Lecture Notes Computational Finance

Sven Karbach

Universiteit van Amsterdam

26th Apr, 2025

1 Financial Markets

We consider a continuous-time **financial market** with time-horizon $[0,T]$ and $d \in \mathbb{N}$ assets. A **financial market** is described by a tuple $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F}, \mathbf{S})$, where:

- Ω represents the set of all **outcomes** $\omega \in \Omega$,
- $\mathcal{F}$ is a σ-algebra on Ω containing all **events** $A \in \mathcal{F}$,
- $\mathbb{P}$ denotes the **physical probability measure** defined on the measurable space $(\Omega, \mathcal{F})$,
- $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is a filtration, i.e. $\mathcal{F}_t \subseteq \mathcal{F}$ for $t \in [0,T]$ being a sub-σ-algebra such that $\mathcal{F}_s \subseteq \mathcal{F}_t$ for all $0 \leq s \leq t \leq T$, and describes the flow of information in the market,
- $\mathbf{S} = (S_t)_{t\in[0,T]} = (S_t^{(1)}, \dots, S_t^{(d)})_{t\in[0,T]}$ denotes a $\mathbb{R}^d$-valued stochastic process on $[0,T]$, describing the time-evolution of some $d \in \mathbb{N}$ asset prices.

1.1 The filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$

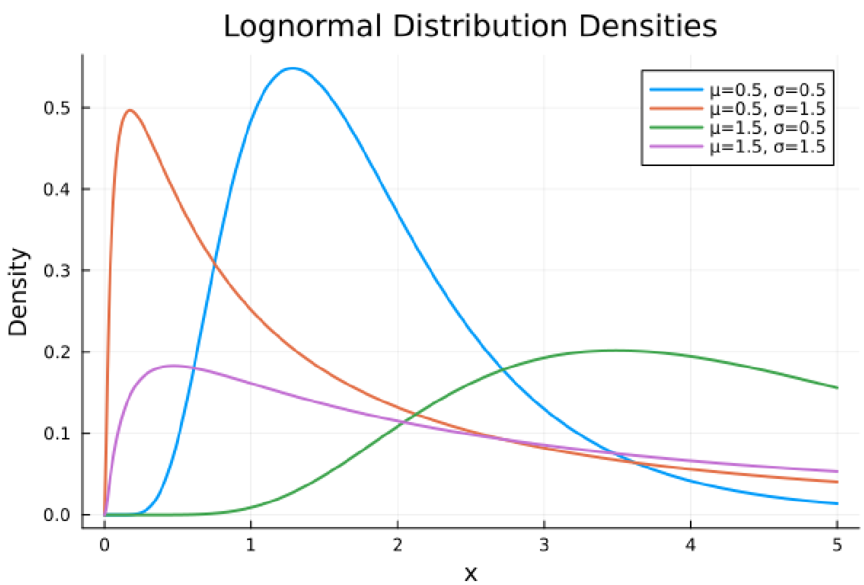
Below, we provide explanations and intuitive interpretations of the abstract definition of a financial market.

Outcomes and Random Variables

- The **set of all outcomes**, denoted by Ω, represents all possible scenarios from which random outcomes can be sampled.
- Drawing samples $\omega_1, \dots, \omega_N$ from Ω according to the probability measure $\mathbb{P}$ allows us to simulate real-world randomness.
- A **random variable** defined on the measurable space $(\Omega, \mathcal{F})$ is a mapping $X\colon \Omega \to \mathbb{R}$ such that $X^{-1}(A) \in \mathcal{F}$ for all Borel sets $A \subseteq \mathbb{R}$, i.e. a measurable function on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.
- The distribution (law) of X is given by $\mathbb{P} \circ X^{-1}$.
- In practice, to simulate the outcomes of the random variable X, we evaluate $X(\omega_1), \dots, X(\omega_N)$. These realizations are typically denoted as samples $x^{(1)}, \dots, x^{(N)}$ and represent the observed data.

Example: Plot of log-Normal Distribution

```
using Distributions
using Plots
# Define the range of x values for plotting
x = 0:0.01:5
# Parameters: arrays of \mu (location) and \sigma (scale) values
mus = [0.5, 1.5]  # Location parameters
sigmas = [0.5, 1.5]  # Scale parameters
# Initialize the plot
# Generate and plot each distribution
plot()
for \mu in mus
    for \sigma in sigmas
        # Create a lognormal distribution with current \mu and \sigma
        dist = LogNormal( \mu, \sigma)
        # Plot the density function
        plot!(x, pdf.(dist, x), label=" \mu=$ \mu, \sigma=$ \sigma", lw=2)
    end
end
plot!(title="Lognormal Distribution Densities", xlabel="x", ylabel="Density")
```



Filtration and the Flow of Information

When modeling time-dependent random phenomena, we must respect the natural chronological ordering implied by the flow of time. Information obtained at an earlier time remains available at later times.

- This structure is captured by the concept of a **filtration**:

$$\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]}. \tag{1}$$

Typically, we assume:

- **All information of today is known to us**: The initial sigma-algebra $\mathcal{F}_0$ is $\mathbb{P}$-trivial, meaning:
  - $\mathbb{P}(A) \in \{0,1\}$ for all $A \in \mathcal{F}_0$.
  - All $\mathcal{F}_0$-measurable random variables are constant almost surely.
- **Terminal information**: The final sigma-algebra contains all events within the time horizon: $\mathcal{F} = \mathcal{F}_T$.

Moreover, a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is said to satisfy the **usual conditions** if:

- The initial sigma-algebra $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets from $\mathcal{F}$.
- The filtration is **right-continuous**:

$$\mathcal{F}_t = \bigcap_{u>t} \mathcal{F}_u \quad \text{for all } t \in [0,T]. \tag{2}$$

Remark: A filtration in practicse can be thought of the information stream, e.g. it carries your trading signals, market news, quote data, or any other information that is relevant for the observed markets.

Stochastic Processes and Adaptedness

- An $\mathbb{R}^d$**-valued stochastic process** $(S_t)_{t\in[0,T]}$ is a family of random variables defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$:

$$\mathbf{S} = (S_t)_{t\in[0,T]} = \{S_t \colon t \in [0,T],\ S_t : \Omega \to \mathbb{R}^d \text{ measurable}\}. \tag{3}$$

To align the stochastic process with the notion of information flow, we introduce the concept of **adaptedness**:

- A stochastic process $\mathbf{S}$ is said to be **adapted** to the filtration $\mathbb{F}$ if each random variable S_t is $\mathcal{F}_t$-measurable for all $t \in [0,T]$. Intuitively, this means the value S_t at time t depends only on information known up to that time.

We call $\omega \mapsto S_t(\omega)$ a **sample at time** t and $t \mapsto S_t(\omega)$ a **sample path** of $(S_t)_{t\in[0,T]}$.

Definition: Martingale

A stochastic process $(M_t)_{t\geq 0}$ adapted to a filtration $(\mathcal{F}_t)$ is called a **martingale** if, for all $s \leq t$, it satisfies:

1. $\mathbb{E}[|M_t|] < \infty$.
2. **Martingale property**:
$$\mathbb{E}[M_t|\mathcal{F}_s] = M_s. \tag{4}$$

Intuitively, a martingale has no predictable trend; its expected future value, given current information, equals its current value. Martingales represent "fair games" or a balanced process in probability and finance.

Example: Brownian Motion

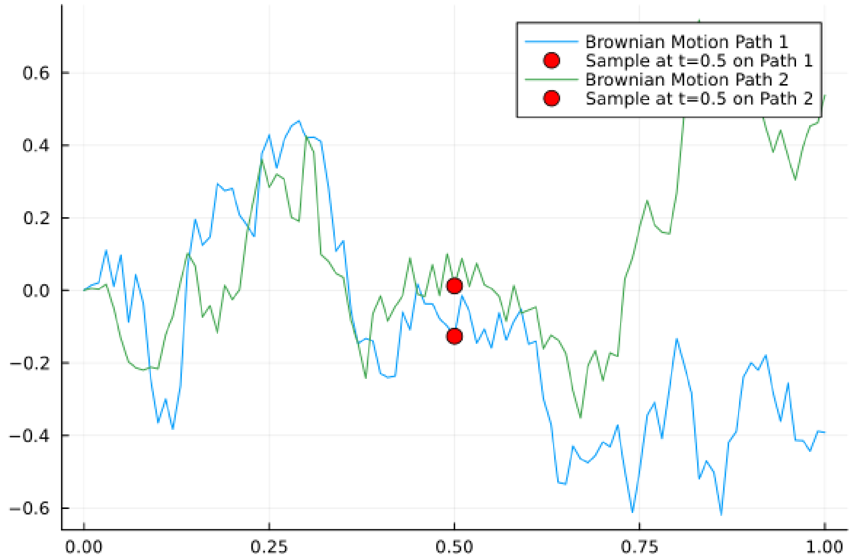
A stochastic process $(W_t)_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ is called a **standard Brownian motion** if it satisfies the following properties:

1. **Starting Point**: $W_0 = 0$ almost surely.
2. **Independent increments**: For all $0 \leq s < t$, the increment $W_t - W_s$ is independent of $\mathcal{F}_s$.
3. **Gaussian increments**: (W**_**t - W**_**s \sim \mathcal{N}(0, t-s)), with mean zero and variance (t-s).
4. **Continuous Paths**: With probability one, the function $t \mapsto W_t$ is continuous.

Next, let us plot as example two sample paths of a Brownian motion. We highlight two sample points at $t = 0.5$.

```
using Plots

# Parameters
T = 1.0  # Total time
N = 100  # Number of steps
dt = T / N  # Time step size
t = 0:dt:T  # Time vector

# Function to generate Brownian motion
function generate_brownian_motion(N, dt)
    W = zeros(N+1)  # Initialize Wiener process array
    for i = 2:length(W)
        W[i] = W[i -1] + sqrt(dt) * randn()
    end
    return W
end
# Generate two paths of Brownian motion
W1 = generate_brownian_motion(N, dt)
W2 = generate_brownian_motion(N, dt)
```

```
# Fixed time point to sample
fixed_time = 0.5
fixed_index = Int(floor(fixed_time / dt)) + 1
plot(t, W1, label="Brownian Motion Path 1", legend=:topright)
scatter!([fixed_time], [W1[fixed_index]], label="Sample at t=0.5 on Path 1", color="red", markersize=6)
plot!(t, W2, label="Brownian Motion Path 2")
scatter!([fixed_time], [W2[fixed_index]], label="Sample at t=0.5 on Path 2", color="red", markersize=6)
```



2 Stochastic Modelling of Financial Markets

Asset prices are represented by a stochastic process taking values in $\mathbb{R}^d$:

$$\mathbf{S} = (S_t)_{t\in[0,T]} = \left(S_t^{(1)}, \ldots, S_t^{(d)}\right)_{t\in[0,T]}. \tag{5}$$

This stochastic process describes how prices evolve over time for $d \in \mathbb{N}$ distinct financial **risky assets**, typically referred to as the **underlyings**. For instance, S_t could denote the price at time t of

- Stocks (shares of publicly traded companies),
- Commodities (gold, oil, agricultural products),
- Foreign exchange rates (currencies),
- Other exchange-traded instruments (observable at continuous or discrete time points).

Consider an asset price process $\mathbf{S} = (S_t)_{t\in[0,T]} = (S_t^{(1)}, \ldots, S_t^{(d)})_{t\in[0,T]}$. In case of a one-asset market $d = 1$, we simply write $(S_t)_{t\in[0,T]}$ instead of $(S_t^{(1)})_{t\in[0,T]}$.

A sample $\omega \mapsto S_t(\omega)$ is **distributed** according to the model specification. For some models, the distribution at each time point is known in closed form, e.g., in the Black-Scholes Model we have:

$$ln(S_t) \sim \mathcal{N}\big(ln(S_0) + (\mu - \frac{\sigma^2}{2})t,\ \sigma^2 t\big). \tag{6}$$

However, since we model financial markets dynamic in time, we are usually interested in the entire dynamics of the asset price process over the interval $[0, T]$. This is done using stochastic differential equations (SDE). For example, in the Black-Scholes model the dynamics of an asset price are modelled as

$$dS_t = \mu S_t\, dt + \sigma S_t\, dW_t, \quad S_0 > 0, \tag{7}$$

where $(W_t)_{t\geq 0}$ denoes a standard Brownian motion, the parameter μ is the drift consisting of the instantaneous interest rate r and a risk premium. In the following, we give provide some preliminaries on stochastic calculus, but only limited to what is needed in this course.

2.1 Stochastic Calculus Tools for Computational Finance

Stochastic calculus is the fundamental mathematical framework used in computational finance to describe and analyze the random evolution of financial variables such as asset prices. At its core is the concept of the Itô integral, which allows us to rigorously define integrals with respect to stochastic processes.

Admissible Integrands and Itô Integrals

To define stochastic integrals, we first specify the class of processes we can integrate:

Definition (Admissible Integrands)
Consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, with a standard Brownian motion $(W_t)_{t\geq 0}$. A real-valued stochastic process $\varphi = (\varphi_t)_{t\geq 0}$ is an **admissible integrand** if:

- φ is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$
- φ is square-integrable, i.e., for all $T > 0$:

$$\mathbb{E}\left[\int_0^T \varphi_t^2\, dt\right] < \infty. \tag{8}$$

The **Itô integral** of φ with respect to the Brownian motion $(W_t)_{t\geq 0}$ is formally defined as:

$$\int_0^T \varphi_t \, dW_t, \tag{9}$$

which under suitable conditions is a square-integrable martingale.

Interpretation of the Itô Integral

The Itô integral can be intuitively understood as the **stochastic limit** of sums involving the integrand φ and increments of the Brownian motion $(W_t)_{t\geq 0}$. Specifically, let us consider a partition $0 = t_0 < t_1 < \cdots < t_n = T$ of the interval $[0, T]$. The Itô integral is defined as the limit (in probability and mean-square sense) of the following sum as the mesh size of the partition tends to zero:

$$\int_0^T \varphi_t \, dW_t = \lim_{\|\pi\| \to 0} \sum_{i=0}^{n-1} \varphi_{t_i} (W_{t_{i+1}} - W_{t_i}), \tag{10}$$

where $\|\pi\| = \max_{0 \leq i \leq n-1} (t_{i+1} - t_i)$.

2.1.1 Itô Processes in Finance

Financial models typically involve processes whose evolution depends on their current state. A general scalar-valued Itô process (S_t) often used in financial modeling can be written as:

$$dS_t = \mu(t, S_t) \, dt + \sigma(t, S_t) \, dW_t, \quad S_0 \in \mathbb{R}, \tag{11}$$

where:

- $\mu(t, S_t)$ is the drift coefficient, representing deterministic trends.
- $\sigma(t, S_t)$ is the volatility coefficient, representing the randomness.
- $(W_t)_{t\geq 0}$ is a standard Brownian motion.

To gain deeper insight, consider a small increment Δt of time. By integrating the SDE approximately over the interval $[t, t + \Delta t]$, we have:

$$S_{t+\Delta t} - S_t \approx \mu(t, S_t)\Delta t + \sigma(t, S_t)\Delta W_t, \tag{12}$$

where $\Delta W_t = W_{t+\Delta t} - W_t \sim \mathcal{N}(0, \Delta t)$. Taking conditional expectation at time t, we obtain the first-order approximation:

$$\mathbb{E}[S_{t+\Delta t} - S_t \,|\, \mathcal{F}_t] \approx \mu(t, S_t)\Delta t. \tag{13}$$

Thus, the drift $\mu(t, S_t)$ directly captures the instantaneous expected rate of change of S_t:

$$\mu(t, S_t) \approx \frac{\mathbb{E}[S_{t+\Delta t} - S_t \,|\, \mathcal{F}_t]}{\Delta t}, \quad \text{as } \Delta t \to 0. \tag{14}$$

Similarly, we compute the conditional variance of the increment:

$$\mathrm{Var}(S_{t+\Delta t} - S_t \,|\, \mathcal{F}_t) = \mathrm{Var}\left(\int_t^{t+\Delta t} \sigma(u, S_u) \, dW_u \middle| \mathcal{F}_t \right). \tag{15}$$

Applying **Itô's isometry**, which states that for any admissible integrand φ_u:

$$\mathbb{E}\left[\left(\int_t^{t+\Delta t} \varphi_u \, dW_u\right)^2 \middle| \mathcal{F}_t\right] = \mathbb{E}\left[\int_t^{t+\Delta t} \varphi_u^2 \, du \middle| \mathcal{F}_t\right], \tag{16}$$

we obtain the approximation for small Δt:

$$\mathrm{Var}(S_{t+\Delta t} - S_t \,|\, \mathcal{F}_t) \approx \sigma(t, S_t)^2 \Delta t. \tag{17}$$

Thus, the volatility coefficient $\sigma(t, S_t)$ directly captures the instantaneous accumulation of variance:

$$\sigma(t, S_t)^2 \approx \frac{\mathrm{Var}(S_{t+\Delta t} - S_t \,|\, \mathcal{F}_t)}{\Delta t}, \quad \text{as } \Delta t \to 0. \tag{18}$$

Quadratic Variation and Variance Swaps

The concept of **quadratic variation** is essential to quantify the intrinsic variability of stochastic processes:

Definition (Quadratic Variation)
The quadratic variation of an Itô process (S_t) is defined by:

$$\langle S \rangle_t = \int_0^t \sigma(u, S_u)^2 \, du. \tag{19}$$

This quantity captures the accumulated variance of the process over time.

2.2 Example: Variance Swap

A **variance swap** is a derivative instrument whose payoff directly depends on the realized variance (quadratic variation) of the underlying asset's returns:

$$\text{Variance Swap Payoff} : h(S) = \langle \ln(S) \rangle_T - K_{\text{var}}, \tag{20}$$

where K_{var} is a pre-specified strike, representing the initially agreed variance level.

2.3 Itô's Formula and Applications

Itô's formula extends the classical chain rule to stochastic processes, providing a powerful tool for analyzing the evolution of nonlinear functions of stochastic variables.

Theorem (Itô's Formula)
Consider an Itô process $(S_t)_{t\geq 0}$ given by:

dS$_t = \mu(t, S_t)\, dt + \sigma(t, S_t)\, dW_t$.(21)

Let $f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$. Then, Itô's formula states:

$$df(t, S_t) = \left(\frac{\partial f}{\partial t}(t, S_t) + \mu(t, S_t)\frac{\partial f}{\partial S}(t, S_t) + \frac{1}{2}\sigma(t, S_t)^2 \frac{\partial^2 f}{\partial S^2}(t, S_t)\right) dt + \sigma(t, S_t)\frac{\partial f}{\partial S}(t, S_t)\, dW_t. \tag{22}$$

Example: Geometric Brownian Motion (GBM)

Consider the **Geometric Brownian Motion**, a classical model in finance for describing stock prices:

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t. \tag{23}$$

To apply Itô's formula to the logarithmic transformation $f(t, S_t) = \ln(S_t)$, we compute the necessary partial derivatives: $\frac{\partial f}{\partial t}(t, S) = 0$, $\frac{\partial f}{\partial S}(t, S) = \frac{1}{S}$ and $\frac{\partial^2 f}{\partial S^2}(t, S) = -\frac{1}{S^2}$.

Applying Itô's formula, we obtain:

$$d \ln(S_t) = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma \, dW_t. \tag{24}$$

Integrating both sides explicitly from 0 to t:

$$\ln(S_t) = \ln(S_0) + \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t. \tag{25}$$

Exponentiating, we recover the explicit solution of the GBM SDE:

$$S_t = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t\right). \tag{26}$$

Multivariate Itô's Formula

Let X be an $\mathbb{R}^n$-valued Itô process and let $f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R})$ be a smooth function. Then, almost surely, for all $t \geq 0$:

$$\begin{aligned} f(t, X_t) = f(0, X_0) &+ \int_0^t \left(\partial_s f_s + \langle \mu_s, \nabla f_s \rangle + \frac{1}{2} \operatorname{Tr}(\sigma_s^\top \Delta f_s \sigma_s)\right) ds \\ &+ \int_0^t \langle \sigma_s dW_s, \nabla f_s \rangle, \end{aligned} \tag{27}$$

where:

- $f_s := f(s, X_s)$,
- ∇f is the **gradient** of f with respect to the spatial variable,
- Δf is the **Hessian** matrix (second derivatives),
- $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product.

Corollary (Product Rule for Itô Processes)

Let X_t and Y_t be two Itô processes. Then:

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s \, dY_s + \int_0^t Y_s \, dX_s + \int_0^t dX_s \, dY_s. \tag{28}$$

Rearranging the product rule yields the stochastic integration by parts formula:

$$\int_0^t X_s \, dY_s = X_t Y_t - X_0 Y_0 - \int_0^t Y_s \, dX_s - \int_0^t dX_s \, dY_s. \tag{29}$$

2.3.1 The Martingale Representation Theorem

A fundamental result in stochastic calculus, particularly important in financial mathematics, is the **Martingale Representation Theorem**, which characterizes how martingales driven by Brownian motion can be represented.

Theorem (Martingale Representation)
Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space, where the filtration $(\mathcal{F}_t)$ is generated by a standard Brownian motion $(W_t)_{t\geq 0}$. Assume $M = (M_t)_{t\geq 0}$ is a square-integrable martingale with respect to $\mathbb{P}$ and adapted to the filtration $(\mathcal{F}_t)$.

Then, there exists a unique adapted and square-integrable process $\varphi = (\varphi_t)_{t\geq 0}$ such that:

$M_t = M_0 + \int_0^t \varphi_u \, dW_u, \quad \text{for all } t \geq 0.$

Theorem (Girsanov Theorem)
Let $(W_t)_{t\geq 0}$ be a Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. Consider an adapted process θ_t, satisfying Novikov's condition:

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \theta_s^2 \, ds\right)\right] < \infty. \tag{30}$$

Define the process Z_t as:

$$Z_t = \exp\left(-\int_0^t \theta_s \, dW_s - \frac{1}{2}\int_0^t \theta_s^2 \, ds\right), \quad 0 \leq t \leq T, \tag{31}$$

which is a strictly positive martingale with expectation one. Using Z_t, we can define a new probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$:

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T} = Z_T. \tag{32}$$

Then, under the new measure $\mathbb{Q}$, the process $(\tilde{W}_t)_{0\leq t\leq T}$ defined by:

$$\tilde{W}_t = W_t + \int_0^t \theta_s \, ds, \quad 0 \leq t \leq T, \tag{33}$$

is a standard Brownian motion.

3 Operations in Financial Markets

In a continuous-time financial market, a **trading strategy** is described by an adapted stochastic process:

$$\boldsymbol{\theta} = (\theta_t)_{t\in[0,T]} = (\theta_t^{(1)}, \dots, \theta_t^{(d)})_{t\in[0,T]}, \tag{34}$$

where:

- $\theta_t^{(i)}$ denotes the number of units of asset i held in the portfolio at time t.
- The process $\boldsymbol{\theta}$ is **adapted** to the market filtration, meaning that trading decisions at time t are made based solely on the information available up to that time (i.e., no foresight or future knowledge is used).

This formalism ensures that trading strategies are realistic and respect the chronological flow of market information.

Definition (Gains Process)

The **gains process** of a trading strategy $\boldsymbol{\theta}$ measures the cumulative gains or losses accrued from trading over time:

$$G_t(\boldsymbol{\theta}) = \int_0^t \boldsymbol{\theta}_u \cdot d\mathbf{S}_u, \quad t \in [0,T], \tag{35}$$

where:

- $\mathbf{S}_u$ is the vector of asset prices at time u.
- The integral accounts for the infinitesimal changes in asset prices weighted by the portfolio holdings, capturing the cumulated incremental gains or losses from continuous trading activity.

Definition (Self-Financing Strategies and the Wealth Process)

A trading strategy $\boldsymbol{\theta}$ is called **self-financing** if all changes in the portfolio's value $V_t(\boldsymbol{\theta}) := \boldsymbol{\theta} \cdot \mathbf{S}_t = \sum_{i=1}^d \theta_t^{(i)} S_t^{(i)}$ are solely due to price movements of the underlying assets—no additional capital is injected or withdrawn during the trading horizon. In mathematical terms, this reads as:

$$dV_t(\boldsymbol{\theta}) = \boldsymbol{\theta}_t \cdot d\mathbf{S}_t, \quad t \in [0,T]. \tag{36}$$

Given a self-financing strategy $\boldsymbol{\theta}$ with initial capital V_0, the **wealth process** is given by:

$$V_t(\boldsymbol{\theta}) = V_0 + \int_0^t \boldsymbol{\theta}_u \cdot d\mathbf{S}_u. \tag{37}$$

This expression states that the total wealth at time t equals the initial investment plus the accumulated gains from trading over time.

3.1 Discounting and the Bank Account

In financial markets, future payments must be adjusted (or **discounted**) to reflect their present value—this reflects the **time value of money**.

We designate one particular asset as special and refer to it as the **bank account** (or **numéraire**). It is often modeled as the 0-th asset in the market and is typically denoted by B_t.

Definition (The Bank Account Process)

The **bank account process** $(B_t)_{t\in[0,T]}$ models the value of a continuously compounded risk-free investment growing at a deterministic interest rate r_t:

$$B_t = \exp\left(\int_0^t r_u \, du\right), \quad t \in [0,T]. \tag{38}$$

Where:

- r_t is the **instantaneous risk-free interest rate** at time t,
- B_t acts as a **numéraire**, i.e., a reference asset used for discounting other cash flows.

Discounted Asset and Wealth Processes

Given the bank account process B_t, we define the **discounted asset price process** as:

$$\tilde{S}_t = \frac{S_t}{B_t}, \quad t \in [0,T]. \tag{39}$$

Similarly, the **discounted wealth process** for a trading strategy $\boldsymbol{\theta}$ is defined by:

$$\tilde{V}_t(\boldsymbol{\theta}) = \frac{V_t(\boldsymbol{\theta})}{B_t}, \quad t \in [0,T]. \tag{40}$$

By discounting asset price and portfolio wealth with respect to the bank account process, one expresses asset prices in terms of units of the bank account and not with respect to some refference currency.

The Discount Factor and its Dynamics

Assuming the trading strategy $\boldsymbol{\theta}$ is **self-financing**, we define the **discount factor** between times s and t as:

$$D_{s,t} := \frac{B_s}{B_t}, \quad \text{with} \quad D_{0,t} = \frac{1}{B_t}. \tag{41}$$

The dynamics of the discount factor are given by:

$$dD_{0,t} = -r_t D_{0,t} \, dt. \tag{42}$$

Dynamics of the Discounted Wealth Process

Using $D_{0,t}$, we express the discounted wealth process as:

$$\tilde{V}_t := D_{0,t} V_t. \tag{43}$$

The differential of $\tilde{V}_t$ is computed as:

$$\begin{aligned} d\tilde{V}_t &= d(D_{0,t}V_t) \\ &= d\big(\theta_t^{(0)} D_{0,t} B_t + \boldsymbol{\theta}_t \cdot D_{0,t}\mathbf{S}_t\big) \\ &= \theta_t^{(0)} \, d(D_{0,t}B_t) + d\big(\boldsymbol{\theta}_t \cdot D_{0,t}\mathbf{S}_t\big) \\ &= \boldsymbol{\theta}_t \cdot d\tilde{\mathbf{S}}_t. \end{aligned} \tag{44}$$

Hence, under the self-financing condition:

$$d\tilde{V}_t = \boldsymbol{\theta}_t \cdot d\tilde{\mathbf{S}}_t. \tag{45}$$

Integrating over time, we obtain:

$$\tilde{V}_t = \tilde{V}_0 + \int_0^t \boldsymbol{\theta}_u \cdot d\tilde{\mathbf{S}}_u. \tag{46}$$

Martingale Representation Interpretation

Suppose now the asset dynamics have **no drift** and satisfy the stochastic differential equation:

$$d\mathbf{S}_t = \boldsymbol{\sigma}_t \, dW_t, \tag{47}$$

where $\boldsymbol{\sigma}_t$ is adapted to the filtration generated by the Brownian motion W.

Then the discounted wealth process evolves as:

$$\tilde{V}_t = \tilde{V}_0 + \int_0^t \left(D_{0,u} \boldsymbol{\theta}_u \cdot \boldsymbol{\sigma}_u \right) dW_u. \tag{48}$$

This is consistent with the **Martingale Representation Theorem**, with the integrand:

$$\varphi_t = D_{0,t} \boldsymbol{\theta}_t \cdot \boldsymbol{\sigma}_t. \tag{49}$$

4 Derivatives

A **European contingent claim** is any non-negative, $\mathcal{F}_T$-measurable random variable H.

If such a claim H is $\sigma(X_0, X_1, \ldots, X_T)$-measurable for some stochastic process $X = (X_t)_{t \in [0,T]}$, it is called a **(European) derivative** of X.

In many models, contingent claims are written as functionals of the terminal value of an underlying asset S:

$$H = h(S_T), \tag{50}$$

where $h : \mathbb{R} \to \mathbb{R}_+$ is a deterministic, measurable function called the **payoff function**.

4.0.1 Linear Derivatives

Linear derivatives are financial instruments whose payoffs depend *linearly* on the price of the underlying asset.

Forward Contract

- A contract to buy or sell an asset at a specified future date T for a predetermined price K.
- **Payoff at maturity**: $h(S_T) = S_T - K$.

Futures Contract

- Similar to a forward contract in payoff structure.
- Standardized and traded on regulated exchanges.
- Settled daily through a mechanism known as **marking-to-market**, which reduces counterparty risk.

Swap Contract

- An agreement between two parties to exchange a series of future cash flows.
- The most common example is an **interest rate swap**, where one party pays a fixed rate and the other a floating rate.
- Payoffs depend linearly on interest rates, making swaps linear derivatives.

4.0.2 Exchange-Traded Options

An **option** gives the holder the right — but not the obligation — to buy or sell an asset at maturity T for a predetermined **strike price** K.

European Call Option

- Grants the right to **buy** the underlying asset at strike price K at time T.
- **Payoff**:

$$h(S_T) = (S_T - K)^+. \tag{51}$$

European Put Option

- Grants the right to **sell** the underlying asset at strike price K at time T.
- **Payoff**:

$$h(S_T) = (K - S_T)^+. \tag{52}$$

These options are actively traded on exchanges with standardized terms across a wide range of underlying assets.

4.0.3 Over-The-Counter (OTC) Derivatives

OTC derivatives are customized, privately negotiated contracts traded directly between two parties, outside of centralized exchanges.

These contracts offer great flexibility and can be tailored to meet specific investment goals or risk exposures. However, they can also lead to highly complex **structured products**.

Examples of Exotic OTC Derivatives:

- **Arithmetic Asian Option**
  Payoff depends on the average price of the asset over time:

$$h(S_T) = \left(\frac{1}{T} \int_0^T S_t \, dt - K \right)^+. \tag{53}$$

- **Barrier Option (Down-and-Out Call)**
  Payoff is conditional on the underlying never breaching a barrier level B:

$$h(S_T) = (S_T - K)^+ \cdot \mathbf{1}_{\{\min_{0 \le t \le T} S_t > B\}}, \quad \text{with } B < S_0. \tag{54}$$

- **Lookback Option**
  Payoff depends on the minimum price reached during the option's lifetime:

$$h(S_T) = S_T - \min_{0 \le t \le T} S_t. \tag{55}$$

- **Spread Call Option:** $H(\mathbf{S}) \triangleq H(S_T^1, S_T^2) \triangleq (S_T^1 - S_T^2 - K)^+$ with strike $K > 0$ written on two asset prices $(S_t^1)_{t\in[0,T]}$ anf $(S_t^2)_{t\in[0,T]}$,
- **Digital Barrier Option:** $H(\mathbf{S}) \triangleq \mathbf{1}_{\{S_t \ge B \text{ for some } t \in [0,T]\}}$ with Barrier $B > 0$.

4.1 Arbitrage Opportunities

A trading strategy

$$(\theta_t^{(1)}, \dots, \theta_t^{(d)})_{t\in[0,T]} \tag{56}$$

is said to be **admissible** over the horizon $[0, T]$ if its associated wealth process V_t is bounded from below $\mathbb{P}$-almost surely at all times. Formally, there exists a constant $M \in \mathbb{R}$ such that:

$$\mathbb{P}\left(\inf_{0\le t\le T} V_t \ge M\right) = 1. \tag{57}$$

We denote the set of all admissible strategies by $\mathcal{A}$.

A self-financing admissible trading strategy is called an **arbitrage opportunity** if:

- Initial wealth is zero:

$$V_0 = 0, \tag{58}$$

- Terminal wealth is non-negative almost surely:

$$V_T \ge 0 \quad \text{a.s.}, \tag{59}$$

- There is a strictly positive probability of a gain:

$$\mathbb{P}(V_T > 0) > 0. \tag{60}$$

Interpretation:
Arbitrage refers to the possibility of making a **riskless profit** without any investments.

Definition: Arbitrage-Free Market

A financial market is said to be **arbitrage-free** if **no** self-financing admissible trading strategy in the set $\mathcal{A}$ constitutes an arbitrage opportunity. That is, $\forall \boldsymbol{\theta} \in \mathcal{A}$, if $\boldsymbol{\theta}$ is self-financing, then it is not an arbitrage opportunity.

Remark:
This condition is fundamental in financial economics. It serves as a foundational **equilibrium assumption** ensuring consistency in pricing and the absence of "free lunches" in the market.

4.2 Arbitrage-Free Pricing and Hedging

4.2.1 Pricing Problem

- Given a contingent claim H maturing at time T, determine its **fair value** W_t^H at some earlier time $t < T$.
- The **fair value** is defined as a price that rules out arbitrage opportunities — that is, it is consistent with an arbitrage-free market.

4.2.2 Hedging Problem

- When a seller issues a contingent claim H, they commit to paying a potentially uncertain amount at time T.
- The **hedging problem** asks:
  - How can the seller construct a trading strategy to mitigate or eliminate this risk?
  - What is the minimum price the seller should charge today to ensure they can meet their liability at time T with no arbitrage?

Pricing a Derivative: Two-log-Asset Financial Market

Consider a financial market consisting of two primary assets:

1. **Bank Account** $(B_t)_{t\in[0,T]}$

- Grows at a deterministic interest rate r.
- Its expected return under the physical measure $\mathbb{P}$ is:

$$\mathbb{E}^{\mathbb{P}}[\Delta B_T] = r. \tag{61}$$

2. **Stock Price** $(S_t)_{t\in[0,T]}$

- Represents the risky asset in the market.
- Its expected return under $\mathbb{P}$ includes a risk premium μ:

$$\mathbb{E}^{\mathbb{P}}[\Delta S_T] = r + \mu. \tag{62}$$

Now, consider adding a **derivative** to this market. Its price at maturity T is defined by:

$$\pi_T = h(S_T), \tag{63}$$

where $h(\cdot)$ is a deterministic **payoff function**.

The expected return of the derivative under the physical measure $\mathbb{P}$ is:

$$\mathbb{E}^{\mathbb{P}}[\Delta \pi_T] = r + \lambda, \tag{64}$$

where λ is the **risk premium** associated with the derivative.

This leads to the following price at time $t = 0$:

$$\pi_0 = \mathbb{E}^{\mathbb{P}}[h(S_T)] - r - \lambda. \tag{65}$$

To prevent **arbitrage**, the risk premiums μ (for the stock) and λ (for the derivative) must be consistent. Recall that this is an **equilibrium assumption**: no riskless profit should be achievable in a functioning market.

Theorem (First Fundamental Theorem of Asset Pricing (FTAP))

The market $(B_t, S_t, \pi_t)_{t\in[0,T]}$ is **arbitrage-free** if and only if there exists an **equivalent probability measure** $\mathbb{Q} \sim \mathbb{P}$ such that:

$$S_0 = \mathbb{E}^{\mathbb{Q}}[S_T] - r, \quad \pi_0 = \mathbb{E}^{\mathbb{Q}}[h(S_T)] - r. \tag{66}$$

Consequences

- If the equivalent measure $\mathbb{Q}$ is **unique**, the market is called **complete**.
- In a **complete market**:
  - $\mathbb{Q}$ is called the **risk-neutral measure**.
  - The price π_0 is the **risk-neutral price** of the derivative.

Lemma: No Arbitrage via Local Martingales

If there exists a probability measure $\mathbb{Q} \sim \mathbb{P}$ under which the **discounted asset prices** $\tilde{\mathbf{S}}$ are $\mathbb{Q}$-local martingales, then the market admits **no arbitrage**.

Proof

1. Suppose an arbitrage strategy $\boldsymbol{\theta} = (\theta_t^{(0)}, \dots, \theta_t^{(n)})$ exists.
2. Then the **discounted wealth process** is:

$$\tilde{V}_t = \tilde{V}_0 + \int_0^t \boldsymbol{\theta}_u \cdot d\tilde{\mathbf{S}}_u, \quad \text{with } \tilde{V}_0 = 0. \tag{67}$$

3. Under measure $\mathbb{Q}$, $\tilde{V}_t$ is a local martingale **bounded from below**, hence a **supermartingale**.
4. Therefore:

$$\mathbb{E}^{\mathbb{Q}}[\tilde{V}_T] \leq \tilde{V}_0 = 0. \tag{68}$$

5. But arbitrage implies $\tilde{V}_T \geq 0$ and $\mathbb{P}(\tilde{V}_T > 0) > 0$, so also:

$$\mathbb{Q}(\tilde{V}_T > 0) > 0 \Rightarrow \mathbb{E}^{\mathbb{Q}}[\tilde{V}_T] > 0, \tag{69}$$

which is a contradiction.

4.3 Absence of Arbitrage and Equivalent Measures

Let $\mathbb{P}$ and $\mathbb{Q}$ be two probability measures defined on the measurable space $(\Omega, \mathcal{F})$.

Absolute Continuity

- $\mathbb{P}$ is said to be **absolutely continuous** with respect to $\mathbb{Q}$, written as:

$$\mathbb{P} \ll \mathbb{Q}, \tag{70}$$

if for every event $A \in \mathcal{F}$, the following implication holds:

$$\mathbb{Q}(A) = 0 \quad \Rightarrow \quad \mathbb{P}(A) = 0. \tag{71}$$

Equivalence of Measures

- $\mathbb{P}$ and $\mathbb{Q}$ are said to be **equivalent**, denoted by:

$$\mathbb{P} \sim \mathbb{Q}, \tag{72}$$

if:

$$\mathbb{P} \ll \mathbb{Q} \quad \text{and} \quad \mathbb{Q} \ll \mathbb{P}. \tag{73}$$

This means that $\mathbb{P}$ and $\mathbb{Q}$ assign zero probability to exactly the same events (i.e., they have the same null sets).

Definition (Equivalent Local Martingale Measure (ELMM))

A probability measure $\mathbb{Q}$ is called an **Equivalent Local Martingale Measure (ELMM)** — or a **risk-neutral measure** — if:

- $\mathbb{Q} \sim \mathbb{P}$ (i.e., $\mathbb{Q}$ is equivalent to the physical measure),
- The **discounted asset price process** $\tilde{\mathrm{S}}$ is a **$\mathbb{Q}$-local martingale**.

Definition (Attainable Payoffs)

A contingent claim H is said to be **attainable** (or **replicable**) if there exists:

- A **self-financing**, **admissible** trading strategy $\boldsymbol{\theta} \in \mathcal{A}$,
- A scalar value $v \in \mathbb{R}$ (the initial cost), such that:

(a) The discounted terminal wealth replicates the claim:

$$v + \int_0^T \boldsymbol{\theta}_t \, d\tilde{S}_t = D_{0,T} H, \quad \mathbb{P}\text{-almost surely.} \tag{74}$$

(b) The discounted portfolio value process:

$$\left(\int_0^t \boldsymbol{\theta}_u \, d\tilde{S}_u \right)_{t \geq 0} \tag{75}$$

is a **true $\mathbb{Q}$-martingale** for some $\mathbb{Q} \sim \mathbb{P}$.

4.4 Market Completeness and the Second Fundamental Theorem

4.4.1 Definition: Complete Market

A financial market is said to be **complete** if **every** contingent claim H can be **attained** — i.e., replicated — by a **self-financing admissible portfolio**.

$$\text{Every } H \text{ is attainable} \iff \text{Market is complete.} \tag{76}$$

Theorem (Second Fundamental Theorem of Asset Pricing (FTAP))

A market is **complete** if and only if there exists a **unique** Equivalent Local Martingale Measure (ELMM):

$$\text{Completeness} \iff \text{Uniqueness of the risk-neutral measure } \mathbb{Q}. \tag{77}$$

Theorem (Completeness Criterion)

A financial market is complete if and only if the **volatility matrix** σ admits a **left inverse** $\mathbb{P}$-almost surely:

$$\text{Completeness} \iff \text{The matrix } \sigma_t \text{ has full column rank a.s. for all } t \in [0, T]. \tag{78}$$

4.5 Pricing and Hedging in the Black–Scholes Model

Step 1: Setup and Dynamics under the Physical Measure ($\mathbb{P}$)

Consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, equipped with a standard Brownian motion W_t. Under the physical measure $\mathbb{P}$, we model the underlying asset price S_t by the stochastic differential equation (SDE):

$$dS_t = \mu S_t\, dt + \sigma S_t\, dW_t, \quad S_0 > 0, \tag{79}$$

where:

- μ is the drift of the asset price under the real-world measure.
- $\sigma > 0$ is a constant volatility parameter.
- W_t is a standard Brownian motion under measure $\mathbb{P}$.

By Itô's Lemma, this SDE has a unique strong solution given explicitly by the exponential formula:

$$S_t = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t\right), \quad t \geq 0. \tag{80}$$

Step 2: Introduction of the Bank Account and Discounting

Consider a risk-free **bank account** B_t, which evolves deterministically at a constant interest rate r:

$$B_t = e^{rt}, \quad t \geq 0. \tag{81}$$

We define the **discount factor** between times 0 and t:

$$D_{0,t} = \frac{1}{B_t} = e^{-rt}. \tag{82}$$

The discounted asset price is thus given by:

$$\tilde{S}_t = D_{0,t} S_t = e^{-rt} S_t. \tag{83}$$

Applying Itô's lemma to the discounted price, we have under measure $\mathbb{P}$:

$$d\tilde{S}_t = \tilde{S}_t[(\mu - r)\, dt + \sigma\, dW_t]. \tag{84}$$

Step 3: Risk-Neutral Measure and Girsanov's Theorem

To obtain a no-arbitrage valuation framework, we introduce a change of measure via **Girsanov's theorem**. Define the **market price of risk** as:

$$\gamma := \frac{\mu - r}{\sigma}. \tag{85}$$

Then define a new probability measure $\mathbb{Q} \sim \mathbb{P}$ by the Radon–Nikodym derivative:

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T} = \exp\left(-\frac{1}{2}\gamma^2 T - \gamma W_T\right). \tag{86}$$

Under the new measure $\mathbb{Q}$, the process defined by:

$$W_t^{\mathbb{Q}} := W_t + \gamma t \tag{87}$$

is a standard Brownian motion, and thus the discounted asset price $\tilde{S}_t$ becomes a $\mathbb{Q}$-local martingale satisfying:

$$d\tilde{S}_t = \tilde{S}_t \sigma \, dW_t^{\mathbb{Q}}, \quad \text{or equivalently,} \quad dS_t = S_t(r\,dt + \sigma\,dW_t^{\mathbb{Q}}). \tag{88}$$

This measure $\mathbb{Q}$ is called the **risk-neutral measure**.

Step 4: Martingale Representation and Replication of Payoffs

Consider a European contingent claim with payoff H_T, which is $\mathcal{F}_T$-measurable and integrable under the measure $\mathbb{Q}$. By the **Martingale Representation Theorem**, since H_T is discounted, there exists an adapted square-integrable process φ_t such that:

$$D_{0,T} H_T = \mathbb{E}^{\mathbb{Q}}[D_{0,T} H_T] + \int_0^T \varphi_t \, dW_t^{\mathbb{Q}}. \tag{89}$$

To replicate this payoff via trading, we construct a self-financing portfolio Π_t composed of:

- A quantity θ_t of shares of the underlying asset.
- A quantity θ_t^0 invested in the bank account.

Matching diffusion terms to replicate the payoff, we identify:

- $\theta_t = \frac{\varphi_t}{\sigma D_{0,t} S_t}$,
- $\theta_t^0 = D_{0,t}(\Pi_t - \theta_t S_t)$.

The initial capital required to form this replicating portfolio must be the no-arbitrage price:

$$V_0 = \mathbb{E}^{\mathbb{Q}}[D_{0,T} H_T]. \tag{90}$$

Step 5: Explicit Black–Scholes Formula for Call Options

For a European Call option with maturity T, strike K, and payoff $H_T = (S_T - K)^+$, solving the PDE explicitly yields the well-known **Black–Scholes formula**:

$$C_{BS}(S_t, K, t, T, \sigma) = S_t N(d_+) - K e^{-r(T-t)} N(d_-), \tag{91}$$

with

$$d_\pm = \frac{\ln \frac{S_t}{K} + (r \pm \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, \tag{92}$$

and $N(\cdot)$ denoting the standard normal cumulative distribution function.

Example: European Call Option in Black-Scholes model

Consider the Black-Scholes model with drift and zero interest rate given by:

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + S_t \sigma \mathrm{d}W_t, \tag{93}$$

where $(W_t)_{t\geq 0}$ is a Brownian motion and $H(x) = (x-K)^+$ for some strike $K > 0$. Then we obtain:

$$\pi_0(H, \mathbf{S}) = \mathbb{E}_{\mathbb{Q}}[(S_T - K)^+)]$$
$$= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (S_0 \exp(\sigma\sqrt{T}x - \frac{1}{2}\sigma^2 T) - K)^+ \mathrm{e}^{-x\check{s}/2} \mathrm{d}x. \quad (1)$$

Note: Here we use that we know the distribution of the stock prices in the Black-Scholes model under the risk-neutral measure explicitly! The above must be computed numerically, but it is a deterministic integral, and therefore the numerical tools available are quick!

Step 6: Simplified Parametrization in Terms of Log-Forward Moneyness

Introduce a simplified and common market notation:

- **Log-forward moneyness**: $k = \ln \frac{S_t}{Ke^{-r(T-t)}}$
- **Total variance**: $v = \sigma^2(T - t)$

Then, the Call price becomes:

$$C_{BS}(S_t, K, t, T, \sigma) = S_t BS(k, v), \tag{94}$$

with the standard normalized pricing function $BS(k, v)$:

$$BS(k, v) = N(d_+(k, v)) - e^k N(d_-(k, v)), \tag{95}$$

where

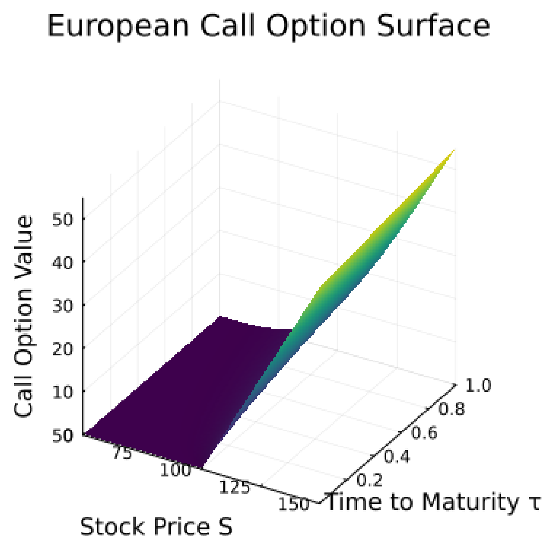
$$d_\pm(k, v) = \frac{-k}{\sqrt{v}} \pm \frac{\sqrt{v}}{2}. \tag{96}$$

This parametrization highlights the explicit dependency on the forward price ratio and total variance.

Below, we plot the European Call Option Surface, i.e. we compute the value of call option prices in the Black-Scholes model for different values of the underlyings stock price and the time-to-maturity.

```
using Plots
using Distributions

# Parameters
K = 100             # Strike price
r = 0.05            # Risk -free rate
 \sigma = 0.2             # Volatility
T = 1.0             # Time to maturity

# Black -Scholes Call Price Function
function call_price(S, \tau; K=K, r=r, \sigma= \sigma)
    if \tau <= 0
        return max(S - K, 0.0)
    end
    d1 = (log(S/K) + (r + 0.5* \sigma^2)* \tau) / ( \sigma*sqrt( \tau))
    d2 = d1 - \sigma*sqrt( \tau)
    N = cdf.(Normal(0, 1), [d1, d2])
    return S*N[1] - K*exp( -r* \tau)*N[2]
end
```

```
# Grid over S and \tau
S_vals = 50:1:150           # Stock prices
 \tau_vals = range(1e -4, T, length=100)  # Avoid \tau = 0 to avoid division by zero

# Compute option prices
Z = [call_price(S, \tau) for S in S_vals, \tau in \tau_vals]

# Plotting
surface(S_vals, \tau_vals, Z', xlabel="Stock Price S", ylabel="Time to Maturity \tau",
        zlabel="Call Option Value", title="European Call Option Surface",
        legend=false, c=:viridis)
```



4.6 The Delta-Hedge

Step 1: the Setup

Suppose we hedge the short option position at time t by:

- Selling the option at price $P(t, S_t)$.
- Buying Δ_t shares of the underlying asset, with Δ_t chosen as:

$$\Delta_t = \frac{\partial P}{\partial S}(t, S_t), \tag{97}$$

to eliminate first-order sensitivity to the underlying price. We reassess and rebalance this position at the next time step $t + \delta t$.

Step 2: Profit and Loss (PnL) Analysis

The total PnL over the short interval δt is composed of two parts:

- **Option's PnL:** due to the change in the option's price and the interest earned on the premium.
- **Hedging instrument's PnL:** due to price changes in the underlying, borrowing costs for shares, and dividends (captured by the repo rate q).

Thus, the total PnL for our short position (short the option, long Δ units of the underlying) is:

$$\text{PnL} = -[P(t+\delta t, S_t+\delta S) - P(t, S_t)] + rP(t, S_t)\,\delta t + \Delta_t(\delta S - rS_t\,\delta t + qS_t\,\delta t), \tag{98}$$

where:

- $\delta S = S_{t+\delta t} - S_t$ is the underlying price increment,
- r is the risk-free interest rate,
- q is the repo rate (including dividends).

Step 3: Taylor Expansion of the Option Price

To analyze this expression clearly, we perform a **Taylor expansion** of the option price around the point (t, S_t) up to first order in δt and second order in δS:

$$P(t+\delta t, S_t+\delta S) = P(t, S_t) + \frac{\partial P}{\partial t}\delta t + \frac{\partial P}{\partial S}\delta S + \frac{1}{2}\frac{\partial^2 P}{\partial S^2}(\delta S)^2 + O\big((\delta t)^2, \delta t\,\delta S, (\delta S)^3\big). \tag{99}$$

Substituting this into our P&L expression yields:

$$\text{PnL} = -\left[\frac{\partial P}{\partial t}\delta t + \frac{\partial P}{\partial S}\delta S + \frac{1}{2}\frac{\partial^2 P}{\partial S^2}(\delta S)^2\right] + rP(t, S_t)\delta t + \Delta_t(\delta S - rS_t\delta t + qS_t\delta t). \tag{100}$$

Step 4: Choosing Δ_t (Delta) to Eliminate First-order Terms

We select the hedge ratio (delta) as previously stated:

$$\Delta_t = \frac{\partial P}{\partial S}(t, S_t), \tag{101}$$

This choice removes the linear term in δS. Thus, after simplification, the PnL expression becomes:

$$\text{PnL} = -\left(\frac{\partial P}{\partial t} - rP(t, S_t) + (r-q)S_t\frac{\partial P}{\partial S}\right)\delta t - \frac{1}{2}\frac{\partial^2 P}{\partial S^2}(\delta S)^2. \tag{102}$$

Step 5: Order-of-Magnitude Considerations

To justify truncation at second order, consider:

- Typically, price increments scale as:

$$\delta S \sim S_t\sigma\sqrt{\delta t}, \tag{103}$$

where σ is volatility. Thus, we have:

- $(\delta S)^2 \sim S_t^2\sigma^2\delta t$, first-order in δt,
- Cross terms (like $\delta S\,\delta t$ or higher orders) scale with higher powers of δt, making them negligible as $\delta t \to 0$.

Thus, our truncation is justified, and the simplified expression is appropriate for daily hedging intervals in practice.

Step 6: Interpreting the Final Simplified Expression

The resulting expression is known as the **carry PnL**:

$$\mathsf{PnL} = -A(t, S_t)\,\delta t - B(t, S_t)\left(\frac{\delta S}{S_t}\right)^2, \tag{104}$$

with:

- $A(t, S_t)$ **(Theta-like term)**: deterministic portion, representing time decay plus financing adjustments:

$$A(t, S_t) = \frac{\partial P}{\partial t} - rP(t, S_t) + (r-q)S_t\frac{\partial P}{\partial S}. \tag{105}$$

- $B(t, S_t)$ **(Gamma term)**: stochastic portion due to curvature in option price:

$$B(t, S_t) = \frac{1}{2}S_t^2\frac{\partial^2 P}{\partial S^2}. \tag{106}$$

Step 7: Economic Meaning and Sign Conditions

- If both A and B have the **same sign** for all t and S_t, we consistently lose (or gain) money regardless of market movements.
  This indicates incorrect model pricing.
- A usable model must have **opposite signs for** A **and** B for all t and S_t to allow both potential gains and losses (no riskless arbitrage).

The breakeven points (where PnL = 0) are given by:

$$\frac{\delta S}{S_t} = \pm\sqrt{-\frac{A(t, S_t)}{B(t, S_t)}\,\delta t}. \tag{107}$$

4.7 From Hedging Arguments to the Black–Scholes Equation

We previously derived conditions to evaluate the usability of a pricing model from a hedging perspective. The condition derived earlier—that the signs of the two main components of the daily PnL, $A(t, S)$ and $B(t, S)$, must differ—is a **necessary condition**. Now we introduce an additional, reasonable requirement based on empirical observations.

Additional Empirical Requirement: Realized Volatility Matching

In practice, daily returns are random, but over time their squares empirically average out to the asset's **realized variance**. Denote by $\hat{\sigma}$ the empirically observed (lognormal) volatility of the underlying asset S, defined such that:

$$\mathbb{E}\left[\left(\frac{\delta S}{S}\right)^2\right] = \hat{\sigma}^2\delta t. \tag{108}$$

A natural **risk management condition** is to require that our hedged portfolio neither consistently gains nor consistently loses money **on average**. Formally, this condition implies:

$$A(t, S) = -\hat{\sigma}^2 B(t, S), \quad \text{for all } t, S. \tag{109}$$

This ensures that, on average, the deterministic part (A) exactly offsets the expected random part (B scaled by volatility squared).

Deriving the Black–Scholes PDE from Hedging Arguments

Substituting the explicit expressions for $A(t, S)$ and $B(t, S)$, we obtain the following important PDE that the option price $P_{\hat{\sigma}}(t, S)$ must satisfy:

$$\frac{\partial P_{\hat{\sigma}}}{\partial t} - rP_{\hat{\sigma}} + (r-q)S\frac{\partial P_{\hat{\sigma}}}{\partial S} = -\frac{1}{2}\hat{\sigma}^2 S^2 \frac{\partial^2 P_{\hat{\sigma}}}{\partial S^2}, \tag{110}$$

where the subscript $\hat{\sigma}$ explicitly indicates dependence of the pricing function on this **break-even volatility level** $\hat{\sigma}$.

Note that the left-hand side is precisely the term $A(t, S)$ we defined earlier. Hence, from our previous P&L equation (1.2), we now have a simplified and intuitive expression:

$$\text{PnL} = -\frac{1}{2}S^2\frac{\partial^2 P_{\hat{\sigma}}}{\partial S^2}\left[\left(\frac{\delta S}{S}\right)^2 - \hat{\sigma}^2\delta t\right]. \tag{111}$$

Interpretation of the Break-even Volatility Condition

The condition for breaking even on average—where the deterministic (theta) and random (gamma) portions exactly offset—is succinctly expressed as a volatility matching condition:

- If the realized squared returns $\left(\frac{\delta S}{S}\right)^2$ exceed the level $\hat{\sigma}^2\delta t$, the position is profitable.
- If they fall below, the position incurs a loss.

Thus, $\hat{\sigma}$ plays the role of a natural hedging volatility. In practice, without a liquid volatility market for the underlying, the chosen $\hat{\sigma}$ should represent the best available estimate of future realized volatility, weighted by the portfolio's gamma.

Implied Volatility and Connection to the Black–Scholes PDE

For actively traded (vanilla) options, we define **implied volatility** as the volatility level $\hat{\sigma}$ that makes the theoretical price $P_{\hat{\sigma}}$ exactly match the observed market price of the option:

- Implied volatility thus emerges naturally from the break-even hedging argument above.

The PDE we derived above is exactly the **Black–Scholes PDE**:

$$\frac{\partial P_{\hat{\sigma}}}{\partial t} + (r-q)S\frac{\partial P_{\hat{\sigma}}}{\partial S} + \frac{1}{2}\hat{\sigma}^2 S^2\frac{\partial^2 P_{\hat{\sigma}}}{\partial S^2} - rP_{\hat{\sigma}} = 0, \tag{112}$$

combined with the terminal condition at maturity:

$$P_{\hat{\sigma}}(T, S) = h(S). \tag{113}$$

Interestingly, a trader employing purely hedging and accounting arguments—without explicit knowledge of Brownian motion or stochastic calculus—would naturally arrive at the Black–Scholes PDE. This derivation rests solely on no-arbitrage and practical hedging considerations, highlighting the intuitive economic basis behind the Black–Scholes framework.

4.7.1 General Criteria for Usable Pricing Models

The Black–Scholes PDE represents a typical example of a market-consistent pricing model used in practice. Such models typically satisfy two key criteria:

1. There is a clearly defined **break-even level** of volatility $\hat{\sigma}$ at which the hedged portfolio's second-order PnL vanishes.
2. This break-even volatility level **does not depend on the specific payoff** of the option under consideration.

This second condition is crucial: if the gamma of an options portfolio becomes zero (making the portfolio locally riskless), its theta must also vanish. If the break-even volatility depended on payoffs, a portfolio with zero gamma but non-zero theta could generate consistent profit or loss, contradicting no-arbitrage principles.

Any pricing model that does not meet these two criteria—especially payoff-independence of the break-even volatility—would not be appropriate for trading or risk management purposes.

4.8 The Greeks: Sensitivities in Option Pricing

In quantitative finance, the **Greeks** are partial derivatives of an option's price with respect to various model parameters. They measure how sensitive the price of an option is to changes in market conditions or model inputs.

These sensitivities are critical for **risk management**, **hedging**, and **trading strategies**. They help traders understand how their portfolio will behave under small changes in key variables like the underlying asset price, volatility, time, or interest rate.

4.9 Common Greeks

Greek	Symbol	Definition	Interpretation
Delta	Δ	$\frac{\partial V}{\partial S}$	Sensitivity to changes in the underlying asset price
Gamma	Γ	$\frac{\partial^2 V}{\partial S^2}$	Sensitivity of delta to changes in the underlying price
Theta	Θ	$\frac{\partial V}{\partial t}$	Time decay — how the option's value decreases over time
Vega	(sometimes ν)	$\frac{\partial V}{\partial \sigma}$	Sensitivity to changes in volatility
Rho	ρ	$\frac{\partial V}{\partial r}$	Sensitivity to changes in the risk-free interest rate

Example (European Call under Black-Scholes)

In the Black-Scholes model for a European call option:

- $\Delta = N(d_1)$
- $\Gamma = \frac{N'(d_1)}{S\sigma\sqrt{T-t}}$
- $\Theta = -\frac{SN'(d_1)\sigma}{2\sqrt{T-t}} - rKe^{-r(T-t)}N(d_2)$
- $\nu = SN'(d_1)\sqrt{T-t}$

- $= K(T-t)e^{-r(T-t)}N(d_2)$

where $N(\cdot)$ is the cumulative distribution function of the standard normal distribution and N' the pdf.

5 Exam Quant Questions Chapter 1:

Financial Markets, Stochastic Processes, Filtrations, and Martingales

Level	Question
Medium	Consider a probability space $(\Omega, \mathcal{F}, P)$. Define precisely what is meant by a **filtration** $\{\mathcal{F}_t\}_{t\geq 0}$. Explain intuitively and formally what it means for a stochastic process $\{X_t\}_{t\geq 0}$ to be **adapted** to this filtration.
Easy	Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ be a filtered probability space. State the mathematical definition of a **martingale** with respect to the filtration $(\mathcal{F}_t)$.
Medium	Formally define a **standard Brownian motion** $\{W_t\}_{t\geq 0}$. List and explain its four key mathematical properties.
Medium	Is a standard Brownian motion W_t a martingale with respect to its natural filtration? Provide a rigorous justification based on the properties of Brownian motion.
Medium	Give an explicit example of a stochastic process adapted to a filtration $\{\mathcal{F}_t\}$, which is not a martingale. Clearly demonstrate, using conditional expectations, why the martingale property fails.
Medium	Consider a geometric Brownian motion given by the SDE: $dX_t = \mu X_t\, dt + \sigma X_t\, dW_t$. Using Itô's lemma, derive the SDE for the stochastic process $Y_t = \ln(X_t)$. Clearly state each step in your derivation.
Hard	Suppose X_t satisfies the following stochastic differential equation (SDE): $dX_t = \alpha(t, X_t)\, dt + \beta(t, X_t)\, dW_t$. Given a twice-differentiable function $f(t, x)$, explicitly state **Itô's formula** for the differential $df(t, X_t)$. Discuss intuitively how Itô's formula extends the classical chain rule from calculus to stochastic processes.
Hard	Let $\{M_t\}_{t\geq 0}$ be a martingale adapted to $\{\mathcal{F}_t\}$, with $M_0 = 0$. Using Itô's isometry, derive an expression for the variance of M_t given its representation as an Itô integral: $M_t = \displaystyle\int_0^t \phi_s\, dW_s$. Explain each step clearly.
Medium	Explain the concept of **quadratic variation** of a stochastic process X_t. Formally define it and calculate the quadratic variation of a Brownian motion W_t. What is the intuitive financial interpretation of quadratic variation?
Medium	Consider the stochastic integral $I_t = \int_0^t W_s\, dW_s$. Compute this integral explicitly using Itô's formula. Provide detailed steps and the final simplified form.

Stochastic Differential Equations and Itô Calculus

Level	Question
Easy	Write down the general form of an **Itô stochastic differential equation (SDE)**: $dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$. Explain the roles of the **drift term** $\mu(t, X_t)$ and the **diffusion term** $\sigma(t, X_t)$. How do they affect the dynamics of the process X_t?
Medium	State **Itô's Lemma** for a function $f(t, X_t)$ where X_t follows the SDE: $dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$. Assume $f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$. Provide the full formula for $df(t, X_t)$.
Medium	Conceptually explain how an **Itô integral** $\int_0^T H_t\,dW_t$ differs from a classical Riemann or Lebesgue integral. Why can't we use standard calculus tools to integrate with respect to Brownian motion? Include a comment on the interpretation of integrands being non-anticipative.
Medium	Consider the **geometric Brownian motion** (GBM) model for an asset price: $dS_t = \mu S_t\,dt + \sigma S_t\,dW_t, \quad S_0 > 0$. Solve this SDE explicitly using Itô's formula applied to $f(S_t) = \ln(S_t)$. Derive a closed-form expression for S_t.
Medium	Consider an Itô process X_t governed by the SDE: $dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$. Derive the **first and second moment approximations** over a small time interval $[t, t + \Delta t]$ using Taylor–Itô expansion.
Medium	Define the concept of an **admissible integrand** in the context of stochastic integration. What are the measurability and integrability conditions required for a process ϕ_t to be integrable with respect to Brownian motion?
Hard	Show that for an Itô process X_t satisfying: $dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$, the **quadratic variation** of X_t is given by: $\langle X \rangle_t = \displaystyle\int_0^t \sigma^2(s, X_s)\,ds$. Explain how this result is used in the context of **realized variance** and **variance swaps** in financial modeling.

Arbitrage, Pricing, Hedging, and Martingale Measures

Level	Question
Easy	What is an **arbitrage opportunity** in a financial market? Formally state the **no-arbitrage condition**, including the requirements for initial capital, terminal payoff, and probability of gain.
Medium	What is a **self-financing portfolio**? Provide the mathematical condition that defines a self-financing trading strategy θ_t in terms of the asset price process S_t. Explain the intuition behind this constraint.
Medium	What does it mean for a market to be **complete**? How does completeness relate to the existence of a **replicating portfolio** for every contingent claim H?
Medium	Define an **equivalent martingale measure** (also called a **risk-neutral measure**) Q. What condition must the **discounted asset price process** $\tilde{S}_t = e^{-rt}S_t$ satisfy under this measure?
Hard	State the **First and Second Fundamental Theorems of Asset Pricing**. How do they relate: 1) the absence of arbitrage to the existence of a risk-neutral measure, and 2) market completeness to the **uniqueness** of this measure?
Hard	Describe how to change the probability measure from the real-world measure P to a risk-neutral measure Q using **Girsanov's theorem**. Specify the form of the **Radon–Nikodym derivative** $\frac{dQ}{dP} = \exp\left(-\int_0^T \theta_s\, dW_s - \frac{1}{2}\int_0^T \theta_s^2\, ds\right)$, where θ_t is the **market price of risk**. What is the impact on the drift term of S_t?
Easy	In the **Black–Scholes model**, under the risk-neutral measure Q, what does the drift of the stock price process S_t become? Justify your answer with the form of the SDE under Q.
Medium	Why do we discount the expected payoff of a derivative at the **risk-free rate** r when using risk-neutral pricing? Explain the connection to the **numéraire** and the **absence of arbitrage**.

The Black–Scholes Model and Greeks

Level	Question
Easy	List the key **assumptions** of the Black–Scholes model. Include aspects such as: frictionless markets, no arbitrage, constant interest rate, constant volatility, continuous trading, and underlying asset following geometric Brownian motion.
Medium	Write down the **SDE** for the stock price S_t in the Black–Scholes model under both the **real-world measure** P and the **risk-neutral measure** Q. Highlight the change in the drift term.
Medium	Derive the **Black–Scholes partial differential equation (PDE)** that the price $V(t,S)$ of a European-style derivative must satisfy: $\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV$. Specify the boundary and terminal conditions.
Hard	State the **Black–Scholes formula** for the price of a European call option: $C(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2)$, where $d_1 = \frac{\ln(S/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$, $d_2 = d_1 - \sigma\sqrt{T-t}$. Define all variables and explain the role of the cumulative distribution function $N(\cdot)$.
Medium	What is the **delta-hedging strategy** in the Black–Scholes model? How is a **delta-neutral position** constructed, and how does this eliminate **local exposure to small changes** in the underlying asset price?
Easy	Define an option's **Delta** (Δ). What does it represent in terms of the sensitivity of the option's price to the underlying asset price?
Easy	What is an option's **Gamma** (Γ)? What does a large Gamma imply about the **stability or instability of Delta** in a hedged portfolio?
Easy	Define **Theta** (Θ), the time decay of an option. What does it mean for an option to have a **highly negative Theta**? How is this relevant for short-dated options?
Easy	What is **Vega** (ν) of an option? With respect to which parameter is the Vega defined, and why is it important for **volatility traders**?
Easy	Define **Rho** (ρ) in the context of option pricing. What does Rho measure, and how does an increase in the risk-free rate affect the value of a European call option?

6 Monte Carlo Methods in Finance

In the second part of our computational finance course, we explore the use of Monte Carlo methods for derivative pricing. Monte Carlo methods are highly regarded in the financial industry for their flexibility (and in some cases their effectiveness) across a broad spectrum of applications. These applications are not limited to derivative pricing but also include areas such as risk and portfolio management. A key advantage of Monte Carlo methods is their ability to efficiently manage high-dimensional settings, such as those involving large portfolios of stocks or derivatives contingent on multiple underlying assets, with relatively few model assumptions. Our emphasis on derivative pricing will lead us to operate within the risk-neutral framework. This approach is distinct from the methods used in risk and portfolio management, which often rely on the physical probability measure to reflect the actual likelihood of events and outcomes. The question which probability measure to use is not a mere theoretic consideration, but it has a profound implication on how we incorporate market data into a pricing model (through calibration) and the distributions we use for sampling (risk-neutral probabilities).

Main Objectives of this Section We aim to solve the following five questions:

1. Understand the stochastic nature of the Monte Carlo Methods

   $\Longrightarrow$ **2.1| Basics of Monte Carlo Methods for Derivative Pricing**

2. Quantify and enhance the speed of convergence of the Monte Carlo Methods

   $\Longrightarrow$ **2.2| Variance Reduction**

3. Simulation of solutions to stochastic differential equations

   $\Longrightarrow$ **2.3| Discretization Schemes for Stochastic Differential Equations**

4. Computing derivatives using Monte Carlo Sampling

   $\Longrightarrow$ **2.4| Sensitivity Analysis**

Q1: How can we compute the risk neutral price $\pi_0(H, \mathbf{S})$, when the exact distribution is unknown?

Assume in the following that $H(\mathbf{S})$ is integrable with respect to $\mathbb{Q}$, i.e. $\pi_0(H, \mathbf{S}) < \infty$. Then by generating a sequence $\mathbf{S}^{(1)}, \mathbf{S}^{(2)}, \ldots$ of independent realizations of $\mathbf{S}$, the law of large numbers assures us that:

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} H(\mathbf{S}^{(i)}) \xrightarrow{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}[H(\mathbf{S})] = \pi_0. \tag{114}$$

This is the essence of the **Monte Carlo Method** for derivatives pricing.

Essential Monte Carlo Method for Option Pricing

1. **Simulate Price Paths**: Generate sample paths $\{(S_t(\omega_i))_{t\in[0,T]} : i = 1, \ldots, N\}$ for the underlying asset's prices over $[0, T]$ following the price model.

2. **Calculate Payoff**: At the end of each sampled path, calculate the option's payoff, e.g. for a plain Vanilla Call option $H(S_T) = (S_T - K, 0)^+$ or an Asian option $H((S_t)_{t\in[0,T]}) = \left(\frac{1}{T}\sum_{i=1}^{M} S_{t_i} - K\right)^+$ for some fixed dates $t_1, \ldots, t_M$.

3. **Average the Payoffs**: Compute the average of all simulated payoffs to estimate the expected payoff:

$$\frac{1}{N}\sum_{i=1}^{N} H(\mathbf{S}(\omega_i)) \tag{115}$$

4. **Discount to Present Value**: Calculate the option's present value by discounting the expected payoff using the risk-free rate:

$$\hat{\pi}_0 = \exp(-rT) \cdot \frac{1}{N} \sum_{i=1}^{N} H(\mathbf{S}(\omega_i)). \tag{116}$$

7 Basics of Monte Carlo Methods for Derivative Pricing

Our objective in this section is to use Monte Carlo methods for computing $\pi_0 = \mathbb{E}_{\mathbb{Q}}[H(\mathbf{S})]$, i.e., the expected value under a risk-neutral probability measure $\mathbb{Q}$ of some pay-off $H(\mathbf{S})$ depending on the underling asset prices $\mathbf{S} \triangleq \mathbf{S}_{t\in[0,T]}$.

7.1 Monte Carlo Integration Error

We define the Monte Carlo integration error ε_N as:

$$\varepsilon_N := \varepsilon_N(H, \mathbf{S}) := \pi_0(H, \mathbf{S}) - \hat{\pi}_0^N(H, \mathbf{S}), \tag{117}$$

where $\hat{\pi}_0^N(H, \mathbf{S})$ denotes the estimate from the first N samples defined as:

$$\hat{\pi}_0^N(H, \mathbf{S}) := \frac{1}{N} \sum_{i=1}^{N} H(\mathbf{S}^{(i)}). \tag{118}$$

Note that $\hat{\pi}_0^N(H, \mathbf{S})$ is an unbiased estimate for $\pi_0(H, \mathbf{S})$, i.e.

$$\mathbb{E}_{\mathbb{Q}}[\hat{\pi}_0^N(H, \mathbf{S})] = \pi_0(H, \mathbf{S}), \quad \text{and thus} \quad \mathbb{E}_{\mathbb{Q}}[\varepsilon_N] = 0. \tag{119}$$

Furthermore, we introduce the **mean square error (MSE)** as measures of error:

$$\mathrm{MSE}(\hat{\pi}_0^N) := \mathbb{E}_{\mathbb{Q}}[\varepsilon_N(H, \mathbf{S})^2], \tag{120}$$

and its square root, the **root mean square error (RMSE)**:

$$\mathrm{RMSE}(\hat{\pi}_0^N) := \sqrt{\mathbb{E}_{\mathbb{Q}}[\varepsilon_N(H, \mathbf{S})^2]}. \tag{121}$$

Proposition: *Let $\sigma = \sigma(H, \mathbf{S}) < \infty$ denote the standard deviation of the random variable $H(\mathbf{S})$. Then the root mean square error satisfies*

$$\mathrm{RMSE}(\hat{\pi}_0^N) = \sqrt{\mathbb{E}_{\mathbb{Q}}[\varepsilon_N(H, \mathbf{S})^2]} = \frac{\sigma}{\sqrt{N}}. \tag{122}$$

Moreover, the random variable $\sqrt{N} \cdot \varepsilon_N$ is asymptotically normally distributed with standard deviation $\sigma(H, \mathbf{S})$, i.e. for every $x_1 < x_2 \in \mathbb{R}$, we have

$$\lim_{N\to\infty} \mathbb{Q}\left(x_1 \frac{\sigma}{\sqrt{N}} < \varepsilon_N < x_2 \frac{\sigma}{\sqrt{N}} \right) = \Phi(x_2) - \Phi(x_1), \tag{123}$$

where Φ denotes the cumulative distribution function of the standard normal random variable.

Proof. Using the independence of the $\mathbf{S}^{(i)}$'s and the fact that $\hat{\pi}_0^N(H, \mathbf{S})$ is an unbiased estimator of $\pi_0(H, \mathbf{S})$, we obtain:

$\mathsf{E}_{\mathbb{Q}}[\varepsilon_N^2] = \mathsf{Var}\left(\frac{1}{N}\sum_{i=1}^N H(\mathbf{S}^{(i)})\right) = \frac{1}{N^2}\sum_{i=1}^N \mathsf{Var}(H(\mathbf{S}^{(i)})) = \frac{\sigma^2}{N}.$(124) In addition, from the central limit theorem, we know that

$1\frac{}{\sigma\sqrt{N}\left(\sum_{i=1}^N H(\mathbf{S}^{(i)})-N\pi_0(H,\mathbf{S})\right)\xrightarrow{d}\mathcal{N}(0,1), \quad \text{as } N\to\infty.}$(125) This yields the asymptotic normality of the error. □

Implications of Proposition 2.1

Proposition 1.1 highlights two crucial aspects of the Monte Carlo simulation error:

1. **Probabilistic Nature of Error:** The simulation error lacks a deterministic bound, meaning that for any given simulation and sample size N, the error magnitude can vary widely. However, the likelihood of encountering large errors diminishes as N increases.
2. **Convergence Rate:** The typical error, such as the root mean square error $\mathrm{MSE}(\hat{\pi}_0^N)$, diminishes at the rate of $\frac{1}{\sqrt{N}}$. To enhance the accuracy tenfold, or to gain an additional significant digit, the sample size N must be increased by a factor of 100. This demonstrates that the Monte Carlo method converges at a rate of the square-root $\frac{1}{2}$.

High-Dimensions and Monte Carlo Integration

For simplicity, consider $\mathbf{S}$ to be a d-dimensional uniform random variable. In that case the fair-price of an European Option would be:

$\pi_0(H,U) = \int_{[0,1]^d} H(x)dx.$(126)

Traditional methods for numerical integration typically rely on a grid defined by $0 \leq x_1 < x_2 < \ldots < x_N \leq 1$, for an arbitrary length N. For a d-dimensional space, this grid expands to $\{x_1,\ldots,x_N\}^d$, encompassing $n = N^d$ points. The integral approximation is derived from evaluating the function f at these grid points and interpolating between them with functions whose integrals can be explicitly calculated.

For a numerical integration method of order k, the error scales as $(1/N)^k$. However, the computational effort scales with n points, making the accuracy, in terms of computational complexity n, proportional to $n^{-k/d}$. Consequently, the effective rate of convergence, considering computational costs, diminishes to k/d with increasing dimension d. This degradation in performance due to rising dimensionality is known as the *curse of dimensionality*.

7.2 Confidence Intervals of the Monte Carlo Method

In Monte Carlo simulations, controlling the error is crucial due to its inherent randomness.

Q3: How large should we choose N to ensure that the probability of the error exceeding a certain tolerance level $\varepsilon > 0$ is below a desired threshold $\delta > 0$?

Mathematically, this is represented as:

$\mathsf{Q}(|\varepsilon_N(H,\mathbf{S})| > \varepsilon) \leq \delta.$(127) This formula sets the stage for understanding how to manage errors within Monte Carlo simulations effectively.

The answer to this question is partially provided by Proposition 1.1, which states:

$$\mathbb{Q}\left(|\varepsilon_N| > \epsilon\right) = 1 - \mathbb{Q}\left(-\frac{\varepsilon}{\sigma\sqrt{N}} \leq \varepsilon_N \leq \frac{\varepsilon}{\sigma\sqrt{N}}\right) \sim 2 - 2\Phi\left(\frac{\sqrt{N}\varepsilon}{\sigma}\right). \tag{128}$$

Note: that the Monte Carlo error is only asymptotically normal. This implies that the equation above accurately represents the error distribution only as N approaches infinity, as indicated by the "$\sim$" symbol.

A3: To align the error probability with a given δ, we solve for N, yielding:

$$N = \big(\Phi^{-1}(1-\frac{\delta}{2})\big)^2\frac{\sigma^2}{\varepsilon^2}. \tag{129}$$

This reveals that the required number of samples, N, scales inversely with the square of the tolerance, ε^2, a principle already noted in earlier discussions. This relationship underscores the increase in sample size needed to refine accuracy within the constraints of Monte Carlo simulations.

Remark: Challenges in Variance Estimation

This analysis has, until now, operated under the assumption that we know $\sigma = \sigma(H, \mathbf{S})$, the variance of the function $H(\mathbf{S})$, which is derived from the mean of $H(\mathbf{S})$. Given that the primary goal of initiating Monte Carlo simulations is to compute the mean of $H(\mathbf{S})$, it is highly improbable that the variance of $H(\mathbf{S})$ would be known a priori. Consequently, in practical applications, $\sigma(H, \mathbf{S})$ must often be substituted with a sample estimate. For guidance on deriving a sample estimator of σ.

It is critical to acknowledge the potential issues this substitution presents, notably regarding the Monte Carlo error associated with approximating $\sigma(H, \mathbf{S})$. The accuracy of Monte Carlo simulations not only depends on the estimator $\hat{\pi}_0^N(H, \mathbf{S})$ but also significantly on our estimation of σ. This introduces an additional layer of uncertainty, as the error in estimating σ directly impacts the confidence we can have in our Monte Carlo results.

Moreover, since the Monte Carlo estimator $\hat{\pi}_0^N(H, \mathbf{S})$ itself behaves as a random variable, merely reporting its value without context is insufficient. The estimator's reliability is inherently linked to the sample size N. Without knowledge of N, the precision of the estimate remains uncertain. Consequently, it is more informative to report not only the estimator's value but also a confidence interval that reflects the sample size and the variability in the estimation process. This approach provides a more comprehensive understanding of the estimator's accuracy and the confidence level of the results obtained from Monte Carlo simulations.

Definition: Let Z be a random variable and consider some level $\alpha \in (0,1)$. The $(1-\alpha)$-level confidence interval is defined by

$\big(-z_{1-\frac{\alpha}{2}}, z_{1-\frac{\alpha}{2}}\big)$ (130) such that the critical number $z_{1-\frac{\alpha}{2}}$ satisfies:

Q$\big(|Z| \le z_{1-\frac{\alpha}{2}}\big) = 1-\alpha$.(131) The critical number $z_{1-\frac{\alpha}{2}}$ for a given level $(1-\alpha)$ can be computed from the inverse cumulative distribution function. Consider, for example, the normal distribution; then we get that

z$_{1-\frac{\alpha}{2}} = \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$.(132) In particular, using the inverse cdf of the normal distribution, we get that for $\alpha = 5\%$ the critical number equals 1.96, while for $\alpha = 1\%$ it equals 2.58.

Now, we can use the asymptotic normality of the Monte Carlo error ε_N to derive confidence intervals for $\hat{\pi}_0^N(H, \mathbf{S})$. Indeed, using Proposition 1.1 and denoting $\varepsilon_N = \pi_0 - \hat{\pi}_0^N$, we have

$$\begin{aligned} 1-\alpha &\approx \mathbb{Q}\left(-\frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{N}} \le \varepsilon_N \le \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{N}}\right) \\ &= \mathbb{Q}\left(\hat{\pi}_0^N - \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{N}} \le \pi_0 \le \hat{\pi}_0^N + \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{N}}\right). \end{aligned}$$

Thus, the $(1-\alpha)$-level confidence interval for π_0 is

CI$_\alpha(\hat{\pi}_0^N) = \left(\hat{\pi}_0^N - \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{N}}, \hat{\pi}_0^N + \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{N}}\right)$.(133)

Example: Monte Carlo Simulation in the Black–Scholes Model

We consider the Black–Scholes model:

$$\begin{cases} dS_t &= rS_t dt + \sigma S_t dW_t, \\ S_0 &= s \in \mathbb{R}^+. \end{cases}$$

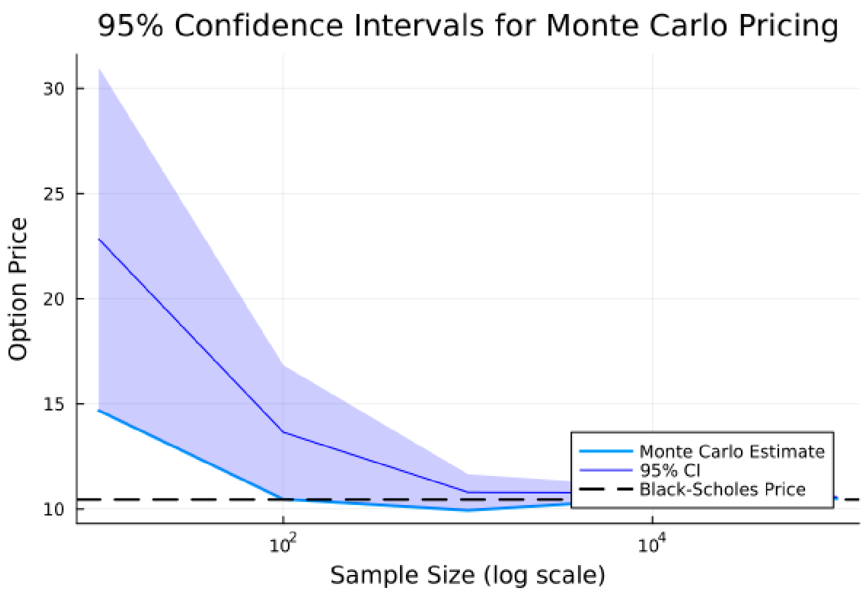
where W is a standard Brownian motion, while we assume we are already under the martingale measure. We want to compute the price of a European call option with payoff function

$$H(S_T) = (S_T - K)^+, \tag{134}$$

together with the 95% and 99% confidence intervals.

```
using Distributions  # For normal distribution
using Plots

# Black -Scholes Formula
function black_scholes(S0, K, T, r, sigma)
    d1 = (log(S0 / K) + (r + sigma^2 / 2) * T) / (sigma * sqrt(T))
    d2 = d1 - sigma * sqrt(T)
    price = S0 * cdf(Normal(0, 1), d1) - K * exp( -r * T) * cdf(Normal(0, 1), d2)
    return price
end

# Monte Carlo pricer with confidence interval
function monte_carlo_bs(S0, K, T, r, sigma, M)
    Z = randn(M)  # Standard normal draws
    ST = S0 * exp.((r - 0.5 * sigma^2) * T .+ sigma * sqrt(T) .* Z)
    payoffs = exp( -r * T) .* max.(ST . - K, 0)
    price = mean(payoffs)
    std_error = std(payoffs) / sqrt(M)
    return price, std_error
end

# Parameters
S0 = 100
K = 100
T = 1
r = 0.05
sigma = 0.2

# Black -Scholes reference
bs_price = black_scholes(S0, K, T, r, sigma)

# Monte Carlo experiments
Ms = [10, 100, 1000, 10000, 100000]
mc_prices = Float64[]
se_values = Float64[]
lower_95 = Float64[]
upper_95 = Float64[]

z_95 = 1.96
```

```
for M in Ms
    price, se = monte_carlo_bs(S0, K, T, r, sigma, M)
    push!(mc_prices, price)
    push!(se_values, se)
    push!(lower_95, price - z_95 * se)
    push!(upper_95, price + z_95 * se)
end

# Plot results with 95% CI ribbons
plot(Ms, mc_prices, xscale = :log10, label = "Monte Carlo Estimate", lw = 2,
     xlabel = "Sample Size (log scale)", ylabel = "Option Price",
     title = "95% Confidence Intervals for Monte Carlo Pricing", legend = :bottomright)
plot!(Ms, upper_95, ribbon = (mc_prices . - lower_95, upper_95 . - mc_prices),
      fillalpha = 0.2, color = :blue, label = "95% CI")
hline!([bs_price], lw = 2, linestyle = :dash, color = :black, label = "Black -Scholes Price")
```



8 Discretization Schemes for SDEs

The simulation of continuous-time financial market models necessitates their discretization into a discrete-time process. We assume that the asset-price is given by an Itô Process of the form:

$$dS_t = \mu(t, S_t)\, dt + \sigma(t, S_t)\, dW_t, \quad S_0 \in \mathbb{R}, \tag{135}$$

where:

- $\mu(t, S_t)$ is the drift coefficient, representing deterministic trends.
- $\sigma(t, S_t)$ is the volatility coefficient, representing the randomness.
- $(W_t)_{t \geq 0}$ is a standard Brownian motion.

In the following we assume a univariate model, where the stock price process $\mathbf{S} = (S_t)_{t\in[0,T]}$ is given by the following SDE:

$\mathrm{dS}_t = \mu(t, S_t)dt + \sigma(t, S_t)dW_t,$(136) where $(W_t)_{t\in[0,T]}$ denotes a Brownian motion, μ is the drift term and σ the instantaneous volatility.

We simulate $(S_t)_{t\in[0,T]}$ over the time interval $[0, T]$, which we assume to be discretized as

$$0 = t_1 < t_2 < \ldots < t_m = T, \tag{137}$$

where the time increments are equally spaced with width Δt.

Remark: Equally-spaced time increments are primarily used for notational convenience, because it allows us to write $t_i - t_{i-1}$ as simply Δt. All the results derived with equally-spaced increments are easily generalized to unequal spacing.

When Do We Need to Discretize?

In derivative pricing, we typically want to evaluate the expected discounted payoff under the risk-neutral measure:

$$\pi_0 = \mathrm{e}^{-rT}\, \mathbb{E}^{\mathbb{Q}} \left[h\left((S_t)_{t\in[0,T]} \right) \right]. \tag{138}$$

The need to **discretize** the stochastic differential equation (SDE) for $(S_t)_{t\in[0,T]}$ typically arises in the following situations:

1. Exotic and Path-Dependent Payoffs

If the payoff h depends on the **entire path** of the asset price — not just the terminal value S_T — then we must approximate/sample the full trajectory of S_t. Examples include:

- **Asian Options** (depend on the average of S_t),
- **Barrier Options** (depend on whether S_t crosses a certain level),
- **Lookback Options**, and many other exotic derivatives.

In these cases, the distribution of $(S_t)_{t\in[0,T]}$ is not known in closed form, making simulation via discretization (e.g. Euler–Maruyama) necessary.

Generation of Sample Paths for Brownian Motions

Generating sample paths from a stochastic process, especially for numerical solutions to SDEs and pricing path-dependent options (like Asian or barrier options), is essential.

We first focus on generating paths for a one-dimensional Brownian motion, B, noting that the techniques apply to multi-dimensional cases as well.

Since we cannot sample the full path $(B_t)_{t\in[0,T]}$ due to its infinite-dimensional nature, we work with a finite-dimensional "skeleton" $(B_{t_1},\ldots,B_{t_n})$ based on a partition $0=t_0<t_1<\ldots<t_n=T$.

Example 3.1: In the Black-Scholes model the stock price process S can be sampled using sampled Brownian motions and transform according to:

$$S_t = S_0\exp\left(\sigma B_t + \left(\mu - \frac{\sigma^2}{2}\right)t\right). \tag{139}$$

8.1 Discretization of Brownian Motion as Random Walk

Input: Time points $t_1, t_2, \ldots, t_n$, where $0=t_0<t_1<t_2<\ldots<t_n=T$ are the discretized time points for sampling Brownian motion.

Output: Sampled values $B_{t_1}, B_{t_2}, \ldots, B_{t_n}$ of Brownian motion at the specified time points.

1. **Initialization:**
   - Set $B_{t_0}=0$ as the starting point of Brownian motion.
   - Calculate time increments $\Delta t_i = t_i - t_{i-1}$ for $i=1,2,\ldots,n$.
2. **For each time increment Δt_i:**
   - Generate an n-dimensional standard normal random variable $X=(X_1,\ldots,X_n)$.
   - Calculate the increment $\Delta B_i = \sqrt{\Delta t_i}X_i$.
3. **Construct Brownian motion path:**
   - For $i=1$ to n: Compute $B_{t_i} = B_{t_{i-1}} + \Delta B_i$.
4. **Return** the sampled values $B_{t_1}, B_{t_2}, \ldots, B_{t_n}$.

8.1.1 Sampling from a Multivariate Normal Distribution using Cholesky Factorization

A method for generating sample paths from the finite-dimensional skeleton $(B_{t_1},\ldots,B_{t_n})$ is based on the following property of Brownian motion:

$$(B_{t_1},\ldots,B_{t_n}) \sim \mathcal{N}(0,\Sigma), \tag{140}$$

with $\Sigma_{i,j} = \min(t_i,t_j)$, for $1\le i,j\le n$.

Given n independent one-dimensional normal random variables $X=(X_1,\ldots,X_n)$, we obtain an n-dimensional normal random vector with covariance matrix Σ by AX, where $\Sigma = AA^T$. In this particular case, it is easy to find the Cholesky factorization A by

$$A = \begin{pmatrix} \sqrt{t_1} & 0 & \ldots & 0 \\ \sqrt{t_2-t_1} & \sqrt{t_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{t_n-t_{n-1}} & \sqrt{t_n-t_{n-1}} & \ldots & \sqrt{t_n} \end{pmatrix}. \tag{141}$$

The matrix A transforms independent standard normal variables into a correlated Brownian motion path with covariance matrix Σ.

9 Discretization Schemes and Convergence

Our goal in this section is to find approximations $\hat{\mathbf{S}}$ to the solution $\mathbf{S}$ of some SDE on a fixed time interval $[0,T]$. These approximations will be based on a time grid $\Delta = \{t_0 = t_0 < t_1 < \ldots < t_N = T\}$ of size N.

Moreover, we denote the **mesh-size** of the grid by

$$|\Delta| := \max_{1\leq i\leq N} |t_i - t_{i-1}|, \tag{142}$$

and we define the increments of time and of any process $\mathbf{S}$ along the grid by $\Delta t_i := t_i - t_{i-1}$ and $\Delta \mathbf{S}_i := S_{t_i} - S_{t_{i-1}}, 1 \leq i \leq N$, respectively.

Moreover, for $t \in [0,T]$, we set

$$\bar{t} = \sup\{t_i | 0 \leq i \leq N, t_i \leq t\}. \tag{143}$$

We will define the discrete approximation $\hat{\mathbf{S}}^{\Delta}$ along the grid as follows:

$$\hat{\mathbf{S}}_i^{\Delta} \triangleq \hat{\mathbf{S}}_{t_i} \quad \text{for } 0 \leq i \leq N. \tag{144}$$

Q: In what sense should $\hat{\mathbf{S}}^{\Delta}$ approximate $\mathbf{S}$?

We introduce two notions to describe the type of approximation of $\hat{\mathbf{S}}^{\Delta}$ to $\mathbf{S}$:

Definition: We say that the scheme $\hat{\mathbf{S}}^{\Delta}$ converges **strongly** to $\mathbf{S}$ if

$$\lim_{|\Delta|\to 0} \mathbb{E}_{\mathbb{Q}}[|\hat{\mathbf{S}}_T^{\Delta} - \mathbf{S}_T|] = 0. \tag{145}$$

Moreover, we say that the scheme $\hat{\mathbf{S}}^{\Delta}$ has strong order γ if (for $|\Delta|$ small enough)

$$\mathbb{E}_{\mathbb{Q}}[|\hat{\mathbf{S}}_T^{\Delta} - \mathbf{S}_T|] \leq C|\Delta|^{\gamma} \tag{146}$$

for some constant $C > 0$, which does not depend on $\gamma > 0$ or $|\Delta| > 0$.

Definition: Given a suitable class $\mathcal{G}$ of functions $f : \mathbb{R}^d \to \mathbb{C}$, we say that the scheme $\hat{\mathbf{S}}^{\Delta}$ converges **weakly** (with respect to $\mathcal{G}$) if for all $f \in \mathcal{G}$ we have:

$\lim_{|\Delta|\to 0} \mathbb{E}_{\mathbb{Q}}[f(\hat{\mathbf{S}}_T^{\Delta})] = \mathbb{E}_{\mathbb{Q}}[f(\mathbf{S}_T)]. \quad (2)$(147)

Moreover, we say that $\hat{\mathbf{S}}^{\Delta}$ has weak order $\gamma > 0$ if for every $f \in \mathcal{G}$ there is a constant $C > 0$ (not depending on $|\Delta|$) such that

$$|\mathbb{E}_{\mathbb{Q}}[f(\hat{\mathbf{S}}_T^{\Delta})] - \mathbb{E}_{\mathbb{Q}}[f(\mathbf{S}_T)]| \leq C|\Delta|^{\gamma} \tag{148}$$

provided that $|\Delta|$ is small enough.

Remark:

1. If your main objective is the pricing of derivatives, as we placed ourself in, then weak convergence is usually enough for our purpose. Indeed, you might think of the class $\mathcal{G}$ appearing in the definition of the weak convergence as a class of pay-off functions. Then Equation (2) tells us that in order to compute $\pi_0(H,\mathbf{S}) = \mathbb{E}_{\mathbb{Q}}[H(\mathbf{S}_T)]$ we can first approximate $\mathbb{E}_{\mathbb{Q}}[H(\mathbf{S}_T)]$ by $\mathbb{E}_{\mathbb{Q}}[H(\hat{\mathbf{S}}_T^{\Delta})]$ for small Δ, and then apply the Monte Carlo Method to approximate $\mathbb{E}_{\mathbb{Q}}[H(\hat{\mathbf{S}}_T^{\Delta})]$ by:

$$\frac{1}{N}\sum_{i=1}^{N} H(\hat{\mathbf{S}}_T^{\Delta,(i)}), \tag{149}$$

where $\mathbf{S}_T^{\Delta,(1)}, \ldots, \mathbf{S}_T^{\Delta,(N)}$ are independent samples of $\mathbf{S}_T^{\Delta}$.

2. Even if the pay-off function H is not in the class $\mathcal{G}$, then weak convergence is often still sufficient for pricing purposes. Indeed, if the class $\mathcal{G}$ consists out of the Fourier basis elements, i.e. $\mathcal{G} = \{x \mapsto \mathrm{e}^{\mathrm{i}ux} \colon u \in \mathbb{R}\}$ then weak convergence means convergence of the characteristic functions and therefore we are talking about convergence in distribution. Since the price π_0 is given as the (discounted) risk-neutral expectation of the pay-off $H(\mathbf{S})$, and the expectation is only dependend on the distribution, this shows that weak convergence is enough for simple pricing purpose.

3. Note that for path-dependent options, such as Asian options, lookback options, and barrier options, the payoff depends not just on the final asset price but also on the path taken by the asset price over the life of the option. In these cases, strong convergence may be necessary to accurately capture the dynamics of the path-dependent features!

9.1 The Euler Scheme

Integrating $dS_t = \mu(t, S_t)dt + \sigma(t, S_t)dW_t$ from t to $t + \Delta t$ yields the following integral equation:

$$S_{t+\Delta t} = S_t + \int_t^{t+\Delta t} \mu(u, S_u)du + \int_t^{t+\Delta t} \sigma(u, S_u)dW_u, \tag{150}$$

which will be our starting point for any discretization scheme. We assume that at time t the value of S_t is known, and we wish to obtain the next value $S_{t+\Delta t}$!

The simplest way to discretize the process in Equation (2) is to use Euler discretization. This is equivalent to approximating the integrals using the **left-point rule**.

Hence the first integral is approximated as the product of the integrand at time t, and the integration range Δt

$\int_t^{t+\Delta t} \mu(S_u, u)du \approx \mu(S_t, t) \int_t^{t+\Delta t} du = \mu(S_t, t)\Delta t.$(151)

We use the left-point rule since at time t the value $\mu(S_t, t)$ is known.

The second integral is approximated as

$\int_t^{t+\Delta t} \sigma(S_u, u)dW_u \approx \sigma(S_t, t)(W_{t+\Delta t} - W_t) = \sigma(S_t, t)\sqrt{\Delta t}Z,$(152) where Z is a standard normal variable.

Note that $W_{t+\Delta t} - W_t$ and $\sqrt{\Delta t}Z$ are identical in distribution.

We therefore define the **Euler discretization** of the SDE takes as follows:

$$\hat{S}_{t+\Delta t} = \hat{S}_t + \mu(\hat{S}_t, t)\Delta t + \sigma(\hat{S}_t, t)\sqrt{\Delta t}Z. \quad (1) \tag{153}$$

9.1.1 Euler Scheme for the Black-Scholes Model

The Black-Scholes stock price dynamics under the risk neutral measure are

dS$_t = rS_tdt + \sigma S_t dW_t$.(154) Following (1) the Euler discretization of the Black-Scholes model is thus given by:

$$\hat{S}_{t+\Delta t} = \hat{S}_t + r\hat{S}_t\Delta t + \sigma\hat{S}_t\sqrt{\Delta t}Z. \tag{155}$$

Remark:

It can be shown that the Euler-Maruyama scheme has a strong convergence rate of $\mathcal{O}(\Delta t)$ in the time step size Δt. This implies that the error of the numerical solution decreases as the square root of the time step size. In the following sections, we explore opportunities for schemes to perform better than that.

As an alternative, we can generate log-stock prices, and exponentiate the result.

Indeed, using Itô's lemma (see Part I of this course) $\ln S_t$ is given by:

$$d \ln S_t = \left(r - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t. \tag{156}$$

Euler discretization via Equation (1) yields:

$$\ln \hat{S}_{t+\Delta t} = \ln \hat{S}_t + \left(r - \frac{1}{2}\sigma^2\right) \Delta t + \sigma \sqrt{\Delta t} Z$$

so that

$$\hat{S}_{t+\Delta t} = \hat{S}_t \exp\left(\left(r - \frac{1}{2}\sigma^2\right) \Delta t + \sigma \sqrt{\Delta t} Z\right).$$

```
using Plots
using Distributions

function simulate_euler_black_scholes(S0, mu, sigma, T, N)
    dt = T/N # Time step
    time = 0:dt:T # Time grid
    S = zeros(length(time))
    S[1] = S0

    for t in 2:length(time)
        dW = rand(Normal(0, sqrt(dt))) # Brownian motion increment
        S[t] = S[t -1] + mu*S[t -1]*dt + sigma*S[t -1]*dW # Euler update
    end

    return time, S
end

# Parameters
S0 = 100 # Initial asset price
mu = 0.05 # Drift
sigma = 0.2 # Volatility
T = 1.0 # Time to maturity
N = 1000 # Number of steps

time, S = simulate_euler_black_scholes(S0, mu, sigma, T, N)
```

```
(0.0:0.001:1.0, [100.0, 100.88435342656506, 100.77729425035623, 101.19462107595083, 101.84650642029837,
```

9.1.2 Euler Scheme for the Heston Model

Recall that the Heston model is described by the bivariate stochastic process for the stock price S_t and its variance V_t as follows:

$$
\begin{aligned}
dS_t &= rS_t\Delta t + \sqrt{V_t}S_t dW_t^{(1)} \\
dV_t &= \kappa(\theta - V_t)\Delta t + \xi\sqrt{V_t}dW_t^{(2)}
\end{aligned}
$$

where $\mathbb{E}_{\mathbb{Q}}[dW_t^{(1)}dW_t^{(2)}] = \rho\Delta t$ for some correlation parameter $\rho \in [-1, +1]$.

Discretization of the variance process V

In contrast to the Black-Scholes model, we also have to approximate the variance process in the Heston model: The SDE for $(V_t)_{t\in[0,T]}$ reads in integral form as

$$
V_{t+\Delta t} = V_t + \int_t^{t+\Delta t} \kappa(\theta - V_u)du + \int_t^{t+\Delta t} \xi\sqrt{V_u}dW_u^{(2)}.
$$

The Euler discretization approximates the integrals using the left-point rule

$$
\begin{aligned}
&\int_t^{t+\Delta t} \kappa(\theta - V_u)du \approx \kappa(\theta - V_t)\Delta t \\
&\int_t^{t+\Delta t} \xi\sqrt{V_u}dW_u^{(2)} \approx \xi\sqrt{V_t}(W_{t+\Delta t} - W_t) = \xi\sqrt{V_t\Delta t}Z_V
\end{aligned}
$$

where Z_V is a standard normal random variable.

Truncating negative variance

The Euler discretization of the variance is therefore given by:

$$
\hat{V}_{t+\Delta t} = \hat{V}_t + \kappa(\theta - \hat{V}_t)\Delta t + \xi\sqrt{\hat{V}_t\Delta t}Z_V.
$$

Note: We can not rule out that $\hat{V}_t$ assumes negative values, which invalids the square-root computation. This motivates the use of truncation schemes. In the **full truncation scheme** we replace the latter $\hat{V}_t$ by $\hat{V}_t^+ \triangleq \max(0, \hat{V}_t)$, whereas in the **reflection scheme** one replaces the latter twp $\hat{V}_t$ by its absolute value $|\hat{V}_t|$.

Discretization of the asset's price process S

The SDE for S_t is written in integral form as

$$
S_{t+\Delta t} = S_t + r\int_t^{t+\Delta t} S_u du + \int_t^{t+\Delta t} \sqrt{V_u}S_u dW_u.
$$

Euler discretization approximates the integrals with the left-point rule

$$
\begin{aligned}
\int_t^{t+\Delta t} S_u du &\approx S_t\Delta t \\
\int_t^{t+\Delta t} \sqrt{V_u}S_u dW_{1,u} &\approx \sqrt{V_t}S_t(W_{t+\Delta t} - W_t) = \sqrt{V_t\Delta t}S_t Z_S
\end{aligned}
$$

where Z_S is a standard normal random variable that has correlation ρ with Z_V. We therefore end up with:

$$
\hat{S}_{t+\Delta t} = \hat{S}_t + r\hat{S}_t\Delta t + \sqrt{V_t\Delta t}\hat{S}_t Z_S.
$$

Discretization of the asset's log-price process $\ln S$

By Itô's lemma $\ln S_t$ follows the diffusion

$$
d\ln S_t = \left(r - \frac{1}{2}V_t\right)\Delta t + \sqrt{V_t}dW_t^{(1)}
$$

or in integral form

$$\ln S_{t+\Delta t} = \ln S_t + \int_0^t \left(r - \frac{1}{2}V_u\right) du + \int_0^t \sqrt{V_u} dW_u^{(1)}.$$

Euler discretization of the process for $\ln S_t$ is thus

$$\ln S_{t+\Delta t} = \ln S_t + \left(r - \frac{1}{2}V_t\right)\Delta t + \sqrt{V_t}\Delta t Z_S.$$

Hence the Euler discretization of S_t is

$$\hat{S}_{t+\Delta t} = \hat{S}_t \exp\left(\left(r - \frac{1}{2}\hat{V}_t\right)\Delta t + \sqrt{\hat{V}_t \Delta t} Z_S\right),$$

where to avoid negative variances we must apply the full truncation or reflection scheme by replacing V_t everywhere by V_t^+ or $|V_t|$.

Algorithm: Euler Scheme for (S, V) or $(\ln S, V)$ with full truncation

1. **Initialization**: Start with the initial values S_0 for the stock price and V_0 for the variance; specify all parameters.
2. **Iteration**: For each time step from t to $t + \Delta t$, perform the following steps:

   a. **Update Variance**: Compute $\hat{V}_{t+\Delta t}$ using the formula:

   $_{t+\Delta t} = \hat{V}_t + \kappa(\theta - \hat{V}_t^+)\Delta t + \xi\sqrt{\hat{V}_t^+ \Delta t} Z_V,$ (157)

   b. **Update Stock Price**: Obtain $S_{t+\Delta t}$ using either of the following formulas:

   For the arithmetic model:

   $_{t+\Delta t} = \hat{S}_t + r\hat{S}_t\Delta t + \sqrt{\hat{V}_t^+ \Delta t}\hat{S}_t Z_S,$ (158)

   or for the geometric model:

   $_{t+\Delta t} = \hat{S}_t \exp\left(\left(r - \frac{1}{2}\hat{V}_t^+\right)\Delta t + \sqrt{\hat{V}_t^+ \Delta t} Z_S\right).$ (159)

3. **Generate Correlated Random Variables**:
   - To generate Z_V and Z_S with correlation ρ, first generate two independent standard normal variables Z_1 and Z_2.
   - Then set $Z_V = Z_1$ and $Z_s = \rho Z_1 + \sqrt{1 - \rho^2} Z_2$.

9.2 Milstein Scheme

The Milstein scheme works for time-homogeneous SDEs, i.e. SDEs with coefficients $\mu_t = \mu(S_t)$ and $\sigma_t = \sigma(S_t)$ depend only on S_t and not explicitly on t.

Therefore, we assume in the following that the stock price S is driven by the SDE:

$$\begin{aligned} dS_t &= \mu(S_t)dt + \sigma(S_t)dW_t \\ &= \mu_t dt + \sigma_t dW_t, \end{aligned}$$

which in integral form is equivalent to:

$$S_{t+\Delta t} = S_t + \int_t^{t+\Delta t} \mu_s ds + \int_t^{t+\Delta t} \sigma_s dW_s \tag{160}$$

The key idea of the Milstein scheme is based on the idea to enhace the accuracy of the discretization by expanding the coefficients $\mu_t = \mu(S_t)$ and $\sigma_t = \sigma(S_t)$ via Itô's lemma.

Assume that μ and σ are sufficiently regular. Then by Itô's formula, we have:

$$d\mu(S_t) = \left(\mu'(S_t)\mu(S_t) + \frac{1}{2}\mu''(S_t)\sigma(S_t)^2\right) dt + (\mu'(S_t)\sigma(S_t))dW_t \tag{161}$$

$$d\sigma(S_t) = \left(\sigma'(S_t)\mu(S_t) + \frac{1}{2}\sigma''(S_t)\sigma(S_t)^2\right) dt + (\sigma'(S_t)\sigma(S_t))dW_t \tag{162}$$

where the prime refers to differentiation in S and where the derivatives in t are zero, as we assumed that μ_t and σ_t have no direct dependence on t!

The integral form of the coefficients at time s (with $t < s < t + dt$)

$$\begin{aligned} \mu(S_s) = \mu(S_t) + \int_t^s \left(\mu'(S_u)\mu(S_u) + \frac{1}{2}\mu''(S_u)\sigma(S_u)^2\right) du \\ + \int_t^s (\mu'(S_u)\sigma(S_u))dW_u \end{aligned}$$

$$\begin{aligned} \sigma(S_s) = \sigma(S_t) + \int_t^s \left(\sigma'(S_u)\mu(S_u) + \frac{1}{2}\sigma''(S_u)\sigma(S_u)^2\right) du \\ + \int_t^s (\sigma'(S_u)\sigma(S_u))dW_u \end{aligned}$$

Substitute for μ_s and σ_s to obtain:

$$\begin{aligned} S_{t+\Delta t} = S_t + \int_t^{t+\Delta t} \Big(\mu(S_s) + \int_t^s (\mu'(S_u)\mu(S_u) \\ + \frac{1}{2}\mu''(S_u)\sigma(S_u)^2)du + \int_t^s (\mu'(S_u)\sigma(S_u))dW_u\Big) ds \\ + \int_t^{t+\Delta t} (\sigma(S_s) + \int_t^s \left(\sigma'(S_u)\mu(S_u) + \frac{1}{2}\sigma''(S_u)\sigma(S_u)^2\right) du \\ + \int_t^s (\sigma'(S_u)\sigma(S_u))dW_u\Big)dW_s. \end{aligned}$$

The terms higher than order one are $dsdu = \mathcal{O}((dt)^2)$ and $dsdW_u = \mathcal{O}((dt)^{3/2})$ and will be ignored.

The term involving $dW_u dW_s$ is retained since $dW_u dW_s = \mathcal{O}(dt)$ is of order one. This leaves us with:

$$\begin{aligned} S_{t+\Delta t} \approx S_t + \int_t^{t+\Delta t} \mu(S_s)ds + \int_t^{t+\Delta t} \sigma(S_s)dW_s \\ + \int_t^{t+\Delta t} \int_t^s (\sigma'(S_u)\sigma(S_u))dW_u dW_s. \end{aligned}$$

Apply Euler discretization to the last term to obtain

$$\begin{aligned}\int_t^{t+\Delta t}\int_t^s \sigma'(S_u)\sigma(S_u)dW_u dW_s &\approx \sigma_t'\sigma(S_t)\int_t^{t+\Delta t}\int_t^s dW_u dW_s\\ &= \sigma'(S_t)\sigma(S_t)\int_t^{t+\Delta t}(W_s - W_t)dW_s\\ &= \sigma'(S_t)\sigma(S_t)\int_t^{t+\Delta t} W_s dW_s - W_t W_{t+\Delta t} + W_t^2.\end{aligned}$$

Now define $dY_t \triangleq W_t dW_t$. Using Itô's Lemma, it is easy to show that Y_t has solution $Y_t = \frac{1}{2}(W_t^2 - t)$ so that

$$\int_t^{t+\Delta t} W_s dW_s = Y_{t+\Delta t} - Y_t = \frac{1}{2}W_{t+\Delta t}^2 - \frac{1}{2}W_t^2 - \frac{1}{2}\Delta t. \tag{163}$$

Substitute this back to obtain

$$\int_t^{t+\Delta t}\int_t^s \sigma'(S_u)\sigma_u dW_u dW_s \approx \frac{1}{2}\sigma'(S_u)\sigma(S_u)\left((\Delta W_t)^2 - \Delta t\right) \tag{164}$$

where $\Delta W_t = W_{t+\Delta t} - W_t$, which is equal in distribution to $\sqrt{\Delta t}Z$ with Z distributed as standard normal. Combining the equations above, the general form of Milstein discretization is therefore:

$$\hat{S}_{t+\Delta t} = \hat{S}_t + \mu(\hat{S}_t)\Delta t + \sigma(\hat{S}_t)\sqrt{\Delta t}Z + \frac{1}{2}\sigma'(\hat{S}_t)\sigma(\hat{S}_t)\Delta t(Z^2 - 1). \tag{165}$$

Remark: Note here that the Milstein scheme includes a correction for the change in the diffusion term, i.e., the stochastic integral term in the SDE. This is particularly suitable for equations where the volatility coefficients depend on the state variable or a stochastic volatility. It can be shown that the Milstein scheme has a strong convergence rate of $\mathcal{O}(\Delta t)$ for SDEs with Lipschitz continuous coefficients. This represents a significant improvement over the Euler-Maruyama scheme which has $\mathcal{O}(\sqrt{\Delta t})$, under the condition that the diffusion coefficient's derivative with respect to the state variable is also Lipschitz continuous.

9.3 Milstein Scheme for the Black-Scholes Model

In the Black-Scholes model we have $\mu(S_t) = rS_t$ and $\sigma(S_t) = \sigma S_t$, so the Milstein scheme becomes:

$$\hat{S}_{t+\Delta t} = \hat{S}_t + r\hat{S}_t\Delta t + \sigma\hat{S}_t\sqrt{\Delta t}Z + \frac{1}{2}\sigma^2\hat{S}_t\Delta t(Z^2 - 1) \tag{166}$$

which adds the correction term $\frac{1}{2}\sigma^2\hat{S}_t\Delta t(Z^2 - 1)$ to the Euler scheme from before.

In the Black-Scholes model for the log-stock price, we have $\mu(S_t) = r - \frac{1}{2}\sigma^2$ and $\sigma(S_t) = \sigma$

so that $\mu'(S_t) = \sigma'(S_t) = 0$. The Milstein scheme is therefore

$\ln\hat{S}_{t+\Delta t} = \ln\hat{S}_t + \left(r - \frac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t}Z$(167) which is identical to the Euler scheme from before. Hence, while the Milstein scheme improves the discretization of S_t in the Black-Scholes model, it does not improve the discretization of $\ln S_t$!

9.4 Milstein Scheme for the Heston Model

Recall the Heston Model from before, given by:

$$dS_t = rS_t\Delta t + \sqrt{V_t}S_t\Delta W_t^{(1)}$$
$$dV_t = \kappa(\theta - V_t)\Delta t + \xi\sqrt{V_t}\Delta W_t^{(2)}$$

Process for V

The coefficients of the variance process are $\mu(V_t) = \kappa(\theta - V_t)$ and $\sigma(V_t) = \xi\sqrt{V_t}$ so an application of the Milstein Scheme for V_t produces

$_{t+\Delta t} = \hat{V}_t + \kappa(\theta - \hat{V}_t)\Delta t + \xi\sqrt{\hat{V}_t\Delta t}Z_V + \frac{1}{4}\xi^2\Delta t(Z_V^2 - 1)$(168) which can be written

$_{t+\Delta t} = \left(\sqrt{\hat{V}_t} + \frac{1}{2}\xi\sqrt{\Delta t}Z_V\right)^2 + \kappa(\theta - \hat{V}_t)\Delta t - \frac{1}{4}\xi^2\Delta t$(169) The Milstein discretization of the variance process produces far fewer negative values for the variance than Euler discretization. Nevertheless, the full truncation scheme or the reflection scheme must be applied to this scheme as well!

Process for S and $\ln S$

The coefficients of the stock price process are $\mu(S_t) = rS_t$ and $\sigma(S_t) = \sqrt{V_t}S_t$ so we obtain:

$_{t+\Delta t} = \hat{S}_t + r\hat{S}_t\Delta t + \sqrt{\hat{V}_t\Delta t}\hat{S}_tZ_S + \frac{1}{2}\hat{V}_t\hat{S}_t\Delta t(Z_S^2 - 1)$(170) We can also discretize the log-stock process, which by Itô's lemma follows the process

d $\ln\hat{S}_t = \left(r - \frac{1}{2}\hat{V}_t\right)\Delta t + \sqrt{\hat{V}_t}\Delta W_t^{(1)}$(171)

Since V_t is known at time t, we can treat it as a constant in the coefficients. An application of the Milstein Scheme yields:

$\ln\hat{S}_{t+\Delta t} = \ln\hat{S}_t + \left(r - \frac{1}{2}\hat{V}_t\right)\Delta t + \sqrt{\hat{V}_t\Delta t}Z_S$(172) which is identical to what we had before. Hence, as in the Black-Scholes model, the discretization of $\ln S_t$ rather than S_t means that there are no higher corrections to be brought to the Euler discretization. The discretization of the stock price is

$$\hat{S}_{t+dt} = \hat{S}_t\exp\left(\left(r - \frac{1}{2}\hat{V}_t\right)dt + \sqrt{\hat{V}_t\Delta t}Z_S\right). \tag{173}$$

Again, it is necessary to apply the full truncation or reflections schemes here!

Algorithm: Milstein Scheme for (S, V) or $(\ln S, V)$ with full truncation

Given a value for $\hat{V}_t$ at time t, the process to update to $\hat{V}_{t+dt}$ and obtain $\hat{S}_{t+dt}$ is as follows:

1. **Update Variance**: Update to $V_{t+\Delta t}$ using the equation:

   $_{t+\Delta t} = \hat{V}_t + \kappa(\theta - \hat{V}_t^+)\Delta t + \xi\sqrt{\hat{V}_t^+\Delta t}Z_V + \frac{1}{4}\xi^2\Delta t(Z_V^2 - 1)$(174)

   where κ is the rate of mean reversion, θ is the long-term variance, ξ is the volatility of volatility, and Z_V is a random variable from the correlated standard normal distribution.

2. **Update Stock Price**: Obtain $\hat{S}_{t+\Delta t}$ using either of the following formulas:
    - For the arithmetic model:

      ${}_{t+\Delta t} = \hat{S}_t + r\hat{S}_t\Delta t + \sqrt{\hat{V}_t^+ \Delta t}\hat{S}_t Z_S + \frac{1}{2}\hat{V}_t^+ \hat{S}_t \Delta t (Z_S^2 - 1)$(175) where r is the risk-free rate.
    - For the geometric model:

      ${}_{t+\Delta t} = \hat{S}_t \exp\left(\left(r - \frac{1}{2}\hat{V}_t^+\right)\Delta t + \sqrt{\hat{V}_t^+ \Delta t} Z_S\right)$.(176)
3. **Generate Correlated Random Variables**:
    - To generate Z_V and Z_S with correlation ρ, first generate two independent standard normal variables Z_1 and Z_2.
    - Set $Z_V = Z_1$ and $Z_S = \rho Z_1 + \sqrt{1-\rho^2} Z_2$.

9.4.1 Comparison between the Euler and Milstein Schemes

The Euler and Milstein schemes are both used for numerically solving stochastic differential equations. We already saw that the Milstein scheme can improve the approximation in some cases (e.g. geometric Brownian motion model). In other settings, however, it coincides with the simpler Euler scheme (e.g. in the arithmetic Brownian motion model) and does not lead to an improvement. In general, the following can be said about the comparison:

Strong Convergence:

Under certain regularity conditions, we have that:

- **Euler Scheme**: Exhibits a strong convergence rate of $\mathcal{O}(\sqrt{\Delta t})$;
- **Milstein Scheme**: Exhibits an improved strong convergence rate of $\mathcal{O}(\Delta t)$.

Weak Convergence:

However, both the Euler and Milstein schemes demonstrate a weak convergence rate of $\mathcal{O}(\Delta t)$!

Conclusion: Since the improvement of the Milstein versus the Euler scheme is only visible in the strong convergence rate, the Milstein scheme is most beneficial in pricing path dependent options!

10 Variance Reduction

The Idea Behind Variance Reduction

Recall tat in Monte Carlo pricing, the **root mean square error (RMSE)** of estimator $\hat{\pi}_0^N$ for the option price is given by:

$$\mathrm{RMSE}(\hat{\pi}_0^N) = \sqrt{\mathbb{E}_{\mathbb{Q}}[\varepsilon_N(H, \mathbf{S})^2]} = \frac{\sigma(H, \mathbf{S})}{\sqrt{N}}, \tag{177}$$

where:

- $H(\mathbf{S})$ is the discounted payoff, possibly depending on the full path of the process,
- $\sigma(H, \mathbf{S}) = \sqrt{\mathrm{Var}(H(\mathbf{S}))}$ is the standard deviation of the estimator,
- N is the number of Monte Carlo samples.

We cannot (in general) reduce the order $N^{-1/2}$ of convergence, but we **can aim to reduce the variance** $\sigma(H, \mathbf{S})^2 = \mathrm{Var}(H(\mathbf{S}))$.

The idea of **variance reduction** is to find — in a systematic way — a new random variable Y and a function g such that:

$$\mathbb{E}_{\mathbb{Q}}[g(Y)] = \mathbb{E}_{\mathbb{Q}}[H(\mathbf{S})], \quad \text{but} \quad \mathrm{Var}(g(Y)) < \mathrm{Var}(H(\mathbf{S})). \tag{178}$$

This ensures a more efficient estimator:

RMSE($^N_{\text{reduced}}$) $= \frac{\sigma(g(Y))}{\sqrt{N}} < \frac{\sigma(H(\mathbf{S}))}{\sqrt{N}}$.(179)

10.1 Antithetic Variates

Antithetic variates is a variance reduction technique that leverages symmetry in distributions to improve Monte Carlo estimators.

Consider a random variable U uniformly distributed on $[0, 1]$. Notice that if $U \sim$ Uniform$(0, 1)$, then also $1 - U \sim$ Uniform$(0, 1)$. Similarly, if B is a d-dimensional standard normal vector, then its negative $-B$ has the same distribution. Such variables $(1 - U)$ or $-B$ are called **antithetic variates** of U or B, respectively.

Unbiasedness:
Since the pairs $\mathbf{S}$ and $\tilde{\mathbf{S}}$ have the same distribution under $\mathbb{Q}$, we have:

$$\mathbb{E}_{\mathbb{Q}}\left[\frac{H(\mathbf{S}) + H(\tilde{\mathbf{S}})}{2}\right] = \mathbb{E}_{\mathbb{Q}}[H(\mathbf{S})]. \tag{180}$$

Variance Reduction:
Antithetic pairing introduces negative correlation between samples, typically resulting in a reduced variance:

$$\mathrm{Var}\left(\frac{H(\mathbf{S}) + H(\tilde{\mathbf{S}})}{2}\right) < \mathrm{Var}(H(\mathbf{S})), \tag{181}$$

provided the payoff H is suitably monotone or symmetrical.

The use of **antithetic variates** improves a Monte Carlo estimator only if the **mean squared error (MSE)** of the antithetic estimator

$$\hat{\pi}_0^{A,N}(H,\mathbf{S}) := \frac{1}{N}\sum_{i=1}^{N}\frac{H(\mathbf{S}^{(i)}) + H(\tilde{\mathbf{S}}^{(i)})}{2} \tag{182}$$

is smaller than the MSE of a standard Monte Carlo estimator using **twice as many independent samples**, i.e., $\hat{\pi}_0^{2N}(H,\mathbf{S})$.

This holds if:

$1\frac{}{N\,\mathrm{Var}\left(\frac{H(\mathbf{S})+H(\tilde{\mathbf{S}})}{2}\right) < \frac{1}{2N}\,\mathrm{Var}(H(\mathbf{S})).(183)}$

By expanding the variance on the left, we get:

$\mathrm{Var}\left(\frac{H(\mathbf{S})+H(\tilde{\mathbf{S}})}{2}\right) = \frac{1}{4}\left[\mathrm{Var}(H(\mathbf{S})) + \mathrm{Var}(H(\tilde{\mathbf{S}})) + 2\,\mathrm{Cov}(H(\mathbf{S}), H(\tilde{\mathbf{S}}))\right].(184)$

Since $\mathbf{S}$ and $\tilde{\mathbf{S}}$ are identically distributed, the variances are equal, and we can simplify:

$\mathrm{Var}\left(\frac{H(\mathbf{S})+H(\tilde{\mathbf{S}})}{2}\right) = \frac{1}{2}\,\mathrm{Var}(H(\mathbf{S})) + \frac{1}{2}\,\mathrm{Cov}(H(\mathbf{S}), H(\tilde{\mathbf{S}})).(185)$

Thus, the antithetic estimator is more efficient **if and only if**

$\mathrm{Cov}(H(\mathbf{S}), H(\tilde{\mathbf{S}})) < 0.(186)$

This condition is typically met when:

- The transformation $\mathbf{S} \mapsto \tilde{\mathbf{S}}$ induces **negative dependence** (e.g., reflection, negation).
- The payoff function H is **monotonic** — for example, standard call or put options.

```
using Distributions, Statistics, Random

# Black-Scholes parameters
S  = 100.0         # initial stock price
K = 105.0          # strike price
T = 1.0            # maturity in years
 \sigma = 0.2            # volatility (20%)
r = 0.05           # risk -free rate (5%)
N = 100_000        # number of simulation samples

# Analytical Black-Scholes formula for comparison
function black_scholes_call(S , K, r, \sigma, T)
    d  = (log(S  / K) + (r + 0.5 * \sigma^2) * T) / ( \sigma * sqrt(T))
    d  = d  - \sigma * sqrt(T)
    return S  * cdf(Normal(), d ) - K * exp( -r * T) * cdf(Normal(), d )
end

# Standard Monte Carlo method
function mc_price_european_call(S , K, r, \sigma, T, N)
    drift = (r - 0.5 * \sigma^2) * T
    diffusion = \sigma * sqrt(T)
```

```
    Z = randn(N)
    ST = S  .* exp.(drift .+ diffusion .* Z)
    payoffs = max.(ST . - K, 0.0)
    discounted = exp( -r * T) .* payoffs
    return mean(discounted), std(discounted) / sqrt(N)
end

# Antithetic variates method
function mc_price_european_call_antithetic(S , K, r, \sigma, T, N)
    drift = (r - 0.5 * \sigma^2) * T
    diffusion = \sigma * sqrt(T)
    Z = randn(N)
    ST1 = S  .* exp.(drift .+ diffusion .* Z)
    ST2 = S  .* exp.(drift . - diffusion .* Z)
    payoff1 = max.(ST1 . - K, 0.0)
    payoff2 = max.(ST2 . - K, 0.0)
    averaged = (payoff1 .+ payoff2) ./ 2
    discounted = exp( -r * T) .* averaged
    return mean(discounted), std(discounted) / sqrt(N)
end

# Run simulations
price_mc, error_mc = mc_price_european_call(S , K, r, \sigma, T, 2N) # Use 2N for fair comparison
price_antithetic, error_antithetic = mc_price_european_call_antithetic(S , K, r, \sigma, T, N)
price_bs = black_scholes_call(S , K, r, \sigma, T)

# Print results
println("European Call Option Pricing Comparison")
println("========================================")
println("Analytical Black-Scholes Price     : \$$(round(price_bs, digits=4))\n")

println("Standard Monte Carlo (2N samples)  : \$$(round(price_mc, digits=4))")
println("  Standard Error                   : \pm$(round(error_mc, digits=5))\n")

println("Antithetic Variates (N pairs)      : \$$(round(price_antithetic, digits=4))")
println("  Standard Error                   : \pm$(round(error_antithetic, digits=5))\n")

println("Efficiency gain (MC Error² / Antithetic Error²) \approx $(round((error_mc^2 / error_antithetic

European Call Option Pricing Comparison
========================================
Analytical Black-Scholes Price     : $8.0214

Standard Monte Carlo (2N samples)  : $7.9804
  Standard Error                   : \pm0.02942

Antithetic Variates (N pairs)      : $7.9892
  Standard Error                   : \pm0.02338

Efficiency gain (MC Error² / Antithetic Error²) \approx 1.58 \times
```

10.2 Control Variates

The **control variates** method aims to reduce the variance of a Monte Carlo estimator by exploiting the known expectation of a related random variable.

Assume there exists:

- A random variable $\mathbf{Y}$,
- A functional $g(\mathbf{Y})$,

such that the expectation $\mathbb{E}_{\mathbb{Q}}[g(\mathbf{Y})]$ is known in closed form.

Then, for any real-valued parameter $\lambda \in \mathbb{R}$, we have the identity:

$\mathrm{E}_{\mathbb{Q}}[H(\mathbf{S})] = \mathbb{E}_{\mathbb{Q}}\left[H(\mathbf{S}) - \lambda\left(g(\mathbf{Y}) - \mathbb{E}_{\mathbb{Q}}[g(\mathbf{Y})]\right)\right]$.(187)

This motivates the following **control variates Monte Carlo estimator**:

$$\hat{\pi}_0^{C,\lambda}(H, \mathbf{S}) := \frac{1}{N}\sum_{i=1}^{N}\left(H(\mathbf{S}^{(i)}) - \lambda\left(g(\mathbf{Y}^{(i)}) - \mathbb{E}_{\mathbb{Q}}[g(\mathbf{Y})]\right)\right), \tag{188}$$

where $(\mathbf{S}^{(i)}, \mathbf{Y}^{(i)})$ are N independent samples of the joint distribution of $(\mathbf{S}, \mathbf{Y})$.

Optimal Choice of λ

The optimal value of λ—i.e., the one that minimizes the variance of the estimator—is given by:

$\lambda^\star = \frac{\mathrm{Cov}(H(\mathbf{S}), g(\mathbf{Y}))}{\mathrm{Var}(g(\mathbf{Y}))}$.(189)

This choice ensures that the variance of the control variates estimator

$$\mathrm{Var}\left(H(\mathbf{S}) - \lambda^\star\left(g(\mathbf{Y}) - \mathbb{E}_{\mathbb{Q}}[g(\mathbf{Y})]\right)\right) \tag{190}$$

Then the variance of the control variates estimator is:

$$\boxed{\mathrm{Var}(H(\mathbf{S})) - \frac{\mathrm{Cov}(H(\mathbf{S}), g(\mathbf{Y}))^2}{\mathrm{Var}(g(\mathbf{Y}))}} \tag{191}$$

is **smaller** than the variance of the naive Monte Carlo estimator, provided that $H(\mathbf{S})$ and $g(\mathbf{Y})$ are **positively correlated**.

In other words, the improvement of the control variates Monte Carlo simulation over the standard Monte Carlo method comes in the form of a multiplicative factor. Assuming that the computational work per realization is c times higher using control variates, the equation above implies that the control variates technique is $\frac{1}{c(1-\rho^2)}$-times faster than standard Monte Carlo. In particular, the improvement in speed from the use of control variates is larger as the correlation between $H(\mathbf{S})$ and $g(\mathbf{Y})$ becomes higher. If, for example, $\rho = 0.8$ and $c = 2$ the speed-up factor equals 1.38, while if $\rho = 0.95$ the speed-up factor equals 5.

10.2.1 Finding Good Control Variates

A natural question is: **how can we find or construct effective control variates?**
Unfortunately, there is no one-size-fits-all answer—control variates are typically tailored to the specific problem.

That said, in **option pricing**, there are several robust strategies:

- **The underlying asset** itself often serves as a powerful control variate. For instance, under the risk-neutral measure and assuming a zero risk-free rate, we have:

  $\mathsf{E}_{\mathbb{Q}}[S_t] = S_0, \quad \forall t \geq 0.$(192) This known expectation makes the asset an excellent candidate for variance reduction.
- **Simple options with closed-form solutions** (such as vanilla calls and puts under the Black–Scholes model) can act as control variates when pricing more complex derivatives (e.g. barrier or Asian options).
- **Simplified models** can serve as control variates when working in more complex environments. For example:
  - Use the **Black–Scholes model** as a control variate when pricing options under stochastic volatility models.
  - Use **models with deterministic volatility** to approximate more general dynamics.

Example: Control Variate in European Call Option Pricing

Assume we want to compute the price of a European option with payoff $H(S_T)$ using Monte Carlo. Let $S_T^{(1)}, \ldots, S_T^{(N)}$ be independent samples from the distribution of S_T.
We use the terminal asset price S_T itself as a control variate, since under the risk-neutral measure:

$$\mathbb{E}_{\mathbb{Q}}[S_T] = S_0 e^{rT}. \tag{193}$$

The control variate estimator takes the form:

$$\hat{\pi}_0^{C,\hat{\lambda}}(H, S_T) = \frac{1}{N} \sum_{i=1}^{N} \left(H(S_T^{(i)}) - \hat{\lambda} S_T^{(i)} \right) + \hat{\lambda} \mathbb{E}_{\mathbb{Q}}[S_T], \tag{194}$$

where $\hat{\lambda}$ is estimated from the sample by:

$$\hat{\lambda} = \frac{\mathsf{Cov}(H(S_T), S_T)}{\mathsf{Var}(S_T)}. \tag{195}$$

For a European call option $H(S_T) = (S_T - K)^+$, the effectiveness of the control variate depends on the strike price K.

```
using Distributions, Statistics, Random

# Parameters
S  = 100.0        # Initial stock price
K = 105.0         # Strike price
T = 1.0           # Time to maturity
 \sigma = 0.2           # Volatility
r = 0.05          # Risk -free rate
```

```
N = 100_000       # Number of simulations

# Generate terminal stock prices under Black -Scholes model
drift = (r - 0.5 * \sigma^2) * T
diffusion = \sigma * sqrt(T)
Z = randn(N)
S_T = S  .* exp.(drift .+ diffusion .* Z)

# Compute call payoffs
payoffs = max.(S_T . - K, 0.0)

# Control variate: use S_T with known E[S_T] = S  * exp(rT)
mean_ST = S  * exp(r * T)

# Estimate lambda
cov_payoff_ST = cov(payoffs, S_T)
var_ST = var(S_T)
 \lambda_hat = cov_payoff_ST / var_ST

# Compute control variate estimator
cv_estimates = payoffs . - \lambda_hat .* S_T .+ \lambda_hat * mean_ST
cv_price = mean(cv_estimates)
cv_std_error = std(cv_estimates) / sqrt(N)

# Compare to crude Monte Carlo estimate
plain_price = mean(payoffs) * exp( -r * T)
plain_std_error = std(payoffs) / sqrt(N) * exp( -r * T)

# Output
println("European Call Option Pricing using Control Variate")
println(" - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
println("Crude Monte Carlo Estimate:      \$$(round(plain_price, digits=4)) \pm $(round(plain_std_error
println("Control Variate Estimate:        \$$(round(cv_price * exp( -r * T), digits=4)) \pm $(round(cv_
println("Estimated \lambda: $(round( \lambda_hat, digits=4))")
```

```
European Call Option Pricing using Control Variate
 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Crude Monte Carlo Estimate:      $7.9899 \pm 0.04151
Control Variate Estimate:        $7.9991 \pm 0.01881
Estimated \lambda: 0.5818
```

10.3 Importance Sampling

Importance sampling is a variance reduction technique that aims to improve sampling efficiency by allocating more samples to regions of the sample space that contribute significantly to the expectation. It is conceptually related to the acceptance-rejection method and to Girsanov's theorem in continuous-time finance.

Assume that the underlying random variable $\mathbf{S} \in \mathbb{R}^d$ has a known density $q(x)$, and let $H(\mathbf{S})$ be a quantity of interest (e.g., a derivative payoff). For any alternative density $p(x) > 0$ on the support of q, we can reweight the integral using the **change of measure**:

$$\mathbb{E}_{\mathbb{Q}}[H(\mathbf{S})] = \int_{\mathbb{R}^d} H(x)q(x)dx = \int_{\mathbb{R}^d} H(x)\frac{q(x)}{p(x)}p(x)dx = \mathbb{E}_{\mathbb{P}}\left[\frac{q(\mathbf{Y})}{p(\mathbf{Y})}H(\mathbf{Y})\right], \tag{196}$$

where $\mathbf{Y} \sim p$ is a random variable sampled under the new (importance) density p, and the ratio $\frac{q(x)}{p(x)}$ is the **likelihood ratio** (also known as the **Radon–Nikodym derivative**).

The corresponding **importance sampling Monte Carlo estimator** is:

$$\hat{\pi}_0^{\text{Imp},N}(H, \mathbf{S}) = \frac{1}{N} \sum_{i=1}^{N} \frac{q(\mathbf{Y}^{(i)})}{p(\mathbf{Y}^{(i)})} H(\mathbf{Y}^{(i)}), \tag{197}$$

where $\mathbf{Y}^{(i)} \sim p$ are independent and identically distributed samples. The **variance** of the weighted payoff

H(Y) $\cdot \frac{q(\mathbf{Y})}{p(\mathbf{Y})}$(198) is cruicial to the efficiency of the method. A good importance distribution $p(x)$ should:

- Be close to the original density $q(x)$,
- Emphasize the regions where $H(x)$ is large (i.e., contributes significantly to the expected value).

Q1: How to choose the importance sampling density p?

The key to successful importance sampling lies in selecting a density p that minimizes the variance of the weighted estimator:

$$\mathbb{E}_p\left[\left(H(\mathbf{Y}) \cdot \frac{q(\mathbf{Y})}{p(\mathbf{Y})}\right)^2\right] - \left(\mathbb{E}_q[H(\mathbf{S})]\right)^2. \tag{199}$$

Ideal Choice (Theoretical Optimum)

The **optimal** density that minimizes variance is:

$$p^*(x) \propto |H(x)| \cdot q(x). \tag{200}$$

This choice minimizes the variance of the estimator to zero. However, this is purely theoretical — it requires knowledge of $\mathbb{E}_q[H(\mathbf{S})]$, which is exactly what we are trying to compute. Thus, it's not implementable in practice.

Since p^* cannot be used directly, we aim for a **good approximation**:

- Choose p such that the ratio $\frac{H(x)\cdot q(x)}{p(x)}$ is approximately constant (i.e., the integrand is **flat**).
- This flattens the variance contribution across the sample space.
- Sampling more from regions where $H(x) \cdot q(x)$ is large helps reduce rare but important contributions.

Example: Importance Sampling: Mean Shift in European Call Option

We want to compute the price of a European call with payoff:

$$H(S_T) = (S_T - K)^+. \tag{201}$$

Assume under the risk-neutral measure $\mathbb{Q}$, the asset follows:

$$\log S_T \sim \mathcal{N}\left(\log S_0 + \left(r - \frac{\sigma^2}{2}\right)T, \sigma^2 T\right). \tag{202}$$

But if $K \gg S_0$, most samples S_T will be below the strike, and the payoff will often be zero — contributing nothing to the estimate but still adding variance.

Shift the mean of the normal distribution upward to better cover the payoff region:

$$Y = \log S_0 + \left(r - \frac{\sigma^2}{2} + \theta\right) T + \sigma\sqrt{T}Z, \quad Z \sim \mathcal{N}(0,1) \tag{203}$$

Then adjust via the Radon-Nikodym derivative:

$$\frac{q(Y)}{p(Y)} = \exp\left(-\theta Z - \frac{1}{2}\theta^2 T\right) \tag{204}$$

so the Monte Carlo estimator becomes:

$$\hat{\pi}_0^{\mathsf{Imp}} = e^{-rT} \cdot \frac{1}{N} \sum_{i=1}^{N} \left((S_T^{(i)} - K)^+ \cdot \exp\left(-\theta Z^{(i)} - \frac{1}{2}\theta^2 T\right)\right) \tag{205}$$

This method then improves accuracy when the call is deep out-of-the-money.

10.4 Concluding Remarks on Variance Reduction Techniques

Variance reduction techniques enhance the efficiency of Monte Carlo estimators by lowering the estimator's variance without increasing the number of samples.

- **Antithetic Variates**
  These are simple to implement and computationally cheap. They offer moderate variance reduction, especially when the payoff function is monotonic. However, the improvement is often limited unless strong negative correlation is present between the original and antithetic samples.
- **Control Variates**
  These can significantly reduce variance by incorporating known expectations of related variables (e.g., the underlying asset or closed-form options). When a suitable control is available and strongly correlated with the target variable, this method is both powerful and efficient. Its effectiveness hinges on the availability and correct identification of a good control.
- **Importance Sampling**
  This method strategically changes the sampling distribution to focus on regions of high contribution to the expectation. It is especially useful for rare-event simulation or out-of-the-money options. The choice of the importance distribution is crucial and problem-specific, making implementation more involved.

11 Estimating Option Sensitivities (Greeks)

In this section, we introduce techniques for estimating the **Greeks**—that is, the derivatives of option prices with respect to various model parameters (e.g., volatility, interest rate) and non-model parameters (e.g., initial asset price).

Greeks play a crucial role in financial risk management, hedging, and portfolio optimization. While they were conceptually introduced in the first part of the course, we now turn to practical estimation techniques, particularly in the context of Monte Carlo simulation.

Problem Formulation

We are interested in estimating the **sensitivities** of a financial derivative's price — commonly referred to as the **Greeks** — with respect to various input parameters.

Let the price of the derivative at time 0 be denoted by:

$$\pi_0 = \mathbb{E}_{\mathbb{Q}}[H(\mathbf{S})] = \mathbb{E}_{\mathbb{Q}}[H(S_0, \sigma, \tau, r, \ldots)], \tag{206}$$

where:

- S_0: current price of the underlying asset,
- σ: volatility of the underlying asset,
- τ: time to maturity,
- r: risk-free interest rate,
- $\ldots$: other model parameters affecting the dynamics of the stochastic process $\mathbf{S}$.

Additionally, the payoff function H often depends on **contract-specific (non-model) parameters**, so that we may also write:

$$\pi_0 = H(S_0, \sigma, \tau, r, \ldots \mid K, B, \ldots), \tag{207}$$

where:

- K: strike price of the option,
- B: barrier level (in case of barrier options),
- $\ldots$: other contractual parameters like rebate levels, averaging windows, etc.

Our goal is to compute:

- **Model sensitivities**: derivatives of π_0 with respect to $S_0, \sigma, \tau, r, \ldots$,
- **Contract sensitivities**: derivatives with respect to $K, B, \ldots$.

These sensitivities enable hedging strategies and risk control in both theoretical modeling and practical trading environments. In the following, we will explore **Monte Carlo techniques** that yield unbiased or low-variance estimators for these derivatives.

Reminder on Greeks

The **Greeks** quantify how sensitive the value of a derivative π_0 is to changes in various parameters. These sensitivities play a fundamental role in hedging, risk management, and understanding the behavior of financial instruments under different market conditions.

First-Order Greeks

For example, the **Delta** measures sensitivity to the initial asset price (S_0):

$$\Delta \triangleq \frac{\partial \pi_0}{\partial S_0} \tag{Delta}$$

Similarly, we define other first-order sensitivities:

- **Vega**: sensitivity to volatility σ

  V $\triangleq \frac{\partial \pi_0}{\partial \sigma}$(208)

- **Theta**: sensitivity to time to maturity τ

$\Theta \triangleq \frac{\partial \pi_0}{\partial \tau}$(209)

Second-Order (Higher-Order) Greeks

These measure the curvature or second-order effects:

- **Gamma**: rate of change of Delta with respect to S_0

$\Gamma \triangleq \frac{\partial \Delta}{\partial S_0} = \frac{\partial^2 \pi_0}{\partial S_0^2}$(210)

- **Vomma (or Volga)**: sensitivity of Vega to changes in volatility

Vomma $\triangleq \frac{\partial \mathcal{V}}{\partial \sigma} = \frac{\partial^2 \pi_0}{\partial \sigma^2}$(211)

11.1 Bump-and-Revalue Method for Estimating Greeks

The **bump-and-revalue** method is one of the most intuitive and widely used approaches for estimating Greeks. It approximates partial derivatives via **finite differences**, relying on re-evaluating the option price after a small perturbation (or "bump") to one of the input parameters.

Estimating Delta (Δ)

To estimate the sensitivity of the option price π_0 with respect to the initial stock price S_0, proceed as follows:

1. **Bump** the underlying asset price S_0 by a small amount ΔS_0.
2. **Revalue** the option using both the original and bumped values:
   - Compute $\pi_0(S_0)$ — the original option price.
   - Compute $\pi_0(S_0 + \Delta S_0)$ — the option price with bumped input.
3. **Approximate Delta** using the finite difference formula:

$$\Delta \approx \frac{\pi_0(S_0 + \Delta S_0) - \pi_0(S_0)}{\Delta S_0}, \quad \text{for small } \Delta S_0. \tag{212}$$

Algorithm: Bump-and-Revalue Estimation of Greeks

The **bump-and-revalue** method approximates option sensitivities (Greeks) by perturbing one model input at a time and observing the impact on the derivative price.

Inputs:

- Initial parameter vector $\mathbf{p}$ (e.g., $S_0, \sigma, r, \tau, \ldots$)
- Pricing function or Monte Carlo estimator $\pi_0(\mathbf{p})$
- Small bump size δ
- Index or name of the parameter to bump

Output:

- Estimated Greek value $\widehat{\text{Greek}}$

Procedure:

1. **Evaluate the base price**:

$$\pi_0^{\text{original}} = \pi_0(\mathbf{p}) \tag{213}$$

2. **Apply the bump** to the chosen parameter:

$$\mathbf{p}_{\text{bump}} = \mathbf{p}, \quad \mathbf{p}_{\text{bump}}[i] \leftarrow \mathbf{p}[i] + \delta \tag{214}$$

3. **Re-evaluate the price** with the bumped parameter:

$$\pi_0^{\text{bump}} = \pi_0(\mathbf{p}_{\text{bump}}) \tag{215}$$

4. **Estimate the Greek** using the finite difference:

$$\widehat{\text{Greek}} = \frac{\pi_0^{\text{bump}} - \pi_0^{\text{original}}}{\delta} \tag{216}$$

5. **Return** the estimated value $\widehat{\text{Greek}}$

Variance Reduction Using the Same Seed in Bump-and-Revalue

An important consideration when applying the **bump-and-revalue** method is the **variance of the Greek estimator**, especially when the Monte Carlo method is used for pricing.

Estimator Setup

The estimator for a Greek—say, **Delta**—is computed as:

$$\widehat{\text{Greek}} = \frac{\hat{\pi}_0^{\text{bump}} - \hat{\pi}_0}{\delta}, \tag{217}$$

where each (\hat{\pi}_0) is a Monte Carlo estimate of the option price before and after bumping a parameter (e.g., the initial asset price).

Variance of the Finite Difference Estimator

The **variance** of this finite difference estimator is:

$$\mathrm{Var}(\widehat{\text{Greek}}) = \frac{1}{\delta^2}\mathrm{Var}(\hat{\pi}_0^{\text{bump}} - \hat{\pi}_0). \tag{218}$$

Using the standard identity for the variance of a difference:

$$\mathrm{Var}(A - B) = \mathrm{Var}(A) + \mathrm{Var}(B) - 2\mathrm{Cov}(A, B), \tag{219}$$

we obtain:

$$\mathrm{Var}(\widehat{\text{Greek}}) = \frac{1}{\delta^2}\left(\mathrm{Var}(\hat{\pi}_0) + \mathrm{Var}(\hat{\pi}_0^{\text{bump}}) - 2\mathrm{Cov}(\hat{\pi}_0, \hat{\pi}_0^{\text{bump}})\right). \tag{220}$$

Use of Common Random Numbers

A powerful variance reduction technique is to use **the same random number seed** for both $\hat{\pi}_0$ and $\hat{\pi}_0^{\text{bump}}$, i.e., simulate the same Brownian paths for both valuations. This induces a **positive correlation** between the two estimators, which increases the covariance term and thus **reduces the overall variance**:

Cov$({}_0, \hat{\pi}_0^{\text{bump}}) \uparrow \quad \Rightarrow \quad \mathrm{Var}(\widehat{\text{Greek}}) \downarrow$.(221)

12 The Pathwise Derivative Method

The **pathwise method** estimates option sensitivities by directly differentiating each simulated sample path with respect to the parameter of interest.

Setup

Let p be a model parameter (e.g., initial asset price S_0, volatility σ, etc.).
Assume we are given the pathwise representation of the underlying price process:

$$(S_t(\omega, p))_{t\in[0,T]}, \tag{222}$$

where each path depends smoothly on p. Then, the price of a derivative with discounted payoff functional $\tilde{H}$ is given by:

$$\pi_0(p) = \mathbb{E}_{\mathbb{Q}}\left[\tilde{H}((S_t(p))_{t\in[0,T]})\right]. \tag{223}$$

If the function $\tilde{H}$ is differentiable with respect to p, and the derivative can be taken inside the expectation (under certain regularity conditions), we obtain:

$$\frac{d}{dp}\pi_0(p) = \mathbb{E}_{\mathbb{Q}}\left[\frac{\partial}{\partial p}\tilde{H}((S_t(p))_{t\in[0,T]})\right]. \tag{224}$$

This is particularly useful for **smooth payoff functions**, where differentiation can be handled analytically on each path.

- The method works best when the payoff is **continuously differentiable** in the parameter p.
- Examples: European call/put options, Asian options with continuous averaging.
- The method **fails** if $\tilde{H}$ is not differentiable in p, e.g. for digital options or discontinuous barrier options.

Digital options, which pay a fixed amount if the underlying asset is above (or below) a certain strike price at maturity, present a challenge to the pathwise method. The payoff function of a digital call option can be represented as:

$$H(S_T) = \begin{cases} 1 & \text{if } S_T > K \\ 0 & \text{otherwise} \end{cases} \tag{225}$$

This function is not continuously differentiable due to the discontinuity at $S_T = K$. Therefore, applying the pathwise method directly to such payoffs is not feasible.

12.0.1 Regularity Assumption:

The payoff function $H(S_T)$ must be continuously differentiable, at least in the regions of interest. This is typically denoted as $H(S_T) \in C^1$, where C^1 represents the class of functions that are continuously differentiable.

Digital options, which pay a fixed amount if the underlying asset is above (or below) a certain strike price at maturity, present a challenge to the pathwise method. The payoff function of a digital call option can be represented as:

$$H(S_T) = \begin{cases} 1 & \text{if } S_T > K \\ 0 & \text{otherwise} \end{cases} \tag{226}$$

This function is not continuously differentiable due to the discontinuity at $S_T = K$. Therefore, applying the pathwise method directly to such payoffs is not feasible.

To address this, we introduce a **smooth approximation** of the digital payoff.

12.1 Smooth Approximation

We define a family of smooth functions $H_\epsilon(S_T)$, parameterized by $\epsilon > 0$, that approximate the digital payoff:

$$H_\epsilon(S_T) = \frac{1}{1 + \exp\left(-\frac{S_T - K}{\epsilon}\right)}. \tag{227}$$

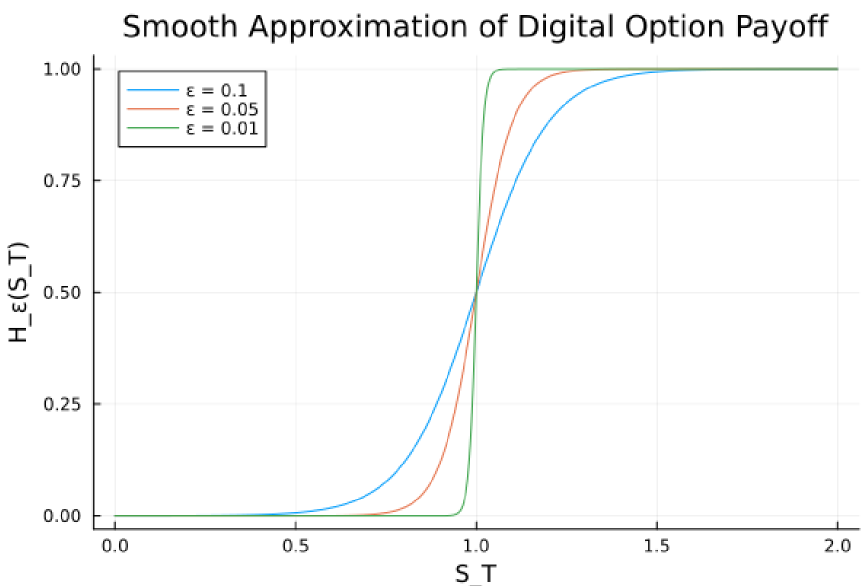
- This function is **infinitely differentiable** for $\epsilon > 0$.
- As $\epsilon \to 0$, $H_\epsilon(S_T) \to H(S_T)$, recovering the digital payoff.
- The slope of the sigmoid becomes steeper as ϵ decreases.

The smoothed function $H_\epsilon(S_T)$ is now suitable for **analytical differentiation** and we can compute sensitivities such as Delta and Vega using the **pathwise method** without introducing discontinuity-related numerical errors. **But** there is a trade-off: A smaller ϵ better mimics the digital payoff, but may increase numerical instability or gradient variance. Therefore choose ϵ based on:

- The scale of the underlying asset (e.g., $\epsilon \approx 0.5\% - 1\%$ of S_0).
- The desired accuracy vs. smoothness trade-off in gradient estimation.

```
using Plots

# Parameters
S_T = range(0, stop=2, length=1000)
K = 1
epsilons = [0.1, 0.05, 0.01]

# Initializing the plot
p = plot(size=(600, 400))

for epsilon in epsilons
    H_epsilon = 1 ./ (1 .+ exp.( -(S_T . - K) ./ epsilon))
    plot!(S_T, H_epsilon, label=" \epsilon = $epsilon")
end

xlabel!("S_T")
ylabel!("H_ \epsilon(S_T)")
```

```
title!("Smooth Approximation of Digital Option Payoff")
legend=:topright

display(p)
```



12.2 Black-Scholes Delta by the Pathwise Method

Assume that $S_t = S_0 e^{(r-\frac{1}{2}\sigma^2)t+\sigma W_t}$ for all $t \in [0,T]$, i.e. $(S_t)_{t\in[0,T]}$ is the Black-Scholes model. To simulate S at T, we insert $t = T$ and W_T by $\sqrt{T}Z$ for $Z \sim \mathcal{N}(0,1)$.

Next, consider a European Call option with strike K, and risk-neutral price given by:

$$\pi_0(S(S_0, r, \sigma)) = \mathbb{E}_{\mathbb{Q}}[e^{-rT}(S_T - K)^+]. \tag{228}$$

Our goal is to estimate

$$\Delta = \frac{\partial \pi_0(S(S_0, r, \sigma))}{\partial S_0}, \tag{229}$$

which by the pathwise method is equivalent to:

$$\Delta = \mathbb{E}_{\mathbb{Q}}\left[e^{-rT}\frac{\partial (S_T - K)^+}{\partial S_0}\right]. \tag{230}$$

This payoff is differentiable with respect to S_T almost everywhere, and its derivative is:

$$\frac{\partial (S_T - K)^+}{\partial S_T} = \mathbf{1}_{\{S_T > K\}}. \tag{231}$$

Applying the **chain rule** for differentiating with respect to (S_0):

$$\frac{\partial (S_T - K)^+}{\partial S_0} = \frac{\partial (S_T - K)^+}{\partial S_T} \cdot \frac{\partial S_T}{\partial S_0} = \mathbf{1}_{\{S_T > K\}} \cdot \frac{S_T}{S_0}. \tag{232}$$

Hence, the **pathwise estimator** of the Delta is given by:

$$\widehat{\Delta} = e^{-rT} \cdot \mathbf{1}_{\{S_T > K\}} \cdot \frac{S_T}{S_0}. \tag{233}$$

Remark:

- The expected value of this expression is indeed the Black-Scholes delta, so the estimator is unbiased, i.e.
$$\Delta = \mathbb{E}_{\mathbb{Q}}[\widehat{\Delta}]. \tag{234}$$
- Note that the relation
$$\frac{\partial S_T}{\partial S_0} = \frac{S_T}{S_0} \tag{235}$$
holds more widely than GBM and is often useful as an approximation.

12.3 Black-Scholes Vega by the Pathwise Method

Assume the same setting as before, but this time we aim to estimate Vega, i.e.,

$$\mathcal{V} = \frac{\partial \pi_0(S(S_0, r, \sigma))}{\partial \sigma} = \mathbb{E}_{\mathbb{Q}}\left[\mathrm{e}^{-rT}\frac{\partial (S_T - K)^+}{\partial \sigma}\right]. \tag{236}$$

We find that:
$$\frac{\partial (S_T - K)^+}{\partial \sigma} = \frac{\partial (S_T - K)^+}{\partial S_T}\frac{\partial S_T}{\partial \sigma} = \mathbf{1}_{\{S_T > K\}} S_T(-\sigma T + \sqrt{T}Z). \tag{237}$$

Hence, we obtain:

$$\widehat{\mathcal{V}} = \mathrm{e}^{-rT}\mathbf{1}_{\{S_T > K\}} S_T(-\sigma T + \sqrt{T}Z). \tag{238}$$

The expected value of this estimator is the Black-Scholes' Vega, so this estimator is unbiased

$$\mathcal{V} = \mathbb{E}_{\mathbb{Q}}[\hat{\mathcal{V}}]. \tag{239}$$

12.3.1 Example 1: Asian Option's Delta

As in the previous example, the underlying asset is modeled by a geometric Brownian motion, but now let the payoff be path dependent. Namely, we look at an Asian option which is written on the discretely monitored arithmetic average

$$\overline{S}_T = \frac{1}{m}\sum_{i=1}^{m} S_{t_i}, \tag{240}$$

with payoff given by $H(S) = (\overline{S}_T - K)^+$.

Following the pathwise delta estimation approach, we have:

$$\widehat{\Delta} = \mathrm{e}^{-rT}\frac{\overline{S}_T}{S_0}1_{\{\overline{S}_T > K\}} \tag{241}$$

Since

$$\frac{\partial \overline{S}}{\partial S_0} = \frac{1}{m}\frac{\partial}{\partial S_0}\sum_{i=1}^{m} S_{t_i} = \frac{1}{m}\sum_{i=1}^{m}\frac{S_{t_i}}{S_0} = \frac{\overline{S}}{S_0}. \tag{242}$$

Note: Due to the fact that there is no analytical formula for the price of an Asian option, this unbiased estimator has genuine practical value. Given that $\bar{S}$ would be simulated anyway in estimating the price of the option, this estimator requires minimal additional effort.

12.3.2 Example 2: Spread Option's Delta

We consider a Spread option with discounted payoff

$$\mathrm{e}^{-rT}\left((S_T^{(1)}-S_T^{(2)})-K\right)^+, \tag{243}$$

where $S^{(1)}$ and $S^{(2)}$ are two asset price processes with Deltas given by

$$\Delta^{(1)}=\mathrm{e}^{-rT}\frac{S_T^{(1)}}{S_0^{(1)}}\mathbf{1}_{S_T^{(1)}-S_T^{(2)}>K}, \tag{244}$$

and

$$\Delta^{(2)}=\mathrm{e}^{-rT}\frac{S_T^{(2)}}{S_0^{(2)}}\mathbf{1}_{S_T^{(1)}-S_T^{(2)}>K}. \tag{245}$$

12.3.3 Example 3: Max-option's Delta

We consider a Max Option's Delta, where the discounted payoff is given as

$$\mathrm{e}^{-rT}\left(\max(S_T^{(1)},\ldots,S_T^{(k)})-K\right)^+. \tag{246}$$

The i-th Delta is then given by:

$$\Delta^{(i)}=\mathrm{e}^{-rT}\frac{S_T^{(i)}}{S_0^{(i)}}\mathbf{1}_{S_T^{(i)}>K,S_T^{(i)}>S_T^{(j)},j\neq i}. \tag{247}$$

More generally, when we denote our portfolio value at time T by $W(S_T^{(1)},\ldots,S_T^{(k)})$, then the i-th Delta of the portfolio is:

$$\Delta_W^{(i)}=\frac{\partial f}{\partial S_T^{(i)}}\frac{\partial S_T^{(i)}}{\partial S_0^{(i)}}. \tag{248}$$

12.4 Option's Sensitivities in Stochastic Volatility Models

The expressions that we have derived in the examples apply more generally!

Indeed, consider an underlying asset price process described by a stochastic volatiliy model of the following form:

$$\frac{dS_t}{S_t}=r_t dt+\sigma_t dW_t^{(1)} \tag{249}$$

$$d\sigma_t=\alpha(\sigma_t)dt+\beta(\sigma_t)dW_t^{(2)} \tag{250}$$

Then, the solution is still given as an geometric process:

$$S_T=S_0\exp\left(\int_0^T(r_t-\frac{1}{2}\sigma_t^2)dt+\int_0^T\sigma_t dW_t^{(1)}\right). \tag{251}$$

Then also in this situation we have:

Standard Call: Delta estimate for $\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-rT}(S_T - K)^+]$ is

$$\widehat{\Delta} = \mathrm{e}^{-rT}\frac{S_T}{S_0}\mathbf{1}_{S_T>K} \tag{252}$$

Asian Option: Delta estimate for $(\overline{S} - K)^+$ is

$$\widehat{\Delta} = \mathrm{e}^{-rT}\frac{\overline{S}}{S_0}\mathbf{1}_{\overline{S}>K} \tag{253}$$

13 The Likelihood Ratio Method

The final method for estimating option sensitivities introduced in this course is the **Likelihood Ratio Method (LRM)**. Unlike the pathwise method, which differentiates through the payoff function, the LRM takes a different route: it **differentiates the probability distribution itself**.

Rather than requiring differentiability of the payoff function $H(S_T)$, the LRM assumes that the **density of the asset price at maturity** $g(x, p)$ is known and differentiable with respect to the parameter of interest p.

This makes the LRM particularly useful in situations where the **payoff function is not smooth** (e.g., digital or barrier options), but the asset price distribution is well understood.

Assume the expected discounted payoff can be written as an integral:

$$\pi_0(p) = \mathbb{E}_{\mathbb{Q}}[H(S_T)] = \int H(x)\, g(x,p)\, dx, \tag{254}$$

where:

- $g(x, p)$ is the **density function** of the terminal asset price S_T,
- and p is the model parameter with respect to which we want to compute the sensitivity.

Note that p **now appears in the density** g, rather than in the payoff function.

13.0.1 Likelihood Ratio Method Derivative Estimates

Take the pricing equation

$$\pi_0(p) = \mathbb{E}_{\mathbb{Q}}[H(S_T)] = \int H(x) g(x,p) dx, \tag{255}$$

and differentiate with respect to p to obtain:

$$\begin{aligned} \frac{\partial \pi_0(p)}{\partial p} &= \int_0^\infty H(x)\frac{\partial g(x,p)}{\partial p} dx \\ &= \int_0^\infty H(x)\frac{\partial \log(g(x,p))}{\partial p} g(x,p) dx \\ &= \mathbb{E}_{\mathbb{Q}}\left[H(S_T)\frac{\partial \log(g(S_T,p))}{\partial p}\right] \end{aligned}$$

This is the same for all payoffs H.

13.1 Black-Scholes Delta by the Likelihood Ratio Method

We now compute the Black-Scholes Delta by the Likelihood Ratio Method. Take:

$$\pi_0 = \mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-rT}(S_T - K)^+]. \tag{256}$$

The log-normal density of S_T is know as

$$g(x, S_0) = \frac{1}{x\sigma\sqrt{2\pi T}} \exp\left[-\frac{1}{2}\left(\frac{\log(x/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right)^2\right], \tag{257}$$

or in other words $g(x, S_0) = \frac{1}{x\sigma\sqrt{T}}\phi(d(x))$, where

$$d(x) = \frac{\log(x/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \tag{258}$$

and $\phi(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$ (CDF of the standard normal distributions).

Note that $\phi'(x) = -x\phi(x)$. Then

$$\begin{aligned}
\frac{\partial \log(g(x,p))}{\partial S_0} &= \frac{\partial \log(g(x,p))}{\partial g(x,p)} \cdot \frac{\partial g(x,p)}{\partial S_0} \\
&= \frac{1}{g(x,p)} \cdot \frac{\partial g(x,p)}{\partial S_0} \\
&= \frac{1}{g(x,p)} \cdot \frac{\partial}{\partial S_0}\left[\frac{1}{x\sigma\sqrt{T}}\varphi(d(x))\right] \\
&= \frac{1}{g(x)} \cdot \frac{1}{x\sigma\sqrt{T}}\varphi'\left(d(x)\frac{\partial d(x)}{\partial S_0}\right) \\
&= \frac{1}{g(x)} \cdot \frac{1}{x\sigma\sqrt{T}}\left(-d(x)\phi(d(x))\frac{\partial d(x)}{\partial S_0}\right) \\
&= -d(x)\frac{\partial d(x)}{\partial S_0} \\
&= \frac{\log(x/S_0) - (r - \frac{1}{2}\sigma^2)T}{S_0\sigma^2 T}.
\end{aligned}$$

Note that $\phi'(x) = -x\phi(x)$. The derivative of $\log(g(x, S_0))$ with respect to S_0 is given by

$$\frac{\partial}{\partial S_0}\log(g(x, S_0)) = \frac{\partial \log(g(x, S_0))}{\partial g(x, S_0)} \cdot \frac{\partial g(x, S_0)}{\partial S_0} \tag{259}$$

This can be simplified as

$$\frac{\partial}{\partial S_0}\log(g(x, S_0)) = \frac{1}{g(x, S_0)} \cdot \frac{\partial g(x, S_0)}{\partial S_0} \tag{260}$$

Further simplifying, we have

$$\frac{\partial}{\partial S_0}\log(g(x, S_0)) = \frac{1}{g(x, S_0)} \cdot \frac{\partial}{\partial S_0}\left[\frac{1}{x\sigma\sqrt{T}}\phi(d(x))\right] \tag{261}$$

This leads to

$$\frac{\partial}{\partial S_0}\log(g(x,S_0)) = \frac{1}{g(x,S_0)}\cdot\frac{1}{x\sigma\sqrt{T}}\phi'\left(d(x)\right)\frac{\partial d(x)}{\partial S_0} \tag{262}$$

Expanding this, we get

$$\frac{\partial}{\partial S_0}\log(g(x,S_0)) = \frac{1}{g(x,S_0)}\cdot\frac{1}{x\sigma\sqrt{T}}\left(-d(x)\phi(d(x))\frac{\partial d(x)}{\partial S_0}\right) \tag{263}$$

Note that

$$-d(x)\frac{\partial d(x)}{\partial S_0} = \frac{\log(x/S_0)-(r-\frac{1}{2}\sigma^2)T}{S_0\sigma^2 T}, \tag{264}$$

and thus when we define Y as

$$Y = \frac{\partial\log(g(S_T,S_0))}{\partial S_0} = \frac{\log(S_T/S_0)-(r-\frac{1}{2}\sigma^2)T}{S_0\sigma^2 T}. \tag{265}$$

Define Y as

$$Y = \frac{\partial\log(g(S_T,S_0))}{\partial S_0} = \frac{\log(S_T/S_0)-(r-\frac{1}{2}\sigma^2)T}{S_0\sigma^2 T}. \tag{266}$$

Then we obtain:

$$\Delta = \frac{\partial\pi_0}{\partial S_0} = \mathbb{E}_{\mathbb{Q}}\left[\mathrm{e}^{-rT}(S_T-K)^+Y\right]. \tag{267}$$

For digital option, just replace $\mathrm{e}^{-rT}(S_T-K)^+$ with $\mathrm{e}^{-rT}\mathbf{1}_{S_T\geq K}$.

13.2 Likelihood Ratio Method for Options on Multiple Assets

Option on $(S1_T,...,Sk_T)$,

$$dSi_t = rSi_t dt + \sigma_i Si_t dW_t^i, i=1,...,k, \tag{268}$$

and

$$E[dW_t^i\cdot dW_t^j] = \rho_{ij}, \tag{269}$$

e.g., basket option discounted payoff $= e^{-rT}\max(\alpha_1 S1_T+...+\alpha_k Sk_T-K,0)$

Covariance matrix $\Sigma_{ij}=\sigma_i\sigma_j\rho_{ij}$.

$$S_{i_T} = S_{i_0}\exp\left((r-\frac{1}{2}\sigma_i^2)T+\sqrt{T}X_i\right) \tag{270}$$

where $X\sim N(0,\Sigma)$.

Density of $(S1_T,...,Sk_T)$ multivariate lognormal $g(x1,...,xk)$ and

$$\frac{\partial \log g}{\partial Si_0} = (X\Sigma^{-1})_i \cdot \frac{1}{Si_0\sqrt{T}}, \tag{271}$$

Likelihood Ratio Method estimate on the ith delta discounted option payoff $\times$

$$[(X\Sigma^{-1})_i \cdot \frac{1}{Si_0\sqrt{T}}]. \tag{272}$$

14 Exam Quant Questions: Monte Carlo Methods, Error Analysis, and Variance Reduction

Level	Question
Easy	What is the basic idea behind **Monte Carlo simulation** in derivative pricing? Explain how **random sampling** is used to estimate the expected payoff $\mathbb{E}^Q[H(S_T)]$.
Medium	Monte Carlo estimates converge at a rate of $O(1/\sqrt{N})$. Explain why the **standard error** of a Monte Carlo price estimate is proportional to $1/\sqrt{N}$, and what this implies about the number of samples required for **increased accuracy**.
Medium	Describe the **Euler–Maruyama method** for simulating paths of an SDE: $dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$. How do you approximate $X_{t+\Delta t}$ given X_t over a discrete timestep Δt?
Hard	Explain the difference between **strong convergence** and **weak convergence** in the numerical simulation of SDEs. Why is this distinction important for different types of Monte Carlo applications in finance?
Medium	What is the purpose of **variance reduction techniques** in Monte Carlo simulations? Why are they particularly important when pricing **complex or path-dependent derivatives**?
Medium	How does the **antithetic variates** method work in Monte Carlo simulations? Describe the mechanism and intuition behind how it reduces estimator variance.
Medium	What is the **control variate** technique in Monte Carlo simulation? Give an example of how a known expectation can be used to improve the accuracy of a new estimate.
Hard	Suppose you are pricing a **deep out-of-the-money option** whose payoff is zero with high probability. Which variance reduction method might be especially helpful, and why? Discuss the use of **importance sampling** in this context.
Medium	How can we construct a **confidence interval** around a Monte Carlo estimate? What is the formula for the 95% confidence interval for an estimated option price $\hat{\pi}_N$ with estimated standard deviation $\hat{\sigma}$?
Medium	Compare Monte Carlo methods with **PDE-based pricing methods**. What is a key advantage of Monte Carlo for high-dimensional problems, and when would you prefer Monte Carlo over analytical or PDE-based approaches?
Easy	State the formula for estimating Delta via the **bump-and-revalue** method. Why is it called a "finite difference" method?
Medium	Explain why **using the same random number seed** (common random numbers) in the bump-and-revalue method reduces variance of the Greek estimator.
Medium	Derive an unbiased estimator for Delta using the **pathwise derivative method** in the Black-Scholes model, and specify for which payoff functions this method is valid.

Level	Question
Medium	Given a European call option, explain how to estimate its Vega via the **pathwise method**. Provide the formula for the Vega estimator.
Medium	What is the advantage of using the **likelihood ratio method** (LRM) over the pathwise method for estimating sensitivities of digital options?
Medium	You're given a payoff $H(S_T) = \mathbf{1}_{\{S_T>K\}}$. Why does the pathwise method fail in this case, and how can you modify the payoff to make it differentiable?
Hard	Derive the estimator for Delta using the **likelihood ratio method** in the Black-Scholes model. What is the key role of the derivative of the log-density?
Hard	Suppose you simulate an Asian option payoff as $H(S) = (\bar{S} - K)^+$, where $\bar{S} = \frac{1}{m}$
Medium	What is the **Milstein scheme** for discretizing an SDE of the form $dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t$? How does it differ from Euler–Maruyama, and in which cases is it more accurate?
Hard	Suppose you simulate an Asian option payoff as $H(S) = (\bar{S} - K)^+$, where $\bar{S} = \frac{1}{m}\sum_{i=1}^{m} S_{t_i}$. Derive the **pathwise estimator** for Delta.
Hard	In the context of **importance sampling**, describe how the **Radon–Nikodym derivative** is used to adjust the Monte Carlo weights. How does this improve variance for deep out-of-the-money options?
Hard	Derive the variance formula of a bump-and-revalue estimator for a Greek using common random numbers. What is the impact of the covariance term?

15 Fourier-Based Option Pricing

Definition: Fourier Transform of a Function

Let $f : \mathbb{R} \to \mathbb{R}$ be an integrable function (typically a payoff function). The **Fourier transform** of f is defined as:

$$\hat{f}(u) := \int_{\mathbb{R}} e^{iux} f(x)\, dx, \quad u \in \mathbb{R}. \tag{273}$$

The inverse transform is:

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iux} \hat{f}(u)\, du, \tag{274}$$

provided f and $\hat{f}$ satisfy suitable integrability conditions, e.g., $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.

15.1 Pricing European Options Using Fourier Methods

Let S_T be the terminal stock price under the **risk-neutral measure** $\mathbb{Q}$. The arbitrage-free price of a European option is:

$$V_0 = \mathbb{E}^{\mathbb{Q}}[e^{-rT} f(S_T)]. \tag{275}$$

Define the **log-price** as $X_T := \log S_T$, and change variables:

$$g(x) := f(e^x), \quad \text{so that} \quad V_0 = \mathbb{E}^{\mathbb{Q}}[e^{-rT} g(X_T)]. \tag{276}$$

If the Fourier transform $\hat{g}(u)$ exists and is integrable, then under Fubini's theorem we can interchange expectation and integration:

$$\begin{aligned} V_0 &= e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{2\pi}\int_{\mathbb{R}} e^{-iuX_T}\hat{g}(u)\, du\right] \\ &= \frac{e^{-rT}}{2\pi}\int_{\mathbb{R}} \mathbb{E}^{\mathbb{Q}}\left[e^{-iuX_T}\right]\hat{g}(u)\, du \\ &= \frac{e^{-rT}}{2\pi}\int_{\mathbb{R}} \varphi_{X_T}(-u)\,\hat{g}(u)\, du, \end{aligned} \tag{277}$$

where $\varphi_{X_T}(u) = \mathbb{E}^{\mathbb{Q}}[e^{iuX_T}]$ is the **characteristic function** of X_T under $\mathbb{Q}$. That means, if we know the distribution of X_T then we can use this information here directly!

15.2 Exponential Damping

Many payoff functions (like call options) are not in $L^1(\mathbb{R})$ and do not have integrable Fourier transforms. To address this, we **exponentially damp** the payoff.

Let $\alpha > 0$ and define:

$$g_\alpha(x) := e^{-\alpha x} f(e^x), \tag{278}$$

so that the **damped Fourier transform** is:

$$\hat{g}_\alpha(u) = \int_{\mathbb{R}} e^{iux} g_\alpha(x)\, dx. \tag{279}$$

Then the **damped price formula** becomes:

$$V_0 = \frac{e^{-rT}}{2\pi} \int_{\mathbb{R}} \varphi_{X_T}(-u - i\alpha)\, \hat{g}_\alpha(u)\, du, \tag{280}$$

assuming $\varphi_{X_T}(u)$ extends analytically into a strip around the real line.

Example: European Call Option

Let $f(S_T) = (S_T - K)^+$. Then $g(x) = (e^x - K)^+$. For $\alpha > 1$, define:

$$g_\alpha(x) = e^{-\alpha x}(e^x - K)^+ = \begin{cases} e^{-(\alpha-1)x} - Ke^{-\alpha x}, & x > \log K \\ 0, & \text{otherwise} \end{cases} \tag{281}$$

This function has an explicit Fourier transform:

$$\hat{g}_\alpha(u) = \int_{\log K}^{\infty} \left(e^{-(\alpha-1)x} - Ke^{-\alpha x}\right) e^{iux}\, dx, \tag{282}$$

which can be evaluated analytically and the call option price is therefore:

$$V_0^{\text{call}} = \frac{e^{-rT}}{2\pi} \int_{\mathbb{R}} \varphi_{X_T}(-u - i\alpha)\, \hat{g}_\alpha(u)\, du. \tag{283}$$

Truncation of the Fourier Integral In practice, we **truncate** the domain of integration to $[-L, L]$ and approximate:

$$\pi_0 \approx \frac{1}{2\pi} \int_{-L}^{L} e^{-iux_0} \widehat{h}(u)\, \varphi_{X_T}(u)\, du. \tag{284}$$

- The truncation parameter L must be large enough so that $\widehat{h}(u)\varphi_{X_T}(u)$ decays sufficiently fast outside $[-L, L]$.
- Often $L \in [50, 100]$ is sufficient for well-behaved models.
- Damping the payoff with an exponential factor helps ensure integrability (see **Carr–Madan** above).

Example: The COS Method

The **COS method** (Fang & Oosterlee, 2008) is a Fourier-cosine series expansion technique based on the following idea:

1. Truncate the log-asset domain to $[a, b]$ such that $X_T \in [a, b]$ with high probability.
2. Expand the density $f_{X_T}(x)$ in a cosine series:

$$f_{X_T}(x) \approx \sum_{k=0}^{N-1} A_k \cos\left(\frac{k\pi(x-a)}{b-a}\right) \tag{285}$$

3. Evaluate the option price as:

$$\pi_0 \approx e^{-rT} \sum_{k=0}^{N-1} \text{Re}\left[\varphi_{X_T}\left(\frac{k\pi}{b-a}\right)\right] V_k \tag{286}$$

where V_k are coefficients depending on the payoff (available in closed form for many cases).

15.3 3.1 Affine Processes and Fourier Pricing

Motivation

In the previous section, we saw that **Fourier-based pricing** relies crucially on the ability to compute the **characteristic function** of the log-asset price under the risk-neutral measure. That is:

$$\varphi_{X_T}(u) = \mathbb{E}^{\mathbb{Q}}\left[e^{iuX_T}\right], \quad \text{where } X_T = \log S_T. \tag{287}$$

For many stochastic models of S_T, including those with stochastic volatility, the distribution of X_T is not known in closed form—but its **characteristic function often is**, particularly for models in the **affine class**.

Definition: Affine Processes

A stochastic process $X_t \in \mathbb{R}^d$ is called **affine** if its conditional characteristic function takes the exponential-affine form:

$$\mathbb{E}\left[e^{u^\top X_T} \mid X_0 = x\right] = \exp\left(\phi(T,u) + \psi(T,u)^\top x\right), \tag{288}$$

for all $u \in \mathbb{C}^d$ and some deterministic functions $\phi(T,u)$ and $\psi(T,u)$ solving **Riccati-type differential equations**.

This structure means that we can often compute $\varphi_{X_T}(u)$ semi-analytically—even when the model itself is complex. As a result, affine models are a natural fit for Fourier-based option pricing.

Example: The Heston Model

One of the most prominent affine models is the **Heston stochastic volatility model**, in which both the asset price and its variance follow stochastic differential equations:

$$\begin{aligned} dS_t &= \mu S_t\,dt + \sqrt{V_t} S_t\,dW_t^S, \\ dV_t &= \kappa(\theta - V_t)\,dt + \sigma\sqrt{V_t}\,dW_t^V, \\ d\langle W^S, W^V\rangle_t &= \rho\,dt. \end{aligned} \tag{289}$$

Here:

- V_t is the **instantaneous variance**,
- κ is the **rate of mean reversion**,
- θ is the **long-run average variance**,
- σ is the **volatility of variance** (also called "vol of vol"),
- ρ is the **correlation** between asset and variance shocks.

Let $X_t = \log S_t$ denote the log-price of the asset. Then, the joint process (X_t, V_t) is a **2-dimensional affine process**.

Under the risk-neutral measure $\mathbb{Q}$, the conditional characteristic function of X_T is of the form:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{iuX_T} \mid X_0 = x, V_0 = v\right] = \exp\left(\phi(T,u) + \psi(T,u)v + iux\right), \tag{290}$$

where $\phi(T,u)$ and $\psi(T,u)$ satisfy the Heston-specific Riccati equations.

15.4 Riccati Equations and Characteristic Function in the Heston Model

In the Heston stochastic volatility model, the joint process (X_t, V_t)—with $X_t = \log S_t$—is affine, and its characteristic function has the form:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{iuX_T} \mid X_0 = x, V_0 = v\right] = \exp\left(\phi(T,u) + \psi(T,u)v + iux\right), \tag{291}$$

for $u \in \mathbb{C}$. The functions ϕ and ψ solve a system of **Riccati differential equations**.

Riccati System (under risk-neutral measure $\mathbb{Q}$)

Let $u \in \mathbb{C}$ be the Fourier argument. Then the Riccati ODEs for $\phi(\tau)$ and $\psi(\tau)$, with $\tau = T - t$, are:

$$\begin{aligned} \frac{d\psi(\tau)}{d\tau} &= \frac{1}{2}\sigma^2\psi(\tau)^2 - (\kappa - iu\rho\sigma)\psi(\tau) + \frac{1}{2}u(u+i), \\ \frac{d\phi(\tau)}{d\tau} &= \kappa\theta\psi(\tau), \end{aligned} \tag{292}$$

Closed-Form Solution

Define the following auxiliary quantities:

- $d := \sqrt{(\rho\sigma iu - \kappa)^2 + \sigma^2(iu + u^2)}$
- $g := \frac{\kappa - iu\rho\sigma - d}{\kappa - iu\rho\sigma + d}$

Then the solutions are:

$$\psi(T,u) = \frac{\kappa - iu\rho\sigma - d}{\sigma^2} \cdot \frac{1 - e^{-dT}}{1 - ge^{-dT}} \tag{293}$$

$$\phi(T,u) = \frac{\kappa\theta}{\sigma^2}\left[(\kappa - iu\rho\sigma - d)T - 2\ln\left(\frac{1 - ge^{-dT}}{1-g}\right)\right] \tag{294}$$

Putting everything together, the conditional characteristic function of $X_T = \log S_T$ under $\mathbb{Q}$ becomes:

$$\varphi_{X_T}(u) = \exp\left(\phi(T,u) + \psi(T,u)V_0 + iu\log S_0\right) \tag{295}$$

16 Calibration to Observed Option Prices

In practice, financial models are used to reproduce and model market prices of traded options. To make a model useful for pricing and risk management, it must first be **calibrated** to observed market prices.

Why and to what calibrating to?

In **incomplete markets** (e.g., with stochastic volatility), the **risk-neutral measure** $\mathbb{Q}$ is **not uniquely determined** by no-arbitrage arguments alone. Therefore, additional structure—usually from market prices of liquid instruments such as vanilla options—is needed to infer a suitable pricing measure.

In market practice, the **risk-neutral pricing measure** $\mathbb{Q}$ is implicitly defined **through calibration**: we adjust the model parameters so that model-implied option prices match observed ones as closely as possible.

The Calibration Problem

Let $\{K_i, T_i, C_i^{\text{mkt}}\}_{i=1}^N$ denote a set of market prices for European call options with strike K_i and maturity T_i, and let $\pi_i^{\text{model}}(\theta)$ be the model price for these options given parameters θ.

Then, the **calibration problem** becomes:

$$\theta^* = \arg\min_{\theta \in \Theta} \sum_{i=1}^{N} w_i \left(\pi_i^{\text{model}}(\theta) - C_i^{\text{mkt}}\right)^2 \tag{296}$$

where:

- $\theta \in \Theta$ are the model parameters (e.g., Heston: $\kappa, \theta, \sigma, \rho, V_0$),
- w_i are **weights** to reflect importance or reliability of each data point.

Choice of Weights w_i

The choice of weights is **crucial** and influences both the stability and interpretability of the calibration.

Some common choices:

- **Uniform weights**: $w_i = 1$. Treats all options equally, regardless of liquidity or scale.
- **Inverse variance weights**: $w_i = 1/\text{Var}(C_i^{\text{mkt}})$. Higher confidence in more liquid or stable quotes.
- **Relative error weights**: $w_i = 1/(C_i^{\text{mkt}})^2$. Minimizes relative pricing errors:

  $\sum_i \left(\frac{\pi_i^{\text{model}}(\theta) - C_i^{\text{mkt}}}{C_i^{\text{mkt}}}\right)^2$ (297)

- **Vega weighting**: Use the option's vega to emphasize regions of high sensitivity: $w_i = 1/\text{Vega}(K_i, T_i)$

Objective: Implied Volatility Fit vs Price Fit

- If the model does **not admit a closed-form** for implied volatility (as is the case for most stochastic volatility models), we usually calibrate **on prices**.
- If we can invert the pricing function easily, we may calibrate **on implied volatilities**:

  $\sum_i \left(\text{IV}^{\text{model}}(K_i, T_i; \theta) - \text{IV}^{\text{mkt}}(K_i, T_i)\right)^2$ (298)

Implementation Notes

- In models like the **Heston model**, option prices $\pi_i^{\text{model}}(\theta)$ are computed via **Fourier inversion** using the **characteristic function**.
- Calibration requires **efficient numerical routines** nd **robust optimization** techniques, as in the optimization process we have to compute option prices very often!

17 Chapter: Finite-Difference Methods – Quant Exam Questions

Level	Question
Easy	What is the main motivation for using finite-difference methods in option pricing? When are they preferable over Monte Carlo methods?
Easy	Describe the difference between the **explicit**, **implicit**, and **Crank–Nicolson** schemes in finite-difference methods.
Easy	What is the **discretized form** of the 1D heat equation $\frac{\partial y}{\partial \tau} = \frac{\partial^2 y}{\partial x^2}$ using an **explicit scheme**?
Medium	Write down the update formula for the explicit finite-difference scheme for the transformed Black–Scholes PDE. Define the role of the **stability parameter** $\lambda = \frac{\Delta\tau}{\Delta x^2}$
Medium	Explain how to impose **Dirichlet** and **Neumann boundary conditions** in a finite-difference scheme. Give examples of each in the context of European call options.
Medium	What are the **initial** and **boundary conditions** used when solving the transformed Black–Scholes PDE numerically using finite-difference methods?
Medium	Derive the **Crank–Nicolson scheme** by averaging the explicit and implicit schemes. Why is this scheme second-order accurate in both time and space?
Hard	Derive the **matrix form** of the finite-difference update: $\vec{y}^{n+1} = A\vec{y}^n + \vec{k}^n$. What do the matrix A and vector $\vec{k}$ represent in this context?
Hard	State the **Lax Equivalence Theorem**. How can it be used to ensure convergence of a finite-difference scheme?
Hard	Describe the **von Neumann stability analysis**. What condition must the **amplification factor** $\rho(\xi)$ satisfy for the scheme to be stable?
Hard	Given a Crank–Nicolson discretization of a PDE, demonstrate the method to check for **consistency** using Taylor expansions.
Hard	For the Crank–Nicolson scheme applied to $\partial_t u = \alpha \partial_x u$, derive the local truncation error and show that it is $O(\Delta t^2 + \Delta x^2)$.