From a Coin-toss to the Greeks

Building the Black-Scholes Model Step by Step with Mathematical and Visual Insights

Milan Peter

November 2025

When studying mathematics in finance, I often encounter a specific barrier: results and formulas appear as if they exist in isolation, without showing the connective tissue between them. Too often, the reasoning that links one step to another is hidden behind implicit assumptions, compact notation, or missing intuition. This leaves the impression that understanding advanced concepts means memorizing rather than truly seeing how they emerge.

This is particularly true for the Black-Scholes model, which sits at the crossroads of stochastic calculus, probability theory, and financial intuition. Every derivation I've seen takes a slightly different route – some emphasize arbitrage arguments, others rely on Itô's Lemma, and others introduce martingale pricing – yet each of them skips something that feels essential. In this study, my goal is to reconstruct the formula step by step, ensuring that each transformation is justified and connected to the economic meaning behind it.

Equally important to me is to compare derivation methods. The Black-Scholes formula is unique in that it can be derived from several theoretical perspectives: through hedging arguments, partial differential equations, or risk-neutral expectations. Contrasting these methods exposes how the same economic logic can take multiple mathematical forms, and how notation or framing can change the way we understand it.

Finally, visualization plays a central role in this work. Equations alone can conceal structure, while graphs and dynamic plots can reveal the geometric and probabilistic intuition beneath them – how volatility, time, and drift shape the evolution of option values. By coupling mathematics with visual explanation, I aim to build a more complete, intuitive understanding of the model.

In short, this study serves an educational purpose: to make the Black-Scholes model clearer, more intuitive, and easier to learn in a connected way.

I. Randomness in Finance: the Coin-toss Analogy

Finance, at its heart, is a story about uncertainty. Every investor, whether consciously or not, plays a probabilistic game: tomorrow's price might rise or fall, and no one knows for sure. To make sense of this, we begin not with complex stochastic calculus, but with something really simple – a fair coin.

I.1. The Coin Flip Model

Imagine flipping a fair coin repeatedly. Each flip has two outcomes:

- Heads → the price *goes up*
- Tails → the price *goes down*

Let's define a simple model of price changes:

$$X_n = X_0 + \sum_{i=1}^{n} \xi_i,$$

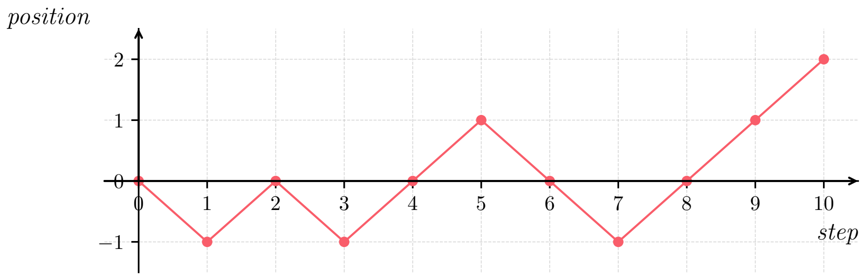
where each increment (ξ_i) represents the result of the i-th coin flip:

$$\xi_i = \begin{cases} +\Delta x & \text{with probability } \frac{1}{2}, \\ -\Delta x & \text{with probability } \frac{1}{2}. \end{cases}$$

Here, (Δx) is the step size – the magnitude of change in one step. For instance, if ($\Delta x = 1$), then each flip moves the price one unit up or down.

This process is known as a **symmetric random walk**, the simplest discrete model of randomness. Despite its simplicity, it captures the essential idea of uncertainty in asset prices: future movements are unpredictable, yet statistically structured.

Figure I.1: **Symmetric Random Walk**



A single realization of a symmetric random walk over 10 steps, starting at zero. Each step is determined by a fair coin flip, moving up or down by one unit.

I.2. Expected Value and Variance

Since each step has an equal chance of going up or down, the expected change per flip is zero:

$$E[\xi_i] = (+\Delta x)\left(\frac{1}{2}\right) + (-\Delta x)\left(\frac{1}{2}\right) = 0.$$

Hence, the expected value of the entire process after (n) steps is also:

$$E[X_n] = X_0.$$

In other words, on average, there is no drift – the price fluctuates around its initial value. But uncertainty accumulates. Because the steps are independent, their variances add up. Each step has variance:

$$\mathrm{Var}(\xi_i) = (\Delta x)^2.$$

After (n) steps, the total variance is simply the sum of all these:

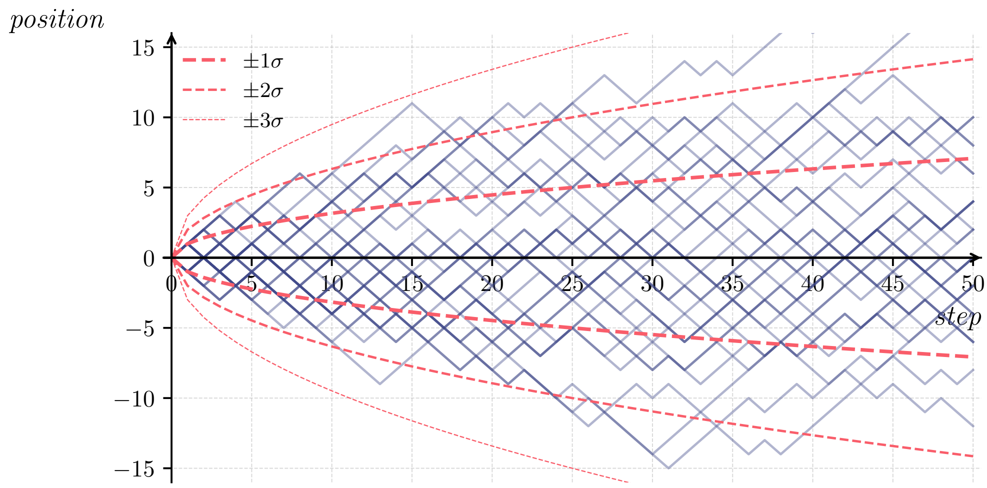
$$\mathrm{Var}(X_n) = n\mathrm{Var}(\xi_i) = n(\Delta x)^2.$$

Variance tells us how spread out the possible outcomes are. To get a sense of scale, we take the square root to obtain the standard deviation:

$$\sigma_n = \sqrt{\mathrm{Var}(X_n)} = \sqrt{n}\Delta x.$$

This describes the "fan-out" shape of uncertainty: as time progresses, outcomes diverge further from the mean, though each individual step remains small and random.

Figure I.2: **Growth of Uncertainty in a Symmetric Random Walk**



Multiple realizations of a symmetric random walk over 50 steps, starting at zero. Each step is determined by a fair coin flip, moving up or down by one unit. The blue dashed lines represent the standard deviation envelopes ($\pm 1\sigma$, $\pm 2\sigma$, $\pm 3\sigma$), illustrating how uncertainty grows with time.

I.3. From Discrete to Continuous: The Central Limit Bridge

If we repeat the random walk many times, each consisting of (n) flips, and look at the distribution of final positions (X_n), we notice something striking: it begins to resemble a normal distribution as (n) grows large.

Mathematically, this is the Central Limit Theorem (CLT) at work. It states that the sum of many independent random variables – no matter their individual distributions – tends toward a Gaussian (normal) distribution, provided each has finite mean and variance.

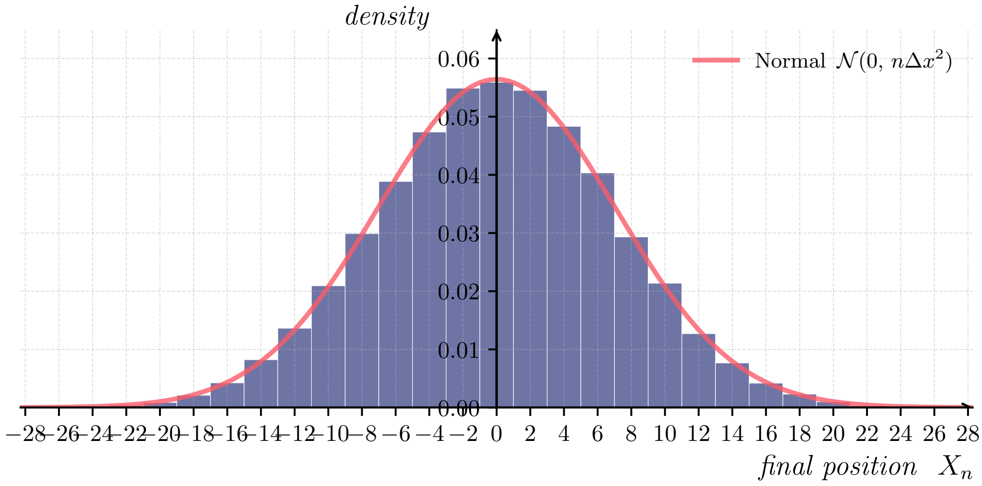
For our symmetric coin flips:

$$\frac{X_n - X_0}{\sqrt{n}\Delta x} \xrightarrow{d} \mathcal{N}(0,1)$$

as $(n \to \infty)$.

This result is profound: it justifies the leap from discrete randomness (coin flips) to continuous randomness (**Brownian motion**). The same logic underpins the stochastic modeling of stock prices in continuous time.

Figure I.3: **Distribution of Final Positions in a Symmetric Random Walk**



Histogram of final positions (X_n) after 50 steps across 10,000 simulated symmetric random walks. The red curve overlays the theoretical normal distribution $\mathcal{N}(0, n\Delta x^2)$, illustrating convergence to normality as predicted by the Central Limit Theorem.

II. From Random Walk to Brownian Motion

In the first section, we built a discrete world of randomness: a fair coin determined whether prices stepped up or down. But financial markets don't move in jumps of fixed size or at fixed intervals – prices evolve continuously, second by second, tick by tick. To model this, we must take the limit of our random walk as steps become infinitely small and frequent. The result of this elegant limiting process is Brownian motion, the cornerstone of continuous-time finance.

II.1. Refining the Random Walk

So far, our random walk advanced one step per "tick", as if each coin flip happened once per unit of time. But markets evolve in continuous time, so let's make the clock more fine-grained.

Suppose one unit of time – say, one day or second – is divided into many small intervals of equal length (Δt). Each coin flip (each random price move) now happens every (Δt) units of time. After (n) flips, the total elapsed time is

$$t = n\Delta t.$$

Now, as we make time finer (smaller (Δt)), we also need to adjust the size of each price move (Δx). If we kept (Δx) fixed while the number of flips per day went to infinity, the cumulative variance would blow up – the process would become infinitely wild. If we shrank (Δx) too quickly, the randomness would disappear entirely and the path would flatten out.

We want a limit that preserves the "texture" of randomness – small fluctuations that, when accumulated over time, produce finite and meaningful uncertainty. The right balance is achieved when the step size shrinks in proportion to the square root of the time step:

$$\Delta x = \sigma\sqrt{\Delta t}.$$

Here ($\sigma > 0$) is a constant representing the volatility, the intensity of the random movement.

This simple scaling rule has deep consequences. Let's check what happens to the variance after (n) steps:

$$\mathrm{Var}[X(t)] = n(\Delta x)^2 = n\sigma^2\Delta t = \sigma^2 t.$$

The variance now grows linearly with time. That means that if you double the time horizon, the expected spread of possible outcomes (the standard deviation) grows by ($\sqrt{2}$). This time-variance relationship is precisely what we observe empirically in financial returns – uncertainty accumulates smoothly over time, not explosively.

This diffusion-style scaling ($\Delta x = \sigma\sqrt{\Delta t}$) is the mathematical tuning fork that keeps the random walk "in tune" as we shrink the time grid. It ensures the process neither diverges nor vanishes but converges to a well-behaved continuous-time limit – the Brownian motion.

II.2. The Limiting Process

Now imagine making the time steps smaller and smaller – ($\Delta t \to 0$), and therefore ($n \to \infty$). The discrete sequence of partial sums (X_n) becomes a continuous curve (W_t).

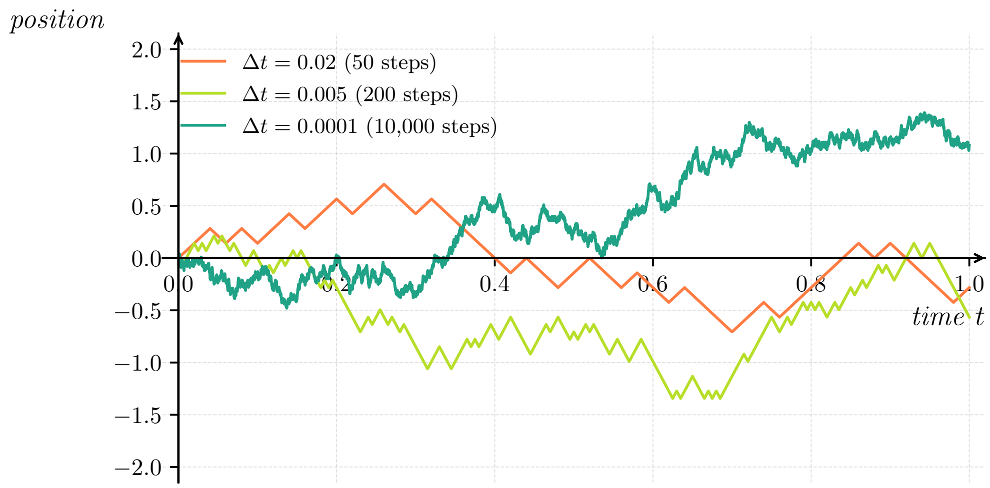
Mathematically, we write:

$$W_t = \lim_{\Delta t \to 0} \sum_{i=1}^{t/\Delta t} \xi_i,$$

where each (ξ_i) has mean zero and variance ($\sigma^2 \Delta t$).

The object (W_t) that emerges from this limit is called a Brownian motion (or Wiener process). Though it looks smooth when plotted, it is an extremely erratic function – continuous everywhere but differentiable nowhere. Its infinite "jaggedness" is what gives continuous-time finance its probabilistic richness.

Figure II.1: **From Discrete to Continuous: Symmetric Random Walks as $\Delta t \to 0$**



Three realizations of symmetric random walks with decreasing time steps ($\Delta t = 0.02$, 0.005, and 0.0001), illustrating how the discrete zig-zag paths converge to a continuous curve as $\Delta t \to 0$. The finest path closely resembles a Brownian motion trajectory.

II.3. Definition and Properties of Brownian Motion

A standard Brownian motion (W_t) is a continuous-time stochastic process with the following properties:

1. **Initial value**: $W_0 = 0$, representing the reference point or origin of uncertainty.
2. **Independent increments**: for any two time points $0 \le s < t$, the increment $W_t - W_s$ is statistically independent of everything that happened before time s.
3. **Normal increments**: $W_t - W_s \sim \mathcal{N}(0, t-s)$, meaning increments are normally distributed with variance equal to elapsed time. This mirrors the Central Limit Theorem we saw earlier: as the number of small random steps increases, their sum tends toward a Gaussian shape.

4. **Continuity**: the paths $(t \mapsto W_t)$ are continuous with probability 1. That means the Brownian path never makes sudden jumps or gaps – it flows smoothly through time.

III. The Stochastic Process with Drift and Diffusion

Up to now, we have seen that the standard Brownian motion W_t represents a martingale (a.1) with zero mean and variance increasing linearly over time. It captures randomness in its purest, unbiased form – an erratic wanderer without a preferred direction. However, real-world quantities such as stock prices, interest rates, or physical systems are rarely so indifferent: they tend to exhibit a systematic trend alongside random fluctuations.

We begin with the idea that over an infinitesimally small time step Δt, the change in X_t can be decomposed into two parts:

$$\Delta X_t = \text{deterministic part} + \text{random part}.$$

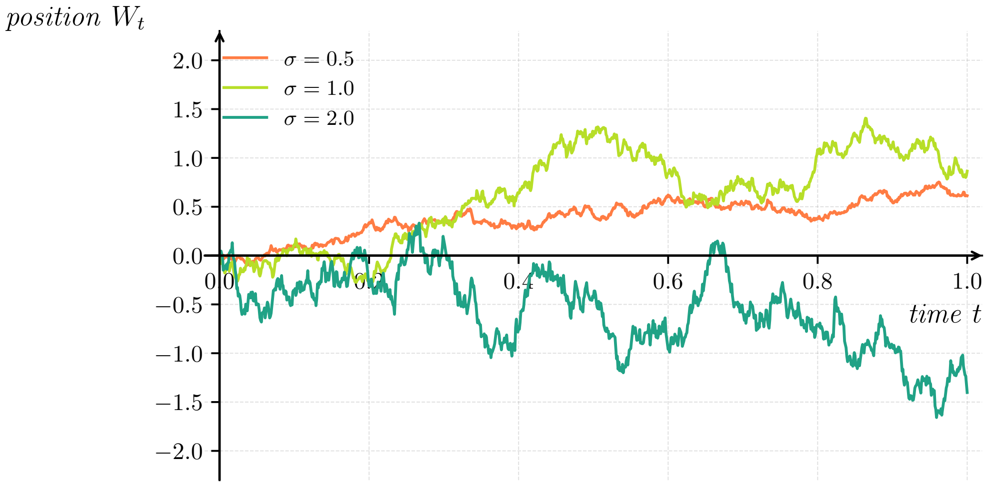
The deterministic part represents a predictable drift, while the random part captures the inherent uncertainty.

III.1. Scaling Brownian Motion: Introducing the Random Component

If we multiply a standard Brownian motion by a constant σ, we obtain a scaled Brownian motion σW_t. This new process has variance $\mathrm{Var}(\sigma W_t) = \sigma^2 t$. The parameter σ acts as the volatility coefficient – it is called the *diffusion coefficient* in mathematical terms. It controls the amplitude of the random fluctuations.

In financial models, σ tells us how turbulent the price process is: a small σ gives gentle fluctuations; a large σ makes the motion wilder. This scaled version σW_t is what we actually use in the Geometric Brownian Motion (GBM) and Black-Scholes model, where it drives the stochastic term in the differential equation for asset prices.

Figure III.1: **Scaled Brownian Motions with Different Volatilities σ**



Three realizations of scaled Brownian motions with different volatilities ($\sigma = 0.5$, 1.0, and 2.0), illustrating how the amplitude of fluctuations changes with σ. The paths demonstrate the impact of volatility on the behavior of the process.

III.2. Introducing Drift: The Deterministic Component

While the diffusion term σW_t introduces randomness, most real-world processes also exhibit a predictable, directional trend over time. This component is called the *drift* and is represented by a constant rate μ.

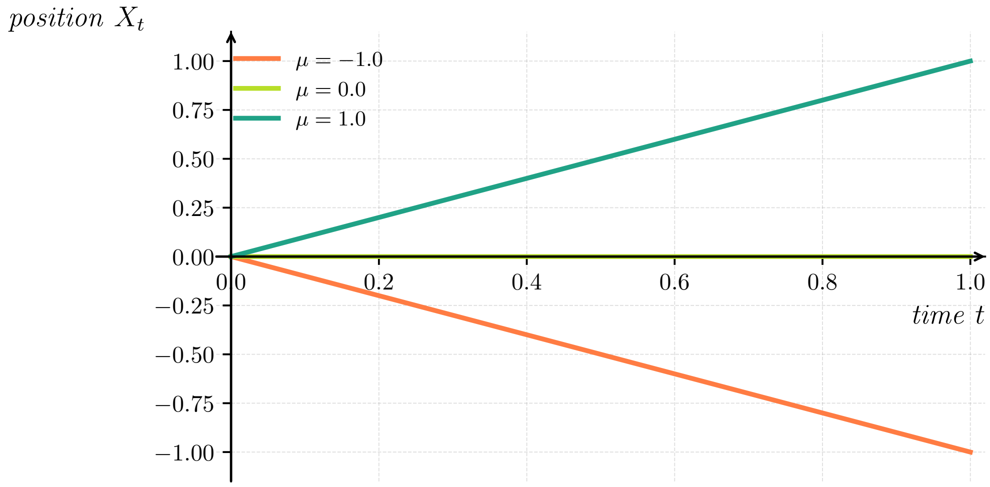
Mathematically, a pure drift process evolves as

$$X_t = X_0 + \mu t,$$

where μ measures the expected rate of change per unit time. Unlike Brownian motion, which oscillates unpredictably around zero, this process increases or decreases linearly and deterministically, depending on the sign of μ.

Economically, the drift μ can represent the expected return of a financial asset, the growth rate of a firm's output, or any steady tendency underlying random fluctuations. A positive drift indicates upward tendency, while a negative drift indicates decline.

Figure III.2: **Pure Drift Processes with Different Drift Parameters μ**



Three realizations of deterministic drift processes with $\mu = -1.0$, $\mu = 0.0$, and $\mu = +1.0$. The paths illustrate how the drift parameter controls the direction and slope of the mean trajectory over time. Unlike Brownian motion, there is no randomness – only a steady trend.

III.3. Combining Drift and Diffusion: The Stochastic Differential Equation

While the diffusion term determines the spread around this path, the drift determines where the center of that spread moves through time. Together, they produce the characteristic shape of a stochastic process that both trends and fluctuates.

Mathematically, we can express this combination using a stochastic differential equation SDE (a.2):

$$dX_t = \mu dt + \sigma dW_t,$$

where:

- dX_t is the infinitesimal change in the process,
- μdt is the deterministic drift component,
- σdW_t is the stochastic diffusion component.

This SDE captures the essence of many financial models, where asset prices are influenced by both predictable trends and random shocks.

III.4. From Finite to Infinitesimal Increments

To show how we arrive at the SDE, let's revisit the finite increment ΔX_t over a small time step Δt. The leap from finite increments ΔX_t to infinitesimal ones dX_t is the transition from a discrete random walk to a continuous-time process.

In discrete time, we write

$$\Delta X_t = \mu \Delta t + \sigma \varepsilon_t \sqrt{\Delta t},$$

where $\varepsilon_t \sim N(0,1)$.

As $\Delta t \to 0$, the scaled sum of these independent Gaussian shocks converges to Brownian motion W_t. By construction of W_t, its infinitesimal increment dW_t represents the limit of these random shocks, which satisfies

$$E[dW_t] = 0, \qquad Var[dW_t] = dt.$$

This reflects that Brownian motion has independent, normally distributed increments with mean zero and variance proportional to elapsed time.

Hence, taking the continuous limit yields

$$dX_t = \mu dt + \sigma dW_t,$$

the infinitesimal evolution law combining deterministic drift and stochastic diffusion.

IV. From Ordinary to Stochastic Calculus

Before we can derive the geometric Brownian motion and ultimately the Black-Scholes model, we must first understand how to handle calculus when randomness is involved. In the ordinary world of smooth curves and predictable functions, the rules of differentiation and Taylor expansion work perfectly well. But once we introduce Brownian motion, the mathematical landscape changes – paths become rough, and small time steps carry unpredictable "jumps" that we can't ignore.

This is why we first need stochastic calculus: a version of calculus adapted to random motion. Its cornerstone is Itô's Lemma, which will soon let us correctly differentiate functions of stochastic processes like asset prices.

IV.1. Why Ordinary Calculus Fails

In ordinary calculus, we study smooth functions – curves that can be approximated locally by straight lines. When a function $f(t,x)$ depends on a smooth variable $x(t)$, we can expand it using the Taylor series:

$$f(t+\Delta t, x+\Delta x) = f(t,x) + f_t \Delta t + f_x \Delta x + \tfrac{1}{2} f_{xx} (\Delta x)^2 + \cdots$$

Now, if $x(t)$ is smooth, its change over a small time step behaves roughly like $(\Delta x \approx x'(t), \Delta t)$. That means

$$(\Delta x)^2 \approx (x'(t))^2 (\Delta t)^2.$$

When we take the limit $\Delta t \to 0$, the term $(\Delta t)^2$ becomes negligible compared to Δt. Therefore, we can safely ignore all higher-order terms, keeping only:

$$df = f_t, dt + f_x, dx.$$

This simplification is what makes ordinary calculus elegant – smooth curves are "locally linear", so their tiny variations behave predictably.

In stochastic calculus, however, that trick breaks. Brownian motion W_t isn't smooth. Its increments don't shrink linearly with time; they shrink only like the square root of time:

$$\Delta W_t \sim \mathcal{N}(0, \Delta t).$$

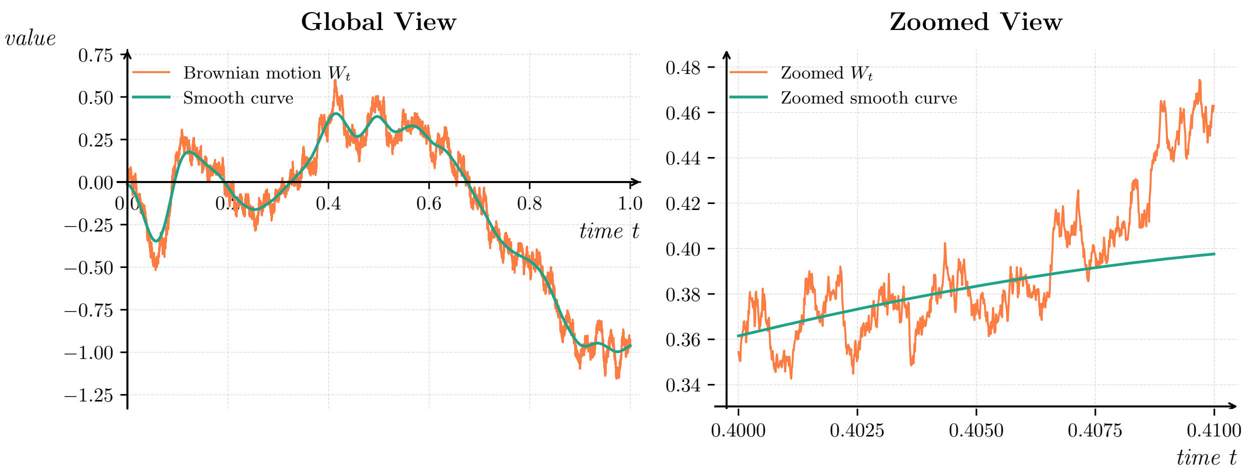
This means that while Δt might be tiny, ΔW_t is roughly of size $\sqrt{\Delta t}$.

Now square it:

$$(\Delta W_t)^2 \approx \Delta t.$$

Unlike the deterministic case, this term no longer vanishes faster than Δt; it's of the same order. So in the stochastic world, the "higher-order" term isn't negligible – it's essential. That's why the standard calculus formula must be modified, and why Itô's Lemma keeps the extra $\frac{1}{2}\sigma^2 f_{xx} dt$ term, as we will soon see.

Figure IV.1: **Smooth Curve vs. Brownian Motion at Different Scales**



At a broad scale (left), both the smooth curve and the Brownian motion appear similarly irregular. However, when we zoom in (right), the difference becomes clear: the smooth curve straightens and becomes locally linear, while the Brownian path remains jagged at every scale.

IV.2. Drift and Diffusion: Two Kinds of Motion

Let's now recall our general stochastic process:

$$dX_t = \mu dt + \sigma dW_t.$$

It contains two parts:

- The drift (μdt): the smooth, deterministic trend (expected change over time).
- The diffusion (σdW_t): the random, unpredictable fluctuation driven by Brownian motion.

The diffusion term introduces "second-order" randomness. It ensures that the squared increment $(dX_t)^2$ is proportional to dt. That nonzero second-order term will be what gives rise to Itô's Lemma.

IV.3. The Taylor Expansion with Randomness

In ordinary calculus, the Taylor expansion (a.3) of a function $f(X_t)$ is

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

and the higher powers of dX_t vanish. Now substitute our stochastic process $dX_t = \mu dt + \sigma dW_t$ into this expansion:

$$(dX_t)^2 = (\mu dt + \sigma dW_t)^2 = \mu^2 \underbrace{(dt)^2}_{1.\ \approx 0} + 2\mu\sigma \underbrace{dt dW_t}_{2.\ \approx 0} + \sigma^2 \underbrace{(dW_t)^2}_{3.\ \sim dt}.$$

As we shrink dt toward zero:

1. $(dt)^2$ becomes negligible compared to dt.
2. $dtdW_t$ also vanishes, since dW_t is of size $\sqrt{dt}$.
3. However, $(dW_t)^2$ behaves like dt, because the variance of dW_t is dt.

Thus the only term that remains is $(dX_t)^2 = \sigma^2 dt$. Plugging this back into the Taylor expansion gives:

$$df = f'(X_t)(\mu dt + \sigma dW_t) + \tfrac{1}{2} f''(X_t)\sigma^2 dt.$$

Note that we only keep terms up to second-order, because higher powers of dW_t shrink much faster than dt. Recall that dW_t behaves like $\sqrt{dt}$. That means:

$$(dW_t)^2 \sim dt, \quad (dW_t)^3 \sim (dt)^{3/2}, \quad (dW_t)^4 \sim (dt)^2, \text{ and so on.}$$

When $dt \to 0$, all terms involving powers higher than $(dW_t)^2$ vanish much faster than dt itself – they have no lasting contribution in the limit. Therefore, the second-order term is the highest one that still matters, because it captures the nonzero "quadratic variation" of Brownian motion. All higher-order terms disappear, leaving only up to the second order in the Itô expansion.

IV.4. Arriving at Itô's Lemma

At this point, we can summarize what we have derived: when randomness enters the picture through Brownian motion, the usual rules of calculus must adapt. The key difference lies in the nonvanishing second-order term – it carries real information about the process's variability.

We now want to know how a smooth function of a stochastic process, say $f(X_t, t)$, evolves over time. In other words, if X_t itself follows a stochastic path, how does this affect the motion of $f(X_t, t)$?

To find out, we extend our expansion to include the time-dependence of f:

$$df = f_t dt + f_x dX_t + \tfrac{1}{2} f_{xx} (dX_t)^2.$$

Here, f_t and f_x denote the partial derivatives of f with respect to time t and the process X_t, respectively. We now substitute our stochastic process:

$$dX_t = \mu dt + \sigma dW_t$$

Expanding step by step gives:

$$\begin{aligned} df &= f_t dt + f_x(\mu dt + \sigma dW_t) + \tfrac{1}{2} f_{xx}(\mu dt + \sigma dW_t)^2 \\ &= f_t dt + f_x(\mu dt + \sigma dW_t) + \tfrac{1}{2} f_{xx}\sigma^2 (dW_t)^2 \\ &= f_t dt + f_x \mu dt + f_x \sigma dW_t + \tfrac{1}{2} f_{xx}\sigma^2 dt, \end{aligned}$$

since $(dW_t)^2 = dt$ and all higher-order terms vanish. Collecting terms by powers of dt and dW_t yields:

$$df = \underbrace{\left(f_t + \mu f_x + \tfrac{1}{2}\sigma^2 f_{xx}\right)}_{\text{drift (deterministic) term}} dt + \underbrace{\sigma f_x}_{\text{diffusion (random) term}} dW_t$$

This is Itô's Lemma – the stochastic equivalent of the chain rule.

The first group of terms describes the average or expected change in f over time, while the second term captures the random fluctuations caused by the Brownian motion. The additional $\frac{1}{2}\sigma^2 f_{xx}, dt$ term is what distinguishes stochastic calculus from ordinary calculus: it represents the persistent influence of randomness on the average drift of the function.

Alternatively, Itô's Lemma is often written in fully expanded partial-derivative notation, which highlights its structure in applications to finance:

$$dF(t, X_t) = \frac{\partial F}{\partial t}\, dt + \mu \frac{\partial F}{\partial X}\, dt + \sigma \frac{\partial F}{\partial X}\, dW + \frac{1}{2} \frac{\partial^2 F}{\partial X^2}\, \sigma^2\, dt.$$

This is exactly the same result as before, but written in a form directly compatible with the Black-Scholes model, where F represents a function of time t and an underlying variable X_t, which evolves according to $dX_t = \mu dt + \sigma dW_t$.

The four terms have a clear interpretation:

- $\frac{\partial F}{\partial t} dt$: explicit time dependence,
- $\mu \frac{\partial F}{\partial X} dt$: deterministic drift,
- $\sigma \frac{\partial F}{\partial X} dW$: random diffusion,
- $\frac{1}{2}\sigma^2 \frac{\partial^2 F}{\partial X^2} dt$: curvature correction from randomness.

This lemma (a.4) will serve as the foundation for everything that follows, especially for deriving the dynamics of the geometric Brownian motion and the Black-Scholes equation.

V. Geometric Brownian Motion

In financial modeling, one of the foundational assumptions is that asset prices evolve continuously over time and exhibit random fluctuations that cannot be perfectly predicted. A mathematical description of such behavior can be obtained through stochastic processes, where uncertainty is explicitly modeled. The most widely used model of this type for asset prices is the Geometric Brownian Motion (GBM), which forms the backbone of the Black-Scholes framework.

V.1. From Arithmetic to Geometric Brownian Motion

To motivate the GBM, consider first the earlier introduced general stochastic process known as the Arithmetic Brownian Motion (ABM):

$$dX_t = \mu dt + \sigma dW_t,$$

where μ represents the deterministic drift (average rate of change), σ denotes the diffusion coefficient (volatility), and W_t is a standard Brownian motion with independent and normally distributed increments.

The Arithmetic Brownian Motion assumes that changes in X_t are additive: the variable evolves by adding a deterministic trend and a random shock. However, this is problematic when modeling financial prices. A stock price cannot take negative values, yet ABM permits such outcomes, as randomness could easily push X_t below zero. Moreover, investors typically think in terms of relative rather than absolute changes in value – for example, a 5% increase has the same proportional meaning regardless of whether the price is €10 or €100.

This reasoning motivates the introduction of the Geometric Brownian Motion (GBM), where percentage changes (or returns) follow a Brownian process:

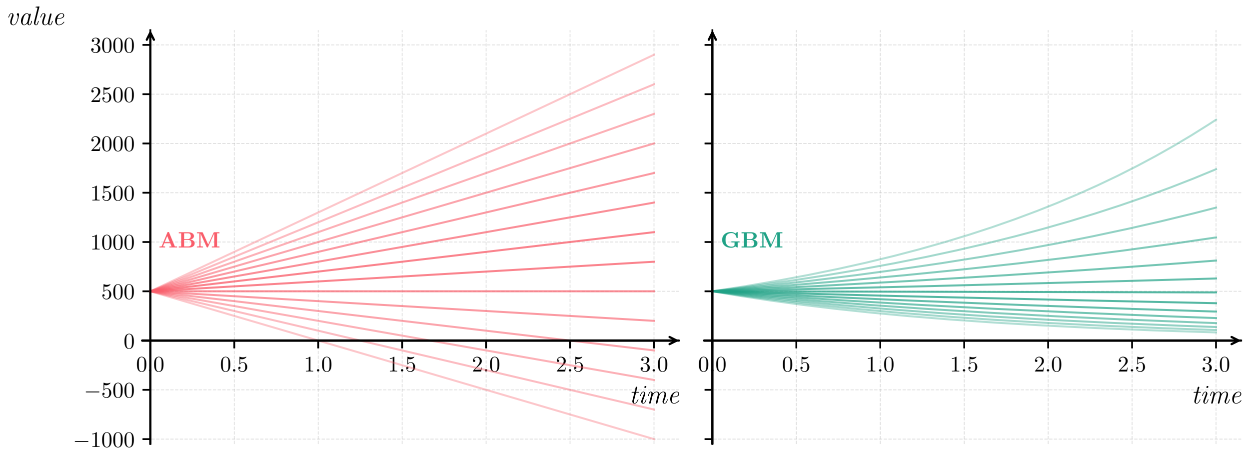
$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t.$$

Multiplying both sides by S_t yields the stochastic differential equation

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

Here, S_t denotes the asset price at time t, μ its expected instantaneous rate of return, and σ its volatility. Unlike the arithmetic case, this process ensures that $S_t > 0$ for all t, as the randomness acts multiplicatively. This property aligns well with financial intuition – price changes compound continuously over time, just like interest or growth rates.

Figure V.1: **Price Paths under Arithmetic vs. Geometric Brownian Motion**



The arithmetic model adds noise linearly, allowing negative values, while the geometric model applies proportional (multiplicative) noise, producing strictly positive, exponentially diverging paths.

V.2. Logarithmic Transformation and the Role of the Exponential Function

The adjective "geometric" in GBM reflects that the process exhibits exponential growth. To see this, it is instructive to take logarithms. Let

$$X_t = \ln S_t.$$

Since $S_t = e^{X_t}$, the dynamics of S_t translate into the dynamics of X_t, which often turn out to be more analytically tractable. The logarithm effectively converts multiplicative growth into additive increments, making it possible to analyze returns using standard linear tools. In finance, this corresponds to working with log returns, which aggregate over time in a simple additive fashion.

The next step is to determine how $X_t = \ln S_t$ evolves when S_t follows a GBM. This transformation cannot be performed using ordinary calculus, since S_t is not a smooth function of time – it contains stochastic noise. Instead, we must apply Itô's Lemma, a fundamental result in stochastic calculus that generalizes the chain rule to random processes.

V.3. Applying Itô's Lemma

Given $dS_t = \mu S_t dt + \sigma S_t dW_t$ and $f(S_t) = \ln S_t$, Itô's Lemma states:

$$df = f'(S_t)dS_t + \tfrac{1}{2}f''(S_t)(dS_t)^2.$$

The derivatives of f are:

$$f'(S_t) = \frac{1}{S_t}, \quad f''(S_t) = -\frac{1}{S_t^2}.$$

We also use the stochastic property $(dW_t)^2 = dt$, implying $(dS_t)^2 = \sigma^2 S_t^2 dt$. Substituting, we obtain:

$$d(\ln S_t) = \frac{1}{S_t}(\mu S_t dt + \sigma S_t dW_t) - \tfrac{1}{2}\frac{1}{S_t^2}(\sigma^2 S_t^2 dt).$$

Simplifying terms gives the differential equation for the logarithm of the price:

$$d(\ln S_t) = (\mu - \tfrac{1}{2}\sigma^2)dt + \sigma dW_t.$$

This is a stochastic process with a deterministic drift $(\mu - \frac{1}{2}\sigma^2)$ and volatility σ.

V.4. Integration and the Closed-Form Solution

Integrating both sides from the initial time 0 to any future time t gives:

$$\int_0^t d(\ln S_u) = \int_0^t (\mu - \tfrac{1}{2}\sigma^2)du + \int_0^t \sigma dW_u.$$

The left-hand side integrates directly to:

$$\ln S_t - \ln S_0.$$

The first integral on the right-hand side is purely deterministic:

$$\int_0^t (\mu - \tfrac{1}{2}\sigma^2)du = (\mu - \tfrac{1}{2}\sigma^2)t.$$

The second integral is a stochastic integral. Since W_t is a standard Brownian motion, this integral is normally distributed with mean 0 and variance $\sigma^2 t$. In particular:

$$\int_0^t \sigma dW_u = \sigma W_t.$$

Here we use the fact that $W_t - W_0 = W_t$ has variance t.

Putting it all together gives the integrated form:

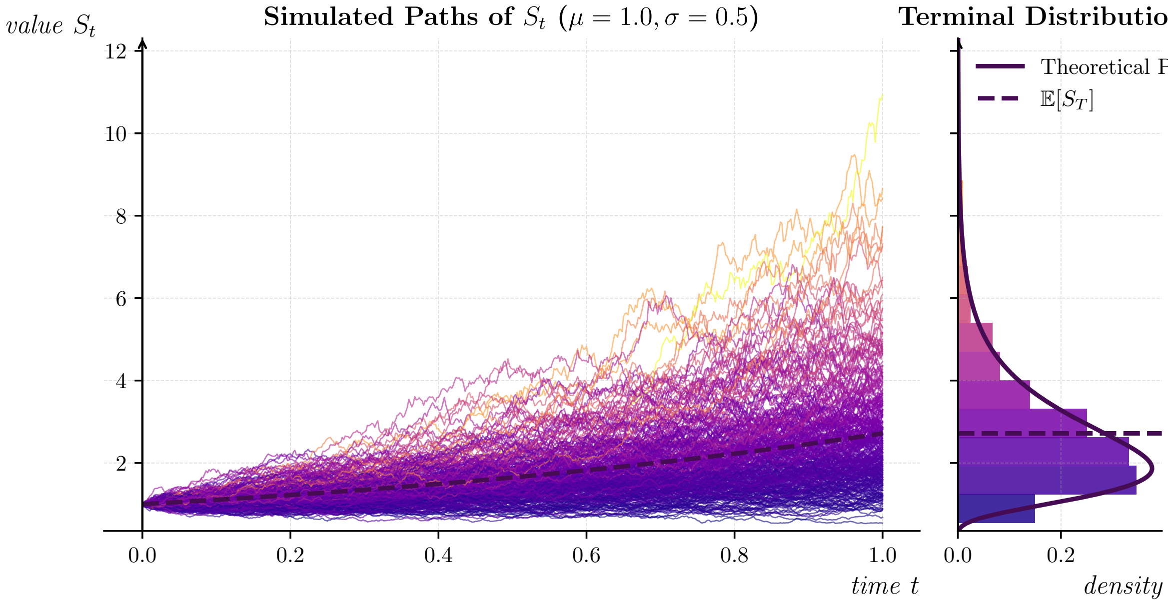
$$\ln S_t - \ln S_0 = (\mu - \tfrac{1}{2}\sigma^2)t + \sigma W_t.$$

To express the process in terms of S_t rather than $\ln S_t$, we exponentiate both sides:

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}.$$

This expression yields the closed-form solution (a.5) for the Geometric Brownian Motion. It clearly illustrates that S_t is lognormally distributed, since $\ln S_t$ is normally distributed.

Figure V.2: **Geometric Brownian Motion Paths Generated by the Closed-Form Solution**



The left panel displays multiple simulated trajectories of S_t under a Geometric Brownian Motion with drift $\mu = 1.0$ and volatility $\sigma = 0.5$. Each path's color corresponds to its terminal value S_T, creating a visual continuity with the right panel, which shows the resulting terminal distribution. The histogram illustrates the lognormal nature of S_T, overlaid with its theoretical probability density and the expected value $E[S_T]$ (dashed line). Together, the panels demonstrate how multiplicative stochastic growth leads to a right-skewed distribution of outcomes.

V.5. Interpretation: Expected Return and Volatility

The GBM neatly decomposes the evolution of prices into two components: a deterministic exponential trend and a stochastic, random fluctuation. The expected rate of return is governed by μ, while σ captures the amplitude of uncertainty.

The term $-\frac{1}{2}\sigma^2$ arises from the curvature correction in Itô's Lemma and has an important economic interpretation. It indicates that higher volatility reduces the expected growth rate of log prices. This phenomenon, often referred to as volatility drag, reflects the fact that uncertainty lowers the geometric (compound) average return even if the arithmetic mean return μ remains unchanged.

Thus, the GBM elegantly captures two fundamental aspects of asset prices:

- Expected return (μ): the central tendency or drift of the price process.
- Volatility (σ): the magnitude of randomness that drives price fluctuations.

The interplay of these parameters defines the entire probabilistic structure of future prices – a relationship that will be pivotal in deriving the Black-Scholes equation in the subsequent section.

VI. Pricing Principles

Up to this point, we have described the behaviour of an asset's price using the geometric Brownian motion (GBM):

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

This tells us how a stock evolves under the influence of drift and randomness. But the next question is fundamentally different: given this uncertain evolution, how should we determine the fair value of a derivative whose payoff depends on S_t? To answer that, we introduce three central ideas in modern financial economics: hedging, no-arbitrage, and risk-neutral valuation.

VI.1. Hedging: Removing Risk through Combination

So far, we have described how the underlying asset, such as a stock, evolves randomly through time. But in modern financial markets, investors often trade not only the asset itself but also contracts whose value depends on it – these are called derivatives.

A derivative is any financial instrument whose price is derived from another underlying variable, typically a stock price, an interest rate, or an index. Among many types of derivatives (futures, forwards, swaps, etc.), the most fundamental for our purpose is the option.

An option gives its holder the right, but not the obligation, to buy or sell the underlying asset at a predetermined price (the strike price) on or before a certain date (the maturity). For example, a European call option allows its holder to buy the stock at the strike price on the expiration date T, while a put option allows selling it at T.

Thus, the option's value $V(S_t, t)$ depends on both the current stock price S_t, which evolves stochastically, and the remaining time to maturity $T - t$.

Because S_t is random, the option's value is random as well. However, as Black and Scholes famously realized, it is possible to construct a combination of the option and the underlying stock that eliminates this randomness over an infinitesimal time interval. That's the idea of hedging.

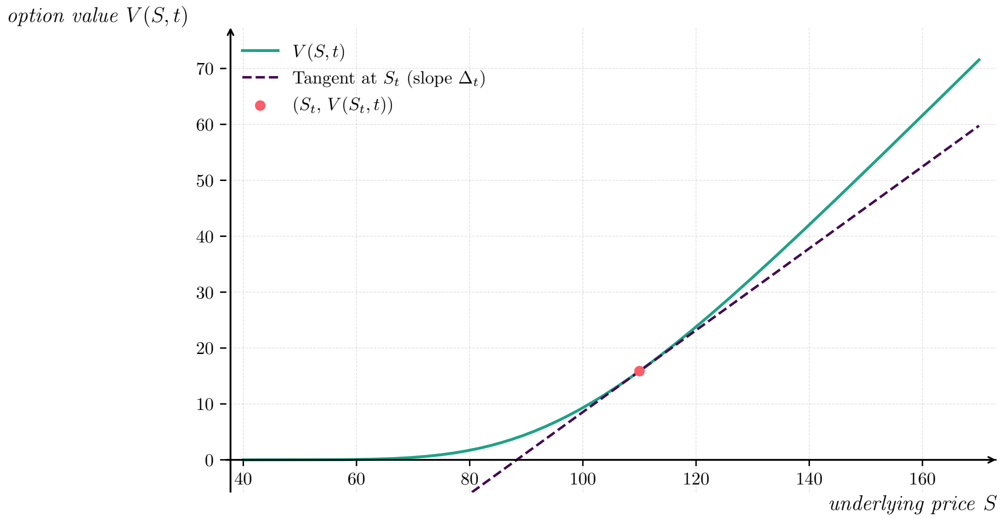
Imagine forming a portfolio Π_t consisting of:

$$\Pi_t = V(S_t, t) - \Delta S_t,$$

where Δ is the number of shares held short (if $\Delta > 0$) or long (if $\Delta < 0$). The goal is to choose Δ such that the random component – the term involving dW_t – disappears from $d\Pi_t$. If this can be achieved, the portfolio becomes locally risk-free over a small time step dt.

This idea – that a combination of risky assets can eliminate risk – is the seed of replicating portfolios and the key to the Black-Scholes construction.

Figure VI.1: **Local Delta-Hedge: Option Value and Tangent at S_t**



This figure illustrates how the option value $V(S_t,t)$ can be locally approximated by a straight line with slope $\Delta_t = \frac{\partial V(S_t,t)}{\partial S_t}$. The curved line represents the non-linear relationship between the option price and the underlying asset price, while the dashed tangent shows how holding Δ_t units of the underlying replicates the option's instantaneous sensitivity to small changes in S_t. By constructing this local linear combination, the random component of the portfolio's value can be neutralized – forming the core idea of delta-hedging in the Black-Scholes framework.

VI.2. No-Arbitrage: Equal Risk, Equal Return

Once the portfolio is made risk-free, logic demands that it must earn the same rate of return as any other risk-free investment. If it didn't – say, if it earned more than the risk-free rate r – investors could borrow money at r, form the same portfolio, and make a guaranteed profit. Such ”free money” opportunities are called arbitrage, and their absence is a cornerstone assumption of rational markets.

Therefore, the value of the hedged portfolio Π_t must satisfy:

$$d\Pi_t = r\Pi_t dt.$$

This condition – equality of returns between the riskless portfolio and the risk-free asset – translates the economic principle of no-arbitrage into a mathematical constraint.

VI.3. Risk-Neutral Valuation: Changing the Measure

So far, everything we've done – the stochastic model, the hedging argument, and the no-arbitrage condition – has been expressed in the real world, the one investors actually live in. In this world, the stock's expected growth rate μ exceeds the risk-free rate r because investors require a risk premium for bearing uncertainty. The GBM we derived earlier reflects that reality (denoted by the physical measure $\mathbb{P}$):

$$dS_t = \mu S_t dt + \sigma S_t dW_t^{\mathbb{P}}.$$

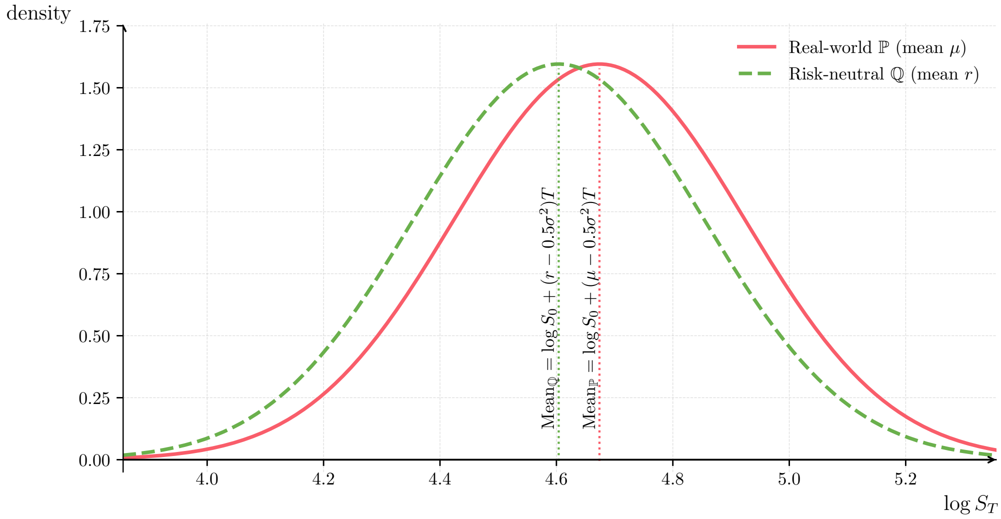
Now imagine a different world – a fictional but very convenient one – where investors are completely indifferent to risk. In this risk-neutral world, nobody demands extra return for uncertainty, because risk can be perfectly hedged away (as we just saw with the replicating portfolio). If risk can be eliminated, then it should no longer be rewarded.

That single idea changes everything: in the risk-neutral world, every asset, whether risky or not, grows on average at the same rate as the risk-free asset. Mathematically, the stock's drift μ is replaced by r:

$$dS_t = rS_t dt + \sigma S_t dW_t^{\mathbb{Q}}.$$

This is not a new stochastic process – it's the same random evolution viewed through a different probabilistic lens, called the risk-neutral measure $\mathbb{Q}$. Under this measure, prices evolve as if all investors were risk-neutral.

Figure VI.2: **Real-World vs. Risk-Neutral Stock Price Dynamics**



This figure illustrates the probability density of the logarithmic terminal price, $\log S_T$*, under the real-world measure* $\mathbb{P}$ *(drift* μ*) and the risk-neutral measure* $\mathbb{Q}$ *(drift* r*). Both distributions share the same variance,* $\sigma^2 T$*, but differ in their means. This horizontal shift visually represents the transition from the real world – where investors demand a risk premium – to the risk-neutral world, where all assets are expected to grow at the risk-free rate (*r*).*

VII. The Black-Scholes Model

In the previous sections, we developed the stochastic foundation of asset price dynamics using the Geometric Brownian Motion (GBM) and explored how randomness affects any function of that process through Itô's Lemma. However, our objective now shifts from describing how prices behave to determining what the fair value of a derivative on that asset should be.

The bridge between these two worlds lies in three central ideas: hedging, no-arbitrage, and risk-neutral valuation. Together, they express a single truth – the fair price of a derivative is the one that prevents any riskless profit. Even though the underlying price follows a random path, it is possible to construct a continuously hedged portfolio that completely eliminates this randomness. Once such a riskless combination exists, logic dictates that it must earn the same rate of return as any other risk-free investment. Otherwise, an arbitrage opportunity would arise.

The Black-Scholes model formalizes this intuition mathematically. By applying Itô's Lemma to the derivative value $V(S_t, t)$, substituting the GBM dynamics for S_t, and enforcing the no-arbitrage condition that the hedged portfolio earns the risk-free rate r, we obtain the Black-Scholes partial differential equation – the mathematical core of modern option pricing theory.

VII.1. Deriving the Black-Scholes Equation

We start from the option value $V(S_t, t)$, which depends on both the underlying price S_t and time t. $V(S_t, t)$ denotes the derivative's value, but it is also often denoted as $C(S_t, t)$ for a call option or $P(S_t, t)$ for a put option. Since S_t follows a stochastic process, the option value will also move randomly. By applying Itô's Lemma, we can express its infinitesimal change as:

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(dS_t)^2.$$

The underlying price evolves according to the geometric Brownian motion:

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

where μ is the expected return, σ the volatility, and dW_t the random shock (the increment of a Wiener process or Brownian motion). Substituting this into the expression for dV gives:

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}(\mu S_t dt + \sigma S_t dW_t) + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(\mu S_t dt + \sigma S_t dW_t)^2,$$

remembering that $(dW_t)^2 = dt$:

$$dV = \frac{\partial V}{\partial t}dt + \mu S_t \frac{\partial V}{\partial S}dt + \sigma S_t \frac{\partial V}{\partial S}dW_t + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S_t^2 dt,$$

then collecting terms yields:

$$dV = \left(\frac{\partial V}{\partial t} + \mu S_t \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S_t \frac{\partial V}{\partial S}dW_t.$$

Both S_t and $V(S_t, t)$ depend on the same random term dW_t. This is what makes both the stock and the option risky – they are exposed to the same underlying source of uncertainty.

VII.2. Constructing a Riskless Portfolio

To remove this randomness, we construct a portfolio that combines the derivative and the underlying stock in such proportions that the random parts cancel. Let this portfolio be:

$$\Pi_t = V(S_t, t) - \Delta S_t,$$

where Δ represents the number of shares of the underlying held short (if $\Delta > 0$) or long (if $\Delta < 0$).

The small change in the portfolio value over time dt is given by:

$$d\Pi_t = dV - \Delta dS_t.$$

Substituting the earlier expressions for dV and dS_t gives:

$$d\Pi_t = \left(\frac{\partial V}{\partial t} + \mu S_t \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S_t \frac{\partial V}{\partial S}dW_t - \Delta(\mu S_t dt + \sigma S_t dW_t),$$

then rearranging terms yields:

$$d\Pi_t = \left(\frac{\partial V}{\partial t} + \mu S_t \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} - \Delta \mu S_t\right)dt + \left(\sigma S_t \frac{\partial V}{\partial S} - \Delta \sigma S_t\right)dW_t.$$

The second bracket contains the random term multiplied by dW_t. We can make the portfolio risk-free by choosing Δ such that this random part disappears. To eliminate it, set:

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

This choice makes the entire coefficient of dW_t equal to zero. In words, by holding exactly $\frac{\partial V}{\partial S}$ shares of the underlying stock for each option, we perfectly offset the instantaneous risk of small price movements.

Substituting this value of Δ back into the expression for $d\Pi_t$, we obtain:

$$d\Pi_t = \left(\frac{\partial V}{\partial t} + \mu S_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S^2} - \mu S_t\frac{\partial V}{\partial S}\right)dt + \left(\sigma S_t\frac{\partial V}{\partial S} - \sigma S_t\frac{\partial V}{\partial S}\right)dW_t,$$

simplifies to:

$$d\Pi_t = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S^2}\right)dt.$$

The randomness has been completely removed – the portfolio now evolves deterministically. It is locally risk-free over an infinitesimal time step.

VII.3. Applying the No-Arbitrage Condition

A portfolio that carries no risk must earn the same rate of return as any other risk-free investment, otherwise arbitrage opportunities would arise. This means that the return on the hedged portfolio must equal the risk-free rate r:

$$d\Pi_t = r\Pi_t dt.$$

Since the portfolio's value is $\Pi_t = V - \Delta S_t$, substituting this in gives:

$$d\Pi_t = r(V - \Delta S_t)dt.$$

Finally, we substitute $\Delta = \frac{\partial V}{\partial S}$ and the earlier deterministic expression for $d\Pi_t$:

$$\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S^2}\right)dt = r\left(V - S_t\frac{\partial V}{\partial S}\right)dt.$$

Since dt appears on both sides, we can divide through by it to obtain the differential equation:

$$\boxed{\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S^2} + rS_t\frac{\partial V}{\partial S} - rV = 0.}$$

This is the Black-Scholes partial differential equation, the mathematical backbone of modern option pricing. It expresses the precise balance required to prevent arbitrage in a continuously hedged market.

The beauty of this result is that the stock's expected return μ – which depends on investors' risk preferences – has disappeared entirely. Once risk is eliminated through hedging, the only rate that matters is the risk-free rate r. This is why the equation is universal: it prices all derivatives that can be perfectly hedged, independent of subjective expectations.

VII.4. Interpreting the Equation's Terms

Each term in Equation has a clear financial meaning:

- $\frac{\partial V}{\partial t}$ represents the time decay of the derivative,
- $\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}$ captures the effect of curvature – how the option's sensitivity changes as the underlying moves,
- $rS\frac{\partial V}{\partial S}$ reflects the risk-neutral drift of the underlying required by replication, growing at the risk-free rate under the risk-neutral measure,
- $-rV$ represents the opportunity cost of capital – the return that could be earned by investing the same amount in a risk-free bond.

The Black–Scholes equation expresses a fundamental equilibrium: the instantaneous expected return on the derivative must equal the return from a risk-free investment. If this balance were violated, investors could construct a self-financing arbitrage strategy.

VII.5. Solving the Equation for European Options

Once we have the Black-Scholes partial differential equation,

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

we can apply it to the valuation of a specific derivative. Because our main object of interest is the European call option, we now denote the value function by $C(S, t)$. The contract pays $\max(S_T - K, 0)$ at maturity, so the terminal condition is:

$$C(S_T, T) = \max(S_T - K, 0).$$

This payoff is the boundary condition from which we must solve the differential equation backwards in time to obtain the fair price at any earlier moment $t < T$.

To solve the PDE, a sequence of transformations is applied: time is reversed via $\tau = T - t$, the stock price is expressed in logarithmic form, and the variables are rescaled so that the PDE becomes the classical heat equation (a.6). This step is technical, but the outcome is elegant: the solution takes the form of a discounted risk-neutral expectation. Converting back to the original variables yields the closed-form price of the European call:

$$C(S_t, t) = S_t N(p_+) - Ke^{-r(T-t)} N(p_-),$$

where $N(\cdot)$ is the cumulative distribution function of the standard normal distribution, and the quantities p_+ and p_- are the expressions that emerge from the change of variables:

$$p_+ = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \qquad p_- = p_+ - \sigma\sqrt{T-t}.$$

p_+ and p_- come from the probability distribution of the log-price under the risk-neutral measure (a.7). In the transformed coordinates, the terminal payoff becomes a kinked function of a normally distributed variable. Undoing those transformations brings exactly these combinations in the numerator and denominator – hence their central role in the final expression.

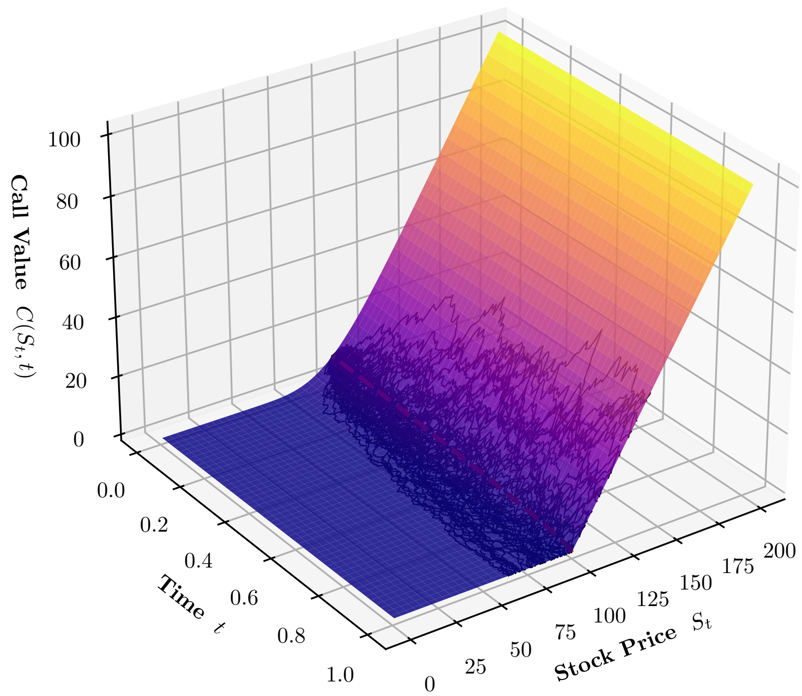
The two terms in the formula have intuitive interpretations:

- $S_t N(p_+)$ represents the expected value, under the risk-neutral measure, of receiving one share of the stock at maturity if the option finishes in the money.
- $Ke^{-r(T-t)} N(p_-)$ represents the present value of paying the strike price K upon exercise.

Their difference gives the fair, arbitrage-free value of the call option at any earlier time $t < T$.

This closed-form solution captures all the key parameters that influence an option's price: the current stock price S_t, strike price K, time to maturity $T - t$, risk-free rate r, and volatility σ. Each enters the equation in a way that reflects its economic role – volatility increases the option's value by expanding the range of possible outcomes, while higher interest rates reduce the present value of the strike payment.

Figure VII.1: **Option Value Surface under the Black–Scholes Model**



The surface shows the arbitrage-free price of a European call option as a function of the underlying price S_t and time t, obtained from the Black-Scholes formula. The simulated geometric Brownian motion paths (in black) represent possible evolutions of the underlying asset. Each path lies on the surface at its corresponding (S_t, t), illustrating that although the underlying moves randomly, the option's value evolves smoothly and deterministically once hedging and no-arbitrage conditions are imposed. The red curve highlights the strike region, where the transition into the payoff at maturity becomes sharp as $t \to T$.

VII.6. The European Put Price via Put-Call Parity

Once we have the expression for the call price $C(S_t, t)$, the price of a European put option follows directly from the put-call parity relationship:

$$C(S_t, t) - P(S_t, t) = S_t - Ke^{-r(T-t)}.$$

Solving for the put price gives:

$$P(S_t, t) = C(S_t, t) - S_t + Ke^{-r(T-t)}.$$

Substituting the call solution into this expression, we obtain:

$$P(S_t, t) = Ke^{-r(T-t)}N(-p_-) - S_t N(-p_+).$$

The same quantities p_+ and p_- appear – for the simple reason that the underlying price process and the transformed normal variable governing the distribution of $\log(S_T)$ are the same for both calls and puts. The only difference lies in the direction of payoff asymmetry.

Thus, with a single PDE and a single solution technique, both the European call and put prices emerge cleanly and consistently.

VIII. The Option Greeks

Once we have analytical expressions for the prices of European call and put options, we can differentiate these prices with respect to the model's inputs. These derivatives – called the Greeks – measure how the option's value responds to small changes in the underlying parameters. They translate mathematical sensitivity into economic meaning. Although they arise formally from calculus, each Greek captures a very concrete idea about how option prices behave.

Because the call price is

$$C(S,t) = SN(p_+) - Ke^{-r(T-t)}N(p_-),$$

and the put price is

$$P(S,t) = Ke^{-r(T-t)}N(-p_-) - SN(-p_+),$$

we can obtain all the Greeks directly by differentiating these formulas. What matters more than the formal derivatives, however, is their interpretation – what they mean and how they relate to risk, hedging, and market behaviour.

VIII.1. Delta

Delta measures how much the option price changes when the underlying price changes by a tiny amount:

$$\Delta_{\text{call}} = \frac{\partial C}{\partial S}, \qquad \Delta_{\text{put}} = \frac{\partial P}{\partial S}.$$

For the European call, differentiation gives:

$$\Delta_{\text{call}} = N(p_+).$$

This is a beautifully simple result. $N(p_+)$ is the risk-neutral probability that the option will finish in the money (a.8), but it also appears as the slope of the option price with respect to the underlying. In geometric terms, Delta tells you how steeply the option value curve rises at the current price S. When the call is far out of the money, Delta is close to 0; deep in the money, it approaches 1. In between, it smoothly transitions, capturing how "option-like" or "stock-like" the payoff is at this instant.

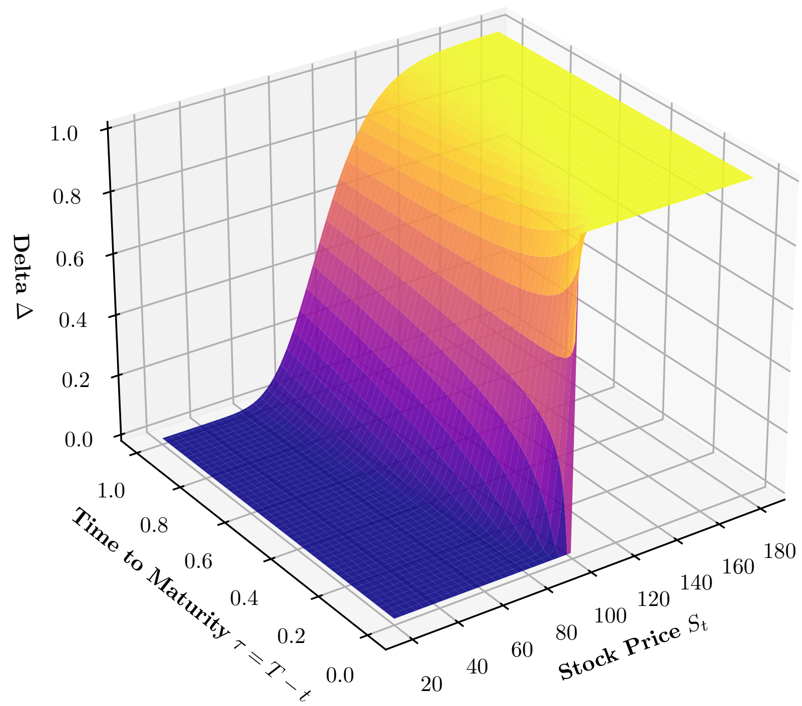
For the put,

$$\Delta_{\text{put}} = N(p_+) - 1 = -N(-p_+),$$

so Delta is always negative: when the asset price rises, the put loses value.

Economically, Delta is the quantity that drives the hedging ratio. It is exactly the Δ used in the construction of the riskless portfolio earlier.

Figure VIII.1: **Delta Surface $\Delta_{\mathbf{call}}(S, \tau)$ of a European Call Option**



This surface shows how the call option's Delta smoothly transitions from 0 to 1 as the stock price increases. The ridge along higher S_t values reflects the option becoming increasingly stock-like. As time to maturity shrinks, the transition becomes sharper, illustrating how Delta collapses to a step function at expiration.

VIII.2. Gamma

Gamma measures how Delta itself changes when the underlying moves:

$$\Gamma = \frac{\partial^2 C}{\partial S^2} = \frac{\partial^2 P}{\partial S^2}.$$

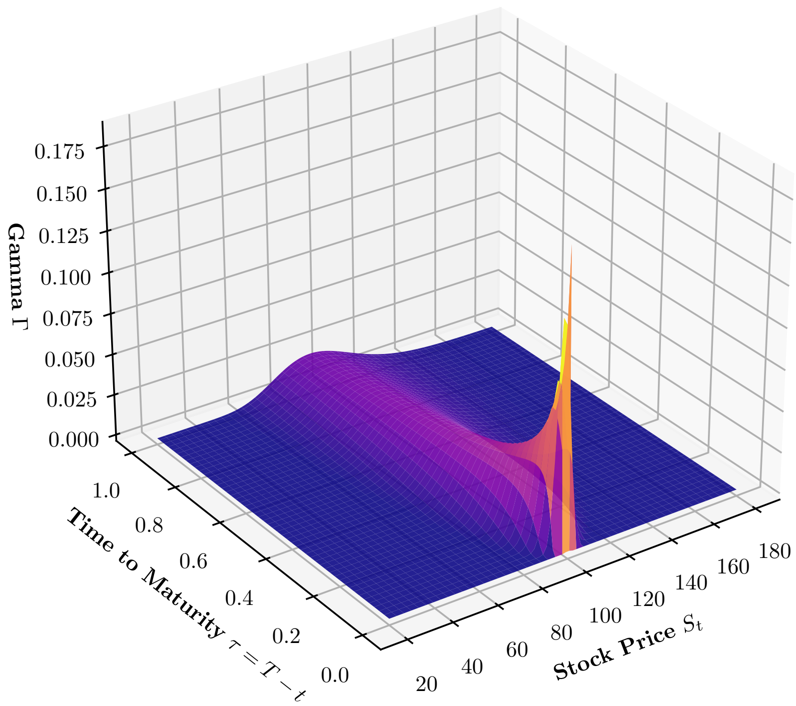
For both calls and puts in the Black-Scholes framework, Gamma is:

$$\Gamma = \frac{N'(p_+)}{S\sigma\sqrt{T-t}},$$

where N' is the standard normal density. Gamma is always positive for both calls and puts, reflecting the convexity of option prices: they benefit from volatility and curvature.

A high Gamma means Delta reacts strongly to small price movements – the option is very sensitive, and the hedge must be rebalanced frequently. Gamma tends to be largest when the option is at the money and close to expiry, because small changes around this region sharply affect whether the payoff will be zero or positive.

Figure VIII.2: **Gamma Surface $\Gamma(S, \tau)$ of a European Call/Put Option**



Gamma peaks sharply when the option is near the strike and close to maturity, forming a thin ridge centered around the at-the-money region. This visualises how the option's curvature – and thus the sensitivity of Delta – becomes extremely concentrated as expiry approaches, while deep in- and out-of-the-money regions remain flat with almost no Gamma.

VIII.3. Theta

Theta measures how the option's price changes as time passes:

$$\Theta = \frac{\partial C}{\partial t}, \qquad \Theta = \frac{\partial P}{\partial t}.$$

For a European call option,

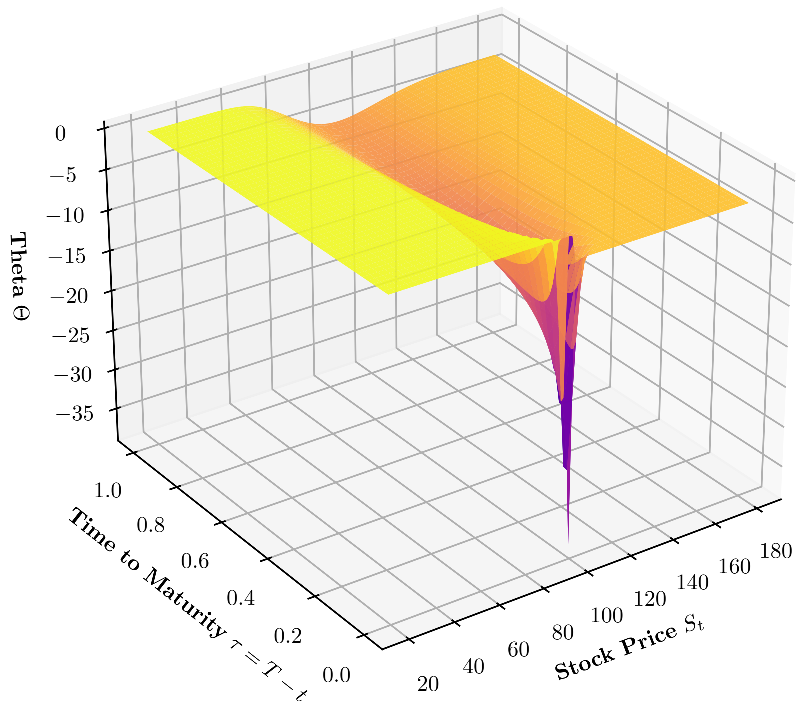
$$\Theta_{\text{call}} = -\frac{SN'(p_+)\sigma}{2\sqrt{T-t}} - rKe^{-r(T-t)}N(p_-).$$

The first term reflects the "decay" of the option's time value. As maturity approaches, the option has fewer opportunities to move into the money, and this erosion is captured in the negative sign. The second term reflects the effect of discounting the strike price.

Theta is usually negative for calls and puts – options lose value as time runs out – although deep in-the-money puts can have positive Theta in certain cases.

In practical trading, Theta represents the "cost" of holding optionality: even if nothing changes in the market, the option's value tends to drift downward with the passage of time.

Figure VIII.3: **Theta Surface $\Theta_{\mathbf{call}}(S, \tau)$ of a European Call Option**



Theta is most negative when the option is near the strike and close to expiration, forming a deep downward spike in the surface. This represents the rapid erosion of time value for near-the-money options approaching maturity, while far in- or out-of-the-money regions lose value much more slowly.

VIII.4. Vega

Vega (not a letter in the Greek alphabet; the name arises from misreading the Greek letter ν) measures how the option price responds to a change in volatility:

$$\nu = \frac{\partial C}{\partial \sigma} = \frac{\partial P}{\partial \sigma}.$$

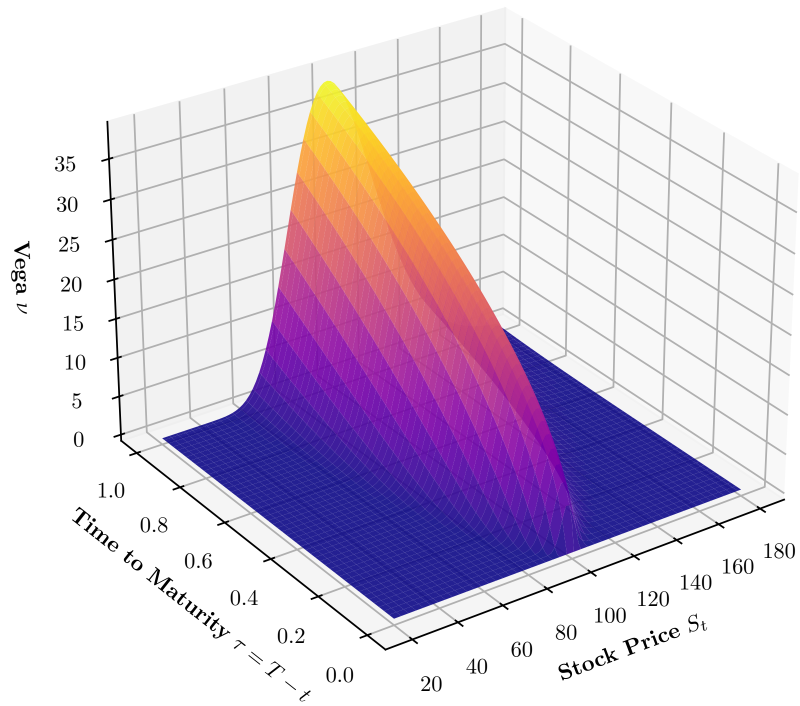
For both calls and puts,

$$\nu = SN'(p_+)\sqrt{T-t}.$$

This expresses the intuitive fact that options become more valuable when volatility increases, because greater volatility means a wider distribution of future outcomes and more chance of landing in the money.

Vega is highest when the option is at the money – exactly the point where the payoff has the most uncertainty – and declines as the option becomes deeply in or out of the money.

Figure VIII.4: **Vega Surface $\nu(S, \tau)$ of a European Call/Put Option**



Vega forms a smooth, bell-shaped ridge centered at the at-the-money region, with height increasing as maturity grows. This shows that options are most sensitive to volatility when they are near the strike and have substantial time remaining, while short-dated or deep ITM/OTM options have very little Vega.

VIII.5. Rho

Rho measures the sensitivity of the option to the risk-free interest rate:

$$\rho_{\text{call}} = \frac{\partial C}{\partial r}, \qquad \rho_{\text{put}} = \frac{\partial P}{\partial r}.$$

For the call,

$$\rho_{\text{call}} = K(T-t)e^{-r(T-t)}N(p_-),$$

which is positive: higher interest rates increase call values because the strike becomes cheaper in present-value terms.

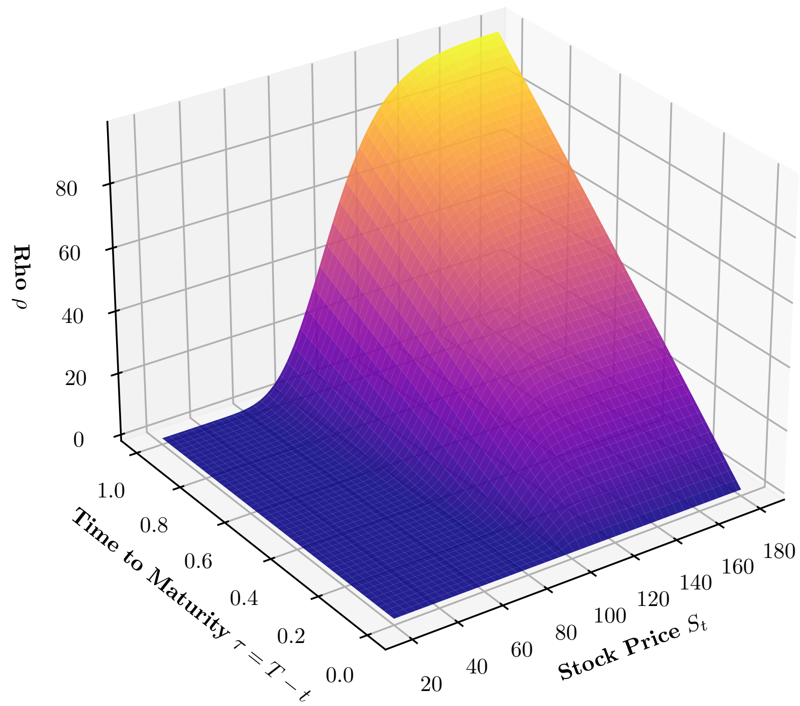
For the put,

$$\rho_{\text{put}} = -K(T-t)e^{-r(T-t)}N(-p_-),$$

which is negative: higher interest rates reduce the value of puts for the same reason.

Rho is usually the least important Greek for short-dated options but becomes meaningful for long-maturity contracts.

Figure VIII.5: **Rho Surface $\rho_{\text{call}}(S, \tau)$ of a European Call Option**



Rho rises with both the stock price and the time to maturity, reflecting how long-dated, deep-in-the-money calls are the most sensitive to interest rate changes. The nearly flat region near short maturities shows that Rho is negligible when little time remains before expiration.

VIII.6. The Greeks as a System

Taken together, the Greeks form a complete dictionary of how an option responds to its environment:

- **Delta** (Δ) measures directional sensitivity,
- **Gamma** (Γ) measures the curvature of that sensitivity,
- **Theta** (Θ) captures the erosion of value through time,
- **Vega** (ν) captures exposure to volatility, and
- **Rho** (ρ) captures exposure to interest rates.

Each Greek corresponds to one dimension of risk that traders must monitor and hedge. In continuous hedging theory, Delta and Gamma play the central role; in practical trading, Theta and Vega matter just as much.

The Greeks turn the seemingly complex behaviour of option prices into a structured, interpretable system – a map of sensitivities that arises naturally from the Black-Scholes framework.

IX. Final Remarks

The progression from a discrete coin-toss model to the continuous-time Black-Scholes framework illustrates how a simple stochastic mechanism, when taken to its diffusion limit, yields a mathematically tractable representation of asset dynamics. By applying Itô's Lemma to the Geometric Brownian Motion and imposing the principle of no-arbitrage through dynamic hedging, we arrive at a pricing equation whose closed-form solution defines the value of European options and their associated sensitivity measures, the Greeks. This pathway highlights the internal coherence of the model: randomness is formalised, hedging neutralises unpriced risk, and valuation emerges from risk-neutral expectations.

IX.1. Scope and Exclusions

The derivation presented applies exclusively to European-style claims. American options, which allow early exercise, fall outside the analytical structure of the Black-Scholes partial differential equation. Their valuation requires optimal stopping theory and numerical techniques such as finite-difference methods, binomial/trinomial trees, or simulation-based approaches. These instruments introduce discontinuities and free boundaries that the classical model does not accommodate.

IX.2. Limitations

The Black-Scholes model relies on a set of restrictive assumptions that create both its elegance and its shortcomings. Volatility is assumed constant, markets are frictionless, trading is continuous, interest rates remain fixed, and the underlying follows a lognormal diffusion without jumps. Empirical evidence – volatility smiles, stochastic volatility effects, price jumps, liquidity frictions, and discrete rebalancing costs – demonstrates that real markets deviate systematically from these idealised conditions. As a result, the formula is best interpreted as a first-order approximation rather than a literal description of market behaviour.

IX.3. Outlook

Despite these limitations, the Black-Scholes framework remains foundational. Its conceptual structure – dynamic replication, risk-neutral valuation, and the decomposition of option value into well-defined sensitivities – continues to inform both theory and practice. Modern developments such as local and stochastic volatility models, jump-diffusion processes, and data-driven volatility surfaces can be viewed as extensions that relax particular assumptions while retaining the core logic of the original approach. The model's future lies not in its ability to perfectly describe markets, but in its enduring role as the baseline against which more sophisticated methods are defined and evaluated.

Attribution

This study was written by **Milán Péter**. You are welcome to share or distribute this document in its original form, but please do not claim authorship or present its contents as your own work.

If you have questions, spot an error, or would like to discuss any part of the study, feel free to reach out: `peter.milan77@gmail.com`.

References

[1] Black, F. and M. Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81(3), pp. 637–654.

[2] Merton, R. C. (1973). Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science* 4(1), pp. 141–183.

[3] Hull, J. C. (2018). *Options, Futures, and Other Derivatives* (10th ed.). Pearson.

[4] Shreve, S. E. (2004). *Stochastic Calculus for Finance I: The Binomial Asset Pricing Model.* Springer.

[5] Shreve, S. E. (2004). *Stochastic Calculus for Finance II: Continuous-Time Models.* Springer.

[6] Björk, T. (2009). *Arbitrage Theory in Continuous Time* (3rd ed.). Oxford University Press.

[7] Øksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications* (6th ed.). Springer.

[8] Wilmott, P., S. Howison, and J. Dewynne (1995). *The Mathematics of Financial Derivatives.* Cambridge University Press.

a. Appendix

a.1. Martingale

A martingale is a stochastic process whose expected future value equals its current value, given all information available now. Formally, a process X_t is a martingale under a measure $\mathbb{Q}$ if

$$\mathbb{E}^{\mathbb{Q}}\left[X_t \mid \mathcal{F}_s\right] = X_s \quad \text{for all } t \geq s.$$

Intuition: it represents a fair game. There is no drift in the conditional expectation – no systematic way to win or lose on average.

In derivative pricing, martingales become crucial because discounted asset prices under the risk-neutral measure form martingales. This ensures that introducing derivatives does not create arbitrage. Once a price process becomes a martingale, its expected growth rate is the risk-free rate, and valuation reduces to taking risk-neutral expectations.

a.2. Stochastic Differential Equation

A stochastic differential equation describes how a random variable evolves in continuous time. The general form is

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t,$$

where μ is the drift term, σ the diffusion (volatility) term, and dW_t represents a Wiener process increment. An SDE defines the microscopic behaviour of a process. From an SDE, applying Itô's Lemma to a function $V(X_t, t)$ produces a partial differential equation (PDE) for the expected evolution of V. For example, applying Itô to a derivative price $V(S, t)$ under GBM leads directly to the Black-Scholes PDE, which must hold to avoid arbitrage.

Thus:

$$\text{SDE for } S_t \quad \xrightarrow{\text{Itô}} \quad \text{PDE for } V(S, t).$$

This is the conceptual bridge between stochastic modelling and pricing theory.

a.3. Taylor Expansion

The Taylor expansion approximates a smooth function by expressing it as a polynomial built from its derivatives. In one dimension:

$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x + \frac{1}{2}f''(x)\Delta x^2 + \cdots$$

In stochastic calculus, this becomes extremely important because the "small increment" Δx is random. Under Brownian motion,

$$(dW_t)^2 = dt,$$

so the usually negligible second-order term becomes significant. This is the reason why Itô's formula contains a second derivative term. Taylor expansion is the underlying idea behind Itô's Lemma.

a.4. What is a lemma?

A lemma is a mathematical statement proven primarily to support the proof of a more significant theorem. Itô's Lemma plays exactly this role: it is not the "main result", but it is the indispensable tool that allows us to derive the Black-Scholes PDE and many results in continuous-time finance.

So a lemma is a supporting brick in the architecture of a theory—small but critical.

a.5. Closed-Form Solution

A closed-form solution is a formula that can be written explicitly using a finite number of operations and standard functions (exponentials, logarithms, normals, etc.).

The Black-Scholes formula is a closed-form solution to the Black-Scholes PDE:

$$C(S,t) = SN(p_+) - Ke^{-r(T-t)}N(p_-).$$

Closed-form solutions are prized because they avoid numerical approximation, reveal structure directly, and allow for immediate calculation of Greeks.

In finance, closed-form solutions are rare – Black-Scholes is one of the few models where the mathematics aligns cleanly enough to permit one.

a.6. Classical Heat Equation

The classical heat equation describes how heat diffuses through a medium:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}.$$

In option pricing, the Black-Scholes PDE can be transformed into this exact same equation through a change of variables (rescaling of time, log-price substitution, and discounting).

The key insight: volatility in finance plays the same mathematical role as thermal diffusivity in physics. Both describe how uncertainty (in markets) or heat (in materials) spreads over time. This equivalence is why solving the Black-Scholes PDE ultimately reduces to solving the heat equation – one of the most well-understood diffusion equations in physics.

a.7. The Origin of p_+ and p_-

$$p_+ \quad \text{and} \quad p_-$$

represent the probabilities that the terminal log-price ends up above or below a certain threshold under the risk-neutral measure.

Under the risk-neutral measure $\mathbb{Q}$, the log-price is normally distributed:

$$\ln S_T \sim \mathcal{N}\left(\ln S_0 + \left(r - \tfrac{1}{2}\sigma^2\right)T, \sigma^2 T\right).$$

Thus,

- $p_+ = \mathbb{Q}\left(\ln S_T > \ln K\right) = N(p_-)$
- $p_- = 1 - p_+ = N(-p_-)$

These are not arbitrary symbols – they come directly from the cumulative distribution function of the log-price under the risk-neutral transformation. They are the probabilistic core of the Black-Scholes formula, translated into the familiar $N(p_+)$ and $N(p_-)$ terms.

a.8. Moneyness: ITM, ATM, OTM

Moneyness describes the relation between the underlying price S_t and the strike K, indicating whether immediate exercise is profitable.

In the Money (ITM): An option has positive intrinsic value.

$$\text{Call: } S_t > K, \qquad \text{Put: } S_t < K.$$

At the Money (ATM): The underlying price is approximately equal to the strike:

$$S_t \approx K.$$

Exercise yields almost zero payoff; time value is typically maximal.

Out of the Money (OTM): Immediate exercise gives zero payoff.

$$\text{Call: } S_t < K, \qquad \text{Put: } S_t > K.$$

Deep In / Out of the Money: *Deep ITM*: underlying far beyond the strike (e.g., $S_t \gg K$ for calls). Delta approaches ± 1; value dominated by intrinsic value. *Deep OTM*: underlying far away in the opposite direction (e.g., $S_t \ll K$ for calls). Low probability of finishing profitable; delta near 0.

Moneyness	**Call Condition**	**Put Condition**	**Meaning**
ITM	$S_t > K$	$S_t < K$	Profitable now
ATM	$S_t \approx K$	$S_t \approx K$	Maximum time value
OTM	$S_t < K$	$S_t > K$	Not profitable now
DITM	$S_t \gg K$	$S_t \ll K$	Option behaves like asset
DOTM	$S_t \ll K$	$S_t \gg K$	Very low payoff probability