Pricing Financial Options Using Monte Carlo and Finite Difference Methods in MATLAB

Dr. Lok Pati Tripathi
June 25, 2024

School of Mathematics and Computer Science
Indian Institute of Technology Goa, India

Faculty Development Program (FDP) on Financial Mathematics, SPDE Theory, Mathematical Modeling and Current Numerical Trends
(FINSMMCNT 2024) June 24 -28, 2024
organized by
the Department of Mathematics,
School of Advanced Sciences,VIT-AP University.

Derivative Security

A financial contract whose value depends upon or derived from one or more underlying assets such as Stocks, Bonds, Foreign Currency, Commodities, etc.

Option

An option gives the right, but not the obligation, to buy or sell a prescribed financial asset for a prescribed price (strike price) by a prescribed expiration date (maturity date).

K

Strike Price

T

Maturity Date

S_T

Asset Price at Maturity

European style call option

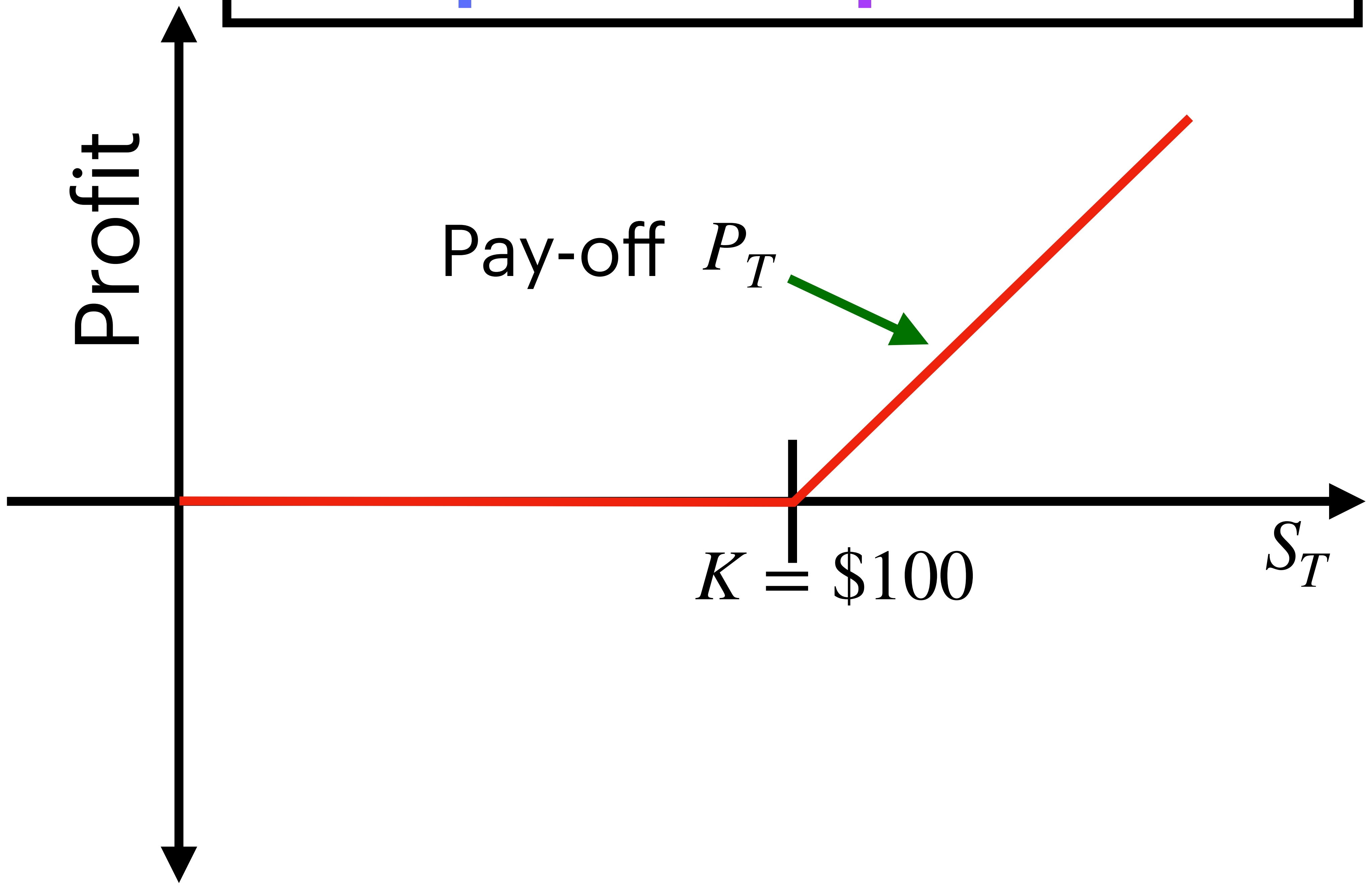
A European style **call** option gives its holder the right (but not the obligation) to **buy** a prescribed asset S_T from the writer of the option for strike price K at the maturity date T.

Thus, the fair price P of European style call option at the maturity date T is given by the pay-off

$$P_T = P(S_T, T) := \max(S_T - K, 0)$$

$$P_0 = P(S_0, 0) = ?$$

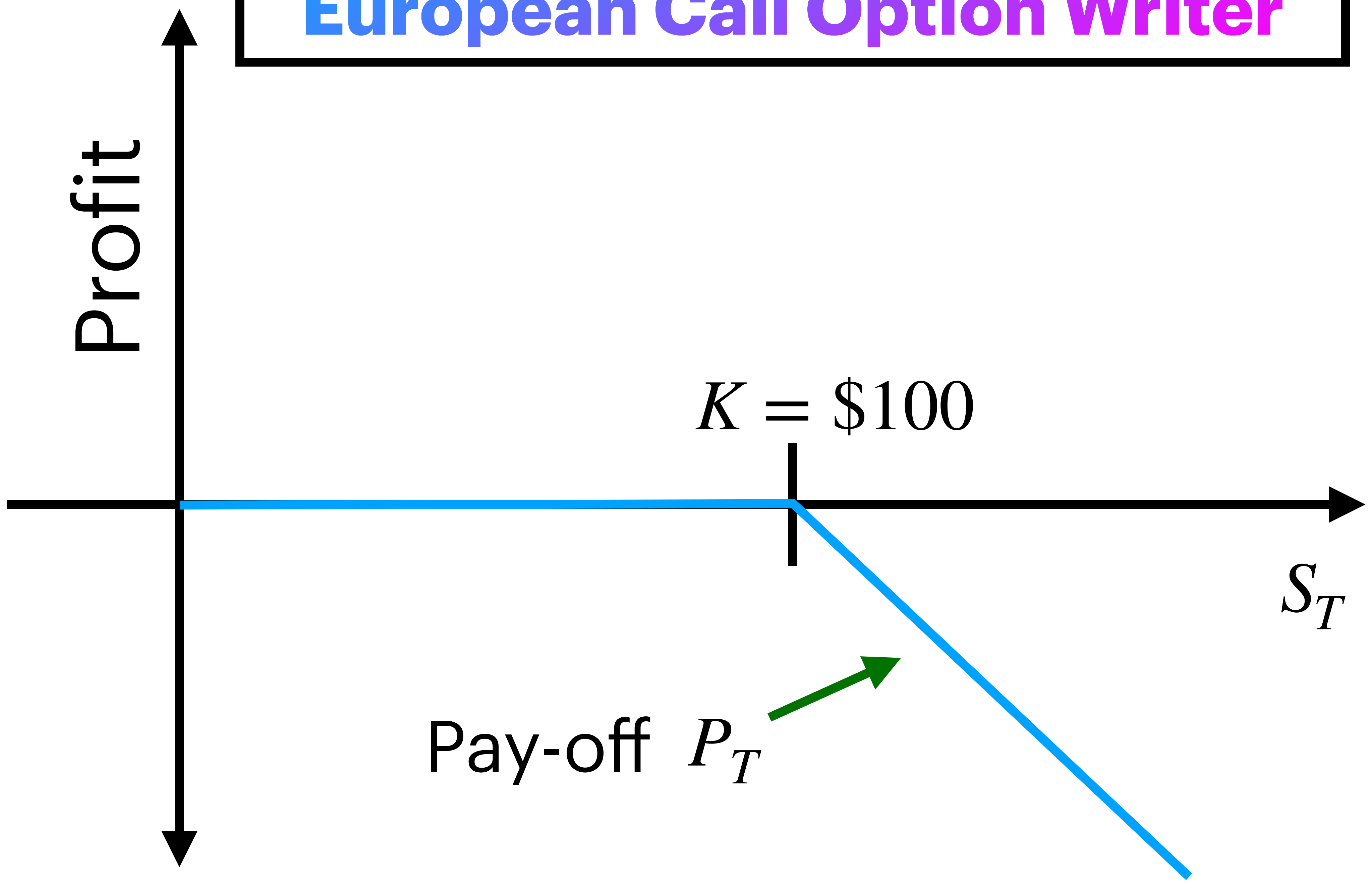
European Call Option Holder



S_T in \$	Profit (Y/N)	Exercise Option (Y/N)
80	N (0)	N
100	N (0)	N/Y
110	Y (+10)	Y
130	Y (+30)	Y
200	Y (+100)	Y

Never Loose!

European Call Option Writer

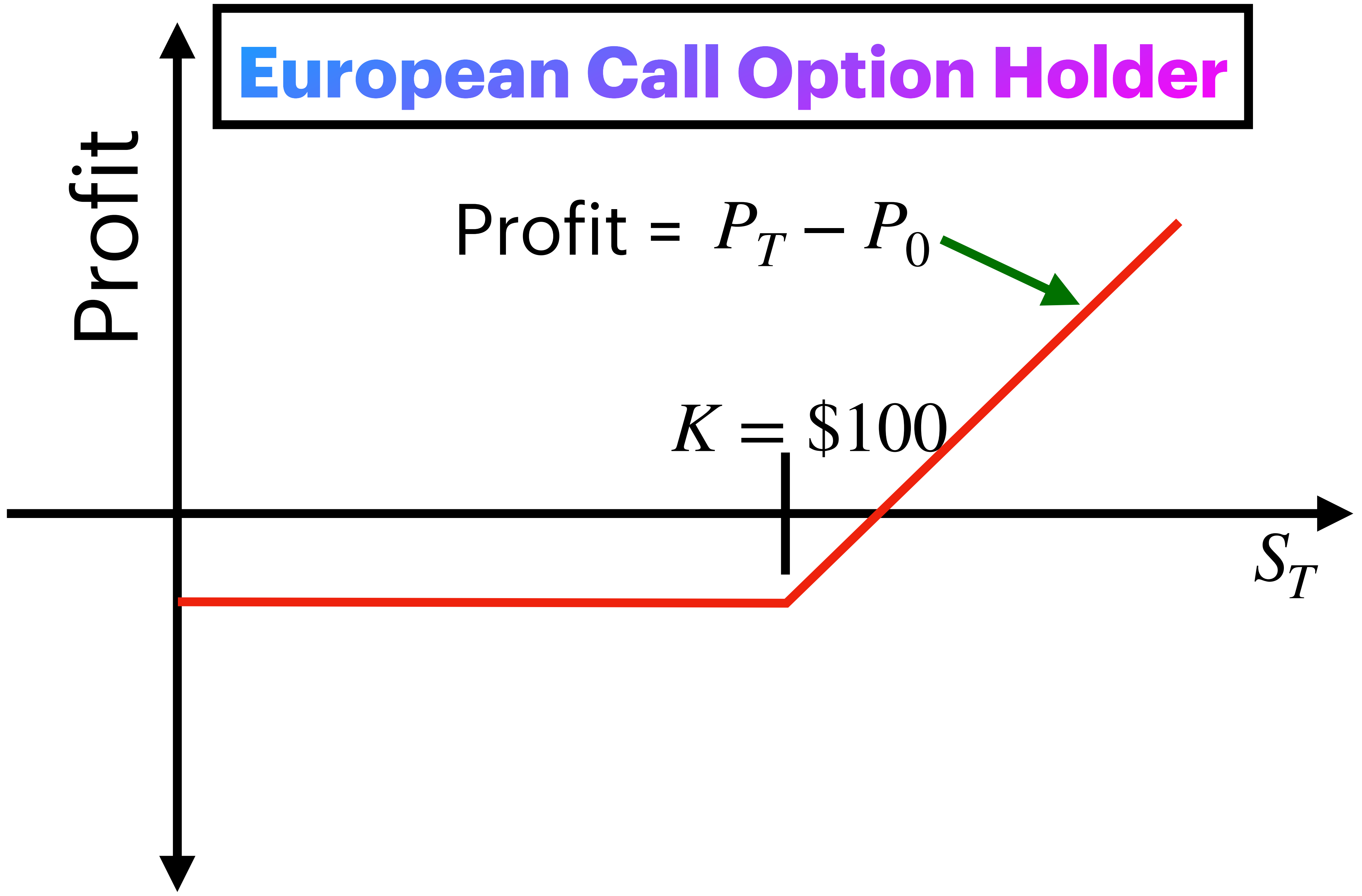


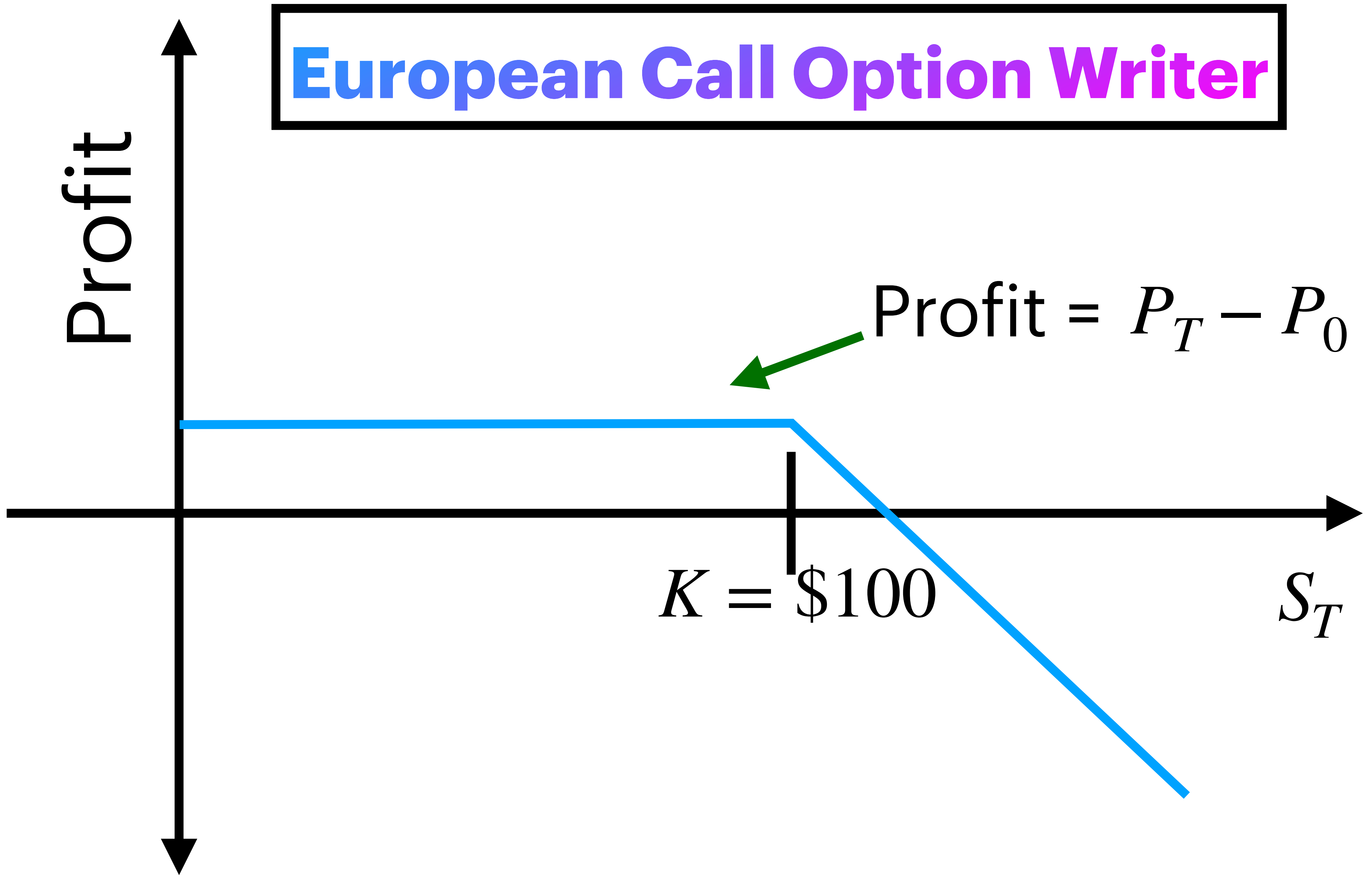
S_T in \$	Profit (Y/N)	Option Exercised (Y/N)
80	N (0)	N
100	N (0)	N/Y
110	N (-10)	Y
130	N (-30)	Y
200	N (-100)	Y

Never Gain!

No profit! Why should one write a Call Option?

Option writer demands a premium $P_0(=\$10)$ from the buyer of the option





S_T in \$	Profit (Y/N)	Exercise Option (Y/N)
80	N (-10)	N
100	N (-10)	N/Y
110	N (0)	Y
130	Y (+20)	Y
200	Y (+90)	Y

S_T in \$	Profit (Y/N)	Option Exercised (Y/N)
80	Y (+10)	N
100	Y (+10)	N/Y
110	N (0)	Y
130	N (-20)	Y
200	N (-90)	Y

Option Pricing

Now the problem is, how to find the fair premium price for the option at present time?

Under risk-neutrality assumption and the efficient market hypothesis (weak form), an option pricing formula also called Black-Scholes formula can be derived.

Risk-Neutrality Assumption

At any time, the average return on a risky investment of an asset is equal to the return on a risk-free investment of that asset.

European style call option: ***Pay-off function*** $P_T := \max(S_T - K, 0)$

Under risk-neutrality assumption,

$$\mathbb{E}[P_T] = P_0 e^{rT} \ , \quad \text{that is,} \quad P_0 = e^{-rT}\, \mathbb{E}[\max(S_T - K, 0)], \quad S_T = ?$$

Risky, and risk-free asset-price models

Risk-free asset-price model

Risk-free asset-price model: A Discrete Model

A Bond issued by the Government can be regarded a risk-free asset.

B_0 : Risk-free investment at time $t = 0$ (i.e., Depositing B_0 amount in a risk-free saving account at time $t = 0$)

r : annual interest rate

$B_1 = (1 + r)B_0$: The value of the investment after one year

$B_m = (1 + 2r)B_{m-2} = \cdots = (1 + mr)B_0$: The value of the investment after $m(>1)$ years (In case of Simple Interest Rate)

$B_m = (1 + r)^2 B_{m-2} = \cdots = (1 + r)^m B_0$: The value of the investment after $m(>1)$ years (In case of Compound Interest Rate)

Risk-free asset-price model: A Discrete Model

$t_n = n\delta t,\ n = 0,1,\ldots,N$: Trading time spots
$(0 = t_0 < t_1 < \cdots < t_N = T,\ \delta t = 1/N)$

N : Compounding frequency

$$\frac{B(t_n) - B(t_{n-1})}{B(t_{n-1})} = r\,\delta t, \quad n = 1,\ldots,N$$

$B(t_n)$: Amount received at trading time spot $t_n \in (0,T]$,
$n = 1,2,\ldots,N$, where t_n is measured in years

Risk-free asset-price model: A Continuous Model

Now, as

$$B(T) = B(t_N) = (1 + r\delta t)^N B_0 = \left[(1 + r\delta t)^{1/\delta t}\right]^T B_0,$$

if the interest being continuously compounded then

$$B(T) = \lim_{\delta t \to 0} \left[(1 + r\delta t)^{1/\delta t}\right]^T B_0 = e^{rT} B_0$$

Risky asset-price model

Risky asset-price model: A discrete model

Efficient market hypothesis (weak form) -

“The current asset price fully reflect all past trading information”. This includes historical price, trading volume, and any other data that could be desired from market trading activity.

With this hypothesis, changes in the asset price are a Markov process.

$$0 = t_0 \quad t_1 \quad t_2 \quad t_3 \qquad t_{n-1} \quad t_n \qquad T = t_N = N\delta t$$

Risk-free investment

Discrete Model:

$B(t_0) \; = \; S_0$, and for $n = 1,2,\ldots,N$,

$$\frac{B(t_n) - B(t_{n-1})}{B(t_{n-1})} \; = \; r\,\delta t$$

Continuous Model:

$$B(T) \; = \; S_0\, e^{rT}$$

Risky investment

Discrete Model:

$S(t_0) \; = \; S_0$, and for $n = 1,2,\ldots,N$,

$$\mathbb{E}\left[\frac{S(t_n) - S(t_{n-1})}{S(t_{n-1})}\right] \; = \; r\,\delta t$$ **(Risk-neutrality assumption)**

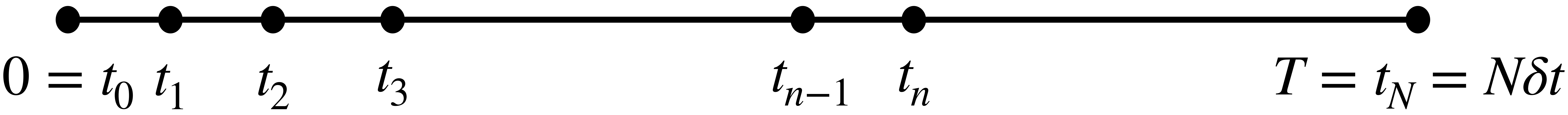
Efficient market hypothesis suggest that

$$\frac{S(t_n) - S(t_{n-1})}{S(t_{n-1})} \; = \; r\,\delta t + \sigma\sqrt{\delta t}\; Y_{n-1}, \quad \sigma > 0$$

where Y_n's are i.i.d. and $\mathbb{P}(Y_n = 1) = 0.5 = \mathbb{P}(Y_n = -1)$.

Continuous Model:

$$S(T) \; = \; S_0\, e^{?\cdots}$$



Risk-free investment

Discrete Model:

$B(t_0) \; = \; S_0$, and for $n = 1,2,\ldots,N$,

$$\frac{B(t_n) - B(t_{n-1})}{B(t_{n-1})} \; = \; r\,\delta t$$

Continuous Model:

$$B(T) \; = \; S_0\, e^{rT}$$

Risky investment

Discrete Model:

$S(t_0) \; = \; S_0$, and for $n = 1,2,\ldots,N$,

$$\mathbb{E}\left[\frac{S(t_n) - S(t_{n-1})}{S(t_{n-1})}\right] \; = \; r\,\delta t$$ **(Risk-neutrality assumption)**

Efficient market hypothesis suggest that

$$\frac{S(t_n) - S(t_{n-1})}{S(t_{n-1})} \; = \; r\,\delta t + \sigma \sqrt{\delta t}\; Y_{n-1}, \quad \sigma > 0$$

where Y_n's are i.i.d. and $\mathbb{P}(Y_n = 1) = 0.5 = \mathbb{P}(Y_n = -1)$.

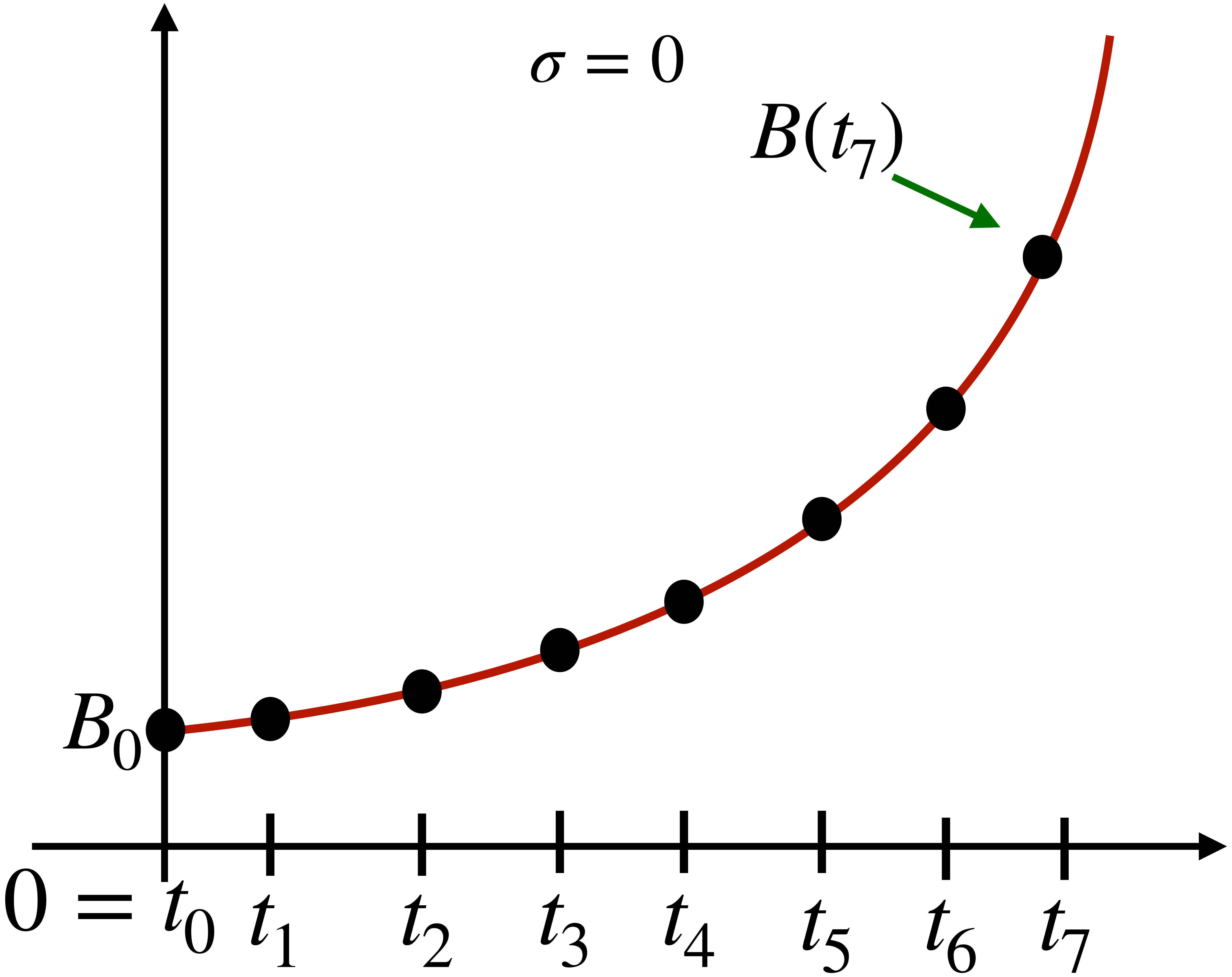
Continuous Model:

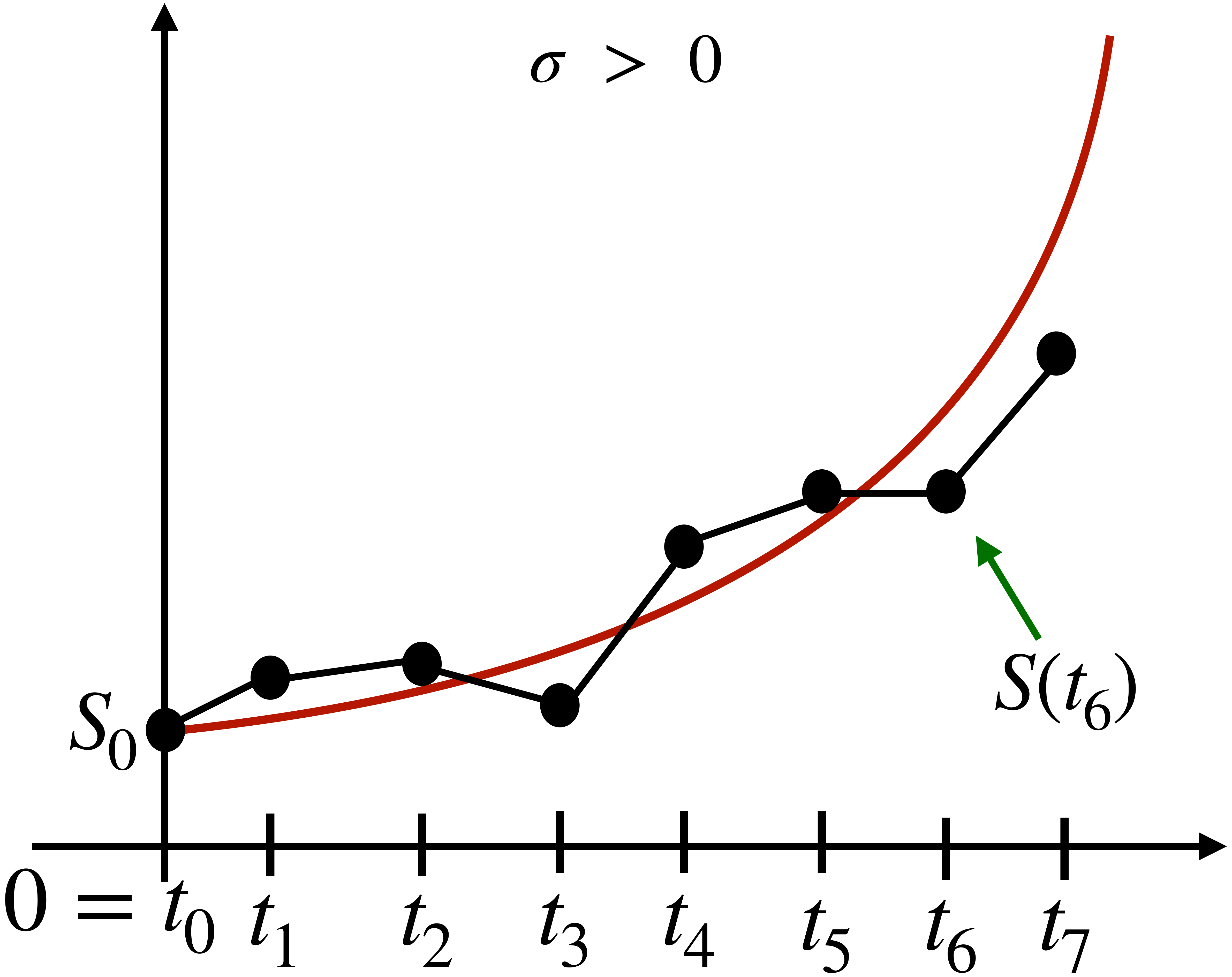
$$S(T) \; = \; S_0\, e^{?\cdots}$$

Dr. Lok Pati Tripathi, SMCS, IIT Goa

$$\frac{B(t_n) - B(t_{n-1})}{B(t_{n-1})} = r\,\delta t$$

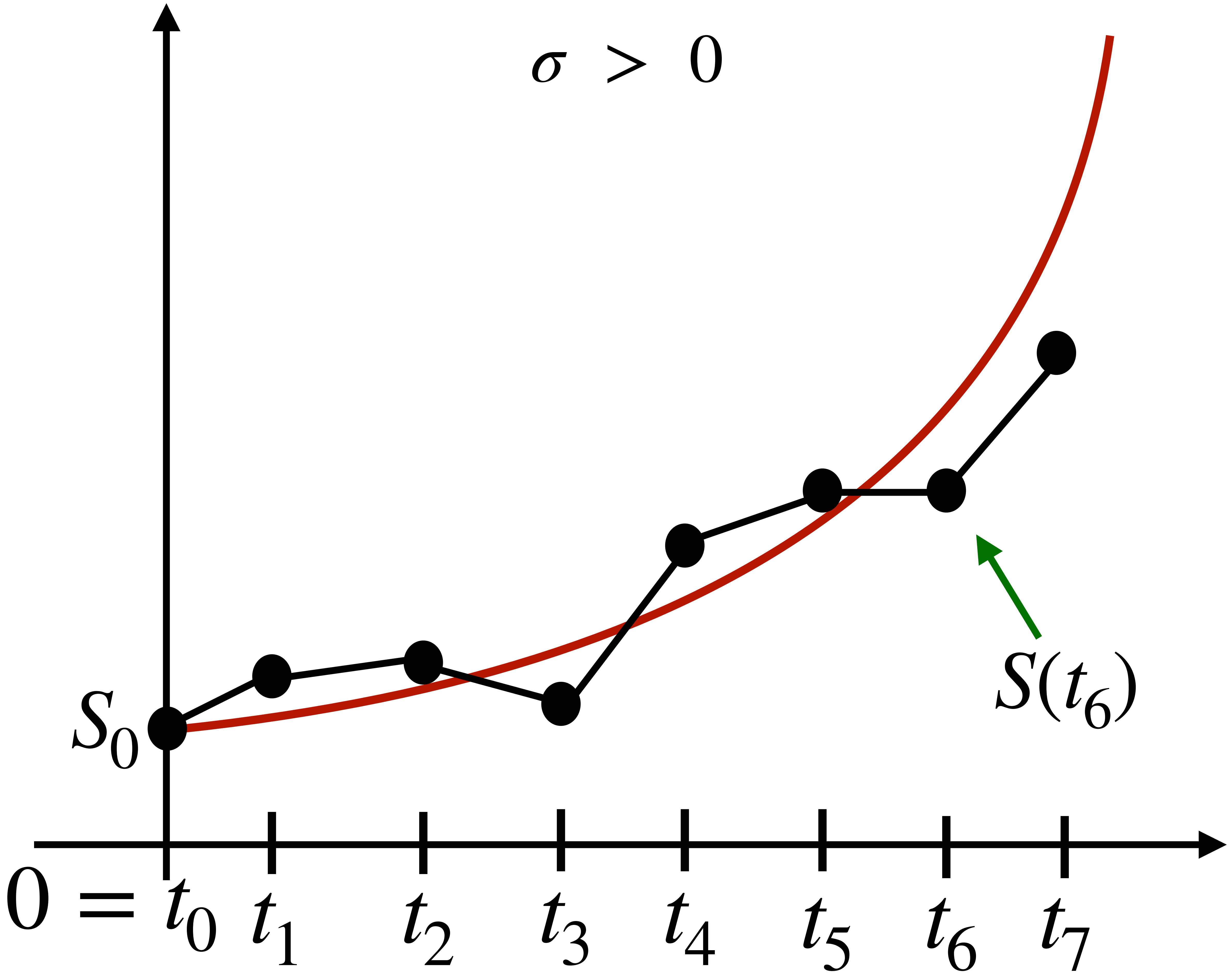
$$\frac{S(t_n) - S(t_{n-1})}{S(t_{n-1})} = r\delta t + \sigma\sqrt{\delta t}\,Y_{n-1}$$

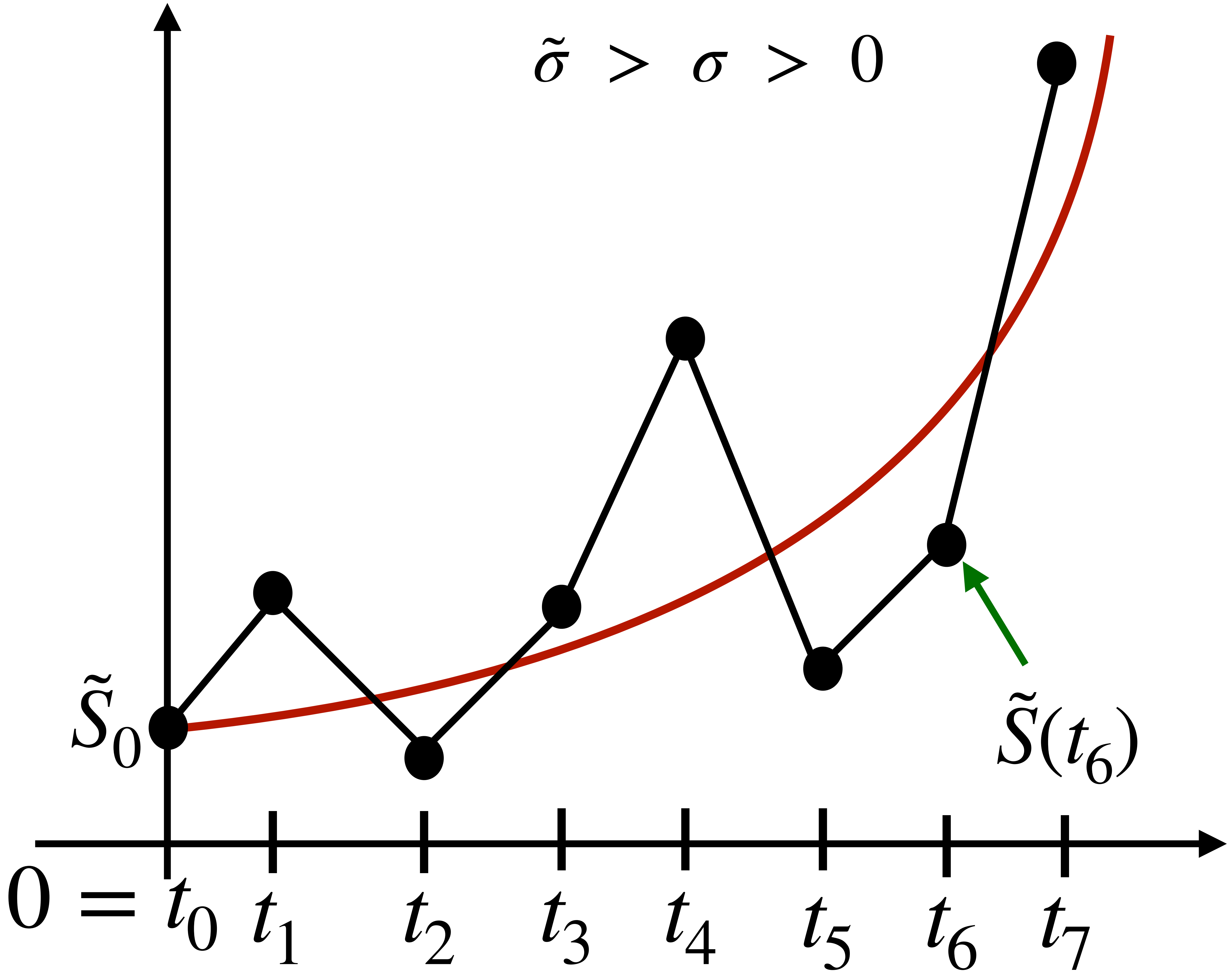




$$\boxed{\frac{S(t_n) - S(t_{n-1})}{S(t_{n-1})} = r\delta t + \sigma\sqrt{\delta t}\,Y_{n-1}}$$

$$\boxed{\frac{\tilde{S}(t_n) - \tilde{S}(t_{n-1})}{\tilde{S}(t_{n-1})} = r\delta t + \tilde{\sigma}\sqrt{\delta t}\,Y_{n-1}}$$





Central Limit Theorem:

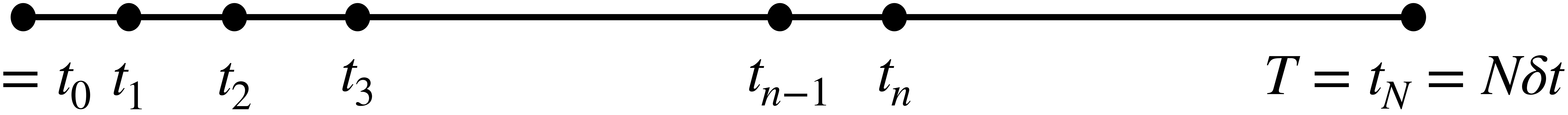
Let $(\Omega, \mathscr{M}, \mathbb{P})$ be a probability space and let $X_1, X_2, X_3, \ldots$ be iid random variables with mean 0 and variance 1. Then the CDF F_n of random variable $Y_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$ converge point-wise to the CDF N of standard normal random variable, that is

$$\lim_{n \to \infty} F_n(y) = N(y) \quad \forall\, y \in \mathbb{R},$$

that is $$\boxed{Y_n \xrightarrow{d} Y \quad \text{as } n \to \infty, \quad Y \sim N(0,1),}$$

where, $$N(y) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-s^2/2}\, ds, \quad y \in \mathbb{R}, \qquad F_n(y) := \mathbb{P}(Y_n \le y), \quad y \in \mathbb{R}$$

Risky asset-price model: A continuous model



The discrete risky asset-price model yields:

$$S(t_n) \;=\; S_0 \prod_{j=0}^{n-1} \left(1 + r\delta t + \sigma\sqrt{\delta t}\; Y_j\right), \quad n = 1,2,\ldots,N$$

$$S(T) \;=\; S(t_N) \;=\; S_0 \prod_{j=0}^{N-1} \left(1 + r\delta t + \sigma\sqrt{\delta t}\; Y_j\right)$$

Risky asset-price model: A continuous model

$$\implies \quad \log_e\left(\frac{S(T)}{S_0}\right) = \sum_{j=0}^{N-1} \log_e\left(1 + r\delta t + \sigma\sqrt{\delta t}\ Y_j\right)$$

For small δt, using $\log_e(1 + \delta t\theta) = \delta t\theta - \frac{1}{2}(\delta t\theta)^2 + O(\delta t^3), \quad \forall\ \theta \in \mathbb{R} : \ 1 + \delta t\theta > 0,$

$$\log_e\left(\frac{S(T)}{S_0}\right) = O\left(\delta t^{1/2}\right) + \sum_{j=0}^{N-1} \left(r\delta t + \sigma\sqrt{\delta t}\ Y_j - \frac{1}{2}\sigma^2\delta t\ Y_j^2\right) \longleftarrow X_j$$

$$= O\left(\delta t^{1/2}\right) + \sum_{j=0}^{N-1} X_j$$

Risky asset-price model: A continuous model

$$\mu_{X_j} := \mathbb{E}(X_j) = \left(r - \frac{1}{2}\sigma^2\right)\delta t,$$

$$\sigma^2_{X_j} := \ Var(X_j) \ = \ \sigma^2\delta t + O(\delta t^{3/2})$$

$$\frac{\log_e \frac{S(T)}{S_0} \ - \ N\mu_{X_j}}{\sigma_{X_j}\sqrt{N}} + O\left(\delta t^{1/2}\right) \ = \ \frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} \frac{X_j - \mu_{X_j}}{\sigma_{X_j}},$$

where we have used $\sigma_{X_j}\sqrt{N} = \sigma\sqrt{t_N} + O(\delta t^{1/4}) = \sigma\sqrt{T} + O(\delta t^{1/4})$.

Risky asset-price model: A continuous model

Now, as $N\mu_{X_j} = \left(r - \frac{1}{2}\sigma^2\right)T$, $\sigma_{X_j}\sqrt{N} = \sigma\sqrt{T} + O(\delta t^{1/4})$, and by CLT

$$\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}\frac{X_j - \mu_{X_j}}{\sigma_{X_j}} \xrightarrow{d} Y \quad \text{as } N \to \infty \text{ (or } \delta t \to 0), \quad Y \sim N(0,1),$$

after taking the limit $\delta t \to 0$, we obtain

$$\frac{\log_e \frac{S(T)}{S_0} - \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} = Y, \quad Y \sim N(0,1)$$

Risky asset-price model: A continuous model

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

Risky asset-price model for a dividend paying asset: A discrete model

Assume that the underlying asset pays out a continuous constant dividend yield q. Then under risk-neutrality assumption

$\mathbb{E}\left[\frac{1}{\delta t}\frac{S(t_n)-S(t_{n-1})}{S(t_{n-1})}\right]+q \;=\; r$. Thus in this case the discrete model is

$$S(t_0) = S_0$$

$$\frac{S(t_n)-S(t_{n-1})}{S(t_{n-1})} \;=\; (r-q)\delta t+\sigma\sqrt{\delta t}\;Y_{n-1}, \quad n=1,2,\ldots,N, \quad \sigma>0,$$

where Y_n's are i.i.d. and $\mathbb{P}(Y_n=+1)=\dfrac{1}{2}=\mathbb{P}(Y_n=-1), \quad n=0,1,2,\ldots$

Risky asset-price model for a dividend paying asset: A continuous model

After passing the limit we obtain the following continuous model for a dividend paying asset.

$$\boxed{S(T) = S_0 e^{\left(r-q-\frac{1}{2}\sigma^2\right)T+\sigma\sqrt{T}\,Y}, \quad Y \sim N(0,1)}$$

Option Pricing: Black-Scholes Formula

Option Pricing: Balck-Scholes formula

European style call option: **Pay-off function** $P_T := \max(S(T) - K, 0)$

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

Under risk-neutrality assumption,

$$\mathbb{E}[P_T] = P_0 e^{rT}\ , \quad \text{that is,} \quad P_0 \ = \ e^{-rT}\,\mathbb{E}[\max(S_T - K, 0)],$$

Thus, the price $P_0 := P(S_0, 0)$ of the call option at $t = 0$ is given by

$$P_0 \;=\; e^{-rT}\,\mathbb{E}[\max(S_T - K, 0)] \;=\; S_0 e^{-qT}\Phi(d_1) - K e^{-rT}\Phi(d_2)$$

where

$$d_1 := \frac{1}{\sigma\sqrt{T}}\left[\ln\left(\frac{S_0}{K}\right) + \left(r - q + \frac{1}{2}\sigma^2\right)T\right]$$

$$d_2 := d_1 - \sigma\sqrt{T}$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{s^2}{2}}\,ds$$

(Standard Normal Cumulative Distribution Function)

Put-call parity

Consider a portfolio whose pay-off is

$$\Pi(S_T, T) = S_T + \max(K - S_T, 0) - \max(S_T - K, 0) - K = 0,$$

Then, under risk-neutrality assumption, the spot value $\Pi(S_0, 0)$ of the portfolio is

$\Pi(S_0, 0) = e^{-rT}\mathbb{E}\left[\Pi(S_T, T)\right] = 0$. This implies

$$e^{-rT}\mathbb{E}[S_T] + e^{-rT}\mathbb{E}[\max(K - S_T, 0)] - e^{-rT}\mathbb{E}[\max(S_T - K, 0)] - e^{-rT}\mathbb{E}[K] = 0$$

That is,

$e^{-rT}e^{(r-q)T}S_0 + P_0^{put} - P_0^{call} - e^{-rT}K = 0$ and hence

$$\boxed{P_0^{call} + e^{-rT}K \; = \; e^{-qT}S_0 + P_0^{put}}$$

Option Pricing: Monte Carlo Method

Stochastic approach

Stochastic Approach: Monte-Carlo Method

As $P_0 = e^{-rT}\,\mathbb{E}[\max(S_T - K, 0)]$, $S_T = S_0 e^{\left(r-q-\frac{1}{2}\sigma^2\right)T+\sigma\sqrt{T}\,Y}$, $Y \sim N(0,1)$,

Law of Large Numbers implies: $P_0 = P(S_0,0) \approx e^{-rT}\dfrac{1}{N}\displaystyle\sum_{n=1}^{N} P_T\left(S^n\right)$

where the samples $\{S^n\}$, $n = 1,2,\ldots,N$, $N \in \mathbb{N}$, are drawn via

$$S^n = S_0 e^{\left(r-q-\frac{1}{2}\sigma^2\right)T+\sigma\sqrt{T}\,Y^n}, \quad Y^n \sim N(0,1)$$

Option Pricing: Finite Difference Method

Deterministic approach

How to Find the Fair Premium Price?

The Classical Black-Scholes Model (1973): A deterministic approach

Under certain assumptions, the price P of the European style option with pay-off $P_T(S)$ is the solution of the following problem-

$$\frac{\partial P}{\partial \tau} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP = 0 \qquad (S, \tau) \in \mathbb{R}_+ \times [0, T)$$

$$P(S, T) = P_T(S) \qquad S \in \mathbb{R}_+$$

where

$$P_T(S) := \begin{cases} \max\{S - K, 0\} & : \text{European call option} \\ \max\{K - S, 0\} & : \text{European put option} \end{cases}$$

Finite Difference Method:

Transformation:

$$\tau = T - t, \quad S = S_0 e^x, \quad P(S,\tau) = P(S_0 e^x, T-t) = u(x,t)$$

$$\frac{\partial u}{\partial t} - \frac{1}{2}\sigma^2 \frac{\partial^2 u}{\partial x^2} + \left(\frac{1}{2}\sigma^2 - r\right)\frac{\partial u}{\partial x} + ru = 0, \quad (x,t) \in \mathbb{R} \times (0,T]$$

$$u(x,0) \; = \; \max(S_0 e^x - K, 0)$$

Reference: Paul Wilmott, Jeff Dewynne, Sam Howison, "Option Pricing: Mathematical Models and Computations." *Oxford Financial Press* (1993)

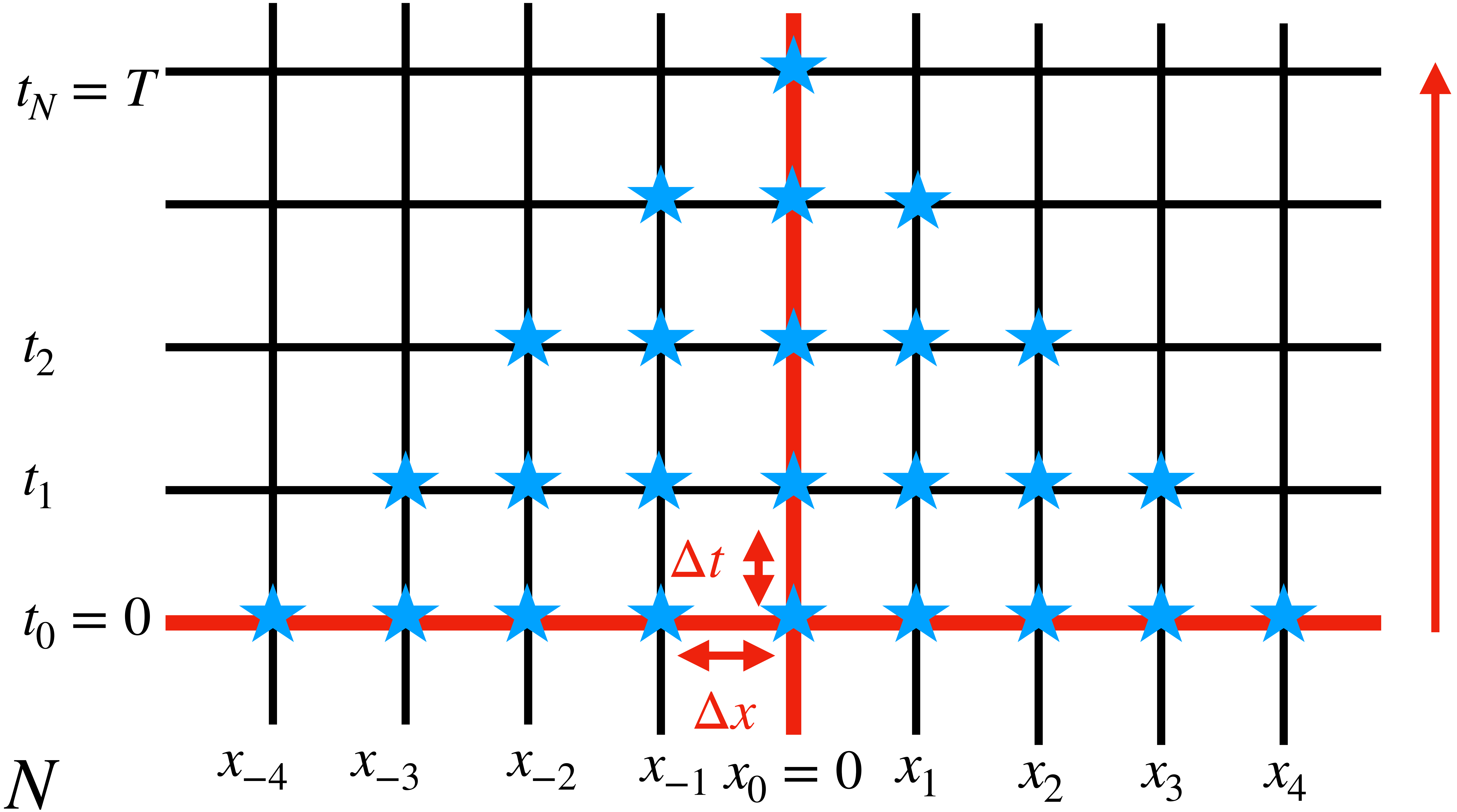
Forward Euler method:

$$x_j = j\Delta x,$$

$$j \in \mathbb{Z}$$

$$t_n = n\Delta t,$$

$$n = 0,1,2,\ldots,N$$



Forward/Explicit Euler method:

$$U_i^0 = \max(S_0 e^{x_i} - K, 0), \quad i \in \mathbb{Z}$$

$$\frac{U_i^n - U_i^{n-1}}{\Delta t} - \frac{1}{2}\sigma^2 \frac{U_{i+1}^{n-1} - 2U_i^{n-1} + U_{i-1}^{n-1}}{\Delta x^2} + \left(\frac{1}{2}\sigma^2 - r\right) \frac{U_{i+1}^{n-1} - U_{i-1}^{n-1}}{2\Delta x} + rU_i^{n-1} = 0$$

$$i \in \mathbb{Z}, \quad n = 1,2,\ldots,N$$

$$P(S_0,0) = u(0,T) \; \approx \; U_0^N$$

Multi-Asset Black-Scholes model:

Under certain assumptions, the price P of the d-asset European style option with pay-off $P_T(\boldsymbol{S})$ is the solution of the following problem-

$$\frac{\partial P}{\partial \tau} + \frac{1}{2}\sum_{i,j=1}^{d} \rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 P}{\partial S_i \partial S_j} + \sum_{j=1}^{d} (r - q_j) S_j \frac{\partial P}{\partial S_j} - rP = 0, \ (\boldsymbol{S}, \tau) \in \mathbb{R}_+^d \times [0, T)$$

$$P(\boldsymbol{S}, T) = P_T(\boldsymbol{S}) \qquad \boldsymbol{S} \in \mathbb{R}_+^d$$

where $\boldsymbol{S} := (S_j)_{j=1}^{d}$.

Reference: B. Tomas, "Arbitrage Theory in Continuous Time." *Oxford Finance Series, Oxford* (2009).

Computational difficulties:

Almost all the grid based numerical methods (FDM, FEM, FVM, etc.) for such high-dimensional ($d >> 2$) PDEs in the literature suffer from the so-called curse of dimensionality in the sense that the computational complexity of such numerical methods grows exponentially with respect to the number of assets.

Reference: R.E. Bellman, "Dynamic Programming." Princeton University Press, Princeton (1957).

Stochastic Approach: Monte-Carlo Method

Under risk-neutrality assumption: $P(\boldsymbol{S}_0, 0) = e^{-rT}\,\mathbb{E}\left[P_T\left(\boldsymbol{S}(T)\right) \mid \boldsymbol{S}(0) = \boldsymbol{S}_0\right]$

What is $\boldsymbol{S}(T)$?

Again, under risk-neutrality assumption and the efficient market hypothesis (weak form), following the derivation of single asset-price dynamics with an appeal to "Multivariate Central Limit Theorem" one can derive

$$S_i(T) = S_{0i}\, \exp\left(\left(r - q_i - \frac{1}{2}\sigma_i^2\right)T + \sqrt{T}\sum_{j=1}^{d} \Sigma_{ij}\, Z_j\right),\ i = 1,2,\ldots,d,$$

$$Z_j \sim N(0,1),\ \ \Sigma = (\Sigma_{ij})_{i,j=1}^{d},\ \ \Sigma\Sigma^T := \left(\rho_{ij}\sigma_i\sigma_j\right)_{i,j=1}^{d}.$$

Multivariate Central Limit Theorem:

Let $(\Omega, \mathcal{M}, \mathbb{P})$ be a probability space and let $X_j : \Omega \to \mathbb{R}^d$, $j = 1,2,3,\ldots$ be iid random vectors with mean $\mu \in \mathbb{R}^d$ and covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$. Then the CDF F_{Y_n} of random vector $Y_n := \dfrac{1}{\sqrt{n}} \sum_{i=1}^{n} \Sigma^{-1/2}(X_i - \mu)$ converge point-wise to the CDF F_Y of standard normal random vector $Y \sim N(0, I)$, that is

$$\lim_{n\to\infty} F_{Y_n}(y) = F_Y(y) \quad \forall\, y \in \mathbb{R}^d,$$

that is $\boxed{Y_n \xrightarrow{d} Y \quad \text{as } n \to \infty, \quad Y \sim N(0, I),}$ where I is the identity matrix,

and, $F_Y(y) := \dfrac{1}{(2\pi)^{d/2}} \displaystyle\int_{-\infty}^{y} e^{-\frac{1}{2}\|x\|^2}\, dx, \quad y \in \mathbb{R}^d, \qquad F_{Y_n}(y) := \mathbb{P}(Y_n \le y), \quad y \in \mathbb{R}^d$

Stochastic Approach: Monte-Carlo Method

Law of Large Numbers implies: $P(0, \boldsymbol{S}_0) \approx e^{-rT}\,\frac{1}{N}\sum_{n=1}^{N} P_T\left(\boldsymbol{S}^n\right)$

where the samples $\{\boldsymbol{S}^n\}$, $n = 1,2,\ldots,N$, $N \in \mathbb{N}$, are drawn via

$$S_i^n = S_{0i}\ \exp\left(\left(r - q_i - \frac{1}{2}\sigma_i^2\right)T + \sqrt{T}\sum_{j=1}^{d}\Sigma_{ij}\ Z_j^n\right),\ i = 1,2,\ldots,d,\ \ Z_j^n \sim N(0,1),$$

$$\Sigma = (\Sigma_{ij})_{i,j=1}^{d},\ \ \Sigma\Sigma^T := \left(\rho_{ij}\sigma_i\sigma_j\right)_{i,j=1}^{d}.$$

Remark: Cholesky factorization allows choosing Σ as a lower triangular matrix, reducing computational and storage costs, particularly in higher dimensions, without sacrificing generality.

A Scientific Machine Learning Approach:

- The Monte-Carlo method provide the option price only at a single point S_0 at a time.
- In practice, we generally need a quick access to the option price at any arbitrary point in a region of interest.
- The simplest approach to achieve it is that to compute the solution at different points in the region of interest and then construct a model (function) via some interpolation methods or neural networks.
- As we need to approximate the option price at several points to train the above model, this approach may not be very efficient in terms of computational cost.

A Scientific Machine Learning Approach:

An equivalent optimization problem:

Let $D \subset \mathbb{R}^d_+$ be a closed and bounded domain of interest. Then, under suitable assumptions, there is a unique continuous function $\phi : D \to \mathbb{R}$ such that

$$\mathbb{E}\left[\, | P_T\left(\boldsymbol{S}(T)\right) - \phi(\boldsymbol{\xi}) \, |^2 \, \middle| \boldsymbol{S}(0) = \boldsymbol{\xi} \right] \; = \; \inf_{\psi \in C(D;\mathbb{R})} \mathbb{E}\left[\, | P_T\left(\boldsymbol{S}(T)\right) - \psi(\boldsymbol{\xi}) \, |^2 \, \middle| \boldsymbol{S}(0) = \boldsymbol{\xi} \right]$$

where $\boldsymbol{\xi} : \Omega \to D$ is a continuous uniformly distributed random variable.

Moreover, $\forall \boldsymbol{x} \in D \quad \phi(\boldsymbol{x}) = \mathbb{E}\left[P_T\left(\boldsymbol{S}(T)\right) | \, \boldsymbol{S}(0) = \boldsymbol{x} \right]$ and thus

$$P(0, \boldsymbol{x}) = e^{-rT} \, \mathbb{E}\left[P_T\left(\boldsymbol{S}(T)\right) | \, \boldsymbol{S}(0) = \boldsymbol{x} \right] \; = \; e^{-rT} \, \phi(\boldsymbol{x}).$$

How to approximate $\phi(\boldsymbol{x})$?

Reference: C. Beck et al., *Journal of Scientific Computing* (2021).

A Scientific Machine Learning Approach:

The Universal Approximation Theorem suggests to approximate the function $\phi : D \to \mathbb{R}$ by a (single) neural network $\phi_\theta : D \to \mathbb{R}$, where $\theta \in \mathbb{R}^N$ collects all unknown parameters of the network. Thus, we can consider the following optimization problem

$$\theta = \arg\min_{\hat{\theta}\in\mathbb{R}^N} \mathbb{E}\left[\left| P_T\left(\boldsymbol{S}(T)\right) - \phi_{\hat{\theta}}(\boldsymbol{\xi}) \right|^2 \middle| \boldsymbol{S}(0) = \boldsymbol{\xi} \right]$$

that can be seen as a discrete approximation of the previously discussed continuous optimization problem.

A Scientific Machine Learning Approach:

Feedforward Neural Network:

Input dimension: d_i Output dimension: d_o

Number of layers: $(M+1)$, $M \in \mathbb{N} \backslash \{1\}$

Number of hidden layers: $M-1$, $M \in \mathbb{N} \backslash \{1\}$

Number of neurons in the n^{th} layer: $m_n \in \mathbb{N}$, $n = 0,1,\ldots,M$

$$(m_0 = d_i \;\&\; m_M = d_o)$$

For simplicity assume that $m_n = m, \quad n = 1,2,\ldots,M-1$

A Scientific Machine Learning Approach:

Feedforward Neural Network:

The neural network:

$\mathcal{N}^{\rho}_{d_i,d_o,M,m} := \{\phi : \mathbb{R}^{d_i} \to \mathbb{R}^{d_o} : \ \phi = A_M \circ \rho \circ A_{M-1} \circ \cdots \circ \rho \circ A_1\}$, where

$$A_n(\boldsymbol{x}) = W_n \boldsymbol{x} + \boldsymbol{\beta}_n,$$

$A_1 : \mathbb{R}^{d_i} \to \mathbb{R}^m$, $A_M : \mathbb{R}^m \to \mathbb{R}^{d_o}$, and $A_n : \mathbb{R}^m \to \mathbb{R}^m$, $n = 2,3,\ldots,M-1$

Weight matrix: W_n Bias vector: $\boldsymbol{\beta}_n$

Activation function: $\rho(\boldsymbol{y}) = (\hat{\rho}(y_1), \hat{\rho}(y_2), \ldots, \hat{\rho}(y_m))$, $\boldsymbol{y} = (y_1, y_2, \ldots, y_m)$, where

$\hat{\rho} : \mathbb{R} \to \mathbb{R}$ is a one-dimensional activation function.

A Scientific Machine Learning Approach:

Examples of a few commonly used 1-D activation functions $\hat{\rho} : \mathbb{R} \to \mathbb{R}$

Sigmoid: $\hat{\rho}(x) = \dfrac{1}{1 + e^{-x}}$

Tanh: $\hat{\rho}(x) = \tanh(x) = \dfrac{2}{1 + e^{-2x}} - 1$ (a scaled sigmoid)

ReLU (Rectified Linear Unit): $\hat{\rho}(x) = \max(x, 0)$

Leaky ReLU: $\hat{\rho}(x) = \max(x, \epsilon x),\ 0 < \epsilon < 1$

A Scientific Machine Learning Approach:

Universal Approximation Theorem:

For each $M \in \mathbb{N} \backslash \{1\}$, the set $\cup_{m \in \mathbb{N}} \mathcal{N}^{\rho}_{d_i, d_o, M, m}$ is dense in $L^2(\mu)$ for any finite measure μ on $\mathbb{R}^{d_i}$, whenever the activation function ρ is continuous and non-constant. Moreover, if ρ is a non-constant C^k function then any C^k function and its derivatives up to order k can be approximated by neural networks in $\cup_{m \in \mathbb{N}} \mathcal{N}^{\rho}_{d_i, d_o, 2, m}$ arbitrarily well on any compact set of $\mathbb{R}^{d_i}$.

References:

- K. Hornik, M. Stinchcombe, and H. White, *Neural Netw.* (1989).
- K. Hornik, M. Stinchcombe, and H. White, *Neural Netw.* (1990).
- C. Bayer, J. Qiu, and Y. Yao, *SIAM J. Financial Math.* (2022).

References:

1. F. Black & M. Scholes, "The pricing of options and corporate liabilities," *Journal of Political Economy*, *81*(3), pp. 637–654, (1973).
2. R. C. Merton, "Theory of rational option pricing," *The Bell Journal of Economics and Management Science*, *4*, pp. 141–183, (1973).
3. Desmond J. Higham "An Introduction to Financial Option Valuation: Mathematics, Stochastics and Computation", Cambridge University Press, Cambridge (2004).
4. Paul Wilmott, Jeff Dewynne, Sam Howison, "Option Pricing: Mathematical Models and Computations." *Oxford Financial Press* (1993).
5. B. Tomas, "Arbitrage Theory in Continuous Time." *Oxford Finance Series, Oxford* (2009).
6. R.E. Bellman, "Dynamic Programming." *Princeton University Press, Princeton* (1957).

References:

7. Christian Beck, Sebastian Becker, Philipp Grohs, Nor Jaafari, and Arnulf Jentzen, “Solving the kolmogorov pde by means of deep learning,” *Journal of Scientific Computing*, 88(3), (2021).
8. K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural networks*, vol. 2, no. 5, pp. 359–366, (1989).
9. K. Hornik, M. Stinchcombe, and H. White, “Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks,” *Neural Networks*, 3, pp 551–560, (1990).
10. C. Bayer, J. Qiu, and Y. Yao, “Pricing options under rough volatility with backward SPDEs,” *SIAM J. Financial Math., 13, 1, pp 179–212* (2022).
11. J. Blechschmidt and O. G. Ernst, “Three ways to solve partial differential equations with neural networks—a review,” *GAMM-Mitteilungen*, 44(2), 2021.

Thank You

lokpati@iitgoa.ac.in