb{E}_{W_{[0,T]}}\left[\exp\left(\theta W_T - \frac{1}{2}\theta^2 T\right) \cdot \left(S_0 \exp(\sigma \max_{0 \leq t \leq T} W_t) - K\right)^+\right] \\
&= \mathbb{E}\left[\exp\left(\theta W_T - \frac{1}{2}\theta^2 T\right) \cdot \left(S_0 \exp(\sigma M_T) - K\right)^+\right] \\
&= \int\int_{\mathbb{R}} \exp\left(\theta x - \frac{1}{2}\theta^2 T\right) \cdot \left(S_0 \exp(\sigma y) - K\right)^+ \cdot f(x,y)\, dx\, dy
\end{aligned}$$

where $f(x,y)$ is the p.d.f. of (W_T, M_T). Hence

$$\begin{aligned}
v &= \mathrm{BLSCall}(S_0, K, T, r, \sigma) + \frac{\sigma^2}{2r} S_0 \left[\Phi(\theta_+) - e^{-rT}\left(\frac{K}{S_0}\right)^{2r/\sigma^2} \Phi(\theta_-)\right] \\
\theta_\pm &= \frac{1}{\sigma\sqrt{T}} \log\frac{S_0}{K} + \left(\frac{\sigma}{2} \pm \frac{r}{\sigma}\right)\sqrt{T}
\end{aligned}$$

Example 1.4.7. [Pricing Exchange Option]
Consider two stocks whose prices are modelled by Geometric Brownian Motion.

$$\begin{aligned}
S_t &= S_0 \exp\left\{\left(r - \frac{1}{2}\sigma_1^2\right) t + \sigma_1 W_t\right\} \\
V_t &= V_0 \exp\left\{\left(r - \frac{1}{2}\sigma_2^2\right) t + \sigma_2 B_t\right\}
\end{aligned}$$

where W_t and B_t are independent standard Brownian motions.
The payoff of an exchange option at maturity T is $X = (S_T - V_T)^+$. Compute price of the option.

[Solution] The price of option is

$$v = \mathbb{E}[e^{-rT}(S_T - V_T)^+] = \mathbb{E}\left[e^{-rT}V_T\left(\frac{S_T}{V_T} - 1\right)^+\right]$$

Using the formula for S_T and V_T, we have

$$\begin{aligned} e^{-rT}V_T &= V_0 \exp\left(-\frac{1}{2}\sigma_2^2 T + \sigma_2 B_T\right) \\ \frac{S_T}{V_T} &= \frac{S_0}{V_0}\exp\left(\frac{1}{2}(\sigma_2^2 - \sigma_1^2)T - \sigma_2 B_T + \sigma_1 W_T\right) \end{aligned}$$

Since W_T and B_T are independent $N(0,T)$ random variables, their joint density function is

$$f(x,y) = \frac{1}{\sqrt{2\pi T}}\exp\left\{-\frac{x^2}{2T}\right\}\cdot\frac{1}{\sqrt{2\pi T}}\exp\left\{-\frac{y^2}{2T}\right\}$$

Express v as an expectation with respect to f, then

$$v = V_0\int\int_{\mathbb{R}^2}\left(\frac{S_0}{V_0}\exp\left(\frac{1}{2}(\sigma_2^2 - \sigma_1^2)T - \sigma_2 y + \sigma_1 x\right) - 1\right)^+ g(x,y)\ dx\ dy$$

where

$$\begin{aligned} g(x,y) &= f(x,y)\exp\left(-\frac{1}{2}\sigma_2^2 T + \sigma_2 y\right) \\ &= \frac{1}{\sqrt{2\pi T}}\exp\left\{-\frac{x^2}{2T}\right\}\cdot\frac{1}{\sqrt{2\pi T}}\exp\left\{-\frac{(y-\sigma_2 T)^2}{2T}\right\} \end{aligned}$$

Note $g(x,y)$ is the joint density function of two independent variables $X \sim N(0,T)$ and $Y \sim N(\sigma_2 T, T)$. Then the price v can be written as an expectation with respect to g:

$$\begin{aligned} v &= V_0\mathbb{E}\left[\left(\frac{S_0}{V_0}\exp\left(\frac{1}{2}(\sigma_2^2 - \sigma_1^2)T - \sigma_2 Y + \sigma_1 X\right) - 1\right)^+\right] \\ &= V_0\mathbb{E}[(U-1)^+] \end{aligned}$$

where U is a lognormal random variable with distribution

$$\log N\left(\log\frac{S_0}{V_0} - \frac{1}{2}(\sigma_2^2 + \sigma_1^2)T,\ (\sigma_2^2 + \sigma_1^2)T\right)$$

It follows that

$$v = S_0\Phi\left(\frac{1}{\sigma\sqrt{T}}\log\frac{S_0}{V_0} + \frac{1}{2}\sigma\sqrt{T}\right) - V_0\Phi\left(\frac{1}{\sigma\sqrt{T}}\log\frac{S_0}{V_0} - \frac{1}{2}\sigma\sqrt{T}\right)$$

where $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$

1.5 Stochastic Integrals

Definition 1.5.1. *[Ito Integral]*
Consider class of stochastic integrals

$$\int_0^T X_t dW_t$$

where W_t is a standard Brownian motion, and X_t is a general stochastic process.
The Ito integral is defined as

$$\int_0^T X_t dW_t = \lim_{n\to\infty} \sum_{i=0}^{n-1} X_{t_i}(W_{t_{i+1}} - W_{t_i})$$

where $0 = t_0 < t_1 < \cdots < t_n = T$ is a partition of the interval $[0, T]$, and

$$\max_{0\le i\le n-1}(t_{i+1} - t_i) \to 0, \quad \text{as } n \to \infty$$

Example 1.5.2. [Compute Stochastic Integral]
Compute stochastic integral $\int_0^T 2W_t dW_t$.
Consider partition $0 = t_0 < t_1 < \cdots < t_n = T$ with $t_i = iT/n$. By definition,

$$\begin{aligned}
\int_0^T 2W_t dW_t &= \lim_{n\to\infty} \sum_{i=0}^{n-1} 2W_{t_i}(W_{t_{i+1}} - W_{t_i}) \\
&= \lim_{n\to\infty} \sum_{i=0}^{n-1} [(W_{t_{i+1}} + W_{t_i}) - (W_{t_{i+1}} - W_{t_i})](W_{t_{i+1}} - W_{t_i}) \\
&= W_T^2 - \lim_{n\to\infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 \\
&= W_T^2 - \lim_{n\to\infty} \frac{T}{n} \sum_{i=0}^{n-1} Z_i^2 = W_T^2 - T
\end{aligned}$$

Remark 1.5.3. *[Properties of Stochastic Integrals]*

i. Linearity: For any constant a, b,

$$\int_0^T (aX_t + bY_t)dW_t = a\int_0^T X_t dW_t + b\int_0^T Y_t dW_t$$

ii. Martingale Property:

$$\mathbb{E}\left[\int_0^T X_t dW_t\right] = 0$$

iii. Ito isometry:

$$\mathbb{E}\left[\int_0^T X_t dW_t\right]^2 = \int_0^T \mathbb{E}[X_t^2]dt$$

iv. Assume $f(t)$ is deterministic function, then stochastic integral $\int_0^T f(t)dW_t$ is normally distributed with mean 0 and variance $\int_0^T f^2(t)dt$.

Remark 1.5.4. *[Differential Notation]*
Let X_t be stochastic process satisfying

$$X_t = X_0 + \int_0^t Y_s ds + \int_0^t Z_s dW_s$$

The differential notation for above equation is

$$dX_t = Y_t dt + Z_t dW_t$$

Note that

$$(dt)^2 = 0, \quad dtdW_t = dW_t dt = 0, \quad (dW_t)^2 = dt, \quad dW_t dB_t = \rho dt$$

where (W_t, B_t) is a 2-dimensional Brownian motion with covariance

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

Definition 1.5.5. *[Single Dimension Ito Formula]*
Suppose $f : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable function, and W_t is standard Brownian motion. Then

$$\begin{aligned} f(W_T) &= f(W_0) + \int_0^T f'(W_t)dW_t + \frac{1}{2}\int_0^T f''(W_t)dt \\ df(W_t) &= f'(W_t)dW_t + \frac{1}{2}f''(W_t)dt \end{aligned}$$

Definition 1.5.6. *[General Ito Formula]*
Suppose $f(t,x)$ is continuously differentiable with respect to $x \in \mathbb{R}^d$. Then for d-dimension process X_t,

$$\begin{aligned} df(t, X_t) &= \frac{\partial}{\partial t}f(t, X_t)dt + (\nabla f(t, X_t))^T dX_t + \frac{1}{2}dX_t^T H(t, X_t)dX_t \\ d(X_t) &= (d(X_t^{(1)}), d(X_t^{(2)}), \ldots, d(X_t^{(d)})) \\ \nabla f &= \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_d}\right) \end{aligned}$$

where H is a $d \times d$ matrix, with ij-th entry given by $\frac{\partial^2 f}{\partial x_i \partial y_j}$

Example 1.5.7. [Simple Example One with Ito Formula]
Suppose $f = f(X_t)$, where X_t is a stochastic process. Then

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2$$

Example 1.5.8. [Simple Example Two with Ito Formula]
Suppose $f = X_t Y_t$, where X_t, Y_t are stochastic processes. Then

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t$$

Example 1.5.9. [Geometric Brownian Motion Ito Formula]
Let W_t be a standard Brownian motion. Consider geometric Brownian motion

$$S_t = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right)$$

Compute differential dS_t.
Observe that $S_t = f(t, W_t)$, where

$$f(t,x) = S_0 \left(\left(r - \frac{1}{2}\sigma^2\right)t + \sigma x\right)$$

Applying Ito formula,

$$\begin{aligned} dS_t &= \frac{\partial}{\partial t}f(t, W_t)dt + \frac{\partial}{\partial x}f(t, W_t)dW_t + \frac{1}{2}\frac{\partial^2}{\partial x^2}f(t, W_t)(dW_t)^2 \\ &= rS_t dt + \sigma S_t dW_t \end{aligned}$$

Example 1.5.10. [Examples of Financial Models]

i. Black-Scholes Model: stock price is modelled by

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

whose solution is a geometric Brownian motion.

ii. CEV Model: stock price is assumed to satisfy

$$dS_t + \mu S_t dt + \sigma S_t^{\gamma} dW_t$$

where γ is a positive constant.

iii. Stochastic Volatility Models: volatility of underlying asset is a diffusion process itself, i.e.,

$$\begin{aligned} dS_t &= \mu S_t d_t + \sqrt{\theta_t} S_t dW_t \\ d\theta_t &= a(b-\theta_t)dt + \sigma\sqrt{\theta_t} dB_t \end{aligned}$$

where $a, b > 0$ and (W_t, B_t) is a two-dimensional Brownian motion.

Example 1.5.11. [Example One with Ito Formula]
Solve the SDE

$$dX_t = rX_t dt + \theta(t)dW_t, \quad X_0 = x$$

where r is a constant and $\theta(t)$ is a given function.
Consider process $Y_t = e^{-rt}X_t$. Apply product rule,

$$\begin{aligned} dY_t &= -re^{-rT}X_t dt + e^{-rT}dX_t = e^{-rT}\theta(t)dW_t \\ \Rightarrow Y_t &= Y_0 + \int_0^t e^{-rs}\theta(s)dW_s \\ \Rightarrow X_t &= e^{rt}Y_t = e^{rt}\left(x + \int_0^t e^{-rs}\theta(s)dW_s\right) \end{aligned}$$

where X_t is a normal r.v. at any time t with mean xe^{rt} and variance $e^{2rt}\int_0^t e^{-2rs}\theta^2(s)ds$.

Example 1.5.12. [Example Two with Ito Formula]
Solve the SDE

$$dX_t = a(b - X_t)dt + \sigma dW_t, \quad X_0 = x$$

where a, b, σ are constants.
Let $Y_t = e^{at}X_t$, apply product rule.

$$\begin{aligned} dY_t &= ae^{at}X_t dt + e^{at}dX_t = abe^{at}dt + \sigma e^{at}dW_t \\ \Rightarrow X_t &= e^{-at}Y_t = e^{-at}\left(x + \int_0^t abe^{as}ds + \int_0^t \sigma e^{as}dW_s\right) \\ &= e^{-at}x + b(1-e^{-at}) + e^{-at}\int_0^t \sigma e^{as}dW_s \end{aligned}$$

For any t, X_t is normally distributed as

$$N\left(xe^{-at} + b(1-e^{-at}), \frac{\sigma^2}{2a}(1-e^{-2at})\right)$$

Example 1.5.13. [Example Three with Ito Formula]
Solve the SDE

$$dX_t = rX_t dt + \theta(t)X_t dW_t, \quad X_0 = x > 0$$

where r is constant and $\theta(t)$ is a given function.

Define process $Y_t = \log X_t$, and apply Ito formula,

$$\begin{aligned} dY_t &= \frac{1}{X_t} dX_t - \frac{1}{2}\frac{1}{X_t^2}(dX_t)^2 \\ &= rdt + \theta(t)dW_t - \frac{1}{2}\theta^2(t)dt \end{aligned}$$

Therefore

$$X_t = e^{Y_t} = x \cdot \exp\left(\int_0^t \left(r - \frac{1}{2}\theta^2(s)\right) ds + \int_0^t \theta(s)dW_s\right)$$

Example 1.5.14. [Risk Neutral Pricing with Ito Formula]
Assume underlying stock price satisfies the SDE

$$\frac{dS_t}{S_t} = rdt + \theta(t)dW_t$$

under the risk-neutral probability measure, where the risk-free interest rate r is a constant. Find the price of call option with maturity T and strike price K.
Price of option is $v = \mathbb{E}[e^{-rT}(S_T - K)^+]$. Stock price at time T is

$$\begin{aligned} S_T &= S_0 \cdot \exp\left(\int_0^T \left(r - \frac{1}{2}\theta^2(t)\right) dt + \int_0^T \theta(t)dW_t\right) \\ &= S_0 \cdot \exp\left(\left(r - \frac{1}{2}\sigma^2\right) T + \sigma\sqrt{T}Z\right) \end{aligned}$$

where $Z \sim N(0,1)$, and

$$\sigma = \sqrt{\frac{1}{2}\int_0^T \theta^2(t)dt}$$

The stock price S_T has same distribution as terminal stock price in Black-Scholes model. Therefore the option price is given by Black-Scholes formula with drift r and volatility σ.

2 Monte Carlo Methods

2.1 Monte Carlo Simulation

Method 2.1.1. *Fundamentals of Monte Carlo Simulation*
Consider problem of estimation expected value of some function of random variable X

$$\mu = \mathbb{E}[h(X)]$$

A Monte Carlo simulation proceeds as follows

1. Generate independent, identically distributed (iid) random numbers for the distribution of X
2. The estimate of expected value μ is defined to be the sample average

$$\hat{\mu}_n = \frac{1}{n}[h(X_1) + \cdots + h(X_n)]$$

Note the Monte Carlo estimate $\hat{\mu}$ is a random variable as it depends on the samples.

Definition 2.1.2. *Error of Monte Carlo Estimate*
Let $\hat{\mu}$ be Monte Carlo estimate of $\mu = \mathbb{E}[H], H = h(x)$.
The central limit theorem (CLT) gives distribution of error $\hat{\mu}_n - \mu$, where $\hat{\mu}_n \sim N(\mu, \frac{\sigma_H^2}{n}), \sigma_H^2 = \mathrm{Var}(h(x))$.
Let σ_H^2 be variance of H. The CLT asserts that as $n \to \infty$,

$$\frac{H_1 + \cdots + H_n - n\mu}{\sigma_H \sqrt{n}} = \frac{\sqrt{n}(\hat{\mu}) - \mu}{\sigma_H}$$

converges to the standard normal distribution, i.e., for any $a \in \mathbb{R}$,

$$\mathbb{P}\left\{\frac{\sqrt{n}(\hat{\mu} - \mu)}{\sigma_H} \leq a\right\} \to \Phi(a), \quad n \to \infty$$

Hence the error $\hat{\mu}_n - \mu$ is approx. normally distributed with mean 0 and variance σ_H^2/n.

Definition 2.1.3. *Confidence Interval of Monte Carlo Estimate*
When n is large,

$$\mathbb{P}\left(-z_{\alpha/2} \leq \frac{\sqrt{n}(\hat{\mu}_n - \mu)}{\sigma_H} \leq z_{\alpha/2}\right) \approx 1 - 2\Phi(-z_{\alpha/2}) = 1 - \alpha$$

i.e., with probability $1 - \alpha$, the interval

$$\mu \in \left(\hat{\mu}_n - z_{\alpha/2}\frac{\alpha_H}{\sqrt{n}}, \hat{\mu}_n + z_{\alpha/2}\frac{\alpha_H}{\sqrt{n}}\right)$$

covers the true value μ. This is the $(1 - \alpha)$ confidence interval for μ.
The standard deviation is usually replaced by sample standard deviation

$$s_H = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(H_i - \hat{\mu}_n)^2}$$

The $1 - \alpha$ confidence interval is then

$$\mu \in \left(\hat{\mu}_n - z_{\alpha/2}\frac{s_H}{\sqrt{n}}, \hat{\mu}_n + z_{\alpha/2}\frac{s_H}{\sqrt{n}}\right)$$

Example 2.1.4. [Bias in Price of Lookback Call Option]
For price of lookback call option with payoff

$$X = \left(\max_{0 \leq t \leq T} S_t - K\right)^+$$

In simulations, the estimate of option price is given by sample mean of

$$e^{-rT}\left(\max_{0 \leq t \leq T} s_{t_i} - K\right)^+$$

This introduces a negative bias since

$$\max_{i=1,\dots,m} s_{t_i} \le \max_{0\le t\le T} S_t$$

The bias cannot be eliminated by increasing sample size.

Example 2.1.5. [Estimate Price of European Call Option with Monte Carlo]
Estimate price of European call option on stock whose price is geometric Brownian motion.
[Solution] The price of call option with strike price K and maturity T is

$$v = \mathbb{E}[e^{-rT}(S_T - K)^+]$$

Expectation is taken with respect to risk-neutral measure, where stock is geometric BM with drift r:

$$S_T = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) T + \sigma W_T\right)$$

where $W_T \sim N(0,T)$. Once can write $W_T = \sqrt{T}Z$ where $Z \sim N(0,1)$.

Algorithm 3: Monte Carlo Algorithm for European Call Option Pricing

for $i = 1$ to n **do**
  Generate Z_i from $N(0,1)$.
  Set $Y_T = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) T + \sigma\sqrt{T}Z_i\right)$
  Set $X_i = e^{-rT}(Y_i - K)^+$
end for
Compute estimate $\hat{v} = (X_1 + X_2 + \cdots + X_n)/n$
Compute standard error S.E.$= \sqrt{\frac{1}{n(n-1)}\left(\sum_{i=1}^{n} X_i^2 - n\overline{v}^2\right)}$

Example 2.1.6. [Estimate Price of Asian Option with Monte Carlo]
Consider call option whose payoff at maturity T is

$$X = \left(\frac{1}{m}\sum_{i=1}^{m} S_i - K\right)^+$$

where $0 < t_1 < \cdots < t_m = T$ are a fixed set of dates. Assume under risk-neutral measure

$$S_t = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right)$$

Estimate the price of the option.
[Solution] Note that

$$\begin{aligned}
S_t &= S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right) \\
S_{t_{i+1}} &= S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) t_{i+1} + \sigma W_{t_{i+1}}\right) \\
S_{t_i} &= S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) t_i + \sigma W_{t_i}\right) \\
\frac{S_{t_{i+1}}}{S_{t_i}} &= \exp\left(\left(r - \frac{1}{2}\sigma^2\right) (t_{i+1} - t_i) + \sigma(W_{t_{i+1}} - W_{t_i})\right)
\end{aligned}$$

Also, $W_{t_{i+1}} = \sqrt{(t_{i+1} - t_i)}Z_{i+1}$ where $Z_{i+1} \sim N(0,1)$. Hence $(W_{t_{i+1}} - W_{t_i}) \sim N(0,(t_{i+1} - t_i))$.
Hence, the discrete path $(S_{t_1}, S_{t_2}, \dots, S_{t_m})$ is simulated sequentially from t_1 to t_m:

$$S_{t_{i+1}} = S_{t_i} \exp\left(\left(r - \frac{1}{2}\sigma^2\right) (t_{i+1} - t_i) + \sigma(W_{t_{i+1}} - W_{t_i})\right)$$

where $(W_{t_1} - W_{t_0}, \dots, W_{t_m} - W_{t_{m-1}})$ are i.i.d. normal random variables.

Algorithm 4: Monte Carlo Algorithm for Asian Option Pricing

for $j = 1$ to n **do**
  for $i = 1$ to m **do**
    Generate Z_i from $N(0,1)$.
    Set $S_{t_i} = S_{t_{i-1}} \exp\left(\left(r - \frac{1}{2}\sigma^2\right)(t_i - t_{i-1}) + \sigma\sqrt{t_i - t_{i-1}} Z_i\right)$
  end for
  Compute discounted payoff $X_j = e^{-rT}\left(\frac{1}{m}\sum_{i=1}^{m} S_{t_i} - K\right)^+$
end for
Compute estimate $\hat{v} = (X_1 + X_2 + \cdots + X_n)/n$
Compute standard error S.E.$= \sqrt{\frac{1}{n(n-1)}\left(\sum_{i=1}^{n} X_i^2 - n\overline{v}^2\right)}$

Example 2.1.7. [Estimate Price of Spread Call Option with Monte Carlo]
Estimate price of spread call option whose payoff at maturity T is

$$X = (X_T - Y_T - K)^+$$

where X_T and Y_T are prices of two underlying assets. Assume under risk-neural measure,

$$X_t = X_0 \exp\left(\left(r - \frac{1}{2}\sigma_1^2\right)t + \sigma_1 W_t\right)$$
$$Y_t = Y_0 \exp\left(\left(r - \frac{1}{2}\sigma_2^2\right)t + \sigma_2 B_t\right)$$

where $(W_T, B_T) \sim N(0, T\Sigma)$ is a two-dimensional BM with covariance matrix

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

[Solution] Let $Z = (Z_1, Z_2)'$ be two-dimensional standard normal random vector.
For any 2×2 matrix C, the random vector

$$R = CZ = \begin{pmatrix} C_{11}Z_1 + C_{12}Z_2 \\ C_{21}Z_1 + C_{22}Z_2 \end{pmatrix} \sim N(0, CC')$$

is a jointly normal with mean 0 and covariance CC'. Let CC' be Cholesky factorization of Σ, i.e.,

$$\Sigma = CC' = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}\begin{pmatrix} 1 & \rho \\ \sqrt{1-\rho^2} & 0 \end{pmatrix}$$

Then $\sqrt{T}R$ has same distribution as (W_T, B_T).

Algorithm 5: Monte Carlo Algorithm for Spread Call Option

Set $C_{11} = 1$, $C_{21} = \rho$, $C_{22} = \sqrt{1-\rho^2}$
for $i = 1$ to n **do**
  Generate Z_i and Z_2 from $N(0,1)$.
  Set $R_1 = C_{11}Z_1$ and $R_2 = C_{21}Z_1 + C_{22}Z_2$
  Set $X_i = X_0 \exp\left(\left(r - \frac{1}{2}\sigma_1^2\right)T + \sigma_1\sqrt{T}R_1\right)$, where $\sqrt{T}R_1 = W_T$
  Set $Y_i = Y_0 \exp\left(\left(r - \frac{1}{2}\sigma_2^2\right)T + \sigma_2\sqrt{T}R_2\right)$, where $\sqrt{T}R_2 = B_T$
  Compute discounted payoff $H_i = e^{-rT}(X_i - Y_i - K)^+$
end for
Compute estimate $\hat{v} = (H_1 + H_2 + \cdots + H_n)/n$
Compute standard error S.E.$= \sqrt{\frac{1}{n(n-1)}\left(\sum_{i=1}^{n} H_i^2 - n\overline{v}^2\right)}$

Example 2.1.8. [Sampling Multivariate Normal Distributions]
Consider an option whose payoff depends on d assets. Under the risk-neutral probability measure, the prices of these assets are geometric Brownian motions,

$$S_t^{(i)} = S_t^{(0)} \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma_i W_t^{(i)}\right\}, \quad i = 1, \dots, d$$

where $W = (W^{(1)}, \dots, W^{(d)})$ is a d-dimensional Brownian motion with covariance matrix $\Sigma_{d\times d}$. Note that

$$\begin{pmatrix} W_t^{(1)} \\ W_t^{(2)} \\ \dots \\ W_t^{(d)} \end{pmatrix} \sim N(0, T\Sigma)$$

The option has maturity T and payoff

$$X = \left(\max\{c_1 S_T^{(1)}, \dots, c_d S_T^{(d)}\} - K\right)^+$$

To sample from $N(\mu, \Sigma)$, find a matrix A such that $AA^T = \Sigma$. Generate independent samples $Z_1, \dots, Z_d$ from $N(0,1)$. Set $Z = (Z_1, \dots, Z_d)^T$ and $X = \mu + AZ \sim N(\mu, \Sigma)$.
For Cholesky factorisation of Σ, set

$$A = \begin{pmatrix} A_{11} & 0 & \dots & 0 \\ A_{21} & A_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{d1} & A_{d2} & \dots & A_{dd} \end{pmatrix}$$

where

$$A_{jj} = \sqrt{\Sigma_{jj} - \sum_{k=1}^{j-1} A_{jk}^2}, \quad A_{ij} = \left(\Sigma_{ij} - \sum_{k=1}^{j-1} A_{ik}A_{jk}\right)/A_{jj}, \quad i > j \geq 1$$

Algorithm 6: Sampling from Multivariate Normal Distributions

Compute Cholesky factorisation $\Sigma = AA^T$
for $i = 1$ to n **do**
  Generate $Z_i, \dots, Z_d$ from $N(0,1)$.
  Set $Z = (Z_1, \dots, Z_d)'$ and $Y = AZ$
  for $k = 1$ to d **do**
    Set $S_T^{(k)} = S_T^{(0)} \exp\left\{\left(r - \frac{1}{2}\sigma_k^2\right)T + \sigma_k \sqrt{T} Y_k\right\}$, where $\sqrt{T}Y_k = W_T^{(k)}$
  end for
  Compute discounted payoff $H_i = e^{-rT}(\max\{c_1 S_T^{(1)}, \dots, c_d S_T^{(d)} - K\})^+$
end for
Compute estimate $\hat{v} = (H_1 + H_2 + \dots + H_n)/n$
Compute standard error S.E.= $\sqrt{\frac{1}{n(n-1)}\left(\sum\limits_{i=1}^{n} H_i^2 - n\overline{v}^2\right)}$

Example 2.1.9. [Value at Risk of Portfolio]
Denote by X_i the daily return of a portfolio. Assume $X = \{X_1, X_2, \dots\}$ is a Markov chain.
Given $X_i = x$, $X_{i+1} \sim N(0, \beta_0 + \beta_1 x^2)$ for some $\beta_0 > 0$ and $0 < \beta_1 < 1$. Total return with m-day period is

$$S = \sum_{i=1}^{m} X_i$$

Assume X_1 is standard normal random variable, estimate value at risk (VaR) at confidence interval $1 - p$,

$$\mathbb{P}(S \leq -\text{VaR}) = p$$

Algorithm 7: Value at Risk

for $j = 1$ to n **do**
  Generate X_1 from $N(0,1)$.
  for $i = 2$ to m **do**
    Generate Z from $N(0,1)$.
    Set $X_i = \sqrt{\beta_0 + \beta_1 X_{i-1}^2} \cdot Z$ (where $\sqrt{\beta_0 + \beta_1 X_{i-1}^2}$ is the standard deviation of X_i)
  end for
  Set $S_j = \sum_{i=1}^m X_i$
end for
Sort $\{S_1, \dots, S_n\}$ in increasing order
The number of samples below or at -VaR $\approx [np]$. Hence $\widehat{\text{VaR}} = -S_k, k = [np]$

Example 2.1.10. [Value at Risk Confidence Interval]
Estimate $(1-\alpha)$ confidence interval. Find integers $k_1 < k_2$ such that

$$\mathbb{P}(S_{k_1} \leq -\text{VaR} < S_{k_2}) = 1 - \alpha$$

[Solution]. Let Y denote number of samples that are less than or equal to -VaR. Notice that

$$\mathbb{P}(S_{k_1} \leq -\text{VaR} < S_{k_2}) = \mathbb{P}(k_1 \leq Y < k_2)$$

Also, notice that $Y \sim \text{Bin}(n,p)$, which can be approximated by normal distribution with mean np and variance $np(1-p)$ (central limit theorem). Therefor

$$\begin{aligned}
\mathbb{P}(S_{k_1} \leq -\text{VaR} < S_{k_2}) &\approx 1 - \Phi\left(\frac{k_1 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(-\frac{k_2 - np}{\sqrt{np(1-p)}}\right) \\
k_1 &= np - \sqrt{np(1-p)} z_{\alpha/2} \\
k_2 &= np + \sqrt{np(1-p)} z_{\alpha/2}
\end{aligned}$$

Example 2.1.11. [Credit Risk]
Consider credit risk model where losses are due to default of obligors on contractual payments. Suppose there are m obligors and the i-th obligor defaults if and only if $X_i \geq x_i$ for some r.v. X_i and given level x_i.
The random variable X_i is assumed to take the form

$$X_i = \rho_i Z + \sqrt{1 - \rho_i^2} \epsilon_i, \quad i = 1, \dots, m$$

where $Z, \epsilon_1, \dots, \epsilon_m$ are independent standard normal r.v., and ρ_i are constant satisfying $-1 < \rho_i < 1$.
Let c_i denote the loss from the default of i obligor. Then the total loss is

$$L = \sum_{i=1}^m c_i 1_{\{X_1 \geq x_i\}} = \begin{cases} 1 & \text{if } X_i \geq x_i \\ 0 & \text{otherwise} \end{cases}$$

Estimate the probability that L exceeds a given threshold h. *[Solution]* Note that

$$\begin{aligned}
\mathbb{P}\{L > h\} &= \int_h^\infty f(x) dx \quad \text{where } f \text{ is the pdf of } L \\
&= \int_{\mathbb{R}} 1_{\{x > h\}} f(x) \, dx \\
&= \mathbb{E}[1_{\{L > h\}}]
\end{aligned}$$

Hence the Monte Carlo method requires drawing n independent samples $L_1, \dots, L_n$, then taking the value

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n 1_{\{L_i > h\}}$$

Algorithm 8: Credit Risk

for $k = 1$ to n **do**

  Generate independent samples $Z_1, \epsilon_1, \ldots, \epsilon_m$ from $N(0,1)$.

  Compute $X_i = \rho_i Z + \sqrt{1-\rho_i^2}\epsilon_i$ for $i = 1, \ldots, m$

  Compute $L = \sum_{i=1}^{m} c_i 1_{\{X_i \geq x_i\}}$

  Set $H_k = 1$ if $L > h$, set $H_k = 0$ otherwise

end for

Compute the estimate $\hat{v} = (H_1 + \cdots + H_n)/n$

Compute the standard error $\sqrt{\frac{1}{n(n-1)}\left(\sum_{k=1}^{n} H_k^2 - n\hat{v}^2\right)}$

2.2 Antithetic Sampling

Remark 2.2.1. [Comparison with Plain Monte Carlo Simulation]
In plain Monte Carlo, generate $2n$ i.i.d. samples $X_1, \ldots, X_{2n}$. The sample average, variance, error is then

$$v_1 = \frac{1}{2n}\sum_{i=1}^{2n} X_i, \quad \mathrm{Var}[v_1] = \frac{\sigma_x^2}{2n}, \quad \text{Error} = \frac{\sigma_x}{\sqrt{2n}}$$

In antithetic sampling, generate n pairs of i.i.d samples $(X_1, Y_1), \ldots, (X_n, Y_n)$, where Y_i has same distribution as X_i. Note X_i and Y_i are dependent. The sample average, variance, error are then

$$\begin{aligned}
v_2 &= \frac{1}{2n}\sum_{i=1}^{n}(X_i + Y_i), \quad \mathbb{E}[v_2] = \frac{1}{2n}\sum_{i=1}^{n}(\mathbb{E}[X_i] + \mathbb{E}[Y_i]) = \mu \\
\mathrm{Var}[v_2] &= \mathrm{Var}\left[\frac{1}{2n}\sum_{i=1}^{n}(X_i + Y_i)\right] = \frac{1}{(2n)^2}\sum_{i=1}^{n}\mathrm{Var}[X_i + Y_i] \\
\mathrm{Var}[X_i + Y_i] &= \mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\mathrm{Cov}(X, Y) \\
&= \sigma_x^2 + \sigma_x^2 + 2\beta\sigma_x\sigma_x = 2(1+\beta)\sigma_x^2 \\
\mathrm{Var}[v_2] &= \frac{1}{(2n)^2}\sum_{i=1}^{n} 2(1+\beta)\sigma_x^2 = \frac{1}{2n}(1+\beta)\sigma_x^2
\end{aligned}$$

Note that if X, Y are negatively correlated ($\beta < 0$), then variance reduction is achieved.
The MC estimate, standard deviation, and standard error is then

$$\begin{aligned}
\hat{v} &= \frac{1}{2n}\sum_{i=1}^{n}(X_i + Y_i) \\
\sigma &= \sqrt{\frac{1}{n}\mathrm{Var}\left[\frac{X_i + Y_i}{2}\right]} \\
\text{S.E.} &= \sqrt{\frac{1}{n(n-1)}\left(\sum_{i=1}^{n}\left(\frac{X_i + Y_i}{2}\right)^2 - n\hat{v}^2\right)}
\end{aligned}$$

Example 2.2.2. [Generating Antithetic Samples]
If $X = h(U_1, \ldots, U_k)$ where $U_1, \ldots, U_k$ are i.i.d uniform on $[0, 1]$ and h is monotone, then the sample is

$$Y = h(1 - U_1, \ldots, 1 - U_k)$$

If $X = h(Z_1, \ldots, Z_k)$ where $Z_1, \ldots, Z_k$ are i.i.d $N(0, 1)$ and h is monotone, then the sample is

$$Y = h(-Z_1, \ldots, -Z_k)$$

Example 2.2.3. [Antithetic Sampling for European Call Option]
Assume European call option has maturity T and strike price K.
Assume underlying stock price is a geometric Brownian motion under risk-neutral probability measure

$$S_t = S_0 \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}$$

The discounted option payoff

$$X = e^{-rT}(S_T - K)^+ = h(W_T)$$

is an increasing function of S_T, thus an increasing function of W_T. Hence the antithetic sample

$$Y = h(-W_T)$$

is expected to reduce the variance.

Algorithm 9: Antithetic Sampling for European Call Option

for $i = 1$ to n **do**
  Generate a sample Z from $N(0,1)$.
  Set $S_i = S_0 \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z\right\}$
  Set $X_i = e^{-rT}(S_i - K)^+$ (This is $X = h(Z)$)
  Set $S_i = S_0 \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)T - \sigma\sqrt{T}Z\right\}$
  Set $Y_i = e^{-rT}(S_i - K)^+$ (This is $Y = h(-Z)$)
end for
Compute the estimate $\hat{v} = \frac{1}{2n}\sum_{i=1}^{n}(X_i + Y_i)$
Compute the standard error $\sqrt{\frac{1}{n(n-1)}\left(\sum_{k=1}^{n}\left(\frac{X_i+Y_i}{2}\right)^2 - n\hat{v}^2\right)}$

Example 2.2.4. [Antithetic Sampling for Barrier Option]
Consider discretely monitored down-and-out barrier option with maturity T and payoff

$$(S_T - K)^+ \cdot 1_{\{\min(S_{t_1},\ldots,S_{t_m}) \geq b\}}$$

where $0 < t_1 < \ldots < t_m = T$ are monitoring dates, and the stock price is modelled by geometric Brownian motion. The payoff is an increasing function of stock price.
[Solution] Note

$$\mu = \mathbb{E}[X]$$
$$X = e^{-rT}(S_T - K)^+ \cdot 1_{\{\min(S_{t_1},\ldots,S_{t_m}) \geq b\}}$$
$$\begin{cases} S_{t_{i+1}} = S_{t_i} \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(t_{i+1} - t_i) + \sigma\sqrt{t_{i+1} - t_i}Z_{i+1}\right\} \\ S_{t_0} = S_0 \end{cases}$$

Hence the sample and antithetic samples are

$$X = h(Z_1, \ldots, Z_m), \quad Y = h(-Z_1, \ldots, -Z_m)$$

Algorithm 10: Antithetic Sampling for Barrier Option

for $i = 1$ to n **do**
  for $j = 1$ to m **do**
    Generate a sample Z from $N(0,1)$.
    Set $S_j = S_{j-1} \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(t_j - t_{j-1}) + \sigma\sqrt{t_j - t_{j-1}}Z\right\}$
    Set $\overline{S}_j = \overline{S}_{j-1} \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(t_j - t_{j-1}) - \sigma\sqrt{t_j - t_{j-1}}Z\right\}$
  end for
  Set $X_i = e^{-rT}(S_m - K)^+ \cdot 1_{\{\min(S_1,\ldots,S_m \geq b\}}$
  Set $Y_i = e^{-rT}(\overline{S}_m - K)^+ \cdot 1_{\{\min(\overline{S}_1,\ldots,\overline{S}_m \geq b\}}$
end for
Compute the estimate $\hat{v} = \frac{1}{2n}\sum_{i=1}^{n}(X_i + Y_i)$
Compute the standard error $\sqrt{\frac{1}{n(n-1)}\left(\sum_{k=1}^{n}\left(\frac{X_i+Y_i}{2}\right)^2 - n\hat{v}^2\right)}$

Example 2.2.5. [Antithetic Sampling for Butterfly Spread]
Estimate price of butterfly spread option with maturity T and payoff

$$(S_T - K_1)^+ + (S_T - K_3)^+ - 2(S_T - K_2)^+$$

where $0 < K_1 < K_3$ and $K_2 = (K_1 + K_3)/2$.
Note the payoff is non-monotone, hence using antithetic sampling may increase the variance.

2.3 Control Variates

Remark 2.3.1. [Comparison with Plain Monte Carlo Simulation]
In plain Monte Carlo, generate n i.i.d. samples $X_1, \dots, X_n$. The sample average, variance, error is then

$$v_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \mathrm{Var}[v_1] = \frac{\sigma_x^2}{n}, \quad \text{Error} = \frac{\sigma_x}{\sqrt{n}}$$

In control variates, generate n pairs of i.i.d samples $(X_1, Y_1), \dots, (X_n, Y_n)$, where $Y_i, \dots, Y_n$ are control variate samples. Y_i has known expected value $\overline{\mu}$. The estimate is then

$$v_2 = \frac{1}{n}\sum_{i=1}^{n} \left[X_i - b(Y_i - \overline{\mu})\right], \quad \overline{\mu} = \mathbb{E}[Y] \text{ is known}$$

$$\mathbb{E}[v_2] = \frac{1}{2}\sum_{i=1}^{n} \left[\mathbb{E}[X_i] - b(\mathbb{E}[Y_i] - \overline{\mu})\right] = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu = \mathbb{E}[X] \quad \text{(unbiased estimate)}$$

Hence

$$\begin{aligned}
\mathrm{Var}[v_2] &= \mathrm{Var}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - b(Y_i - \overline{\mu})\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}[X_i - b(Y_i - \overline{\mu})] \\
&= \frac{1}{n}\mathrm{Var}[X - b(Y - \overline{\mu})] \\
&= \frac{1}{n}[\sigma_X^2 - 2b\mathrm{Cov}(X, Y) + b^2\sigma_Y^2] \\
&= \frac{1}{n}(\sigma_Y^2 b^2 - 2b\beta\sigma_X\sigma_Y + \sigma_X^2)
\end{aligned}$$

The optimised value for b is then

$$b^* = \frac{\beta\sigma_X\sigma_Y}{\sigma_Y^2} = \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(Y)}$$

Set $b = b^*$, then

$$\mathrm{Var}[v_2] = \frac{1}{n}(1 - \beta^2)\sigma_X^2, \quad -1 \leq \beta \leq 1$$

Using control variate estimate reduces variance by factor of β^2. Therefore the control variate Y should be chosen so that it has strong correlation with X.
The MC estimate, standard deviation, and standard error of control variate is then

$$\begin{aligned}
\overline{v} &= \frac{1}{n}\sum_{i=1}^{n} H_i, \quad H_i = X_i - b(Y_i - \overline{\mu}) \\
\sigma &= \sqrt{\frac{1}{n}\mathrm{Var}[H]} \\
\text{S.E.} &= \sqrt{\frac{1}{n(n-1)}\left(\sum_{i=1}^{n} H_i^2 - n\hat{v}^2\right)}
\end{aligned}$$

The optimal coefficient b^* can be estimated using sample variance and covariance,

$$b^* = \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}[Y]} \approx \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2}$$

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

Example 2.3.2. [Control Variate for European Call Option]
Estimate price of European call option with maturity T and strike price K.
Under risk-neutral measure, stock price is modelled by geometric Brownian motion:

$$S_t = S_0 \exp\left\{\left(r - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right\}$$
$$\mu = \mathbb{E}[X], \quad X = e^{-rT}(S_T - K)^+, \quad Y = e^{-rT} S_T$$

Use discounted stock price $e^{-rT}S_T$ as control variate. Since

$$\mathbb{E}[Y] = \mathbb{E}[e^{-rT}S_T] = S_0 = \hat{\mu}$$

The control variate estimate of the price of call option is the sample average of i.i.d copies of

$$H = e^{-rT}(S_T - K)^+ - b(e^{-rT}S_T - S_0)$$

Algorithm 11: Control Variate for European Call Option

for $i = 1$ to n **do**
  Generate a sample Z from $N(0,1)$.
  Set $S = S_0 \exp\left\{\left(r - \frac{1}{2}\sigma^2\right) T + \sigma\sqrt{T}Z\right\}$
  Set $X_i = e^{-rT}(S - K)^+$ and $Y_i = e^{-rT}S - S_0$
end for
Set $b = \hat{b}^*$
for $i = 1$ to n **do**
  Set $H_i = X_i - bY_i$
end for
Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} H_i$
Compute the standard error $\sqrt{\frac{1}{n(n-1)}\left(\sum_{k=1}^{n} H_i^2 - n\hat{v}^2\right)}$

Example 2.3.3. [Control Variate for Asian Option]
Estimate price of call price with maturity T and payoff $(\overline{S} - K)^+$, where $\overline{S}$ is arithmetic mean of stock prices

$$\overline{S} = \frac{1}{m}\sum_{i=1}^{m} S_{t_i}$$

where $0 < t_1 < \cdots < t_m = T$ are given dates. Stock price is modelled by geometric Brownian motion.
The control variate estimate for price of call option is sample average of i.i.d copies of

$$H = e^{-rT}(\overline{S} - K)^+ - b(Y - \mathbb{E}[Y])$$

where Y is the control variate. Use average price call option with geometric mean as control variate

$$Y = e^{-rT}(\overline{S}_G - K)^+, \quad \overline{S}_G = \left(\prod_{i=1}^{m} S_{t_i}\right)^{1/m}$$

The evaluate price of the call option is

$$p = \mathbb{E}[Y] = e^{-rT}\left\{\exp\left(\overline{\mu} + \frac{1}{2}\overline{\sigma}^2\right)\Phi(\overline{\sigma} - \theta) - K\Phi(-\theta)\right\}$$
$$\overline{\mu} = \log S_0 + \left(r - \frac{1}{2}\sigma^2\right)\overline{t}, \quad \overline{t} = \frac{1}{m}\sum_{i=1}^{m} t_i$$
$$\overline{\sigma}^2 = \frac{\sigma^2}{m^2}\sum_{i=1}^{m}(2m - 2i + 1)t_i, \quad \theta = \frac{\log K - \overline{\mu}}{\overline{\sigma}}$$

Algorithm 12: Control Variate for Asian Option

for $i = 1$ to n **do**
  for $j = 1$ to m **do**
    Generate a sample Z from $N(0,1)$.
    Set $S_k = S_{k-1} \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(t_k - t_{k-1}) + \sigma\sqrt{t_k - t_{k-1}}Z\right\}$
    Set $X_i = e^{-rT}(S-K)^+$ and $Y_i = e^{-rT}S - S_0$
  end for
  Set $X_i = e^{-rT}(\overline{S}_A - K)^+$ and $Y_i = e^{-rT}(\overline{S}_G - K)^+ - p$
end for
Set $b = \hat{b}^*$
for $i = 1$ to n **do**
  Set $H_i = X_i - bY_i$
end for
Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} H_i$
Compute the standard error $\sqrt{\frac{1}{n(n-1)}\left(\sum_{k=1}^{n} H_i^2 - n\hat{v}^2\right)}$

2.4 Importance Sampling

Remark 2.4.1. [Comparison with Plain Monte Carlo Simulation]
In plain Monte Carlo, generate n i.i.d. samples $X_1, \ldots, X_n$. The sample average, variance, error is then

$$v_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \text{Var}[v_1] = \frac{\sigma_x^2}{n}, \quad \text{Error} = \frac{\sigma_x}{\sqrt{n}}$$

In importance sampling, assuming X has density $f(x)$, and the goal is to estate expected value $\mu = \mathbb{E}[h(X)]$. For an arbitrary density $g(x)$, we have

$$\mu = \int_{\mathbb{R}} h(x)f(x)\ dx = \int_{\mathbb{R}} h(x)\frac{f(x)}{g(x)} \cdot g(x)dx = \mathbb{E}_g\left[h(Y)\frac{f(Y)}{g(Y)}\right]$$

where Y is a random variable with density g. Draw i.i.d samples from density g, then compute the sample average of $h(Y)f(Y)/g(Y)$. The weight $f(Y)/g(Y)$ is the *likelihood ratio*. Note that

$$\begin{aligned}
\hat{\mu}_n &= \sum_{i=1}^{n} h(Y_i)\frac{f(Y_i)}{g(Y_i)} \\
\text{Var}[\hat{\mu}_n] &= \text{Var}\left[\frac{1}{n}\sum_{i=1}^{n} h(Y_i)\frac{f(Y_i)}{g(Y_i)}\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \text{Var}\left[h(Y_i)\frac{f(Y_i)}{g(Y_i)}\right] \\
&= \frac{1}{n}\text{Var}\left[h(Y)\frac{f(Y)}{g(Y)}\right]
\end{aligned}$$

Assume $h \geq 0$, $g(Y) = c \cdot h(Y) \cdot f(Y)$. Then

$$\text{Var}[\hat{\mu}_n] = \frac{1}{n}\text{Var}\left[\frac{1}{c}\right] = 0$$

$$1 = \int_{\mathbb{R}} ch(y)f(y)\ dy \Rightarrow c = \left[\int_{\mathbb{R}} ch(y)f(y)\ dy\right]^{-1} = \frac{1}{\mu}$$

Note that when choosing g, this should be chosen to mimic $h(x)f(x)$,

Definition 2.4.2. [Mode Matching]
Consider estimating expected value $\mu = \mathbb{E}[h(X)]$, where $X \sim N(0,1)$ with density f.
Choose alternative distribution $g \sum N(\theta, 1)$ for some $\theta \in \mathbb{R}$. The likelihood ratio is

$$\frac{f(x)}{g(x)} = \exp\left(-\theta x + \frac{1}{2}\theta^2\right)$$

The mode of g is θ, hence choose

$$\theta = \arg\max_x h(x)f(x)$$

Example 2.4.3. [Mode Matching for Tail Probability]
Estimate tail probability $P(X > a)$, where $X \sim N(0,1)$ and a is large.
[Solution] Note that

$$\begin{aligned}
\mu &= \mathbb{P}\{X > a\} \\
&= \mathbb{E}_f[1_{\{X>a\}}] \\
&= \mathbb{E}_g\left[1_{\{Y>a\}} \cdot \frac{f(Y)}{g(Y)}\right]
\end{aligned}$$

Note that $Y \sim N(\theta, 1)$ and $\theta = \arg\max_X h(X)f(X)$, and

$$h(X)f(X) = 1_{\{X>a\}} \cdot f(x) = \begin{cases} f(X), & X > a \\ 0, & X \leq a \end{cases}$$

Thus $\theta = a$ (for $a > 0$). Hence we generate i.i.d samples from $N(\theta, 1)$, where $\theta = \max(a, 0)$.
Next compute the sample average of

$$1_{\{X>a\}} \cdot \frac{f(X)}{f_\theta(X)} = 1_{\{X>a\}} \cdot \exp\left(-\theta X + \frac{1}{2}\theta^2\right)$$

Example 2.4.4. [Mode Matching for Binary Call Option]
Estimate price of binary call option with maturity T and payoff

$$1_{\{S_T \geq K\}}$$

where under the risk-neutral probability measure, the stock price is a geometric Brownian motion

$$S_t = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right)$$

[Solution] The price of option can be written as

$$\begin{aligned} v &= \mathbb{E}[e^{-rT} 1_{\{S_T \geq K\}}] = \mathbb{E}[e^{-rT} 1_{\{X \geq b\}}] = \mathbb{E}[h(X)] \\ S_T &= S_0 \exp\left(\left(\sigma - \frac{1}{2}\sigma^2\right) T + \sigma\sqrt{T} X\right) \\ S_T \geq K \Leftrightarrow X &\geq \frac{1}{\sigma\sqrt{T}}\left(\log\left(\frac{K}{S_0}\right) - \left(r - \frac{1}{2}\sigma^2\right) T\right) \end{aligned}$$

where

$$X \sim N(0,1), \quad b = \frac{1}{\sigma\sqrt{T}}\left(\log\left(\frac{K}{S_0}\right) - \left(r - \frac{1}{2}\sigma^2\right) T\right)$$

Let f be the probability density of $N(0, 1)$. Then

$$h(x)f(x) = \begin{cases} 0, & \text{if } x < b \\ e^{-rT} f(x), & \text{if } x \geq b \end{cases}$$

whose mode (maxima) is $\theta = \max\{b, 0\}$.
Using $N(\theta, 1)$ as alternative sampling distribution, the importance sampling estimate is the sample average of

$$h(Y) \cdot \frac{f(Y)}{g(Y)} = \begin{cases} 0, & \text{if } Y < b \\ \exp\left(-rT - \theta Y + \frac{1}{2}\theta^2\right), & \text{if } Y \geq b \end{cases}$$

Algorithm 13: Mode Matching for Binary Call Option

Set $\theta = \max\{b, 0\}$
for $i = 1$ to n **do**
  Generate a sample Y from $N(\theta, 1)$.
  Set $H_i = \begin{cases} 0, & \text{if } Y < b \\ \exp\left(-rT - \theta Y + \frac{1}{2}\theta^2\right) & \text{if } Y \geq b \end{cases}$
end for
Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} H_i$
Compute the standard error $\sqrt{\frac{1}{n(n-1)}\left(\sum_{k=1}^{n} H_i^2 - n\hat{v}^2\right)}$

Example 2.4.5. [Mode Matching for European Call Option]
Estimate price of European call option with maturity T and strike price K. The price of call option is

$$v = \mathbb{E}[h(X)]$$

where $X \sim N(0, 1)$, and

$$h(x) = \left(S_0 \exp\left(-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}x\right) - \exp(-rT) \cdot K\right)^+$$

Let f be the probability density of $N(0,1)$ and g be the alternate sampling density $N(\theta,1)$, where θ is the maximiser of $h(x)f(x)$, which solves

$$\frac{\partial}{\partial x}(h(x)f(x)) = S_0 \exp\left(-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}\theta\right)(\sigma\sqrt{T} - \theta) + \exp(-rT)k\theta = 0$$

Algorithm 14: Mode Matching for European Call Option

Solve for θ
for $i = 1$ to n **do**
  Generate a sample Y from $N(\theta, 1)$.
  Set $H_i = h(Y)\exp\left(-\theta Y + \frac{1}{2}\theta^2\right)$
end for
Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} H_i$
Compute the standard error $\sqrt{\frac{1}{n(n-1)}\left(\sum_{k=1}^{n} H_i^2 - n\hat{v}^2\right)}$

Remark 2.4.6. [Extensions to General Distributions - Standard Random Variable]
If $X \sim N(0,1)$ with density f, choose the alternate sampling distribution from family of $N(\theta,1)$, i.e.,

$$g(x) = \exp\left(-\frac{1}{2}\theta^2\right)\exp(\theta x)f(x), \quad f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right)$$
$$g(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(x-\theta)^2\right)$$

Remark 2.4.7. [Extensions to General Distributions - Exponential Tilt Family]
For a general random variable, the alternative sampling distribution is chosen from the exponential tilt family. If X is continuous, assuming f is the density of X, the family consists of probability distributions with density

$$f_\theta(x) = \frac{1}{\mathbb{E}[e^{\theta X}]}e^{\theta x}f(x), \quad \theta \in \mathbb{R}$$

(i.e., if X is normally distributed, the family consists of all normal distributions with the same variance)
If X is discrete, assuming probability mass function is $p(X = x) = p(x)$, the family consists of probability distributions with probability mass function

$$p_\theta(x) = \frac{1}{\mathbb{E}[e^{\theta X}]}e^{\theta x}p(x), \quad \theta \in \mathbb{R}$$

Example 2.4.8. [Importance Sampling on Credit Risk Model]
Consider credit risk model with m independent obligors. Denote p_k by probability that the k-th obligor defaults, and c_k be the loss resulting from the default.
Assuming $c_k = 1$ for every k, and the total loss is

$$L = \sum_{k=1}^{m} c_k X_k = \sum_{k=1}^{m} X_k$$
$$X_k = \begin{cases} 1 & \text{if defaults} \\ 0 & \text{otherwise} \end{cases}, \quad \mathbb{P}(X_k = 1) = p_k, \quad \mathbb{P}(X_k = 0) = 1 - p_k$$

and the X_k's are independent Bernoulli random variables.
Use importance sampling to estimate the tail probability $\mathbb{P}(L > x)$, where x is a large threshold.
[Solution] Choose alternative sampling distribution from the exponential tilt family. Let Y_k's be independent Bernoulli random variables such that

$$\overline{p}_k = \mathbb{P}(Y_k = 1) = \frac{1}{\mathbb{E}[e^{\theta X_k}]}e^{\theta} \cdot \mathbb{P}(X_k = 1) = \frac{p_k e^\theta}{1 + p_k(e^\theta - 1)}$$
$$\mathbb{P}(Y_k = 0) = 1 - \overline{p}_k, \quad k = 1, \ldots, m$$

The likelihood ratio is

$$R(y_1, \ldots, y_m) = \prod_{k=1}^{m} \left(\frac{p_k}{\overline{p}_k}\right)^{y_k} \left(\frac{1-p_k}{1-\overline{p}_k}\right)^{1-y_k}$$

The importance sampling estimate is the sample average of i.i.d. copies of

$$H = 1_{\{\overline{L}>x\}} \cdot R(Y_1, \ldots, Y_m), \quad \text{where } \overline{L} = \sum_{k=1}^{m} Y_k$$

Choose θ to make the variance of estimate as small as possible,

$$\begin{aligned}
\mathrm{Var}[H] &= \mathbb{E}[H^2] - (\mathbb{E}[H])^2 = \mathbb{E}[H^2] - \mu^2 \\
\mathbb{E}[H^2] &= \mathbb{E}\left[1_{\{\overline{L}>x\}} \prod_{k=1}^{m} \left(\frac{p_k}{\overline{p}_k}\right)^{2Y_k} \left(\frac{1-p_k}{1-\overline{p}_k}\right)^{2(1-Y_k)}\right] \\
&= \mathbb{E}[1_{\{\overline{L}>x\}} \exp(-2\theta\overline{L} + 2\phi(\theta))], \quad \Rightarrow \mathbb{E}[H^2] \leq \exp(-2\theta x + 2\phi(\theta))
\end{aligned}$$

where

$$\phi(\theta) = \sum_{k=1}^{m} \log(1 + p_k(e^{\theta} - 1))$$

Minimise the upper bound to determine θ:

$$\min_{\theta} -2\theta x + \phi(\theta) \ \Rightarrow \phi'(\theta) = x$$

Algorithm 15: Importance Sampling on Credit Risk Model

Solve $\phi'(\theta) = x$ for θ
Compute the corresponding $\overline{p}_k$ for $k = 1, \ldots, m$
for $i = 1$ to n **do**
  Generate Y_k from Bernoulli with parameter $\overline{p}_k$ for $k = 1, \ldots, m$
  Set $L = Y_1 + \cdots + Y_m$
  Set $H_i = \sum_{k=1}^{m} \left(\frac{p_k}{\overline{p}_k}\right)^{Y_k} \left(\frac{1-p_k}{1-\overline{p}_k}\right)^{1-Y_k}$ if $L > x$, otherwise $H_i = 0$
end for
Compute the estimate $\hat{v} = \frac{1}{n} \sum_{i=1}^{n} H_i$
Compute the standard error $\sqrt{\frac{1}{n(n-1)} \left(\sum_{k=1}^{n} H_i^2 - n\hat{v}^2\right)}$

2.5 Cross Entropy Method

Remark 2.5.1. [Comparison with Importance Sampling]
In importance sampling, the estimate is based on sampling density

$$g^*(x) = \frac{1}{\mu} h(x) f(x)$$

which has zero variance. However, using this density as the alternative sampling distribution is impractical. The cross-entropy method chooses the density $f_\theta(x)$ that is closest to $g^*(x)$ as the sampling density, i.e.,

$$\min_\theta R(g^* || f_\theta) = \int_{\mathbb{R}} \log \frac{g^*(x)}{f_\theta(x)} \cdot g^*(x) dx$$

where $R(g^*||f_\theta)$ is the *Kullback-Leibler Cross Entropy*, which measures how close the two distributions are. Note that

$$\begin{aligned} R(g^*||f_\theta) &= \int_{\mathbb{R}} g^*(x) \log g^*(x) dx - \int_{\mathbb{R}} g^*(x) \log f_\theta(x) dx \\ &= \int_{\mathbb{R}} g^*(x) \log g^*(x) dx - \frac{1}{\mu} \int_{\mathbb{R}} h(x) f(x) \log f_\theta(x) dx \end{aligned}$$

Therefore

$$\begin{aligned} \min_\theta R(g^*||f_\theta) &\Leftrightarrow \max_\theta \int_{\mathbb{R}} h(x) f(x) \log f_\theta(x) dx \\ &\Leftrightarrow \max_\theta \mathbb{E}[h(X) \log f_\theta(X)] \end{aligned}$$

where the random variable X has density f. The maximiser is the solution to equation

$$0 = \frac{\partial}{\partial\theta} \mathbb{E}[h(X) \log f_\theta(X)] = \mathbb{E}\left[h(X) \frac{\partial}{\partial\theta} \log f_\theta(X)\right]$$

Replace the expected value by sample average

$$0 = \frac{1}{N} \sum_{k=1}^{N} h(X_k) \frac{\partial}{\partial\theta} \log f_\theta(X_k)$$

where X_k's are i.i.d copies of X.
Solving above equation for normal distribution, the solution is then

$$\begin{aligned} X \sim N(0,1), \quad Y \sim N(0,1), \quad f_\theta(x) &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x-\theta)^2\right) \\ 0 &= \frac{1}{N} \sum_{k=1}^{N} h(X_k) \frac{\partial}{\partial\theta} \left[\log \frac{1}{\sqrt{2\pi}} - \frac{1}{2}(X_k - \theta)^2\right] \\ \theta &= \left[\sum_{k=1}^{N} h(X_k) X_k\right] \Bigg/ \left[\sum_{k=1}^{N} h(X_k)\right] \in \mathbb{R}^m \end{aligned}$$

Algorithm 16: Cross Entropy Method

Generate N i.i.d. pilot samples $X_1, \ldots, X_N$ from density $f(x)$
Compute the optimal tilting parameter $\hat{\theta}$
for $i = 1$ to n **do**
    Generate Y_i from the alternative sampling density $f_{\hat{\theta}}(x)$
    Set $H_i = h(Y_i) \cdot f(Y_i)/f_{\hat{\theta}}(Y_i)$
end for
Compute the estimate $\hat{v} = \frac{1}{n} \sum_{i=1}^{n} H_i$
Compute the standard error $\sqrt{\frac{1}{n(n-1)} \left(\sum_{k=1}^{n} H_i^2 - n\hat{v}^2\right)}$

Example 2.5.2. [Cross Entropy Method for European Call Option]
Compute price of European call option with strike price K and maturity T.
Under risk-neutral measure, stock price is a geometric Brownian motion

$$S_t = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right)$$

The price of call option is

$$\begin{aligned} v &= \mathbb{E}[h(X)], \quad X \sim N(0,1) \\ h(x) &= \left(S_0 \exp\left(-\frac{1}{2}\sigma^2 T \sigma \sqrt{T} x\right) - \exp(-rT)K\right)^+ \end{aligned}$$

The alternative sampling density is chosen from family $N(\theta, 1)$. The likelihood ratio is

$$l_\theta(x) = \frac{f(x)}{f_\theta(x)} = \exp\left(-\theta x + \frac{1}{2}\theta^2\right)$$

Algorithm 17: Cross Entropy Method for European Call Option

Generate N i.i.d. pilot samples $X_1, \dots, X_N$ from density $f(x)$
Set $\hat{\theta} = \left[\sum_{k=1}^{N} h(X_k) X_k\right] \Big/ \left[\sum_{k=1}^{N} h(X_k)\right]$
for $i = 1$ to n **do**
  Generate Y_i from the alternative sampling density $N(\hat{\theta}, 1)$
  Set $H_i = h(Y_i) \cdot \exp\left(-\hat{\theta} Y_i + \frac{1}{2}\hat{\theta}^2\right)$
end for
Compute the estimate $\hat{v} = \frac{1}{n} \sum_{i=1}^{n} H_i$
Compute the standard error $\sqrt{\frac{1}{n(n-1)} \left(\sum_{k=1}^{n} H_i^2 - n\hat{v}^2\right)}$

When K is large, this method does not produce meaning result, due to poor estimation of tilting parameter.

Example 2.5.3. [Cross Entropy Method for Asian Call Option]
Compute price of a discretely monitored average call option with payoff $(\overline{S} - K)^+$ and maturity T, where $\overline{S}$ is the arithmetic mean

$$\overline{S} = \frac{1}{m} \sum_{k=1}^{m} S_{t_k}$$

for a given set of dates $0 < t_1 < \cdots < t_m = T$. Assume under risk-neutral probability measure, the underlying asset is a geometric Brownian motion

$$S_t = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right)$$

For $k = 1, \dots, m$,

$$\begin{aligned} S_{t_k} &= S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) t_k + \sigma W_{t_k}\right) \\ &= S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) (t_{k-1} + (t_k - t_{k-1})) + \sigma(W_{t_{k-1}} + \sqrt{t_k - t_{k-1}}) Z_k\right) \\ &= S_{t_{k-1}} \exp\left(\left(r - \frac{1}{2}\sigma^2\right) (t_k - t_{k-1}) + \sigma \sqrt{t_k - t_{k-1}} Z_k\right) \end{aligned}$$

where $Z_1, \dots, Z_m$ are i.i.d standard normal random variables.
The discounted option payoff can be written as a function of the m-dimensional standard normal random vector

$$h(X) = e^{-rT}(\overline{S} - K)^+$$

Denote the joint density function of X by f,

$$f(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^m \exp\left(-\frac{1}{2}\sum_{k=1}^{m} x_i^2\right)$$

Choose the alternative sapling distribution from the family of $N(\theta, I_m)$, where $\theta \in \mathbb{R}^m$, i.e.,

$$f_\theta(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^m \exp\left(-\frac{1}{2}\sum_{k=1}^{m} (x_i - \theta_i)^2\right)$$

The importance sampling estimate is the sampling average of i.i.d copies of

$$h(Y) \cdot \frac{f(Y)}{f_\theta(Y)} = h(Y) \exp\left(-\sum_{k=1}^{m} \theta_i Y_i + \frac{1}{2}\sum_{k=1}^{m} \theta_i^2\right)$$

where Y has distribution $N(\theta, I_m)$.
Let $X_1, \ldots, X_N$ be i.i.d. pilot samples from the original distribution $N(0, I_m)$, then θ is given by

$$\hat{\theta} = \left[\sum_{i=1}^{N} h(X_i) X_i\right] \Big/ \left[\sum_{i=1}^{N} h(X_i)\right]$$

Algorithm 18: Tilting Parameter Computation for Asian Call Option

for $k = 1$ to N **do**
  for $j = 1$ to m **do**
    Generate Z_j from $N(0, 1)$
    Set $S_{t_j} = S_{t_{j-1}} \exp\left(\left(r - \frac{1}{2}\sigma^2\right)(t_j - t_{j-1}) + \sigma\sqrt{t_j - t_{j-1}} Z_j\right)$
  end for
  Set $H_k = \exp(-rT)(\overline{S} - K)^+$
end for
Set $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_m) = \left[\sum_{k=1}^{N} H_k X_k\right] \Big/ \left[\sum_{k=1}^{N} H_k\right]$

Algorithm 19: Cross Entropy Method for Asian Call Option

Compute the tilting parameter using pilot samples
for $i = 1$ to n **do**
  for $j = 1$ to m **do**
    Generate Y_j from $N(\hat{\theta}_j, 1)$
    Set $S_{t_j} = S_{t_{j-1}} \exp\left(\left(r - \frac{1}{2}\sigma^2\right)(t_j - t_{j-1}) + \sigma\sqrt{t_j - t_{j-1}} Y_j\right)$
  end for
  Compute discounted payoff x (likelihood ratio) $H_i = \exp(-rT)(\overline{S} - K)^+ \cdot \exp\left(-\sum_{j=1}^{m} \hat{\theta}_j Y_j + \frac{1}{2}\sum_{j=1}^{m} \hat{\theta}_j^2\right)$
end for
Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} H_i$
Compute the standard error $\sqrt{\frac{1}{n(n-1)}\left(\sum_{k=1}^{n} H_i^2 - n\hat{v}^2\right)}$

2.6 Stratified Sampling

Remark 2.6.1. [Comparison with Antithetic Sampling and Control Variates Sampling]
In antithetic sampling, generate n pairs of i.i.d samples $(X_1, Y_1), \ldots, (X_n, Y_n)$ where X_i and Y_i has the same distributions, and both are negatively correlated ($\beta < 0$). The estimate and variance is then

$$\hat{\mu} = \frac{1}{2n}\sum_{i=1}^{n}(X_i + Y_i), \qquad \mathrm{Var}[\mu^2] = (1+\beta)\cdot\frac{\sigma_x^2}{2n}$$

In control variate, generate n pairs of i.i.d. samples $(X_1, Y_1), \ldots, (X_n, Y_n)$ where X_i and Y_i are correlated, and Y_i has known value $\overline{\mu} = \mathbb{E}[Y_i]$. The estimate and variance is then

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}(X_i - b(Y_i - \overline{\mu})), \qquad \mathrm{Var}[\hat{\mu}] = (1-\beta^2)\cdot\frac{\sigma_x^2}{n}$$

where b is a constant.
In stratified sampling, consider estimating $\mu = \mathbb{E}[X]$, where $X = h(U)$, and U is a random variable, h is a given function. The sample space of U is partitioned into k subsets (stratums). Let Y be the stratification variable which takes values from $\{y_1, \ldots, y_k\}$. Let

$$p_i = P(Y = y_i) = \mathbb{P}(U \in \text{Stratum } i)$$

Assuming U is uniformly distributed on $[0,1]$, and h is an arbitrary function.
Partition $[0,1]$ into k strata of equal length,

$$I_1 = [0, 1/k), \ldots, I_k = [(k-1)/k, 1)$$

Define Y so that $Y = i$ if and only if $U \in I_i$. Then for each i,

$$p_i = \mathbb{P}(Y = i) = \mathbb{P}(U \in I_i) = 1/k$$

Let the conditional expectation of X in each stratum be $\mu_i = \mathbb{E}[X|Y = y_i], i = 1, \ldots, k$.
By tower property of expectations,

$$\mu = \mathbb{E}[X] = \mathbb{E}_Y[\mathbb{E}[X|Y]] = \sum_{i=1}^{k} p_i\mu_i$$

The conditional expectation is then

$$\hat{\mu}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}, \qquad X_{ij} = h(U_{ij})$$

where $U_{ij}, j = 1, \ldots, n_i$ are i.i.d. samples of U from the i-th stratum. The stratified sampling estimate of μ is

$$\hat{\mu} = \sum_{i=1}^{k} p_i\hat{\mu}_i$$

Since $\mathbb{E}[X_{ij}] = \mathbb{E}[X|Y = y_i] = \mu_i$, we have

$$\mathbb{E}[\hat{\mu}] = \sum_{i=1}^{k} p_i\mathbb{E}[\hat{\mu}_i] = \sum_{i=1}^{k} p_i\cdot\frac{1}{n_i}\sum_{j=1}^{n_i}\mathbb{E}[X_{ij}] = \sum_{i=1}^{k} p_i\mu_i = \mu_i$$

hence the stratified sampling estimate is unbiased. The variance of $\hat{\mu}$ is

$$\mathrm{Var}[\hat{\mu}] = \sum_{i=1}^{k} p_i^2\cdot\frac{1}{n_i^2}\sum_{j=1}^{n_i}\mathrm{Var}[X_{ij}] = \sum_{i=1}^{k} p_i^2\cdot\frac{1}{n_i}\sigma_i^2$$

where $\sigma_i^2 = \text{Var}[X|Y = y_i]$. Replace the variance σ_i^2 by sample variance, then we have

$$s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2, \quad \overline{X}_i = \frac{1}{n_i} = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$$

The standard error of $\hat{\mu}$ is then

$$\text{S.E.} = \sqrt{\sum_{i=1}^{k} p_i^2 \cdot \frac{1}{n_i} s_i^2}$$

Remark 2.6.2. [Allocation of Samples]
Let n be total number of samples. The number of samples in each stratum is a fraction of n: $n_i = nq_i$, where

$$q_i > 0, \quad q_1 + \cdots + q_k = 1$$

The choice of q_i depends on proportional allocation ($q_i = p_i$) and minimising the variance of $\hat{\mu}$:

$$\min_{\{q_1,\ldots,q_k\}} \text{Var}[\hat{\mu}] = \frac{1}{n}\sum_{i=1}^{k} p_i^2 \cdot \frac{1}{q_i}\sigma_i^2, \qquad \text{s.t.} \sum_{i=1}^{k} q_i = 1$$

$$q_i^* = (p_i\sigma_i) / \left(\sum_{j=1}^{k} p_j\sigma_j\right)$$

Remark 2.6.3. [Variance Decomposition]
Let μ_i and σ_i^2 be mean and variance of X in stratum i:

$$\mu_i = \mathbb{E}[X|Y = y_i], \qquad \sigma_i^2 = \text{Var}[X|Y = y_i]$$

Then

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[\mathbb{E}[X|Y]] = \sum_{i=1}^{k} p_i\mu_i \\
\mathbb{E}[X^2] &= \mathbb{E}_Y[\mathbb{E}[X^2|Y]] = \sum_{i=1}^{k} p_i\mathbb{E}[X^2|Y = y_i] = \sum_{i=1}^{k} p_i(\sigma_i^2 + \mu_i^2) \\
\text{Var}[X] &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\
&= \left[\sum_{i=1}^{k} p_i\sigma_i^2 + \sum_{i=1}^{k} p_i\mu_i^2\right] - \left(\sum_{i=1}^{k} p_i\mu_i\right)^2 \\
&= \mathbb{E}[\text{Var}[X|Y]] + \text{Var}[\mathbb{E}[X|Y]]
\end{aligned}$$

which is the variance within strata plus variance between strata.
Note that the variance of plan MC estimate is

$$\text{Var}[\hat{\mu}] = \frac{1}{n}\text{Var}[X] = \frac{1}{n}(\mathbb{E}[\text{Var}[X|Y]] + \text{Var}[\mathbb{E}[X|Y]])$$

The variance of stratified sampling with proportional allocation is

$$\text{Var}[\hat{\mu}] = \sum_{i=1}^{k} p_i^2 \cdot \frac{1}{n_i}\sigma_i^2 = \frac{1}{n}\sum_{i=1}^{k} p_i\sigma_i^2 = \frac{1}{n}\mathbb{E}[\text{Var}[X|Y]]$$

Note that stratified sampling eliminates the variance between strata.

Example 2.6.4. [Stratified Sampling for European Call]
Compute price of call option with maturity T and strike price K. Under risk-neutral measure, stock price is modelled by geometric Brownian motion

$$S_t = S_0 \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}$$

Note that

$$S_T = S_0 \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z\right\}, \quad Z \sim N(0,1)$$
$$\mu = \mathbb{E}[X], \quad X = \exp\{-rT\}(S_T - K)^+ = h(U), \quad U \sim \text{Uniform}[0,1]$$

Let Φ be c.d.f. of Uniform[0, 1]. Then

$$\mathbb{P}\{\Phi^{-1}(U) \le z\} = \mathbb{P}\{U \le \Phi(z)\} = \Phi(z)$$

Hence $\Phi^{-1}(U) \sim N(0,1)$, and $W_T = \sqrt{T}\Phi^{-1}(U)$. The price of call option is

$$v = \mathbb{E}[h(U)]$$

where

$$h(U) = \left(S_0 \exp\left(-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}\Phi^{-1}(u)\right) - e^{-rT}K\right)^+$$

Partition $[0,1]$ into k strata of equal length:

$$I_1 = [0, 1/k), \ldots, I_k = [(k-1)/k, k]$$

Each stratum has probability $p_i = 1/k, i = 1, \ldots, k$.
Sample in each stratum. Conditional on $U \in I_i$, U is uniformly distributed on I_i. Let V_{ij} be i.i.d. samples from the uniform distribution on $[0,1]$, then

$$U_{ij} = (i-1)/k + V_{ij}/k, \quad j = 1, \ldots, n_i$$

are i.i.d. samples from uniform distribution on I_i. The MC estimate is given by

$$\hat{\mu} = \sum_{i=1}^{k} \frac{1}{k} \cdot \frac{1}{n_i} \sum_{j=1}^{n_i} h(U_{ij})$$

Algorithm 20: Stratified Sampling for European Call Option

Specify sample size $n_i = \frac{1}{k} \cdot n$ for each stratum
for $i = 1$ to k **do**
  Generate i.i.d. samples $\{V_{ij} : j = 1, \ldots, n_i\}$ uniformly from $[0,1]$.
  Set $U_{ij} = \frac{i-1}{k} + \frac{V_{ij}}{k}$
  Compute sample mean and standard deviation in stratum i

$$\hat{\mu}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} h(U_{ij}), \quad s_i = \sqrt{\frac{1}{n_i - 1} \sum_{j=1}^{n_i} (h(U_{ij}) - \hat{\mu}_i)^2}$$

end for
Compute the estimate $\hat{\mu} = \frac{1}{k} \sum\limits_{i=1}^{k} \hat{\mu}_i$
Compute the standard error S.E.$= \frac{1}{k} \sqrt{\sum\limits_{k=1}^{n} \frac{1}{n_i} s_i^2}$

Example 2.6.5. [Stratified Sampling for Spread Call Option]
Compute price of spread call option with maturity T and payoff $(X_T - Y_T - K)^+$, where X and Y are the prices of two underlying assets. Assume under risk-neutral measure,

$$X_t = X_0 \exp\left(\left(r - \frac{1}{2}\sigma_1^2\right)t + \sigma_1 W_t\right)$$
$$Y_t = Y_0 \exp\left(\left(r - \frac{1}{2}\sigma_2^2\right)t + \sigma_2 B_t\right)$$

where (W, B) is a two dimensional Brownian motion with covariance matrix

$$\begin{pmatrix} 1 & p \\ p & 1 \end{pmatrix}$$

Note that

$$\begin{pmatrix} W_T \\ B_T \end{pmatrix} \sim N(0, T\Sigma), \quad \Sigma = cc^T, \quad c = \begin{pmatrix} 1 & 0 \\ e & \sqrt{1-e^2} \end{pmatrix}$$

Then

$$\begin{pmatrix} W_T \\ B_T \end{pmatrix} = \sqrt{T}c \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \quad \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim N(0, I)$$

Let

$$\begin{pmatrix} \theta \\ \eta \end{pmatrix} := \frac{1}{\sqrt{T}} \begin{pmatrix} W_T \\ B_T \end{pmatrix} = c \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

then (θ, η) is a jointly normal random vector with mean 0 and covariance Σ. The payoff is function of (θ, η). Consider two stratification strategies which use proportional allocation:

1. Stratification of θ alone
   Since θ is a standard normal r.v., it can be written as $\Phi^{-1}(U)$, then stratified by partitioning $[0, 1]$
   Given $\theta = x$, η is normally distributed with mean ρx and variance $1 - \rho^2$.

2. Stratification of random vector (θ, η)
   Let U, V be two independent, uniformly distributed r.v. on $[0, 1]$.
   The vector

   $$Z = \begin{pmatrix} \Phi^{-1}(U) \\ \Phi^{-1}(V) \end{pmatrix}$$

   is a standard random r.v.
   Let CC' be the Cholesky factorization of Σ, then $C\Sigma$ is jointly normal with mean 0 and covariance Σ.

   $$C = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}$$

   The payoff is a function of U and V, where $X = h(U, V)$.
   Partition the unit square $[0, 1] \times [0, 1]$ into k^2 strata:

   $$I_{ij} = \left[\frac{i-1}{k}, \frac{i}{k}\right) \times \left[\frac{j-1}{k}, \frac{j}{k}\right), \quad i, j = 1, \ldots, k$$

   Sample from stratum I_{ij}. Generate u and v uniformly from $[0, 1]$ and let

   $$U = \frac{i-1}{k} + \frac{u}{k}, \quad V = \frac{j-1}{k} + \frac{v}{k}$$

   Set

   $$\begin{pmatrix} \theta \\ \eta \end{pmatrix} = C \begin{pmatrix} \Phi^{-1}(U) \\ \Phi^{-1}(V) \end{pmatrix}$$

Algorithm 21: Stratified Sampling for Spread Call Option

Specify sample size n_i for each stratum I_i
for $i = 1$ to k **do**
  for $j = 1$ to n_i **do**
    Generate sample V_{ij} from uniform distribution on $[0, 1]$
    Set $U_{ij} = \frac{i-1}{k} + \frac{V_{ij}}{k}$
    Set $\theta_{ij} = \Phi^{-1}(U_{ij}) = z_1$
    Generate sample η_{ij} from $N(\rho\theta_{ij}, 1 - \rho^2)$
    Compute discounted payoff H_{ij} given (θ_{ij}, η_{ij})
  end for
  Compute the sample mean and standard deviation in stratum I_i

$$\hat{\mu}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} H_{ij}, \quad s_i = \sqrt{\frac{1}{n_i - 1}\sum_{j=1}^{n_i}(H_{ij} - \hat{\mu}_i)^2}$$

end for
Compute the estimate $\hat{\mu} = \frac{1}{k}\sum\limits_{i=1}^{k}\hat{\mu}_i$
Compute the standard error S.E.$= \frac{1}{k}\sqrt{\sum\limits_{k=1}^{n}\frac{1}{n_i}s_i^2}$

2.7 Simulation of SDEs

Method 2.7.1. *[Euler-Maruyama Method]*
Consider solving the SDE

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad 0 < t \le T$$

where $X_0 = x$. Discretise time interval into m sub-intervals:

$$0 = t_0 < t_1 < \cdots < t_m = T$$

and denote time step by

$$\Delta t_i = t_{i+1} - t_i, \quad i = 0, 1, \ldots, m$$

Integrate the SDE from t_i to t_{i+1}, then we have

$$X_{t_{i+1}} = X_{t_i} + \int_{t_i}^{t_{i+1}} b(t, X_t)dt + \int_{t_i}^{t_{i+1}} \sigma(t, X_t)dW_t$$

Approximate the integrands using their values at the left end point of the interval,

$$\begin{aligned} \hat{X}_{t_{i+1}} &= \hat{X}_{t_i} + b(t_i, \hat{X}_{t_i})\Delta t_i + \sigma(t_i, \hat{X}_{t_i})(W_{t_{i+1}} - W_{t_i}) \\ &= \hat{X}_{t_i} + b(t_i, \hat{X}_{t_i})\Delta t_i + \sigma(t_i, \hat{X}_{t_i})\sqrt{\Delta t_i} Z_{i+1} \end{aligned}$$

for $i = 0, 1, \ldots, m-1$ where $\{Z_1, \ldots, Z_m\}$ are i.i.d. standard normal random variables.
$\{\hat{X}_{t_1}, \hat{X}_{t_2}, \ldots, \hat{X}_{t_m}\}$ are approximate solutions to the SDE at the discrete times $\{t_1, t_2, \ldots, t_m\}$.
The numerical scheme also applies to higher dimensions.

Example 2.7.2. [Geometric Brownian Motion]
Suppose X_t is a geometric Brownian motion with drift r and volatility σ. It satisfies the SDE

$$dX_t = rX_t dt + \sigma X_t dW_t, \quad 0 < t \le T$$

Use Euler-Maruyama scheme to approximate X_T.
Use uniform time step $\Delta t = T/m$, then the Euler-Maruyama scheme is given by

$$\hat{X}_0 = X_0, \ \hat{X}_{ti+1} = \hat{X}_{t_i} + r\hat{X}_{t_i}\Delta t_i + \sigma \hat{X}_{t_i}\sqrt{\Delta t_i} Z_{i+1}$$

where $\{Z_1, \ldots, Z_m\}$ are i.i.d. standard normal random variables.
Compute the error $\mathbb{E}[\hat{X}_T - X_T]$ where X_T is the exact solution computed using

$$X_{t_{i+1}} = X_{t_i} \exp\left(\left(r - \frac{1}{2}\sigma^2\right)\Delta t_i + \sigma\sqrt{\Delta t_i} Z_{i+1}\right)$$

Remark 2.7.3. *[Bias of Euler-Maruyama Scheme]*
Consider problem of estimating $v = \mathbb{E}[h(X_T)]$ for some function h.
Let $\hat{X}_{T,1}, \ldots, \hat{X}_{T,n}$ be terminal values of n i.i.d. sample paths computed from the Euler scheme. Let $\hat{v}$ be the sample average,

$$\hat{v} = \frac{1}{n}\sum_{i=1}^{n} h(\hat{X}_{T,i})$$

In general, since distribution of $\hat{X}_T$ is different from that of X_T due to discretisation error,

$$\mathbb{E}[\hat{v}] = \mathbb{E}[h(\hat{X}_T)] \neq \mathbb{E}[h(X_T)]$$

i.e., the estimate $\hat{v}$ is biased.
Error of estimate consists of error due to time discretisation (bias) and error due to finite sample size (variance).

Example 2.7.4. [No Discretisation Error]
Consider discretely monitored lookback call option with maturity T and floating strike price. Option payoff is

$$X = S_T - \min_{i=1,\ldots,m} S_{t_i}$$

where $0 = t_0 < t_1 < \ldots < t_m = T$ are the given dates.

Assume risk-free interest rate r is a constant and the stock price S satisfies the SDE

$$dS_t = rS_t dt + \theta(t)S_t dW_t$$

under the risk-neutral probability measure. Estimate the option price.
The process of stock price can be simulated exactly without discretisation error.
Define $Y_t = \log S_t$. Then it follows from Ito formula that

$$dY_t = \frac{1}{S_t}dS_t - \frac{1}{2S_t^2}(dS_t)^2 = \left(r - \frac{1}{2}\theta^2(t)\right)dt + \theta(t)dW_t$$

Therefore

$$\begin{aligned} Y_{t_{i+1}} &= Y_{t_i} + \int_{t_i}^{t_{i+1}} \left(r - \frac{1}{2}\theta^2(t)\right)dt + \int_{t_i}^{t_{i+1}} \theta(t)dW_t \\ &= Y_{t_i} + r(t_{i+1} - t_i) - \frac{1}{2}\sigma_i^2 + \sigma_i Z_{i+1} \end{aligned}$$

where

$$\sigma_i^2 = \int_{t_i}^{t_{i+1}} \theta^2(t)dt$$

and $\{Z_1, \ldots, Z_m\}$ are i.i.d. standard normal random variables.

Algorithm 22: SDE Estimation for Discretely Monitored Lookback Call Option

---

for $k = 1$ to n **do**
  Set $Y_0 = \log S_0$
  for $i = 0$ to $m - 1$ **do**
    Generate Z_{i+1} from $N(0, 1)$
    Set $Y_{i+1} = Y_i + r(t_{i+1} - t_i) - \sigma_i^2/2 + \sigma_i Z_{i+1}$
    Set $S_{i+1} = \exp\{Y_{i+1}\}$
  end for
  Set $H_k = e^{-rT}(S_m - \min(S_1, \ldots, S_m))$
end for
Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} H_i$
Compute the standard error S.E.= $\sqrt{\frac{1}{n(n-1)}\left(\sum_{i=1}^{n} H_i^2 - n\hat{v}^2\right)}$

Example 2.7.5. *[Method of Conditioning]*
Suppose that under risk-neutral probability measure, the stock price satisfies the SDE

$$dS_t = r_t S_t dt + \sigma S_t dW_t$$

The short rate r_t satisfies the SDE

$$dr_t = a(b - r_t)dt + \theta r_t dB_t$$

for some constants a, b and θ. Assume (W_t, B_t) is a 2-dimensional Brownian motion with covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

Estimate the price of a call option with maturity T and strike price K.
Let $Y_t = \log S_t$, then it follows from Ito formula that

$$dY_t = \frac{1}{S_t}dS_t - \frac{1}{2S_t^2}(dS_t)^2 = \left(r_t - \frac{1}{2}\sigma^2\right)dt + \sigma dW_t$$
$$\Rightarrow S_T = S_0 \exp\left\{\left(\overline{r} - \frac{1}{2}\sigma^2\right)T + \sigma W_T\right\}$$

where $\overline{r} = \frac{1}{T}\int_0^T r_t dt$ is the average interest rate.
The price of the call option is $v = \mathbb{E}[e^{-\overline{r}T}(S_T - K)^+]$.
To compute the price, need to generate $\overline{r}$ and W_T, which are correlated random numbers.

Conditioning, we have

$$v = \mathbb{E}_{B_t}[\mathbb{E}[e^{-\bar{r}T}(S_T - K)^+ \mid B_t]] = \mathbb{E}_{B_t}[h(B_t)]$$

Let $h(B_t)$ be the option price conditioning on B_t:

$$h(B_t) = \mathbb{E}[e^{-\bar{r}T}(S_T - K)^+ \mid B_t]$$

Note that $h(B_t)$ can be computed using Black-Scholes formula.
By tower property, the option price is $v = \mathbb{E}[h(B_t)]$.
The stock price S does not need to he simulated. Only the process r_t needs to be approximated (i.e., with Euler schemes). Let (Z_1, Z_2) be 2-dimensional standard normal random vector, then

$$\begin{pmatrix} B_T \\ W_T \end{pmatrix} = \sqrt{T}\begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \Rightarrow \begin{array}{l} B_T = \sqrt{T}Z_1 \\ W_T = \sqrt{T}(\rho Z_1 + \sqrt{1-\rho^2}Z_2) = \rho B_T + \sqrt{1-\rho^2}\sqrt{T}Z_2 \end{array}$$

$$\begin{aligned} S_T &= S_0 \exp\left\{\left(\bar{r} - \frac{1}{2}\sigma^2\right)T + \sigma W_T\right\} \\ &= S_0 \exp\left(\rho\sigma B_T - \frac{1}{2}\sigma^2\rho^2 T\right)\exp\left\{\left(\bar{r} - \frac{1}{2}\sigma^2(1-\rho^2)\right)T + \sigma\sqrt{1-\rho^2}\sqrt{T}Z_2\right\} \\ &= X_0 \exp\left\{\left(\bar{r} - \frac{1}{2}\sigma^2(1-\rho^2)\right)T + \sigma\sqrt{1-\rho^2}\sqrt{T}Z_2\right\} \end{aligned}$$

Conditional on B_t, S_t is a geometric Brownian motion with initial value X_0, drift $\bar{r}$ and volatility $\sigma\sqrt{1-\rho^2}$. Then $h(B_t)$ can be computed using the Black-Scholes formula.

Algorithm 23: SDE Estimation with Direct Simulation

---

for $k = 1$ to n **do**
  for $i = 0$ to $m-1$ **do**
    Generate independent samples Z_1, Z_2 from $N(0,1)$
    Set $r_{i+1} = r_i + a(b - r_i)\Delta t_i + \theta r_i\sqrt{\Delta t_i}Z_i$
    Set $W_{i+1} = W_i + \sqrt{\Delta t_i}(\rho Z_1 + \sqrt{1-\rho^2}Z_2)$
  end for
  Set $R = \sum_{i=0}^{m-1} \Delta t_i r_i / T$
  Set $S = S_0 \exp\left(\left(R - \frac{1}{2}\sigma^2\right)T + \sigma W_m\right)$
  Set $H_k = \exp(-RT)(S - K)^+$
end for
Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} H_i$
Compute the standard error S.E.= $\sqrt{\frac{1}{n(n-1)}\left(\sum_{i=1}^{n} H_i^2 - n\hat{v}^2\right)}$

Algorithm 24: SDE Estimation with Method of Conditioning

---

for $k = 1$ to n **do**
  Set $B_T = 0$
  for $i = 1$ to $m-1$ **do**
    Generate sample Y_{i+1} from $N(0,1)$
    Set $r_{i+1} = r_i + a(b - r_i)\Delta t_i + \theta r_i\sqrt{\Delta t_i}Y_{i+1}$
    Set $B_T = B_T + \sqrt{\Delta t_i}Y_{i+1}$
  end for
  Set $R = \sum_{i=0}^{m-1} \Delta t_i r_i / T$
  Set $X_0 = S_0 \exp\left(\rho\sigma B_T - \rho^2\sigma^2 T/2\right)$
  Set $H_k = \mathrm{BLS}_{\mathrm{Call}}(X_0, K, T, R, \sigma\sqrt{1-\rho^2})$
end for
Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} H_i$
Compute the standard error S.E.= $\sqrt{\frac{1}{n(n-1)}\left(\sum_{i=1}^{n} H_i^2 - n\hat{v}^2\right)}$

Method 2.7.6. *[Milstein Scheme]*
Consider the SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad 0 < t \leq T$$

where $b(x)$ and $\sigma(x)$ are twice continuously differentiable functions.
Let $0 = t_0 < t_1 < \ldots < t_m = T$ be a time discretisation. Integrate the SDE from t_i to t_{i+1},

$$X_{t_{i+1}} = X_{t_i} + \int_{t_i}^{t_{i+1}} b(X_t)dt + \int_{t_i}^{t_{i+1}} \sigma(X_t)dW_t$$

To get a discretisation scheme more accurate than Euler scheme, apply Ito formula to $b(X_t)$ and $\sigma(X_t)$. For any twice continuously differentiable function f, it follows from Ito formula that

$$\begin{aligned} df(X_t) &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2 \\ &= \left(f'(X_t)b(X_t) + \frac{1}{2}f''(X_t)\sigma^2(X_t)\right)dt + f'(X_t)\sigma(X_t)dW_t \\ &= L^0 f(X_t)dt + L^1 f(X_t)dW_t \end{aligned}$$

Then

$$f(X_t) = f(X_{t_i}) + \int_{t_i}^{t} L^0 f(X_s)ds + \int_{t_i}^{t} L^1 f(X_s)dW_s$$

Apply Ito formula to $b(X_t)$ and $\sigma(X_t)$ yields

$$\begin{aligned} X_{t_{i+1}} = {} & X_{t_i} + b(X_{t_i})\Delta t_i + \sigma(X_{t_i})(W_{t_{i+1}} - W_{t_i}) \\ & + \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^{t} L^0 b(X_s)ds + \int_{t_i}^{t} L^1 b(X_s)dW_s\right) ds \\ & + \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^{t} L^0 \sigma(X_s)ds + \int_{t_i}^{t} L^1 \sigma(X_s)dW_s\right) dW_t \end{aligned}$$

Neglecting all the double integrals gives the Euler-Maruyama scheme.
Neglecting the double integrals except the last one (leading order term) gives the Milstein scheme.
Approximating the double integral:

$$\begin{aligned} \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} L^1\sigma(X_s)dW_s dW_t &\approx L^1\sigma(X_{t_i}) \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} dW_s dW_t \\ &= L^1\sigma(X_{t_i}) \int_{t_i}^{t_{i+1}} (W_t - W_{t_i})dW_t \\ &= \frac{1}{2}L^1\sigma(X_{t_i})((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i)) \end{aligned}$$

Adding the above term to the Euler scheme, we obtain the Milstein scheme:

$$\begin{aligned} \hat{X}_{t_{i+1}} = {} & \hat{X}_{t_i} + b(\hat{X}_{t_i})\Delta t_i + \sigma(\hat{X}_{t_i})\sqrt{\Delta t_i} Z_{i+1} \\ & + \frac{1}{2}\sigma'(\hat{X}_{t_i})\sigma(\hat{X}_{t_i})\Delta t_i(Z_{i+1}^2 - 1) \end{aligned}$$

where $\{Z_1, \ldots, Z_m\}$ are i.i.d. standard normal random variables.

Example 2.7.7. [SDE for Control Variate]
Consider stochastic volatility model, where under the risk-neutral probability measure the stock price and volatility are modelled by

$$\begin{aligned} dS_t &= rS_t dt + \theta_t S_t dW_t \\ d\theta_t &= a(\Theta - \theta - t)dt + \beta dB_t \end{aligned}$$

for some positive constants r, a, Θ, β. (W_t, B_t) is a two-dimensional Brownian motion with covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

Estimate the price of call option with strike price K and maturity T.

[Plain Monte Carlo] method:

i. Let $Y_t = \log S_t$, then by Ito formula,

$$dY_t = \left(r - \frac{1}{2}\theta_t^2\right) dt + \theta_t dW_t$$

ii. The Euler-Maruyama scheme for (Y_t, θ_t) is then

$$\hat{Y}_{i+1} = \hat{Y}_i + \left(r - \frac{1}{2}\hat{\theta}_i^2\right) \Delta t_i + \hat{\theta}_i \sqrt{\Delta t_i} Z_{i+1}$$
$$\hat{\theta}_{i+1} = \hat{\theta}_i + a(\Theta - \hat{\theta}_i)\Delta t_i + \beta \sqrt{\Delta t_i} R_{i+1}$$

where $(\hat{Y}_i, \hat{\theta}_i)$ are approximations to (Y_{t_i}, θ_{t_i}), $\Delta t_i = t_{i+1} - t_i$, (Z_{i+1}, R_{i+1}) are i.i.d. jointly normal random vectors with distribution $N(0, \Sigma)$.

iii. The MC estimate for the option price is the sample average of

$$X = e^{-rT}(e^{\hat{Y}_T} - K)^+$$

[Monte Carlo with Control Variate] method:

i. Introduce a virtual stochastic process $\overline{Y}_t$:

$$d\overline{Y} = \left(r - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t$$

with initial condition $\overline{Y}_0 = Y_0 = \log S_0$.

ii. The discounted payoff of the call option for this virtual stock is

$$V = e^{-rT}(e^{\overline{Y}_T} - K)^+$$

whose expected value is given by the Black-Scholes formula. Hence V cam be used as control variate.

iii. The control variate estimate for the option price is the sample average of i.i.d. copies of

$$X - b(V - \mathbb{E}[V])$$

iv. The same sequence of $\{Z_i\}$ is used in solving for Y_t and $\overline{Y}_t$.

v. The sample estimate of the optimal coefficient is used for b:

$$\hat{b}^* = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(V_i - \overline{V})}{\sum_{i=1}^{n} (V_i - \overline{V})^2}, \quad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \overline{V} = \frac{1}{n}\sum_{i=1}^{n} V_i$$

vi. Choice of σ in the virtual dynamics: note that the volatility θ_t approaches Θ in the long run. Therefore, choose $\sigma = \Theta$.

Algorithm 25: Monte Carlo with Control Variate

for $k = 1$ to n **do**
  Set $\hat{Y}_0 = \log S_0$, $\overline{Y}_0 = \log S_0$, $\hat{\theta}_0 = \theta_0$
  for $i = 0$ to $m - 1$ **do**
    Generate i.i.d. samples Z_{i+1}, U_{i+1} from $N(0, 1)$
    Set $R_{i+1} = \rho Z_{i+1} + \sqrt{1 - \rho^2} U_{i+1}$
    Set $\hat{Y}_{i+1} = \hat{Y}_i + (r - \hat{\theta}_i^2/2)\Delta t_i + \hat{\theta}_i \sqrt{\Delta t_i} Z_{i+1}$
    Set $\overline{Y}_{i+1} = \overline{Y}_i + (r - \sigma^2/2)\Delta t_i + \sigma\sqrt{\Delta t_i} Z_{i+1}$
    Set $\hat{\theta}_{i+1} = \hat{\theta}_i + a(\Theta - \hat{\theta}_i)\Delta t_i + \beta\sqrt{\Delta t_i} R_{i+1}$
  end for
  Set $X_k = e^{-rT}(\exp(\hat{Y}_m) - K)^+$
  Set $V_k = e^{-rT}(\exp(\overline{Y}_m) - K)^+$

end for
Compute $\hat{b}^*$ and

$$H_k = X_k - \hat{b}^*(V_k - \mathrm{BLS}_{\mathrm{Call}}(S_0, K, T, r, \sigma)), \quad k = 1, \dots, n$$

Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} H_i$

Compute the standard error S.E.= $\sqrt{\frac{1}{n(n-1)}\left(\sum_{i=1}^{n} H_i^2 - n\hat{v}^2\right)}$

Example 2.7.8. *[SDE for Cross-Entropy Method]*
Assume under risk-neutral probability measure, the stock price is modelled by constant elasticity of variance (CEV) process

$$dS_t = rS_t dt + \sigma S_t^{\gamma} dW_t$$

where $0.5 \le \gamma < 1$. Estimate price of call option with maturity T and strike price K.
Let $X_t = e^{-rt}S_t$. Then $X_0 = S_0$, and by Ito's formula,

$$\begin{aligned} dX_t &= d(e^{-rt})S_t + e^{-rt}dS_t \\ &= -re^{-rt}S_t dt + e^{-rt}(rS_t dt + \sigma S_t^{\gamma} dW_t) \\ \Rightarrow dX_t &= \sigma e^{-r(1-\gamma)t} X_t^{\gamma} dW_t \end{aligned}$$

The price of the call option is

$$v = \mathbb{E}[e^{-rT}(S_T - K)^+] = \mathbb{E}[(X_T - e^{-rT}K)^+]$$

The Euler-Maruyama scheme for X_t is

$$\hat{X}_{i+1} = \hat{X}_i + \sigma e^{-r(1-\gamma)t_i} \hat{X}_i^{\gamma} \sqrt{\Delta t_i} Z_{i+1}, \quad i = 0, 1, \dots, m-1$$

where $\hat{X}_i$ is approximation to X_{t_i}, $\Delta t_i = t_{i+1} - t_i$, and $\{Z_1, \dots, Z_m\}$ are i.i.d. standard normal random variables.
The scheme is modified so that $\hat{X}$ remains non-negative:

$$\hat{X}_{i+1} = \max\{0, \hat{X}_i + \sigma e^{-r(1-\gamma)t_i} \hat{X}_i^{\gamma} \sqrt{\Delta t_i} Z_{i+1}\}$$

The plain Monte Carlo estimate is the sample average of i.i.d. copies of

$$h(Y) = (\hat{X}_m - e^{-rT}K)^+$$

where $Y = (Z_1, \dots, Z_m)$ is a jointly normal random vector with distribution $N(0, I_m)$.
The alternative distribution is $N(\theta, I_m)$ where $\theta = (\theta_1, \dots, \theta_m)$ is determined by

$$\theta = \frac{\sum_{k=1}^{N} h(Y_k)Y_k}{\sum_{k=1}^{N} h(Y_k)}$$

where Y_k's are i.i.d. jointly normal random vectors with distribution $N(0, I_m)$.

Algorithm 26: SDE for Cross-Entropy Method

Generate i.i.d. pilot samples $Y_1, \dots, Y_N$ from $N(0, I_m)$.
Compute the tilting parameter θ
for $k = 1$ to n **do**
  for $i = 0$ to $m-1$ **do**
    Generate Z_{i+1} from $N(\theta_{i+1}, 1)$
    Set $\hat{X}_{i+1} = \max\{0, \hat{X}_i + \sigma e^{-r(1-\gamma)t_i} \hat{X}_i^{\gamma} \sqrt{\Delta t_i} Z_{i+1}\}$
    Set $\hat{Y}_{i+1} = \hat{Y}_i + (r - \hat{\theta}_i^2/2)\Delta t_i + \hat{\theta}_i \sqrt{\Delta t_i} Z_{i+1}$
  end for

Compute the discounted payoff × the likelihood ratio:

$$H_k = (\hat{X}_m - e^{-rT}K)^+ \cdot \exp\left(-\sum_{i=1}^{m} \theta_i Z_i + \frac{1}{2}\sum_{i=1}^{m} \theta_i^2\right)$$

end for

Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} H_i$

Compute the standard error S.E.$= \sqrt{\frac{1}{n(n-1)}\left(\sum_{i=1}^{n} H_i^2 - n\hat{v}^2\right)}$

2.8 Sensitivity Analysis

Definition 2.8.1. *[The Greeks]*
Consider option on stock with maturity T. Let $V(S, r, \sigma, t)$ be the value of the option at time $t \in [0, T]$. S is the stock price at time t, r is the risk-free interest rate, σ is the volatility. Then

i. Delta: $\Delta = \frac{\partial V}{\partial S}$

ii. Gamma: $\Gamma = \frac{\partial^2 V}{\partial S^2}$

iii. Rho: $\rho = \frac{\partial V}{\partial r}$

iv. Theta: $\Theta = \frac{\partial V}{\partial t}$

v. Vega: $\nu = \frac{\partial V}{\partial \sigma}$

Example 2.8.2. [Delta of Option]
Suppose stock price is a geometric Brownian motion,

$$S_t = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right)$$

The price of a call option with maturity T and strike price K is given by Black-Scholes formula

$$V = S_0 \Phi(\alpha) - K e^{-rT} \Phi(\alpha - \sigma\sqrt{T})$$

where

$$\alpha = \frac{1}{\sigma\sqrt{T}} \log \frac{S_0}{K} + \left(\frac{\sigma}{2} + \frac{r}{\sigma}\right)\sqrt{T}$$

Then the delta of the option is $\Delta = \frac{\partial V}{\partial S_0} = \Phi(\alpha)$.

Method 2.8.3. *[Methods of Finite Difference]*
Approximate the derivative using centred finite difference:

$$V'(\theta) \approx \frac{V(\theta + h) - V(\theta - h)}{2h}$$

where h is small. The finite difference estimate is biased:

$$\text{Bias} = \frac{V(\theta + h) - V(\theta - h)}{2h} - V'(\theta) \approx \frac{1}{3!} V'''(\theta) h^2$$

Example 2.8.4. [Delta of Call Option]
Estimate delta at $t = 0$ of call option with maturity T and strike price K, assuming underlying stock is a geometric Brownian motion with drift r and volatility σ under risk-neutral probability measure.
The finite difference approximation of delta is

$$\Delta = V'(S_0) \approx \frac{V(S_0 + h) - V(S_0 - h)}{2h}$$

where $V(S_0 \pm h)$ are estimated by sample average of i.i.d. copies of

$$f^{\pm}(Z) = e^{-rT}\left\{(S_0 \pm h) \exp\left[\left(r - \frac{1}{2}\sigma^2\right) T + \sigma\sqrt{T} Z\right] - K\right\}^+$$

The sample average using n i.i.d. samples is

$$\Delta \approx \frac{1}{2h} \cdot \frac{1}{n} \sum_{i=1}^{n} [f^+(Z_i) - f^-(Z_i)]$$

Note that the same random numbers are used in computation of $f^{\pm}(Z)$. This is to reduce the variance of the estimate, since $f^{\pm}(Z)$ are positively correlated.

Algorithm 27: Delta of Call Option with Finite Difference Method

for $i = 1$ to n **do**
  Generate Z_i from $N(0,1)$
  Set $X_i = (S_0 + h) \cdot \exp((r - \sigma^2/2)^T + \sigma\sqrt{T}Z_i)$
  Set $Y_i = (S_0 - h) \cdot \exp((r - \sigma^2/2)^T + \sigma\sqrt{T}Z_i)$
  Set $H_i = \frac{1}{2h}(e^{-rT}(X_i - K)^+ - e^{-rT}(Y_i - K)^+)$
end for
Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} H_i$
Compute the standard error S.E.= $\sqrt{\frac{1}{n(n-1)}\left(\sum_{i=1}^{n} H_i^2 - n\hat{v}^2\right)}$

Method 2.8.5. *Method of Pathwise Differentiation*
Exchange the order of expectation and differentiation:

$$V'(\theta) = \frac{\partial}{\partial\theta}\mathbb{E}[X(\theta)] = \mathbb{E}\left[\frac{\partial}{\partial\theta}X(\theta)\right]$$

The MC estimate is the sample average of i.i.d. copies of $\frac{\partial}{\partial\theta}X(\theta)$. Often an explicit formula of derivative exists. Validity of method (second equality) hinges on interchangeability of the order of differentiation and expectation. It can be justified if $X(\theta)$ is everywhere continuous and almost everywhere continuously differentiable.

Example 2.8.6. [Delta for Binary Option]
Consider the delta at $t = 0$ for binary option with maturity T and discounted payoff

$$X = e^{-rT}1_{\{S_T \geq K\}}$$

X is discontinuous at $S_T = K$, and $\frac{\partial X}{\partial S_0} = 0$ almost everywhere, but delta of option is clearly not zero.
Use method of pathwise differentiation to estimate Δ at $t = 0$ for call option with maturity T, strike price K.
Assume stock price is a geometric Brownian motion

$$S_t = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right)T + \sigma W_T\right)$$

The discounted payoff of the option is

$$X = e^{-rT}(S_T - K)^+$$

It follows that

$$\frac{\partial X}{\partial S_0} = e^{-rT}1_{\{S_T \geq K\}}\frac{\partial S_T}{\partial S_0} = \frac{e^{-rT}S_T}{S_0}1_{\{S_T \geq K\}}, \quad \Delta = \mathbb{E}\left[\frac{e^{-rT}S_T}{S_0}1_{\{S_T \geq K\}}\right]$$

Example 2.8.7. [SDE, Solve for Delta]
Assume that under the risk-neutral probability measure, the underlying stock price satisfies SDE

$$dS_t = rS_t dt + \sigma(S_t)dW_t$$

where r is a constant, $\sigma(S)$ is a given function.
Estimate delta at $t = 0$ of an option with maturity T and payoff $h(S_T)$.
The price of option is

$$V = \mathbb{E}[e^{-rT}h(S_T)] \Rightarrow \Delta = \mathbb{E}\left[e^{-rT}\frac{\partial}{\partial S_0}h(S_T)\right] = \mathbb{E}\left[e^{-rT}h'(S_T)\frac{\partial S_T}{\partial S_0}\right]$$

Note that $\frac{\partial S_T}{\partial S_0}$ cannot be calculated explicitly. Consider Euler scheme for S,

$$S_{i+1} = S_i + rS_i(t_{i+1} - t_i) + \sigma(S_i)\sqrt{t_{i+1} - t_i}Z_{i+1}$$

for $i = 0, 1, \ldots, m-1$. Take derivative w.r.t. S_0 on both sides of above equation, and denote $D_i = \frac{\partial S_i}{\partial S_0}$:

$$D_{i+1} = D_i + rD_i(t_{i+1} - t_i) + \sigma'(S_i)D_i\sqrt{t_{i+1} - t_i}Z_{i+1}, \quad D_0 = 1$$

The MC estimate of the delta is the sample average of i.i.d. copies of

$$e^{-rT}h'(S_m)D_m$$

Example 2.8.8. [Stochastic Volatility Model with Pathwise Differentiation]
Consider stochastic volatility model where under the risk-neutral measure,

$$\begin{aligned} dS_t &= rS_t dt + \theta_t S_t dW_t \\ d\theta_t &= a(\Theta - \theta_t)dt + \beta dB_t \end{aligned}$$

where r, a, Θ, β are positive constants, (W_t, B_t) is a 2-dimensional brownian motion with covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \begin{pmatrix} W_T \\ B_T \end{pmatrix} \sim N(0, T\Sigma)$$

Estimate the delta of the call option with maturity T and strike price K.
[Plain Monte Carlo] method:

i. Let $S_t = S_0 e^{Y_t}$, then it follows from Ito formula that

$$dY_t = \left(r - \frac{1}{2}\theta_t^2\right) dt + \theta_t dW_t, \quad Y_0 = 0$$

ii. Since Y_t is independent of S_0, then

$$\frac{\partial S_T}{\partial S_9} = e^{Y_T}$$

iii. It follows that the delta of the call option is

$$\Delta = \mathbb{E}\left[e^{-rT}\frac{\partial S_T}{\partial S_0} 1_{\{S_T \geq K\}}\right] = \mathbb{E}[e^{-rT}e^{Y_T}1_{\{Y_T \geq y\}}], \quad y = \log\left(\frac{K}{S_0}\right)$$

iv. The SDEs for Y_t and θ can be solved using Euler scheme:

$$\begin{aligned} Y_{i+1} &= Y_i + \left(r - \frac{1}{2}\theta_i^2\right) h_i + \theta_i \sqrt{h_i} Z_{i+1} \\ \theta_{i+1} &= \theta_i + a(\Theta - \theta_i)h_i + \beta\sqrt{h_i} R_{i+1} \end{aligned}$$

where $h_i = t_{i+1} - t_i$ and $\{(Z_1, R_1), \ldots, (Z_m, R_m)\}$ are i.i.d. jointly normal r.v. with distribution $N(0, \Sigma)$.

v. The plain Monte Carlo estimate for Δ is the sample average of i.i.d. copies of

$$X = e^{-rT}e^{Y_m}1_{\{Y_m \geq y\}}$$

[Control Variate] method:

i. Introduce artificial stochastic process

$$d\overline{Y}_t = \left(r - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t, \quad \overline{Y}_0 = 0$$

ii. The control variate is defined as $\overline{X} = e^{-rT}e^{\overline{Y}_T}1_{\{\overline{Y}_T \geq y\}}$.

iii. The delta of option with maturity T and strike price K is $\mathbb{E}[\overline{X}]$, where the underlying stock price is a geometric brownian motion with drift r, volatility σ. $\mathbb{E}[\overline{X}]$ is known:

$$\mathbb{E}[\overline{X}] = \Phi(\alpha), \quad \alpha = \frac{1}{\sigma\sqrt{T}}\log\frac{S_0}{K} + \left(\frac{\sigma}{2} + \frac{r}{\sigma}\right)\sqrt{T}$$

iv. The control variate estimate for the delta is the sample average of

$$X - b(\overline{X} - \Phi(\alpha))$$

Algorithm 28: Stochastic Volatility Model Pathwise Differentiation with Plain Monte Carlo

Set $y = \log(K/S_0)$
for $k = 1$ to n **do**
  Set $Y_0 = 0$
  for $i = 0$ to $m-1$ **do**
    Generate i.i.d. samples Z_{i+1} and U_{i+1} from $N(0,1)$
    Set $R_{i+1} = \rho Z_{i+1} + \sqrt{1-\rho^2}U_{i+1}$
    Set $Y_{i+1} = Y_i + (r - \theta_i^2/2)h_i + \theta_i\sqrt{h_i}Z_{i+1}$
    Set $\theta_{i+1} = \theta_i + a(\Theta - \theta_i)h_i + \beta\sqrt{h_i}R_{i+1}$
  end for
  Set $X_k = 1_{\{Y_m \geq y\}} \exp(Y_m - rT)$
end for
Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} X_k$

Algorithm 29: Stochastic Volatility Model Pathwise Differentiation with Control Variate

Set $\sigma = \Theta$, $y = \log(K/S_0)$, $\alpha = \frac{1}{\sigma\sqrt{T}}\log\frac{S_0}{K} + \left(\frac{\sigma}{2} + \frac{r}{\sigma}\right)\sqrt{T}$
for $k = 1$ to n **do**
  Set $Y_0 = 0$, $\overline{Y} = 0$
  for $i = 0$ to $m-1$ **do**
    Generate i.i.d. samples Z_{i+1} and U_{i+1} from $N(0,1)$
    Set $R_{i+1} = \rho Z_{i+1} + \sqrt{1-\rho^2}U_{i+1}$
    Set $Y_{i+1} = Y_i + (r - \theta_i^2/2)h_i + \theta_i\sqrt{h_i}Z_{i+1}$
    Set $\overline{Y}_{i+1} = \overline{Y}_i + (r - \sigma^2/2)h_i + \sigma\sqrt{h_i}Z_{i+1}$
    Set $\theta_{i+1} = \theta_i + a(\Theta - \theta_i)h_i + \beta\sqrt{h_i}R_{i+1}$
  end for
  Set $X_k = 1_{\{Y_m \geq y\}} \exp(Y_m - rT)$, $\overline{X}_k = 1_{\{\overline{Y}_m \geq y\}} \exp(\overline{Y}_m - rT)$
end for
Compute b^* and $H_k = X_k + b^*(\overline{X}_k) - \Phi(\alpha)$, $k = 1, \ldots, n$
Compute the estimate $\hat{v} = \frac{1}{n}\sum_{i=1}^{n} X_k$

Method 2.8.9. *[Method of Score Function]*
Consider problem of estimating $\frac{\partial V}{\partial \theta}$, where $V = \mathbb{E}[X(\theta)]$.
Perform change of variable so that $X(\theta) = H(Y)$ for some function H that has no dependence on θ, and some random variable Y that has density $f(y;\theta)$. Then

$$
\begin{aligned}
V &= \mathbb{E}[H(Y)] = \int H(y) f(y;\theta) dy \\
\Rightarrow \frac{\partial V}{\partial \theta} &= \int H(y) \frac{\partial f}{\partial \theta} dy = \int H(y) \frac{\partial \log f}{\partial \theta} \cdot f(y;\theta) dy \\
\Rightarrow \frac{\partial V}{\partial \theta} &= \mathbb{E}\left[H(y) \frac{\partial \log f}{\partial \theta}(Y;\theta)\right]
\end{aligned}
$$

The score function is then $\frac{\partial \log f}{\partial \theta}$.
Method requires density function to be smooth in θ (milder than regularity condition on payoff function).

Example 2.8.10. [Score Function for Estimating Delta of Option]
Consider estimating delta of option with maturity T and discounted payoff $h(S_T)$. Stock price is assumed to be geometric Brownian motion with drift r and volatility σ:

$$S_T = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z\right)$$

where Z is a standard normal random variable. Define $Y = Z + \frac{1}{\sigma\sqrt{T}}\log S_0$; then stock price at $t = T$ is

$$S_T = \exp\left(\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Y\right) = \hat{S}_T(Y)$$

The option price is

$$V = \mathbb{E}_Y[h(\hat{S}_T(Y))] = \int_{-\infty}^{\infty} h(\hat{S}_T(y)) \cdot f(y; S_0)dy$$

where f is the density of Y:

$$f(y; S_0) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(y - \frac{1}{\sigma\sqrt{T}}\log S_0\right)^2\right)$$

Note that the function $h(\hat{S}_T(y))$ has no dependence on S_0. The delta of option is (let $H(y) = h(\hat{S}_T(y))$):

$$V'(S_0) = \int H(y)\frac{\partial f}{\partial S_0}(y; S_0)dy = \int H(y)\frac{\partial \log f}{\partial S_0} \cdot f(y; S_0)dy = \mathbb{E}\left[H(Y)\frac{\partial \log f}{\partial S_0}\right]$$

where

$$\frac{\partial \log f}{\partial S_0} = \frac{Y}{\sigma\sqrt{T}S_0} - \frac{1}{\sigma^2 T S_0}\log S_0$$

By definition of Y, we have

$$\frac{\partial \log f}{\partial S_0} = \frac{Z}{\sigma\sqrt{T}S_0}$$

Hence,

$$\Delta = \mathbb{E}\left[h(S_T(Z))\frac{Z}{\sigma\sqrt{T}S_0}\right]$$

Note that in this formulation, the payoff function does not have to be continuous (i.e., binary option).

Example 2.8.11. [Score Function for Estimating Vega of Option]
Estimate Vega at $t = 0$ of option with maturity T and payoff $h(S_T)$. Stock price is a geometric Brownian motion with drift r and volatility σ.
The Vega of the option is $\nu = \frac{\partial V}{\partial \sigma}$. The stock price is

$$S_T = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z\right) = S_0 \exp(rT + Y)$$

where

$$Y = -\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}Z$$

The price of the option is

$$V = \mathbb{E}[H(Y)], \quad H(Y) = e^{-rT}h(S_T) = e^{-rT}h(S_0 e^{rT+Y})$$

The distribution of Y is $N(-\frac{1}{2}\sigma^2 T, \sigma^2 T)$, its density satisfies

$$\begin{aligned}
\frac{\partial \log f}{\partial \sigma}(y; \sigma) &= \frac{y^3}{\sigma^3 T} - \frac{\sigma T}{4} - \frac{1}{\sigma} \\
\Rightarrow \frac{\partial \log f}{\partial \sigma}(Y; \sigma) &= \frac{Y^2}{\sigma^3 T} - \frac{\sigma T}{4} - \frac{1}{4} \\
&= \frac{1}{\sigma^3 T}\left(\sigma\sqrt{T}Z - \frac{1}{2}\sigma^2 T\right)^2 \frac{\sigma T}{4} - \frac{1}{\sigma}
\end{aligned}$$

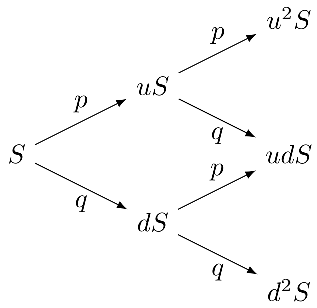
It follows that

$$\begin{aligned}
\nu &= \mathbb{E}\left[H(Y)\frac{\partial \log f}{\partial \sigma}(Y)\right] \\
&= \mathbb{E}\left[e^{-rT}h(S_T)\left(\frac{1}{\sigma}(Z^2 - 1) - \sqrt{T}Z\right)\right]
\end{aligned}$$

2.9 Method for American Options

At any given time, if value of exercising is larger than value of waiting, then it is preferable to exercise. Otherwise, it is preferable to wait.

Definition 2.9.1. *[Binomial Tree Model]*
The stock price is initially S. In each time step, it either moves up to u times its current value, or moves down to d times its current value.



The risk-neutral probability measure is

$$p = \frac{e^{r\Delta t} - d}{u - d}, \quad q = 1 - p$$

Method 2.9.2. *[European Options]*
Model stock price with N-period binomial tree.
Work backwards through the tree from the end to the beginning.

$$\begin{aligned} V_N(x) &= H(x) \\ V_n(x) &= e^{-r\Delta t}\mathbb{E}[V_{n+1}] \\ &= e^{-r\Delta t}[pV_{n+1}(u \cdot x) + qV_{n+1}(d \cdot x)], \quad n = N-1, \ldots, 1, 0 \end{aligned}$$

where $V_n(x)$ is value of option at n-th time step and x is stock price, $H(x)$ is payoff from exercising the option.

Method 2.9.3. *[American Options]*
At each node, the value of the option is grater of the payoff from immediate exercise, and the value of waiting.

$$\begin{aligned} V_N(x) &= H(x) \\ V_n(x) &= \max\{H(x), e^{-r\Delta t}[pV_{n+1}(u \cdot x) + qV_{n+1}(d \cdot x)]\}, \quad n = N-1, \ldots, 1, 0 \end{aligned}$$

The optimal strategy is to exercise if the immediate payment is larger than the expected future payments.

Method 2.9.4. *[Least Square Approach]*
Generate M paths for underlying stock price.
At time t, the expected future payment is computed by fitting the future payoffs for the M maths using quadratic functions for example:

$$\mathbb{E}[Y \mid S] = \alpha + \beta S + \gamma S^2$$

where Y is the future payoff, and S is the stock price at time t.
The coefficients α, β and γ are obtained from solving the least squares problem

$$\begin{gathered} \min_x \|Ax - Y\|_2 \\ f(x) = \|Ax - Y\|_2^2 = x^T A^T Ax - 2x^T A^T y + Y^T Y \\ \nabla f(x) = 0 \Rightarrow A^T Ax = A^T Y \Rightarrow x = (A^T A)^{-1} A^T Y \end{gathered}$$

By comparing expected future payoff and the payoff from immediate exercise, decide if it is preferable to exercise the option immediately or wait one more period.
Work backward from maturity date to time zero. The price of option is average of all discounted payoffs.

Remark 2.9.5. *Extensions of Least Squares Approach*
The method can be extended in a number of ways:

i. If option can be exercised at any time, approximate value by considering large number of exercise points.

ii. The relationship between expected future payoff and S can be assumed to be more complicated, i.e., with a cubic function rather than a quadratic function.

iii. The method can also be used when there are multiple underlying assets:

   (a) A functional form for relationship between expected future payoff and asset price is assumed

   (b) The parameters are estimated using least squares approach

Example 2.9.6. *[American Call Option with Three Underlying Assets]*
American call option written on three underlying assets. The payoff function at time t is

$$\max(0, S_1(t) - K_1, S_2(t) - K_2, S_3(t) - K_3)$$

Assume the expected future payoff is a quadratic function of stock prices:

$$\mathbb{E}[Y \mid S_1, S_2, S_3] = \alpha + \beta_1 S_1 + \beta_2 S_2 + \beta_3 S_3 + \gamma_1 S_1^2 + \gamma_2 S_2^2 + \gamma_3 S_3^2 + \gamma_4 S_1 S_2 + \gamma_5 S_1 S_3 + \gamma_6 S_2 S_3$$

The parameters are computing using least squares fitting.