Computational Finance with C++

Monte Carlo Methods and Numerical Methods for SDEs

Panos Parpas
Imperial College London
p.parpas@imperial.ac.uk

Monte Carlo Methods and Numerical Methods for SDEs

Part I: Monte Carlo Methods

- Mathematical Foundations
- Monte Carlo and option valuation
- Calculation of Greeks
- Variance Reduction via Antithetic Variables
- Variance Reduction via Control Variates
- Importance Sampling

Part II: Numerical Methods for SDEs:

- Euler-Maruyama
- Weak and Strong Convergence
- Implicit Methods and Numerical Stability
- Mean Exit Times
- Numerical Methods for Systems

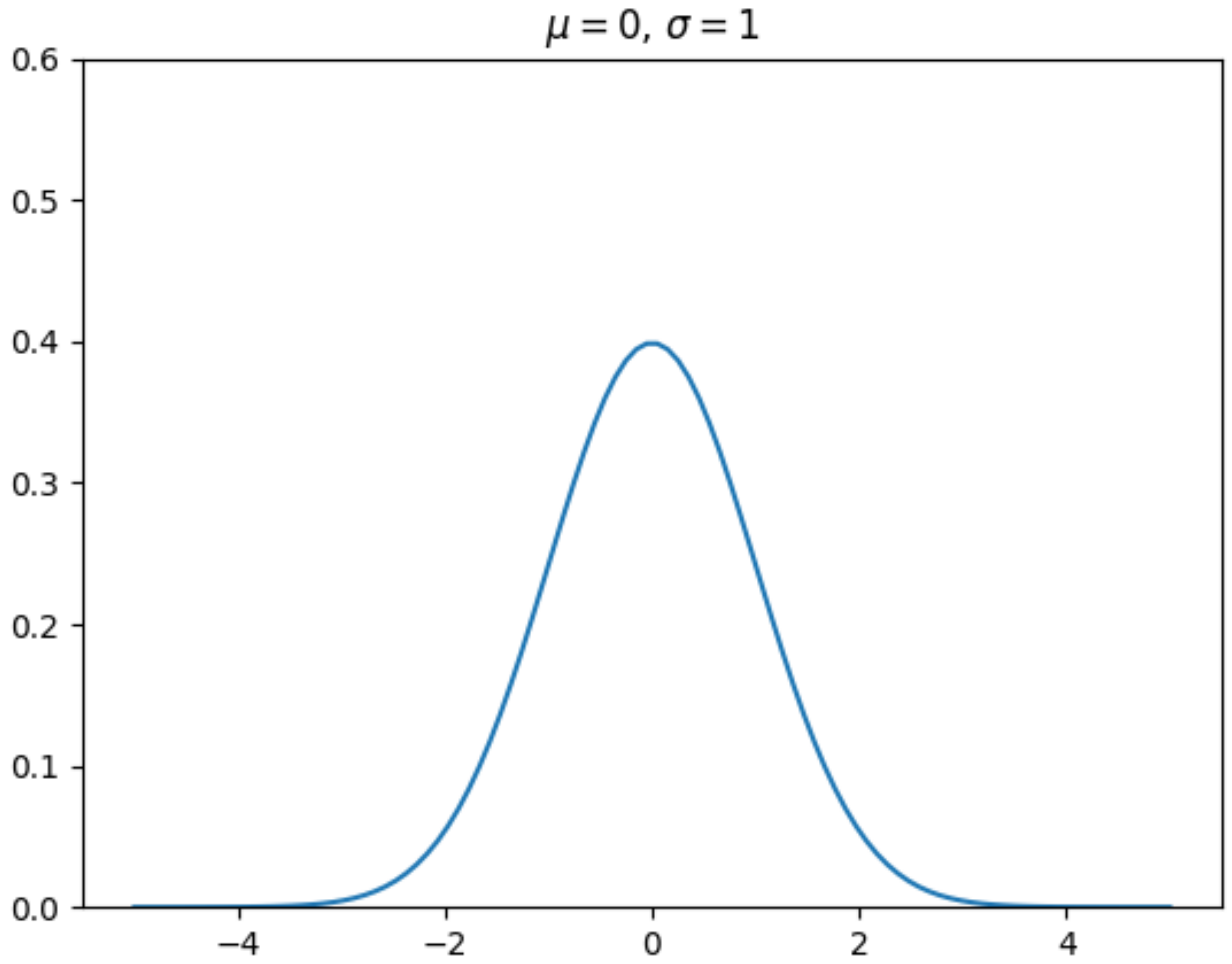
Mathematical Foundations of Monte Carlo Methods

- **Strong Law of Large Numbers** tells us about the average of i.i.d random variables
- **Central Limit Theorem** tells us about the fluctuations around the mean
- **Combine the two** to justify Monte Carlo Methods

Normal Random Variables

If X has density: $p(x) = \dfrac{1}{\sqrt{2\sigma^2\pi}} e^{-(x-\mu)^2/2\sigma^2}$ then we say that X is a normal random vairable.

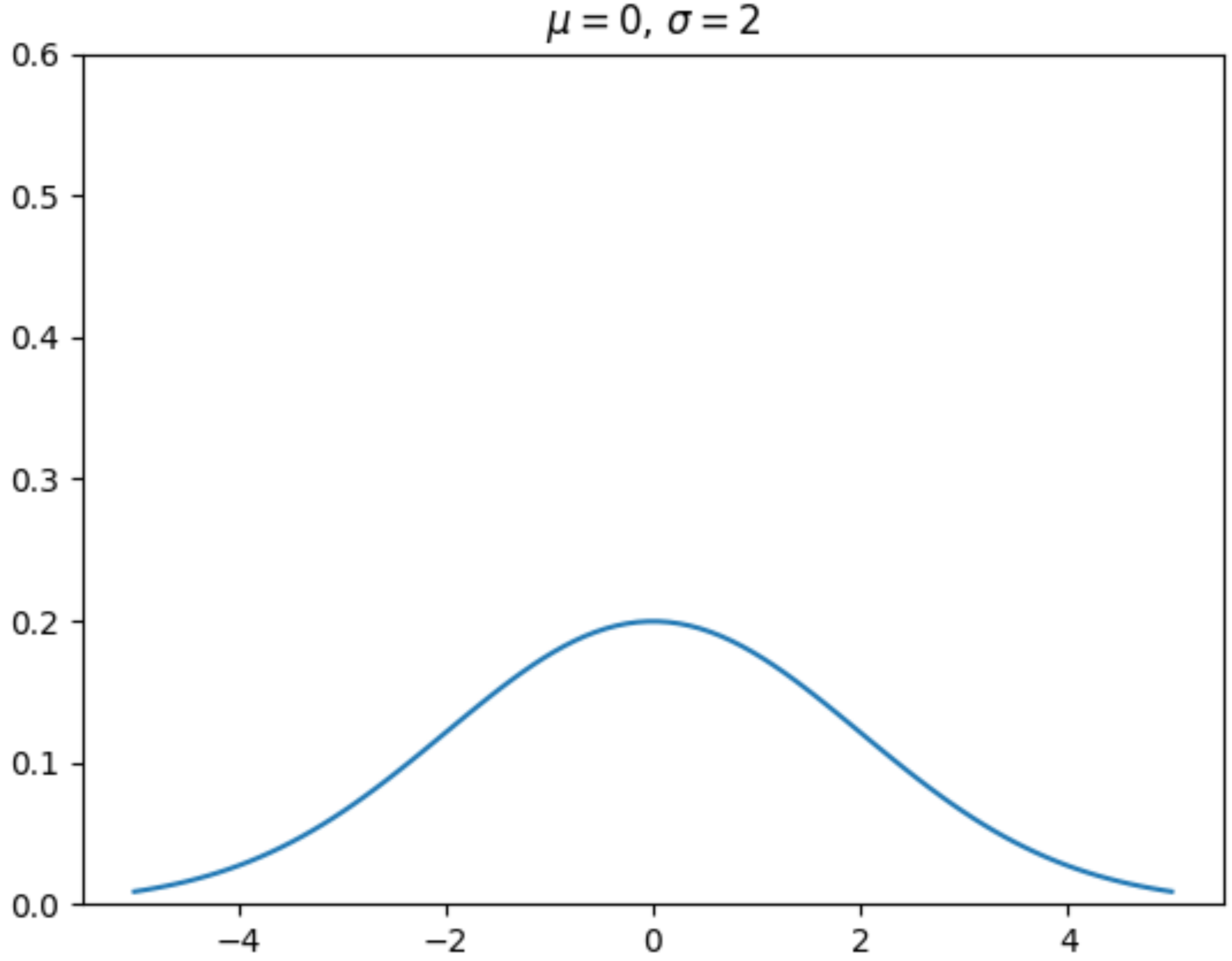
We write $X \sim N(\mu, \sigma)$ to denote a nornal random variable with mean μ and variance σ^2.



Normal Random Variables

If X has density: $p(x) = \dfrac{1}{\sqrt{2\sigma^2\pi}} e^{-(x-\mu)^2/2\sigma^2}$ then we say that X is a normal random vairable.

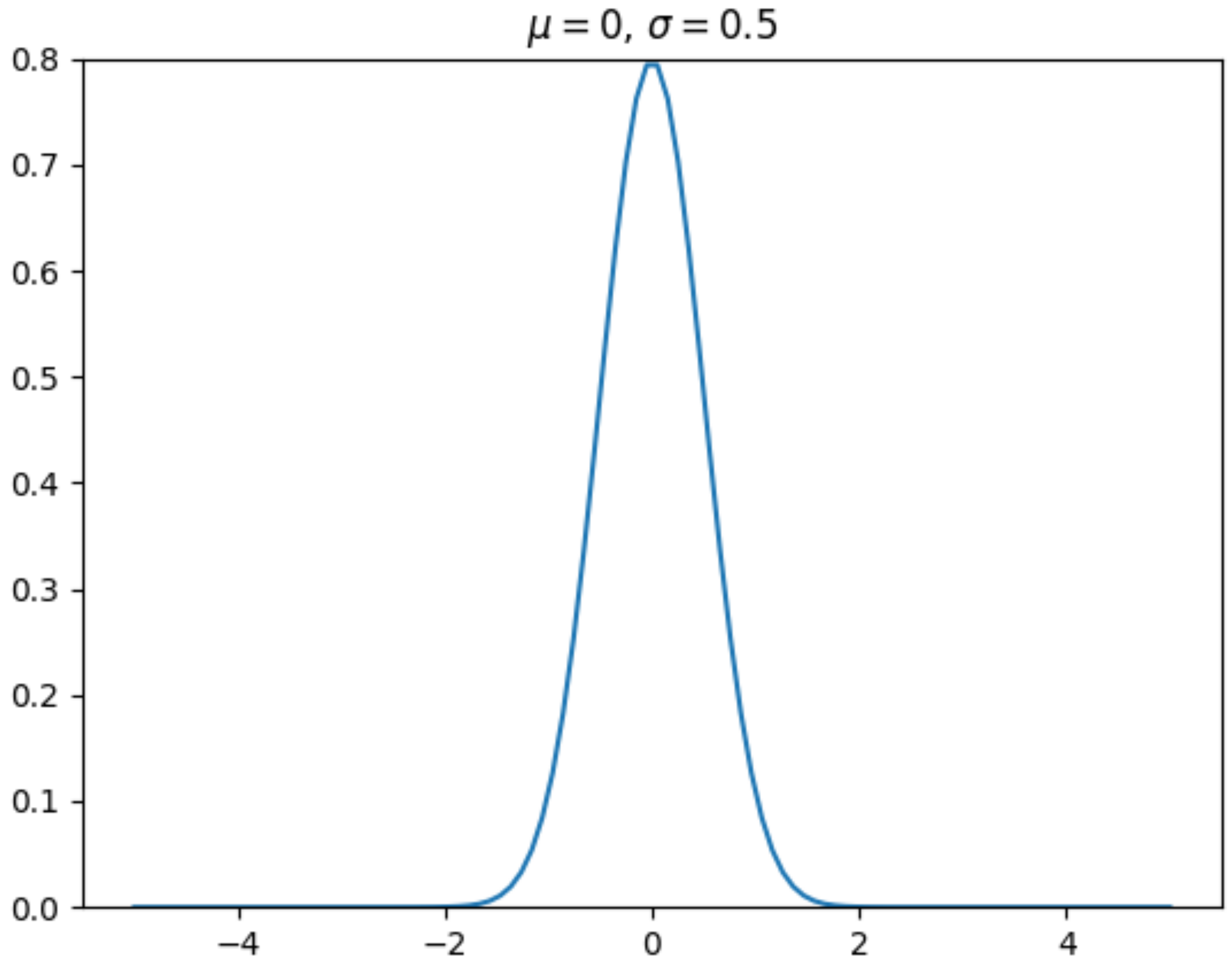
We write $X \sim N(\mu, \sigma)$ to denote a nornal random variable with mean μ and variance σ^2.



Normal Random Variables

If X has density: $p(x) = \dfrac{1}{\sqrt{2\sigma^2\pi}} e^{-(x-\mu)^2/2\sigma^2}$ then we say that X is a normal random vairable.

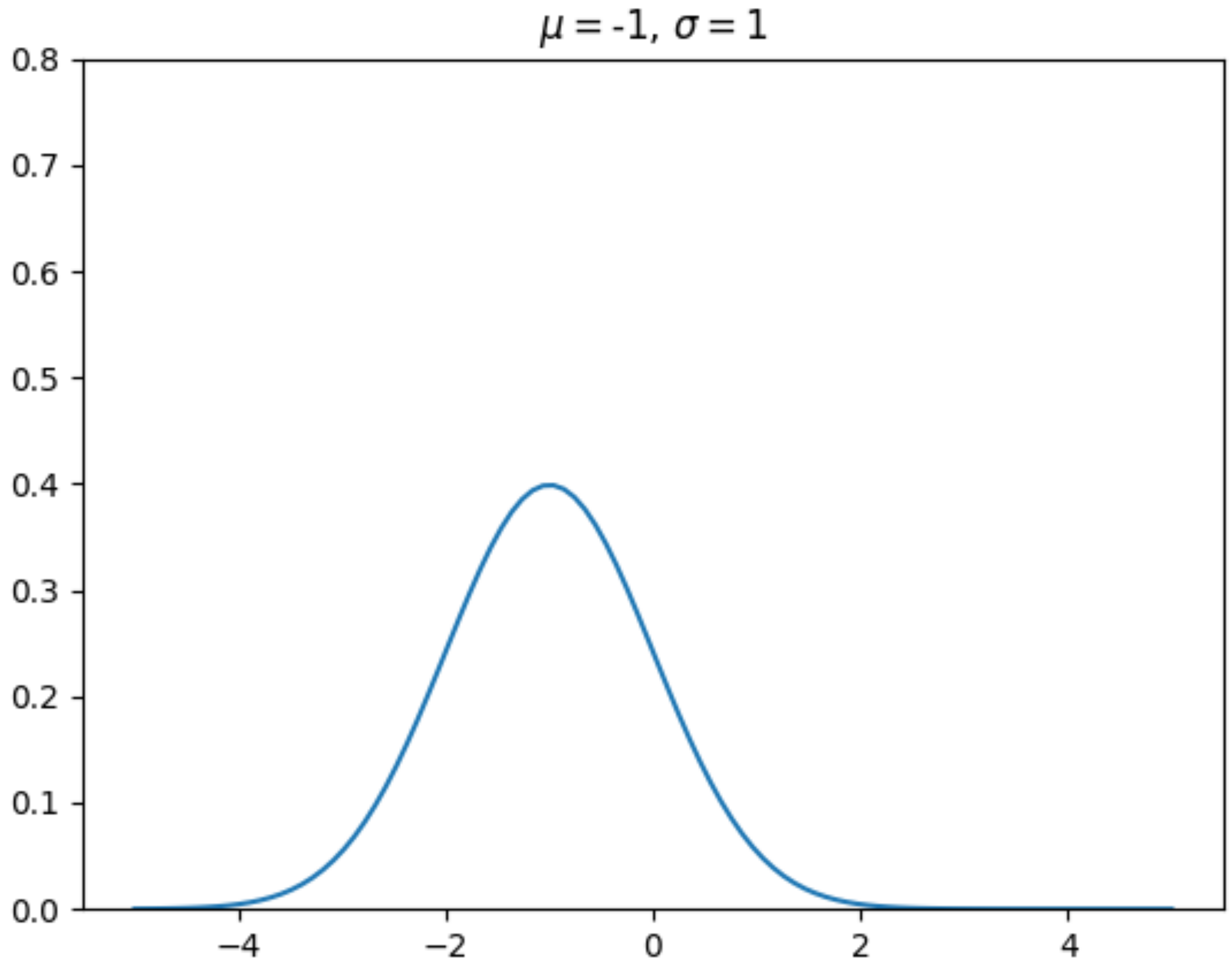
We write $X \sim N(\mu, \sigma)$ to denote a nornal random variable with mean μ and variance σ^2.



Normal Random Variables

If X has density: $p(x) = \frac{1}{\sqrt{2\sigma^2\pi}}e^{-(x-\mu)^2/2\sigma^2}$ then we say that X is a normal random vairable.

We write $X \sim N(\mu, \sigma)$ to denote a nornal random variable with mean μ and variance σ^2.

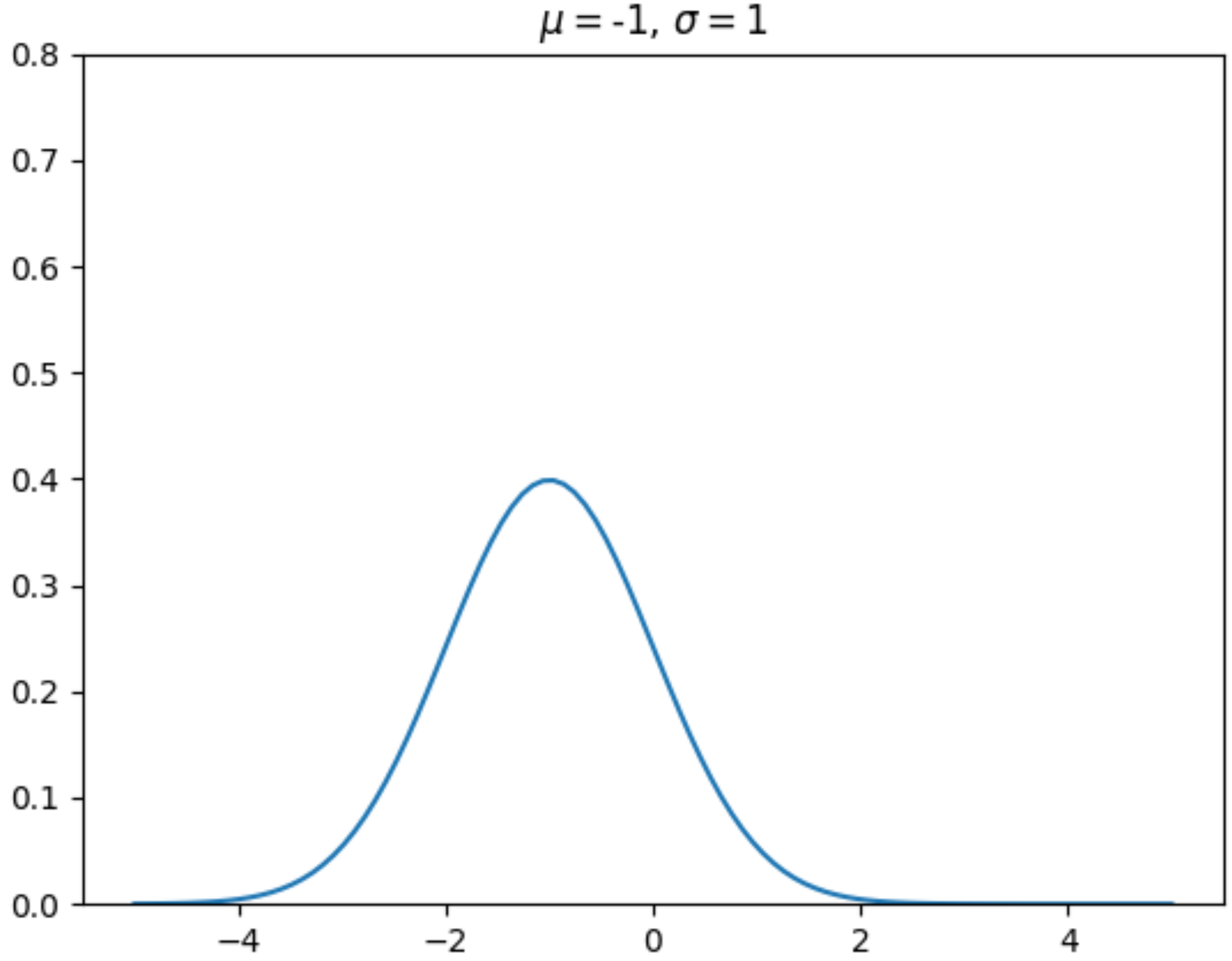


Normal Random Variables

If X has density: $p(x) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-(x-\mu)^2/2\sigma^2}$ then we say that X is a normal random vairable.

We write $X \sim N(\mu, \sigma)$ to denote a nornal random variable with mean μ and variance σ^2.

1. If $X \sim N(\mu, \sigma)$ then $(X - \mu)/\sigma \sim N(0,1)$

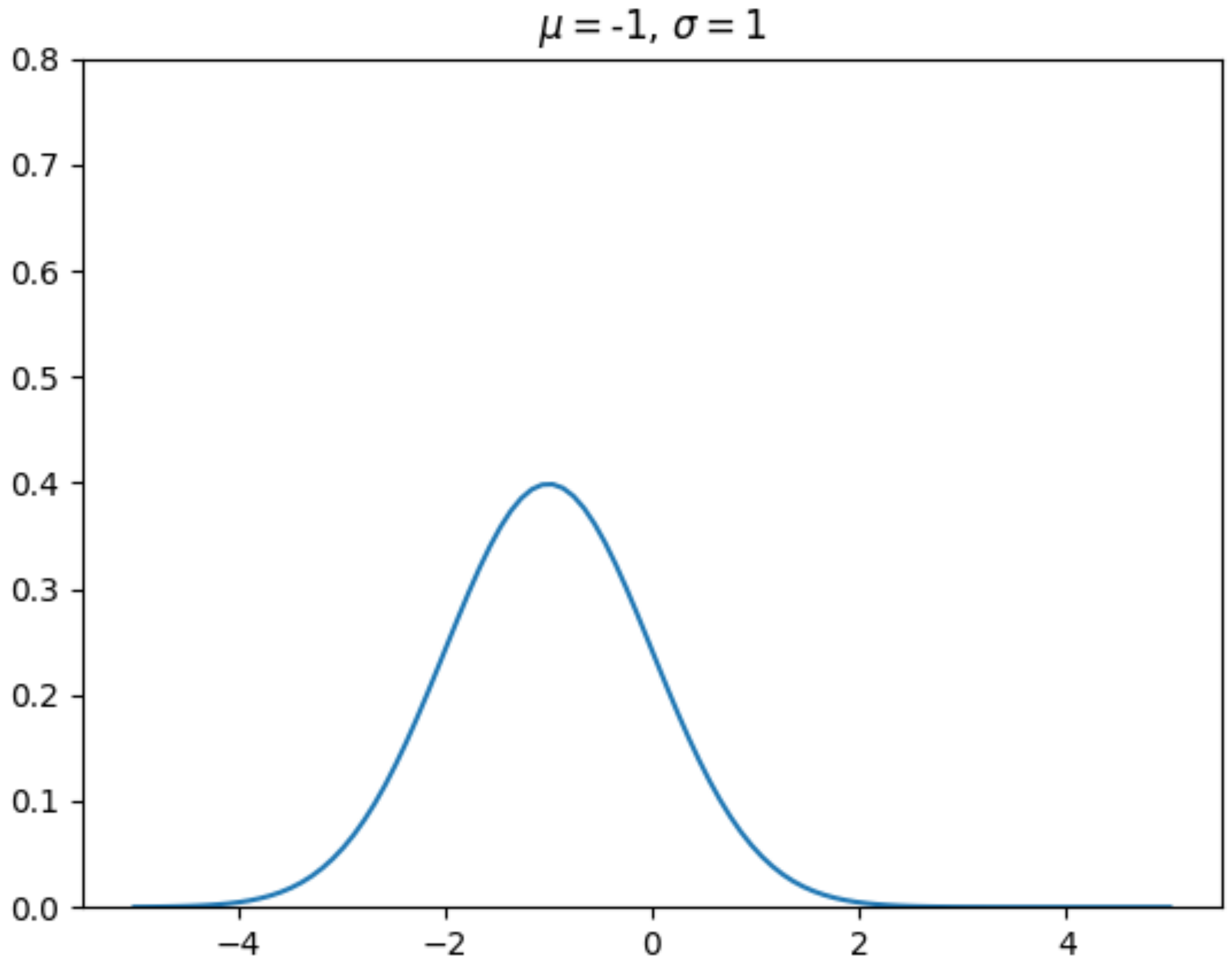


Normal Random Variables

If X has density: $p(x) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-(x-\mu)^2/2\sigma^2}$ then we say that X is a normal random vairable.

We write $X \sim N(\mu, \sigma)$ to denote a nornal random variable with mean μ and variance σ^2.

1. If $X \sim N(\mu, \sigma)$ then $(X - \mu)/\sigma \sim N(0,1)$
2. If $Y \sim N(0,1)$ then $\sigma Y + \mu \sim N(\mu, \sigma^2)$

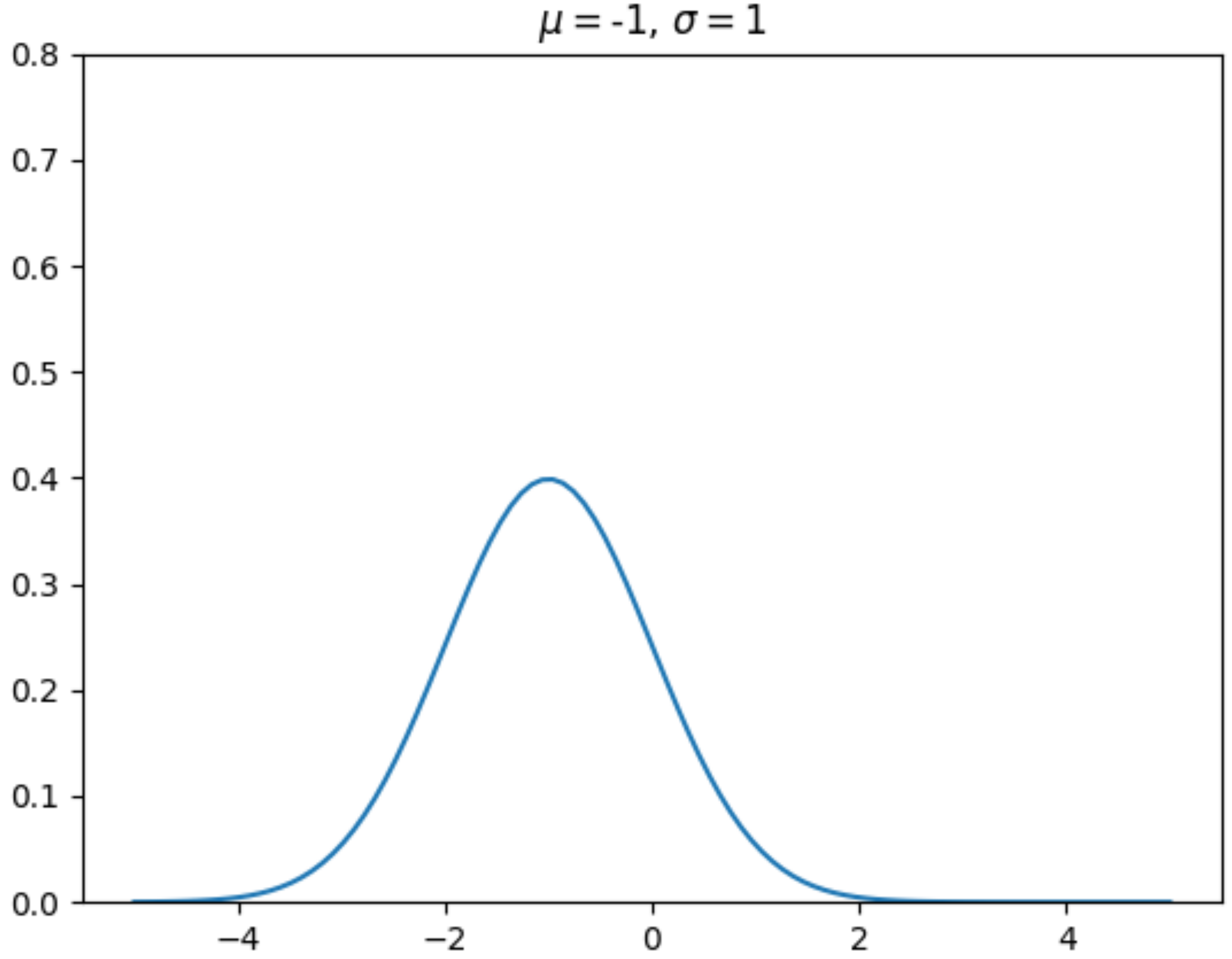


Normal Random Variables

If X has density: $p(x) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-(x-\mu)^2/2\sigma^2}$ then we say that X is a normal random vairable.

We write $X \sim N(\mu, \sigma)$ to denote a nornal random variable with mean μ and variance σ^2.

1. If $X \sim N(\mu, \sigma)$ then $(X - \mu)/\sigma \sim N(0,1)$
2. If $Y \sim N(0,1)$ then $\sigma Y + \mu \sim N(\mu, \sigma^2)$
3. If $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$ and X, Y are independent then $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

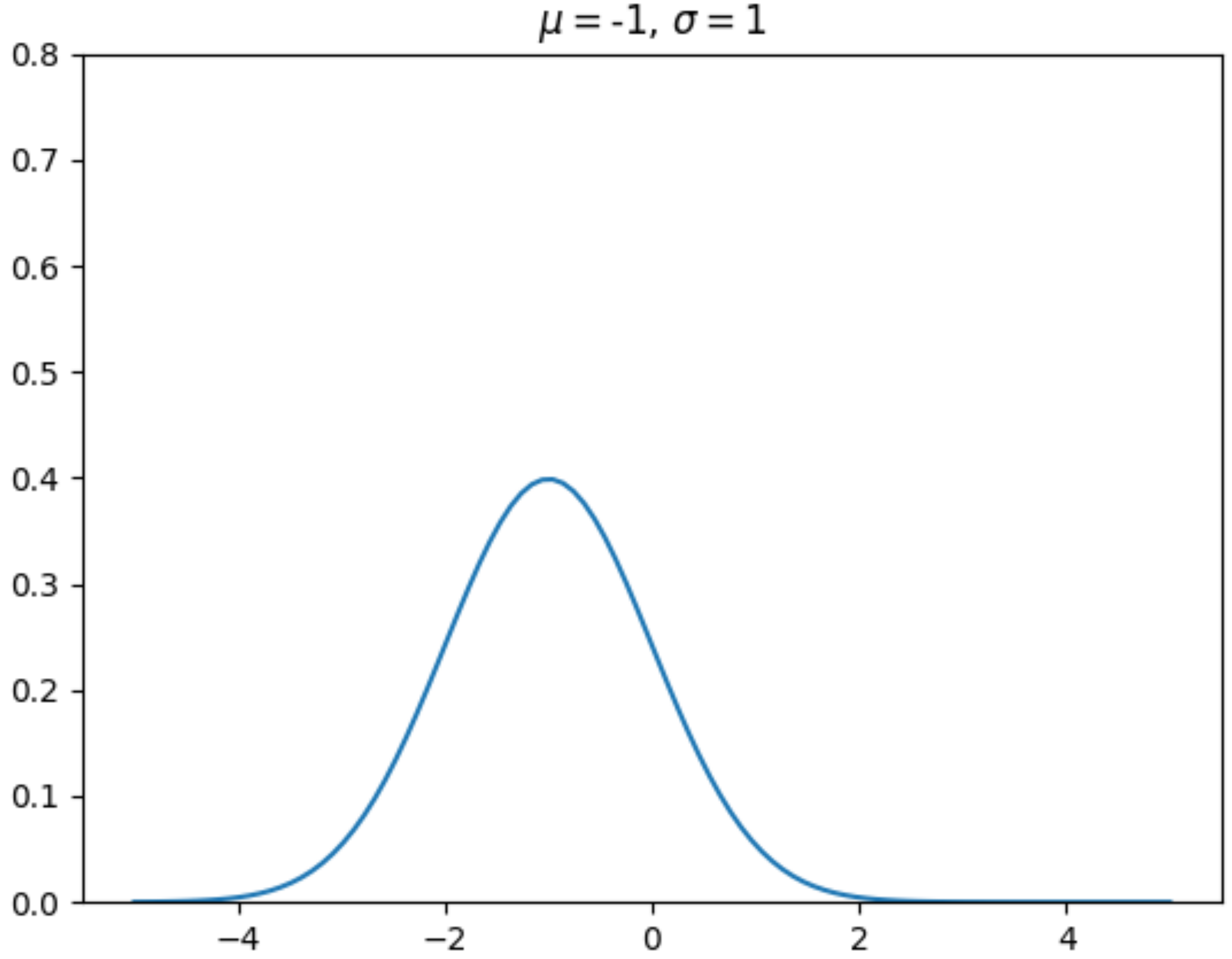


Normal Random Variables

If X has density: $p(x) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-(x-\mu)^2/2\sigma^2}$ then we say that X is a normal random vairable.

We write $X \sim N(\mu, \sigma)$ to denote a nornal random variable with mean μ and variance σ^2 .

1. If $X \sim N(\mu, \sigma)$ then $(X - \mu)/\sigma \sim N(0,1)$
2. If $Y \sim N(0,1)$ then $\sigma Y + \mu \sim N(\mu, \sigma^2)$
3. If $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$ and
   X, Y are independent then
   $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
4. If X and Y are normal RVs, then they are
   independent if and only if $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$



Central Limit Theorem I

If X and Y are independent then we have:

Central Limit Theorem I

If X and Y are independent then we have:

$$\begin{aligned}\mathrm{var}[X+Y] &= \mathbb{E}[(X+Y)^2] - (\mathbb{E}[(X+Y)])^2 \\ &= \mathbb{E}[X^2+Y^2+2XY] - (\mathbb{E}[(X+Y)])^2 \\ &= \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\mathbb{E}[X]\mathbb{E}[Y] - (\mathbb{E}[X])^2 - (\mathbb{E}[(Y])^2 - 2\mathbb{E}[X]\mathbb{E}[(Y] \\ &= \mathrm{var}[X] + \mathrm{var}[Y]\end{aligned}$$

Central Limit Theorem I

If X and Y are independent then we have:

$$\begin{aligned}
\text{var}[X+Y] &= \mathbb{E}[(X+Y)^2] - (\mathbb{E}[(X+Y)])^2 \\
&= \mathbb{E}[X^2 + Y^2 + 2XY] - (\mathbb{E}[(X+Y)])^2 \\
&= \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\mathbb{E}[X]\mathbb{E}[Y] - (\mathbb{E}[X])^2 - (\mathbb{E}[(Y])^2 - 2\mathbb{E}[X]\mathbb{E}[(Y] \\
&= \text{var}[X] + \text{var}[Y]
\end{aligned}$$

Applying the same logic to an i.i.d sequence of RVs with mean a and variance b^2 :

Central Limit Theorem I

If X and Y are independent then we have:

$$\begin{aligned}
\text{var}[X+Y] &= \mathbb{E}[(X+Y)^2] - (\mathbb{E}[(X+Y)])^2 \\
&= \mathbb{E}[X^2 + Y^2 + 2XY] - (\mathbb{E}[(X+Y)])^2 \\
&= \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\mathbb{E}[X]\mathbb{E}[Y] - (\mathbb{E}[X])^2 - (\mathbb{E}[(Y])^2 - 2\mathbb{E}[X]\mathbb{E}[(Y] \\
&= \text{var}[X] + \text{var}[Y]
\end{aligned}$$

Applying the same logic to an i.i.d sequence of RVs with mean a and variance b^2 :

$$\mathbb{E}[\sum_{i=1}^{N} X_i] = Na \text{ and } \text{var}[\sum_{i=1}^{N} X_i] = Nb^2$$

Central Limit Theorem I

If X and Y are independent then we have:

$$\begin{aligned}\mathrm{var}[X+Y] &= \mathbb{E}[(X+Y)^2] - (\mathbb{E}[(X+Y)])^2 \\ &= \mathbb{E}[X^2 + Y^2 + 2XY] - (\mathbb{E}[(X+Y)])^2 \\ &= \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\mathbb{E}[X]\mathbb{E}[Y] - (\mathbb{E}[X])^2 - (\mathbb{E}[(Y])^2 - 2\mathbb{E}[X]\mathbb{E}[(Y] \\ &= \mathrm{var}[X] + \mathrm{var}[Y]\end{aligned}$$

Applying the same logic to an i.i.d sequence of RVs with mean a and variance b^2 :

$$\mathbb{E}[\sum_{i=1}^{N} X_i] = Na \text{ and } \mathrm{var}[\sum_{i=1}^{N} X_i] = Nb^2$$

i.e. $\dfrac{\sum_{i=1}^{N} X_i - Na}{\sqrt{N}b}$ has mean 0 and variance 1

Central Limit Theorem II

Theorem (Central Limit Theorem): Let $X_i,\ i = 1,\ldots,N$ be i.i.d with mean a and variance b^2.
Then for all $\alpha < \beta$:

$$\lim_{N\to\infty} \mathbb{P}\left(\alpha \leq \frac{\sum_{i=1}^{N} X_i - Na}{\sqrt{N}b} \leq \beta\right) = \frac{1}{\sqrt{2\pi}}\int_{\alpha}^{\beta} e^{-\frac{1}{2}x^2}dx$$

- CLT applies to **any** i.i.d with finite mean and variance
- Justification for modelling many sources of noise as a single normal random variable

Strong Law of Large Numbers

If X_i are i.i.d observations of an event then what can we say about:

$\frac{1}{N}\sum_{i=1}^{N} X_i$ and the true mean of X?

Strong Law of Large Numbers

If X_i are i.i.d observations of an event then what can we say about:

$\frac{1}{N}\sum_{i=1}^{N} X_i$ and the true mean of X?

Take $\alpha = -N^{-1/4}$ and $\beta = N^{1/4}$ in CLT to get,

$$\mathbb{P}\left(\frac{-b}{N^{1/4}} \leq \frac{1}{N}\sum_{i=1}^{N} X_i - a \leq \frac{b}{N^{1}/4}\right) = \frac{1}{\sqrt{2\pi}}\int_{-N^{1/4}}^{N^{1/4}} e^{-\frac{1}{2}x^2}dx \to 0 \text{ as } N \to \infty$$

Strong Law of Large Numbers

If X_i are i.i.d observations of an event then what can we say about:

$\frac{1}{N}\sum_{i=1}^{N} X_i$ and the true mean of X?

Take $\alpha = -N^{-1/4}$ and $\beta = N^{1/4}$ in CLT to get,

$$\mathbb{P}\left(\frac{-b}{N^{1/4}} \leq \frac{1}{N}\sum_{i=1}^{N} X_i - a \leq \frac{b}{N^1/4}\right) = \frac{1}{\sqrt{2\pi}} \int_{-N^{1/4}}^{N^{1/4}} e^{-\frac{1}{2}x^2} dx \to 0 \text{ as } N \to \infty$$

Theorem (Strong Law of Large Numbers): Let $X_i,\ i = 1,\ldots,N$ be i.i.d with mean a then:

$$\mathbb{P}\left(\lim_{N\to\infty} \sum_{i=1}^{N} X_i = a\right) = 1$$

Monte Carlo Method & Option Valuation

From Arbitrage Pricing Theory we know that the price if an option is:

$$V(S_0) = \mathbb{E}^Q e^{-rT} h(S_T)$$

Where : S_0 is the current stock price

$\mathbb{E}^Q$ denote expection w.r.t the risk-neutral mesure

r is the risk free rate

T is the expiry date

h is the payoff function

- Suppose we could generate samples from Q (more on this later)
- Then option valuation simply reduces to calculating an expectation

Monte Carlo Method — Basic Idea

Suppose that X is a random variable with mean a and variance b^2

Monte Carlo Method: To estimate the mean of X take a large number of samples $\{\xi_i\}_{i=1}^N$

and compute: $a_N = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \xi_i$

a_N is known as the sample mean

b_N the sample variance is: $b_N^2 = \dfrac{1}{N-1}\displaystyle\sum_{i=1}^{N} (\xi_i - a_N)^2$

Exercise: Show that b_N^2 defined above is an unbiased estimate of the true variance b^2.

Monte Carlo Method — Error Estimates

Suppose that $Y \sim N(\mu, \sigma^2)$ then $\mathbb{P}(\left|\frac{Y-\mu}{\sigma}\right| \leq 1.96) \approx 0.95$

In words: 95% of normal samples lie within two standard deviations (approx)

CLT suggests that for large N the sample mean a_N is approximetly $\sim N(a, b^2/N)$ i.e.

$$\mathbb{P}(a_N - 1.96b/\sqrt{N} \leq a \leq a_N + 1.96b/\sqrt{N}) \approx 0.95$$

We don't know b in practice so we tend to use the sample variance instead and regard the interval:

$$\left[a_N - 1.96b_N/\sqrt{N}, a_N + 1.96b_N/\sqrt{N}\right]$$

as the 95% confidence interval for the true mean. i.e. if we ran MC simulation for a large number of times then we expect 95% of the estimates to lie in this interval

Monte Carlo Method — Error Estimates

$$\left[a_N - 1.96 b_N/\sqrt{N}, a_N + 1.96 b_N/\sqrt{N}\right]$$

The error estimate above is approximate since:

- CLT is only valid for large N
- We replaced the exact variance with the sample variance

Observations:

- The accuracy of the sample mean is proportional to the inverse square root of the number o samples. i.e to get one more decimal point we need to take 100 more samples.
- Monte Carlo is suited to applications that require less accuracy.
- The accuracy is proportional to the square root of the variance (more on this later)

Monte Carlo for Option Valuation

- Monte Carlo is well-suited for pricing **path dependent options**
- Often the **only** option when pricing high-dimensional options that have no closed form solution and when traditional methods suffer from the **curse of dimensionality**
- Monte Carlo methods are more challenging to implement for options with decision features (e.g. American Options)

Monte Carlo for Option Valuation

- Monte Carlo is well-suited for pricing **path dependent options**
- Often the **only** option when pricing high-dimensional options that have no closed form solution and when traditional methods suffer from the **curse of dimensionality**
- Monte Carlo methods are more challenging to implement for options with decision features (e.g. American Options)

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

Risk-Neutral Expectation

$$= e^{-rT} \int h(y, T) p(y, T \mid S_0) dy$$

Integration w.r.t. conditional density

$$\approx e^{-rT} \frac{1}{N} \sum_{i=1}^{N} h(y_i, T)$$

Monte Carlo Estimate

The Monte Carlo Algorithm for Option Valuation

1. **Partition** the time interval:

$$[0,T] \text{ to } 0 = t_0 < t_1 < \ldots < t_i < \ldots < t_m = T$$

2. **Generate samples** from the risk neutral measure (part 2)

$$y_{t_j,i} \quad i = 1,\ldots,N,\ j = 1,\ldots,m$$

3. Compute the average **(approx. option value)**:

$$V(S_0) \approx \bar{V}_N = e^{-rT}\frac{1}{N}\sum_{i=1}^{N} h(y_{T,i}, T)$$

4. Determine the **variance and standard error**

$$\bar{\sigma}_N^2 = \frac{1}{N-1}\sum_{i=1}^{N}(h(y_{T,i}) - \bar{V}_N)^2 \qquad \epsilon_N = \frac{1}{\sqrt{N}}\sigma_N$$

Monte Carlo for Calculation of Greeks I

- “Greeks’ are important for risk management e.g. hedging
- We discuss delta but others can be computed using similar principles

$$\Delta = \frac{\partial V(S,t)}{\partial S} = \lim_{h\to 0} \frac{V(S+h,t) - V(s,t)}{h}$$

$$\approx \frac{V(S+h,t) - V(s,t)}{h} \quad \text{for h small enough}$$

Monte Carlo for Calculation of Greeks I

- “Greeks’ are important for risk management e.g. hedging
- We discuss delta but others can be computed using similar principles

$$\Delta = \frac{\partial V(S,t)}{\partial S} = \lim_{h \to 0} \frac{V(S+h,t) - V(s,t)}{h}$$
$$\approx \frac{V(S+h,t) - V(s,t)}{h} \quad \text{for h small enough}$$

- Equation for delta suggests we can calculate it using two MC simulations
- Standard error for the two MC simulations is roughly $1/\sqrt{N}$
- For delta the standard error is $1/(h\sqrt{N})$
- Since h needs to be small we need a very big N to get the standard error small!

Monte Carlo for Calculation of Greeks II

- We can get some accuracy back by noting that the two paths S_0 and $S_0 + h$ are highly correlated
- The idea is to generate two simulations with the **same** random numbers but **different** initial conditions
- More on this when we discuss **variance reduction** techniques (we will also implement this in C++)

Monte Carlo & Stochastic Integration

- Stochastic integration of: $\int_0^T g(t)dW(t)$
- Where $g(t)$ is some determisitic function and $W(t)$ is the standard Wiener process

Monte Carlo & Stochastic Integration

- Stochastic integration of: $\int_0^T g(t)dW(t)$
- Where $g(t)$ is some determisitic function and $W(t)$ is the standard Wiener process

Approximate the integral by splitting the time interval into m “small” intervals $\Delta t = t_{j+1} - t_j$

$$\int_0^T g(t)dW(t) \approx \sum_{j=0}^{m-1} g(t_j)(W(t_{j+1}) - W(t_j))$$

Where $W(t_{j+1}) - W(t_j)$ has the same distribution as $Z\sqrt{\Delta t}$ with $Z \sim N(0,1)$

Monte Carlo & Stochastic Integration

- Stochastic integration of: $\int_0^T g(t)dW(t)$
- Where $g(t)$ is some determisitic function and $W(t)$ is the standard Wiener process

Approximate the integral by splitting the time interval into m "small" intervals $\Delta t = t_{j+1} - t_j$

$$\int_0^T g(t)dW(t) \approx \sum_{j=0}^{m-1} g(t_j)(W(t_{j+1}) - W(t_j))$$

Where $W(t_{j+1}) - W(t_j)$ has the same distribution as $Z\sqrt{\Delta t}$ with $Z \sim N(0,1)$

MC integration is based on the LLN and CLT and we need to simulate this many times

Variance Reduction for Monte Carlo Methods

$$\text{error} \approx \text{variance}/\sqrt{N}$$

- To reduce error without increasing N we should reduce variance (we need to increase samples x100 to get one more decimal point!)
- **Basic idea**: Generate another set of random samples with the same mean but lower variance

We will discuss:

- Variance Reduction by **antithetic variables**
- Variance Reduction by **control variates**
- **Importance Sampling**

Many techniques to achieve variance reduction (active research field, and beyond the scope of this course, please ask for references if you are interested)

Variance Reduction by Antithetic Variables

$$\text{error} \approx \text{variance}/\sqrt{N}$$

The key to variance reduction is dependent random variables

If X and Y are independent R.Vs then $\mathbb{E}(XY) = \mathbb{E}X\mathbb{E}X$

A measure of dependence is covariance: $\mathrm{cov}(X, Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y$

Variance Reduction by Antithetic Variables

$$\text{error} \approx \text{variance}/\sqrt{N}$$

The key to variance reduction is dependent random variables

If X and Y are independent R.Vs then $\mathbb{E}(XY) = \mathbb{E}X\mathbb{E}X$

A measure of dependence is covariance: $\text{cov}(X, Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y$

The Uniform Case: $V = \mathbb{E}(f(X))$, where $X \sim U[0,1]$

Variance Reduction by Antithetic Variables

$\text{error} \approx \text{variance}/\sqrt{N}$

The key to variance reduction is dependent random variables

If X and Y are independent R.Vs then $\mathbb{E}(XY) = \mathbb{E}X\mathbb{E}X$

A measure of dependence is covariance: $\text{cov}(X, Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y$

The Uniform Case: $V = \mathbb{E}(f(X))$, where $X \sim U[0,1]$

MC Estimate: $V_N = \dfrac{1}{N}\sum_{i=1}^{N}(f(X_i))$, where X_i is i.i.d from $U[0,1]$

Variance Reduction by Antithetic Variables

error $\approx$ variance$/\sqrt{N}$

The key to variance reduction is dependent random variables

If X and Y are independent R.Vs then $\mathbb{E}(XY) = \mathbb{E}X\mathbb{E}X$

A measure of dependence is covariance: $\text{cov}(X, Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y$

The Uniform Case: $V = \mathbb{E}(f(X)),$ where $X \sim U[0,1]$

MC Estimate: $V_N = \dfrac{1}{N}\sum_{i=1}^{N}(f(X_i)),$ where X_i is i.i.d from $U[0,1]$

Antithetic MC Estimate: $\hat{V}_N = \dfrac{1}{N}\sum_{i=1}^{N}\dfrac{1}{2}(f(X_i) + f(1 - X_i)),$ where X_i is i.i.d from $U[0,1]$

Variance Reduction by Antithetic Variables

Antithetic MC Estimate: $\hat{V}_N = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}(f(X_i) + f(1 - X_i))$, where X_i is i.i.d from $U[0,1]$

When/Why will this work?

$$\mathrm{var}\left(\frac{f(X_i) + f(1 - X_i)}{2}\right) = \frac{1}{4}\mathrm{var}(f(X_i)) + \frac{1}{4}\mathrm{var}(f(1 - X_i)) + \frac{1}{2}\mathrm{cov}(f(X_i), f(1 - X_i))$$

$$= \frac{1}{2}\mathrm{var}(f(X_i)) + \frac{1}{2}\mathrm{cov}(f(X_i), f(1 - X_i)) \qquad \text{since } X_i \text{ and } 1 - X_i \text{ are } U[0,1]$$

Variance Reduction by Antithetic Variables

Antithetic MC Estimate: $\hat{V}_N = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}(f(X_i) + f(1 - X_i)),$ where X_i is i.i.d from $U[0,1]$

When/Why will this work?

$$\mathrm{var}\left(\frac{f(X_i) + f(1 - X_i)}{2}\right) = \frac{1}{4}\mathrm{var}(f(X_i)) + \frac{1}{4}\mathrm{var}(f(1 - X_i)) + \frac{1}{2}\mathrm{cov}(f(X_i), f(1 - X_i))$$

$$= \frac{1}{2}\mathrm{var}(f(X_i)) + \boxed{\frac{1}{2}\mathrm{cov}(f(X_i), f(1 - X_i))}$$

since X_i and $1 - X_i$ are $U[0,1]$

The success of the method depends on making this term as negative as possible

Variance Reduction by Antithetic Variables

Antithetic Variance: $$\mathrm{var}\left(\frac{f(X_i)+f(1-X_i)}{2}\right)=\frac{1}{2}\mathrm{var}(f(X_i))+\frac{1}{2}\mathrm{cov}(f(X_i),f(1-X_i))$$

Conclusion: If f is monotonic then **antithetic variables** are likely to be useful

A function is monotonically increasing if $x \leq y \implies f(x) \leq f(y)$

A function is monotonically decreasing if $x \leq y \implies f(x) \geq f(y)$

Note: Most pay-offs in finance are monotonic

Exercise: Extend the analysis above to the Gaussian case

Variance Reduction by Control Variates

Antithetic variates relied on anti-correlated RVs whereas the **Control Variate (CV)** method relies on **known correlations**

Suppose we need to estimate: $\mathbb{E}[X]$ and suppose we knew that another r.v Y that is 'close' to X with a known mean $\mathbb{E}[Y]$ then

$$Z = X + \mathbb{E}[Y] - Y$$
$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y] - \mathbb{E}[Y] = \mathbb{E}[X]$$

Basic idea: Apply Monte Carlo to the expectation of Z instead of X.

Variance Reduction by Control Variates

Basic idea: Apply Monte Carlo to the expectation of Z instead of X.

$$Z = X + \mathbb{E}[Y] - Y$$
$$\mathrm{var}(Z) = \mathrm{var}(X + \mathbb{E}[Y] - Y) = = \mathrm{var}(X - Y)$$

Conclusion: CV should work if $X - Y \approx 0$ and $\mathbb{E}Y$ is easy to compute

If $\mathrm{var}(Z) = R_1 \mathrm{var}(X)$ for some $R_1 < 1$ and the cost of Z is R_2 times the cost of sampling from X the overall gain is $R_1 R_2$ (why?) so CV makes sense if $R_1 R_2 < 1$.

Variance Reduction by Control Variates

Generalisation:

$$Z_\theta = X + \theta(\mathbb{E}[Y] - Y)$$

$$\mathrm{var}(Z_\theta) = \mathrm{var}(X - \theta Y) = \mathrm{var}(X) + \theta^2 \mathrm{var}(Y) - 2\theta \mathrm{cov}(X, Y)$$

Optimal Theta:

Find θ so that $\dfrac{\partial \mathrm{var}(Z_\theta)}{\partial \theta} = 0$

Solution: $\theta_{min} = \dfrac{\mathrm{cov}(X, Y)}{\mathrm{var}(Y)}$

Importance Sampling

Aim to calculate the expectation of $g(\cdot)$ for some r.v. X with pdf $f_X(x)$:

$$\mathbb{E}^X[g(X)] = \int_{\mathbb{R}} g(x) f_X(x) dx = \int_{\mathbb{R}} g(x) dF_X(x)$$

Rewrite the expectation above as follows:

$$\begin{aligned}\int_{\mathbb{R}} g(x) dF_X(x) = \int_{\mathbb{R}} g(x) f_X(x) dx &= \int_{\mathbb{R}} g(x) f_Y(x) \frac{f_X(x)}{f_Y(x)} dx \\ &= \int_{\mathbb{R}} g(x) \frac{f_X(x)}{f_Y(x)} dF_Y(x)\end{aligned}$$

Importance Sampling

Rewrite the expectation above as follows: $\mathbb{E}^X[g(X)] = E^Y[g(Y)L(Y)]$ where $L(x) = \dfrac{f_X(x)}{f_Y(x)}$

The function $L(x)$ is usually refered to as: score function/likelyhood ratio or Radon-Nikodym derivative .

$\mathbb{E}^X[g(X)] \approx \dfrac{1}{N}\sum g(x_i)$ where x_i has p.d.f f_X standard MC

$\mathbb{E}^X[g(X)] \approx \dfrac{1}{N}\sum g(y_i)L(y_i)$ where y_i has p.d.f f_Y Importance Sampling MC

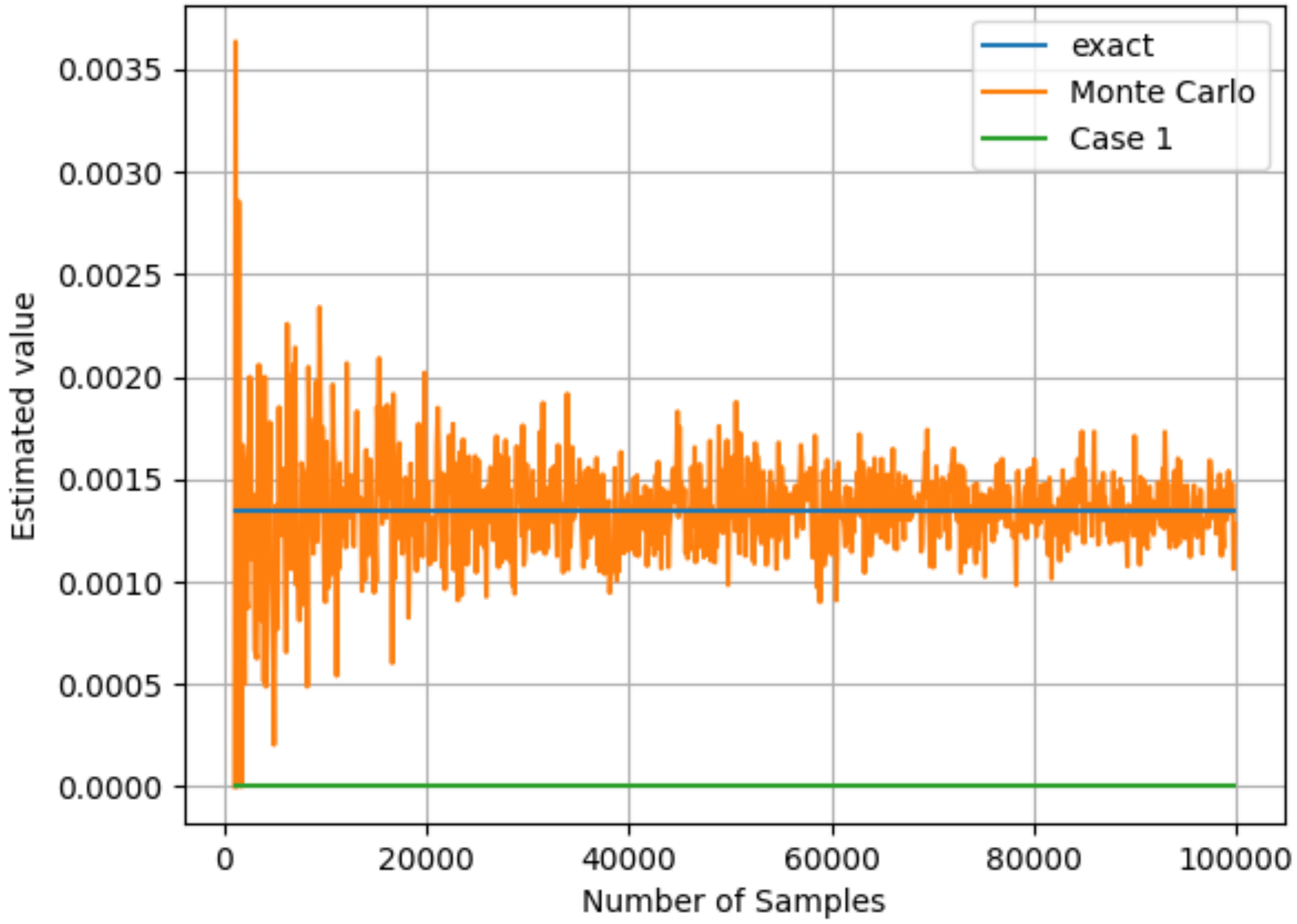
Note: For importance sampling to be valid then both densities need to have the same support

Importance Sampling

Example: Perform MC to approximate $\mathbb{P}[X > 3]$ with $X \sim N(0,1)$.

We compare importance sampling with the following four options:

1. $U[0,1]$
2. $U[0,4]$
3. $N(0,0.25)$
4. $N(0,9)$
5. $N(3,1)$

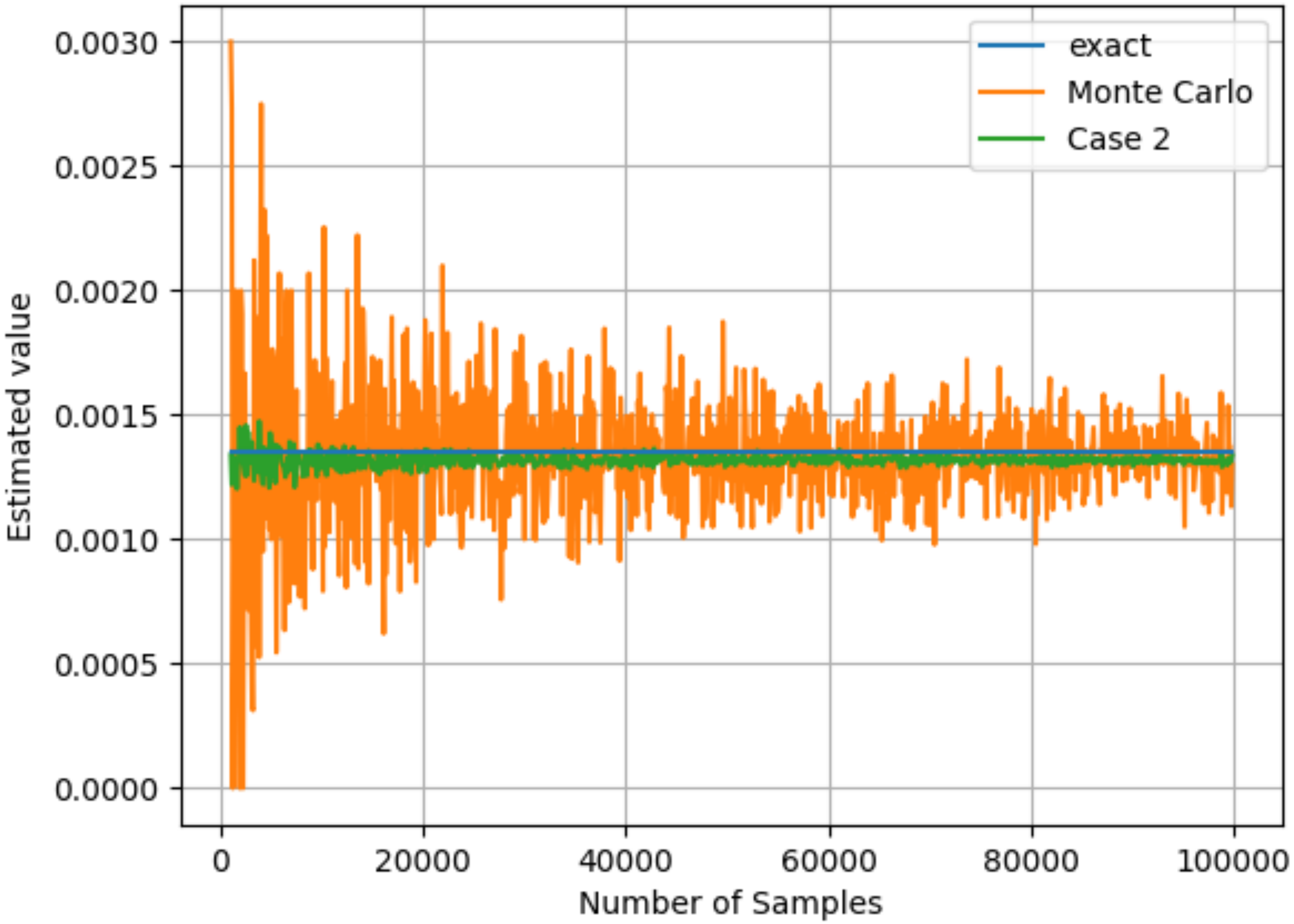


Importance Sampling

Example: Perform MC to approximate $\mathbb{P}[X > 3]$ with $X \sim N(0,1)$.

We compare importance sampling with the following four options:

1. $U[0,1]$
2. $U[0,4]$
3. $N(0,0.25)$
4. $N(0,9)$
5. $N(3,1)$

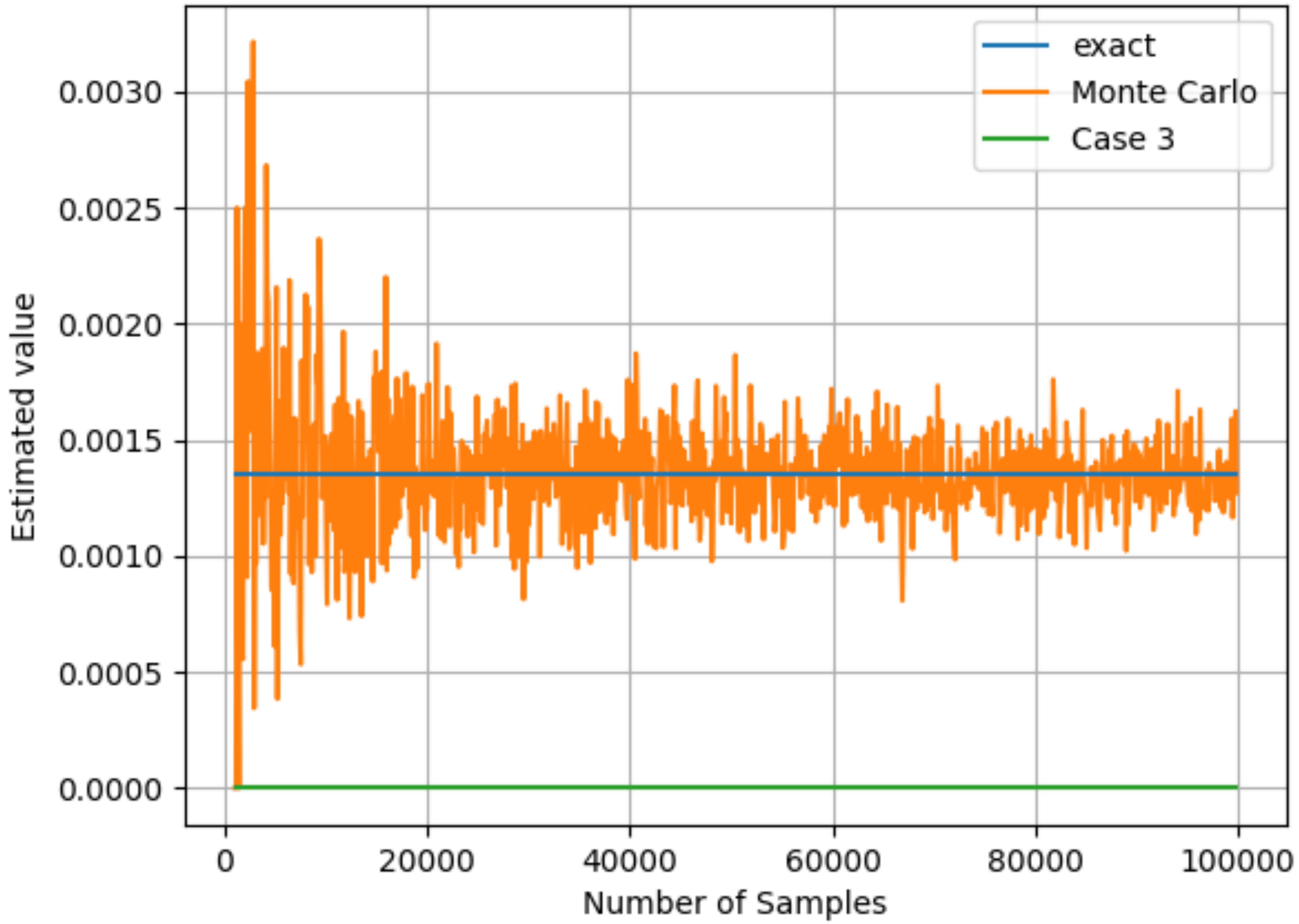


Importance Sampling

Example: Perform MC to approximate $\mathbb{P}[X > 3]$ with $X \sim \ N(0,1)$.

We compare importance sampling with the following four options:

1. $U[0,1]$
2. $U[0,4]$
3. $N(0,0.25)$
4. $N(0,9)$
5. $N(3,1)$

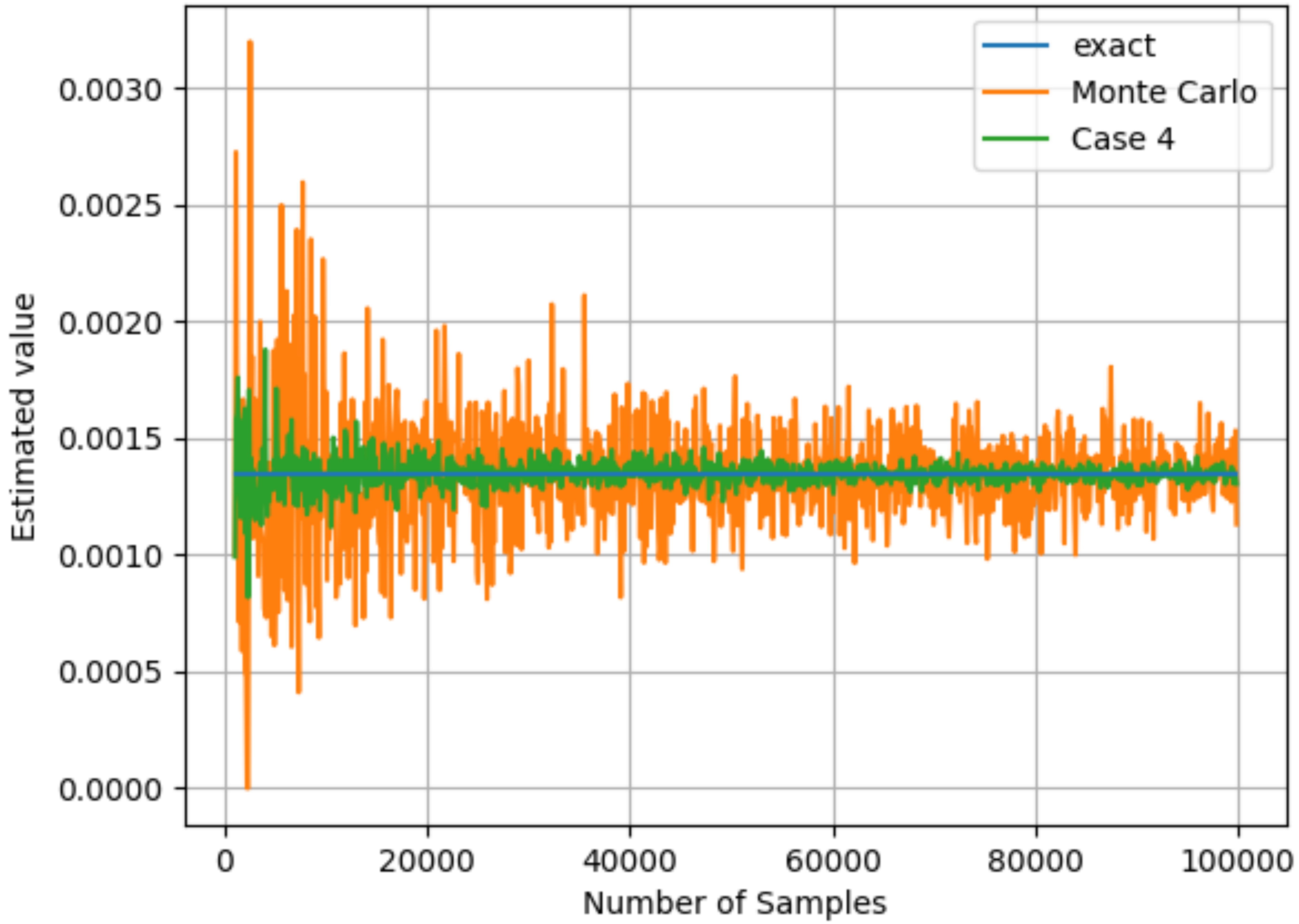


Importance Sampling

Example: Perform MC to approximate $\mathbb{P}[X > 3]$ with $X \sim N(0,1)$.

We compare importance sampling with the following four options:

1. $U[0,1]$
2. $U[0,4]$
3. $N(0,0.25)$
4. $N(0,9)$
5. $N(3,1)$

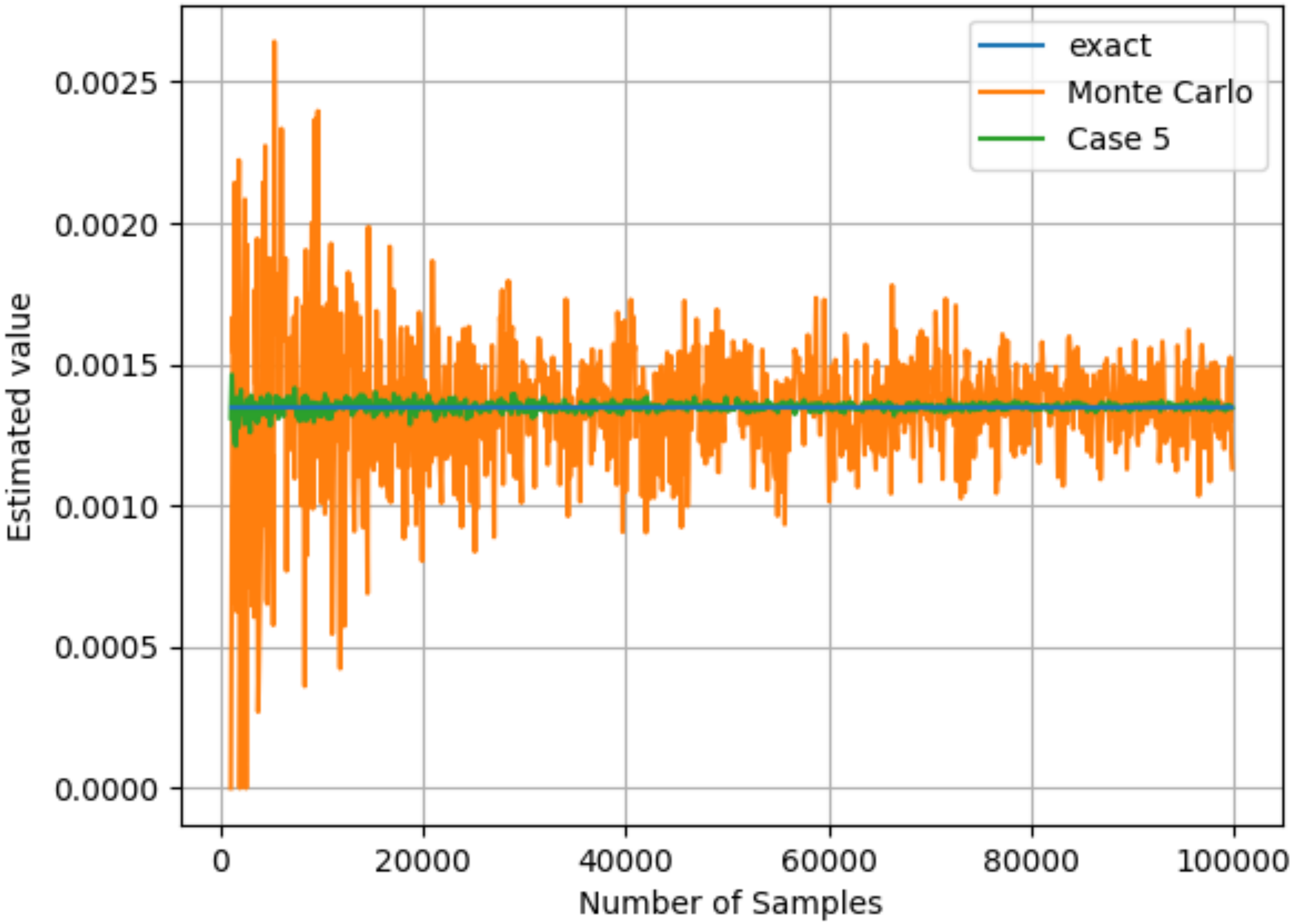


Importance Sampling

Example: Perform MC to approximate $\mathbb{P}[X > 3]$ with $X \sim \ N(0,1)$.

We compare importance sampling with the following four options:

1. $U[0,1]$
2. $U[0,4]$
3. $N(0,0.25)$
4. $N(0,9)$
5. $N(3,1)$



Monte Carlo Methods and Numerical Methods for SDEs

Part I: Monte Carlo Methods

- Mathematical Foundations
- Monte Carlo and option valuation
- Calculation of Greeks
- Variance Reduction via Antithetic Variables
- Variance Reduction via Control Variates
- Importance Sampling

Part II: Numerical Methods for SDEs:

- Euler-Maruyama
- Weak and Strong Convergence
- Implicit Methods and Numerical Stability
- Mean Exit Times
- Numerical Methods for Systems

Euler-Maruyama Scheme

Numerical Simulation of Stocahstic Differential Equation (SDE) :

$$dX(t) = f(X(t))dt + g(X(t))dW(t), \quad 0 \le t \le T, \text{ with } X(0) \text{ given}$$

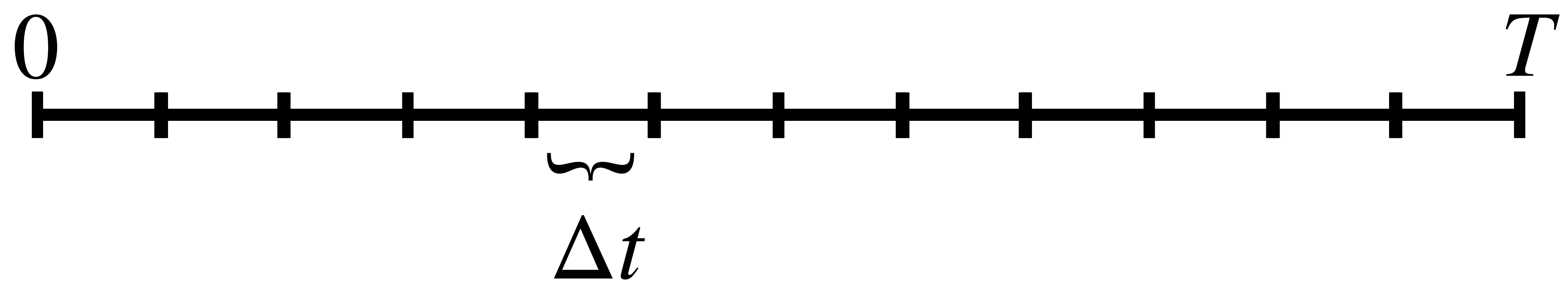
Define a stepsize $\Delta t = T/N$ where N is the number of intervals we divide $[0,T]$

Euler-Maruyama Scheme

Numerical Simulation of Stocahstic Differential Equation (SDE) :

$$dX(t) = f(X(t))dt + g(X(t))dW(t), \quad 0 \le t \le T, \text{ with } X(0) \text{ given}$$

Define a stepsize $\Delta t = T/N$ where N is the number of intervals we divide $[0,T]$

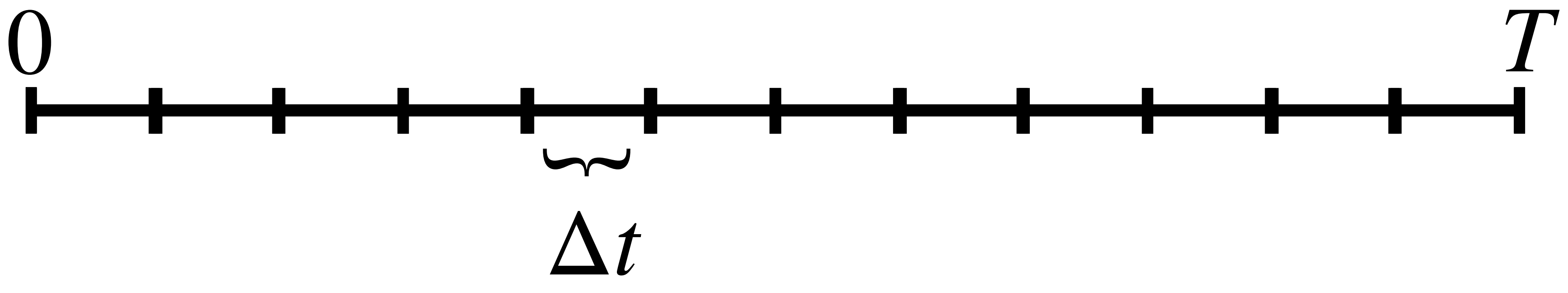


Euler-Maruyama Scheme

Numerical Simulation of Stocahstic Differential Equation (SDE) :

$$dX(t) = f(X(t))dt + g(X(t))dW(t), \quad 0 \leq t \leq T, \text{ with } X(0) \text{ given}$$

Define a stepsize $\Delta t = T/N$ where N is the number of intervals we divide $[0,T]$



Write SDE in integral form:

$$X(t_{n+1}) = X(t_n) + \int_{t_n}^{t_n+1} f(X(s))ds + \int_{t_n}^{t_n+1} g(X(s))dW(s), \quad \text{where } t_n = n\Delta t$$

In the EM scheme we 'pretend' that f and g are consant in each time interval.

This assumption should be valid if Δt is 'small'

Euler-Maruyama Scheme

SDE in integral form:

$$X(t_{n+1}) = X(t_n) + \int_{t_n}^{t_n+1} f(X(s))ds + \int_{t_n}^{t_n+1} g(X(s))dW(s), \quad \text{where } t_n = n\Delta t$$

If f and g are consant in each time interval:

$$X_{n+1} = X_n + f(X_n)\Delta t + g(X_n)\Delta W_n, \quad \text{where } \Delta W_n \sim N(0,\Delta t) \text{ is a Brownian path increment}$$

1. Set stepsize $\Delta t = T/N$, assume X_0 is given.
2. for $n = 0\ldots, N-1$

   Compute $\xi_n \sim N(0,1)$

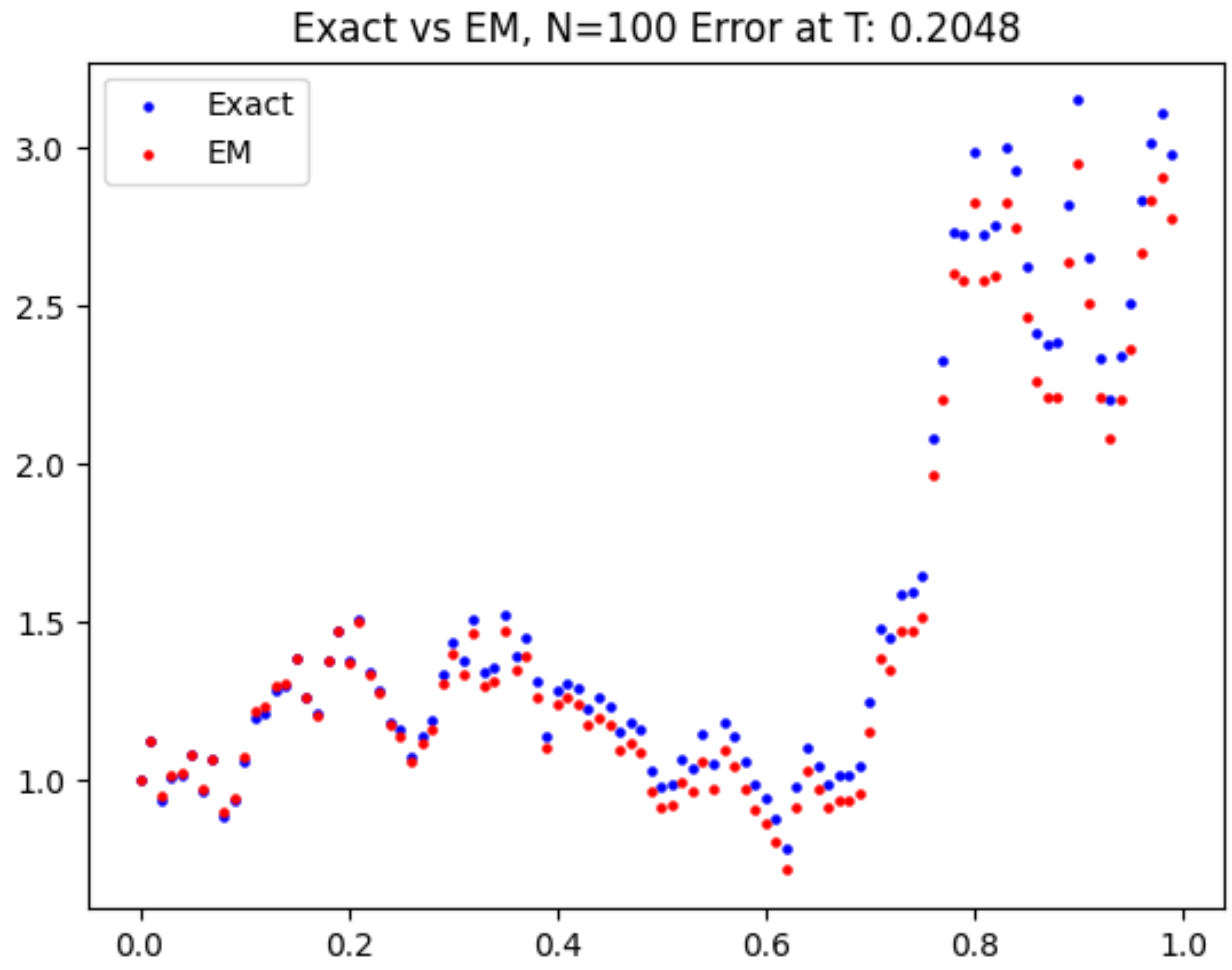
   $X_{n+1} = X_n + f(X_n)\Delta t + g(X_n)\xi_n\sqrt{\Delta t}$

EM is a generalisation of the Euler scheme for deterministic ODEs. We look at its convergence properties next.

Numerical Example

$dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$ (Geometric BM)

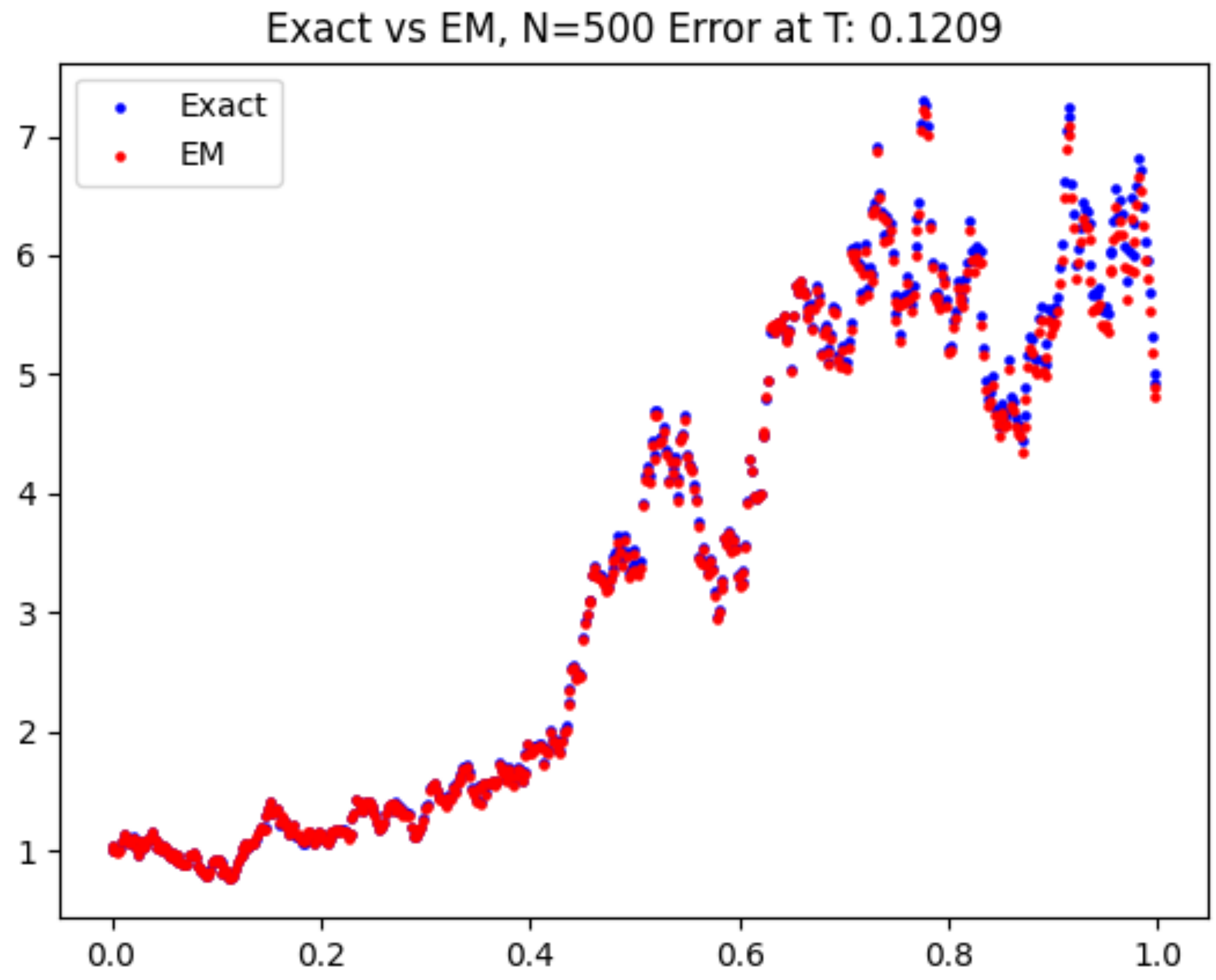
$X(t_n) = X(t)\exp((\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma\xi_n\sqrt{\Delta t})$ (Closed form solution)



Numerical Example

$dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$ (Geometric BM)

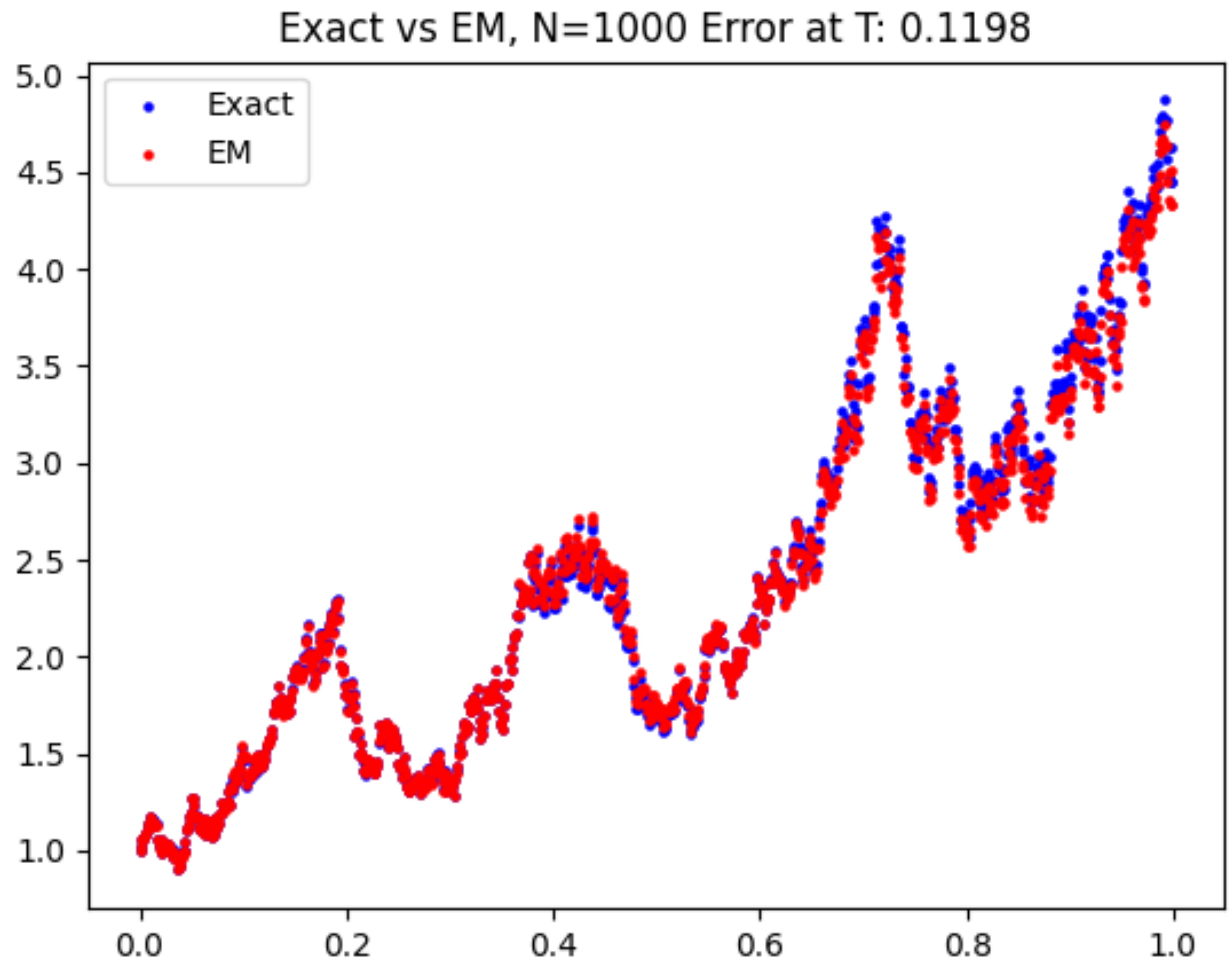
$X(t_n) = X(t)\exp((\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma \xi_n \sqrt{\Delta t})$ (Closed form solution)



Numerical Example

$dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$ (Geometric BM)

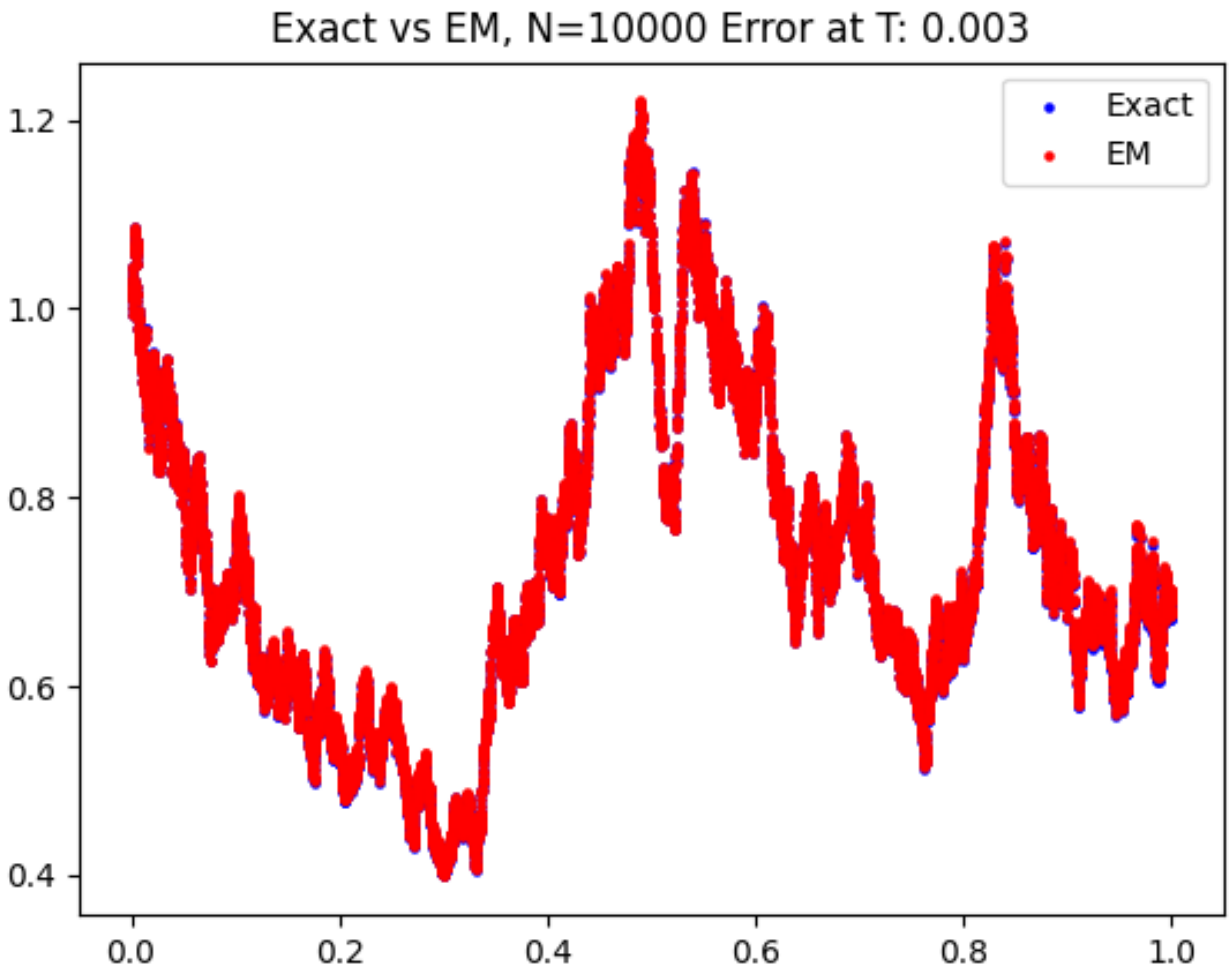
$X(t_n) = X(t)\exp((\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma \xi_n \sqrt{\Delta t})$ (Closed form solution)



Numerical Example

$dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$ (Geometric BM)

$X(t_n) = X(t)\exp((\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma\xi_n\sqrt{\Delta t})$ (Closed form solution)



Weak Convergence of EM

Many applications we intrested in the average solution behaviour

Weak error is defined as:

$$e^w_{\Delta t} = \sup_{0 \leq t_n \leq T} |\mathbb{E}\Phi(X_n) - \mathbb{E}\Phi(X(t_n))| \quad \text{where } \Phi \text{ is a given function}$$

Φ is usually constrained to some class of functions e.g. polynomilas of a certain degree

Weak convergence means that for any Φ within a class we get

$$e^w_{\Delta t} \to 0 \text{ as } \Delta t \to 0$$

We say a method has weak order of convergence p if for some K and $\overline{\Delta t}$

$$e^w_{\Delta t} \leq K \Delta t^p \quad \forall 0 < \Delta t \leq \overline{\Delta t}$$

Weak Convergence Numerical Example

Same example from previous slide (geometric Brownian motion)

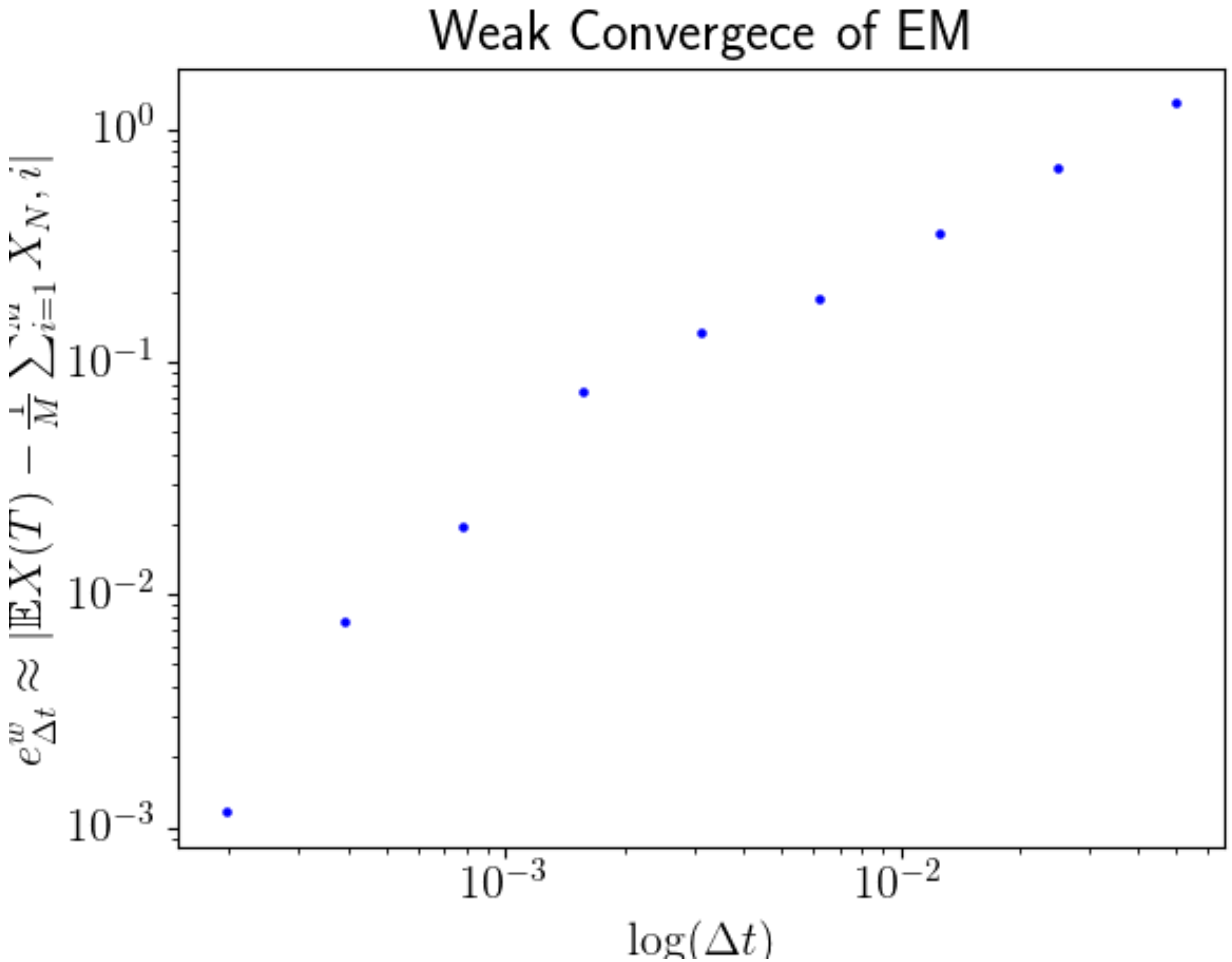
$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t) \qquad \text{(Geometric BM)}$$

$$\mathbb{E}X(T) = \exp(\mu T) \qquad \text{(Closed form solution)}$$

$$e^{w}_{\Delta t} \approx |\,\mathbb{E}X(T) - \frac{1}{M}\sum_{i=1}^{M} X_N, i\,| \qquad \text{(Error estimation with MC)}$$

Weak Convergence Numerical Example

Weak Convergence Numerical Example



Numerical experiment suggest EM converges weakly with order p=1

This can be shown rigorously (ask for references if interested)

Weak Euler-Maruyama Method

- To achieve weak convergence with the EM scheme, exact simulation of Brownian Motion is not necessary
- The following algorithm is called the *weak Euler-Maruyama method*:

$$X_{n+1} = X_n + f(X_n)\Delta t + g(X_n)\overline{\Delta W_n}$$

Where $\overline{\Delta W_n}$ are i.i.d random variables taking two possible values

$$\mathbb{P}[\overline{\Delta W_n} = \sqrt{\Delta t}] = \mathbb{P}[\overline{\Delta W_n} = -\sqrt{\Delta t}] = \frac{1}{2}$$

The advantage of the weak EM scheme is that has the same convergence properties as the EM method but it is cheaper to compute, especially good for large simulations on GPUs

Strong Convergence of EM

Weak error measures the error of the means, **strong convergence measures the mean of the errors:**

$$e^s_{\Delta t} = \sup_{0 \leq t_n \leq T} \mathbb{E}\,|\,X_n - X(t_n)\,|$$

Strong convergence means that

$$e^s_{\Delta t} \to 0 \text{ as } \Delta t \to 0$$

We say a method has strong order of convergence p if for some K and $\overline{\Delta t}$

$$e^s_{\Delta t} \leq K \Delta t^p \qquad \forall 0 < \Delta t \leq \overline{\Delta t}$$

Strong Convergence of EM

1. It can be shown that EM converges strongly with $p = \dfrac{1}{2}$
2. Note the sharp contrast between the deterministic and stochastic case
3. If the noise term is constant it can be shown that EM converges strongly with $p = 1$
4. Strong convergence also has implications about individual paths:

$$\mathbb{E}[\,|X_n - X(t_n)|\,] \leq K\Delta t^{\frac{1}{2}} \implies \mathbb{P}(\,|X_n - X(t_n)| \geq \Delta t^{\frac{1}{4}}] \leq K\Delta t^{\frac{1}{4}} \implies \mathbb{P}(\,|X_n - X(t_n)| < \Delta t^{\frac{1}{4}}] \geq K\Delta t^{\frac{1}{4}}$$

Where we used the Markov Inequality with $a = \dfrac{1}{4}$

$$\mathbb{P}(\,|X| \geq a] \leq \frac{\mathbb{E}[\,|X|\,]}{a}$$

Along any path the error will be small with probability close to 1

Implicit Methods and Numerical Stability

Weak and strong convergence are asymptotic results and are valid for small enough step-size i.e. finite interval [0,T], and stepsize $\Delta t \to 0$

We are also intrested in the case when Δt is fixed and T is large.

Stability analysis address propagation of errors: i.e. if a small error is committed at one step (e.g. large time step) are the errors magnified over time?

Implicit Methods and Numerical Stability

Weak and strong convergence are asymptotic results and are valid for small enough step-size i.e. finite interval [o,T], and stepsize $\Delta t \to 0$

We are also intrested in the case when Δt is fixed and T is large.

Stability analysis address propagation of errors: i.e. if a small error is committed at one step (e.g. large time step) are the errors magnified over time?

The Euler method is known to be inferior to implicit schemes in deterministic settings, how about for SDEs?

Stochastic Theta-Method

Given $\theta \in [0,1]$ the stochstic θ – method is

$$X_{n+1} = X_n + (1-\theta)f(X_n)\Delta t + \theta f(X_{n+1})\Delta t + g(X_n)\Delta W_n$$

When $\theta = 0$ we get the Euler method

When $\theta > 0$ the method is called implicit

1. Set stepsize $\Delta t = T/N,$ assume X_0 is given.
2. for $n = 0\ldots, N-1$

   Compute $\xi_n \sim N(0,1)$

   Solve for X_{n+1} the following equation

   $X_{n+1} = X_n + (1-\theta)f(X_n)\Delta t + \theta f(X_{n+1})\Delta t + g(X_n)\xi_n\sqrt{\Delta t}$

Because of the non-linear equation that needs to be solved in every iteration implicit methods are harder to analyse. In this course we will study the linear case

Stochastic Theta-Method — The Linear Case

We study the stochstic θ – method for the following linear SDE

$$dX(t) = \mu X(t) + \sigma X(t) dW(t)$$

$$X(t) = X(0) e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)} \qquad \text{(Closed-form solution)}$$

$$\mathbb{E} X(t)^2 = X(0)^2 e^{(2\mu + \sigma^2)t} \qquad \text{(Exercise in Tutorial)}$$

Stochastic Theta-Method — The Linear Case

We study the stochstic θ – method for the following linear SDE

$$dX(t) = \mu X(t) + \sigma X(t) dW(t)$$

$$X(t) = X(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)} \qquad \text{(Closed-form solution)}$$

$$\mathbb{E}X(t)^2 = X(0)^2 e^{(2\mu + \sigma^2)t} \qquad \text{(Exercise in Tutorial)}$$

It follows that:

$$\lim_{t \to \infty} \mathbb{E}[X(t)^2] = 0 \iff \mu + \frac{1}{2}\sigma^2 < 0$$

Mean Square

Stability Condition

Stochastic Theta-Method — The Linear Case

We study the stochastic θ – method for the following linear SDE

$$dX(t) = \mu X(t) + \sigma X(t) dW(t)$$

$$X(t) = X(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)} \qquad \text{(Closed-form solution)}$$

$$\mathbb{E}X(t)^2 = X(0)^2 e^{(2\mu+\sigma^2)t} \qquad \text{(Exercise in Tutorial)}$$

Stochastic Theta-Method — The Linear Case

We study the stochastic θ – method for the following linear SDE

$$dX(t) = \mu X(t) + \sigma X(t) dW(t)$$

$$X(t) = X(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)} \qquad \text{(Closed-form solution)}$$

$$\mathbb{E}X(t)^2 = X(0)^2 e^{(2\mu + \sigma^2)t} \qquad \text{(Exercise in Tutorial)}$$

It follows that:

$$\lim_{t\to\infty} \mathbb{E}[X(t)^2] = 0 \iff \mu + \frac{1}{2}\sigma^2 < 0$$

Mean Square

Stability Condition

Asymptotic Stability

A stochastic process is asymptotically stable if: $\lim_{t \to \infty} |X(t)| = 0$ with probablity 1

It can be shown that for the GBM SDE:

$$\lim_{t \to \infty} |X(t)| = 0 \text{ with probablity } 1 \iff \mu - \frac{1}{2}\sigma^2 < 0$$

Asymptotic Stability

A stochastic process is asymptotically stable if: $\lim_{t \to \infty} |X(t)| = 0$ with probablity 1

It can be shown that for the GBM SDE:

$$\lim_{t \to \infty} |X(t)| = 0 \text{ with probablity } 1 \iff \mu - \frac{1}{2}\sigma^2 < 0$$

- For GBM mean square stable => asymptotically stable
- This implication does not generally hold
- Mean-square stability analysis more common (easier to perform)
- Asymptotic stability more useful in applications

Mean-Square Stability of the theta method

Mean-Square Stability of the theta method

We need to show that the EM scheme satisfies: $\lim_{n\to\infty} \mathbb{E}|X_n|^2 = 0$

Mean-Square Stability of the theta method

We need to show that the EM scheme satisfies: $\lim_{n\to\infty} \mathbb{E}|X_n|^2 = 0$

Stochastic θ – method for GBM where, $\xi_n \sim N(0,1)$:

$$X_{n+1} = X_n + (1-\theta)\mu X_n \Delta t + \theta \mu X_{n+1}\Delta t + \sigma\sqrt{\Delta t}\xi_n$$

Mean-Square Stability of the theta method

We need to show that the EM scheme satisfies: $\lim_{n\to\infty} \mathbb{E}\,|X_n|^2 = 0$

Stochastic θ – method for GBM where, $\xi_n \sim N(0,1)$:

$$X_{n+1} = X_n + (1-\theta)\mu X_n \Delta t + \theta\mu X_{n+1}\Delta t + \sigma\sqrt{\Delta t}\xi_n$$

Rewrite: $$(1-\theta\mu\Delta t)X_{n+1} = X_n(1 + (1-\theta)\mu\Delta t + \sigma\sqrt{\Delta t}\xi_n)$$

Mean-Square Stability of the theta method

We need to show that the EM scheme satisfies: $\lim_{n\to\infty} \mathbb{E}\,|X_n|^2 = 0$

Stochastic θ – method for GBM where, $\xi_n \sim N(0,1)$:

$$X_{n+1} = X_n + (1-\theta)\mu X_n \Delta t + \theta\mu X_{n+1}\Delta t + \sigma\sqrt{\Delta t}\xi_n$$

Rewrite: $(1-\theta\mu\Delta t)X_{n+1} = X_n(1 + (1-\theta)\mu\Delta t + \sigma\sqrt{\Delta t}\xi_n)$

Sqaure and take expectations:

$$(1-\theta\mu\Delta t)^2\mathbb{E}[X_{n+1}^2] = ((1+(1-\theta)\mu\Delta t)^2 + \sigma^2\Delta t)\mathbb{E}X_n^2$$

Mean-Square Stability of the theta method

We need to show that the EM scheme satisfies: $\lim_{n\to\infty} \mathbb{E}|X_n|^2 = 0$

Stochastic θ – method for GBM where, $\xi_n \sim N(0,1)$:

$$X_{n+1} = X_n + (1-\theta)\mu X_n \Delta t + \theta \mu X_{n+1} \Delta t + \sigma \sqrt{\Delta t}\xi_n$$

Rewrite: $(1-\theta\mu\Delta t)X_{n+1} = X_n(1 + (1-\theta)\mu\Delta t + \sigma\sqrt{\Delta t}\xi_n)$

Sqaure and take expectations:

$$(1-\theta\mu\Delta t)^2\mathbb{E}[X_{n+1}^2] = ((1+(1-\theta)\mu\Delta t)^2 + \sigma^2\Delta t)\mathbb{E}X_n^2$$

Re-arrange (assuming) $1 - \theta\Delta t\mu \neq 0$:

$$\mathbb{E}[X_{n+1}^2] = \frac{(1+(1-\theta)\mu\Delta t)^2 + \sigma^2\Delta t}{(1-\theta\mu\Delta t)^2}\mathbb{E}X_n^2$$

Mean-Square Stability of the theta method

Since: $\mathbb{E}[X_{n+1}^2] = \dfrac{(1+(1-\theta)\mu\Delta t)^2 + \sigma^2\Delta t}{(1-\theta\mu\Delta t)^2}\mathbb{E}X_n^2$

$$\lim_{n\to\infty}\mathbb{E}[X_n^2] = 0 \iff \frac{(1+(1-\theta)\mu\Delta t)^2 + \sigma^2\Delta t}{(1-\theta\mu\Delta t)^2} < 1$$

Can be re-arranged to: $\lim_{n\to\infty}\mathbb{E}[X_n^2] = 0 \iff \Delta t(1-2\theta)\mu^2 < -2(\mu + \frac{1}{2}\sigma^2)$

Mean-Square Stability of the theta method

$$\lim_{n\to\infty} \mathbb{E}[X_n^2] = 0 \iff \Delta t(1-2\theta)\mu^2 < -2(\mu + \frac{1}{2}\sigma^2)$$

Case 1: $0 \leq \theta < \frac{1}{2}$, θ – method is stable for $\Delta t < \frac{2|\mu + \sigma^2/2|}{(1-2\theta)\mu^2}$

and SDE not stable $\implies$ θ – method is not stable for any $\Delta t > 0$

Case 2: and SDE stable $\iff$ θ – method is stable for all $\Delta t > 0$

Case 3: $\frac{1}{2} < \theta \leq 1$, SDE stable $\implies$ θ – method is stable for all $\Delta t > 0$

SDE not stable $\implies$ θ – method is stable for $\Delta t < \frac{2|\mu + \sigma^2/2|}{(2\theta-1)\mu^2}$

It is possible to perform the asymptotic stability analysis of the theta method but we skip this in the interest of time

Mean Exit Times

Examples of exit times problems in finance:

- How long we have to wait before a stock exits a certain band?
- How long before a stock reaches the barrier of a barrier option?

Exit times are closely related to optimal stopping problems that form the mathematical foundations of algorithmic trading

Mean Exit Times

Examples of exit times problems in finance:

- How long we have to wait before a stock exits a certain band?
- How long before a stock reaches the barrier of a barrier option?

Exit times are closely related to optimal stopping problems that form the mathematical foundations of algorithmic trading

$$dX(t) = f(X(t))dt + g(X(t))dW(t), \qquad \text{with } X(0) \text{ given}$$

$$T_{\text{exit}} = \inf\{t : X(t) = a \text{ or } X(t) = b\}$$

i.e. the first time X leaves the interval (a, b)

Goal is to compute:

$$\bar{T}_{\text{exit}} = \mathbb{E}T_{\text{exit}}$$

Monte Carlo for Mean Exit Time

Monte Carlo for Mean Exit Time

Choose a step-size Δt

Choose a number of paths M

for $s = 1\ldots, M$

While $a < X_n < b$

Compute a $\xi_n \sim N(0,1)$

$$X_{n+1} = X_n + f(X_n)\Delta t + g(X_n)\xi_n\sqrt{\Delta t}$$
$$t_{n+1} = t_n$$

$$T^s_{\text{exit}} = t_n - \frac{1}{2}\Delta t$$

$$a_M = \frac{1}{M}\sum_{s=1}^{M} T^s_{\text{exit}}$$

$$b_M^2 = \frac{1}{M-1}\sum_{s=1}^{M}(T^s_{\text{exit}} - a_M)^2$$

Monte Carlo for Mean Exit Time

Choose a step-size Δt

Choose a number of paths M

for $s = 1\ldots, M$

While $a < X_n < b$

Compute a $\xi_n \sim N(0,1)$

$$X_{n+1} = X_n + f(X_n)\Delta t + g(X_n)\xi_n\sqrt{\Delta t}$$

$$t_{n+1} = t_n$$

$$T^s_{\text{exit}} = t_n - \frac{1}{2}\Delta t$$

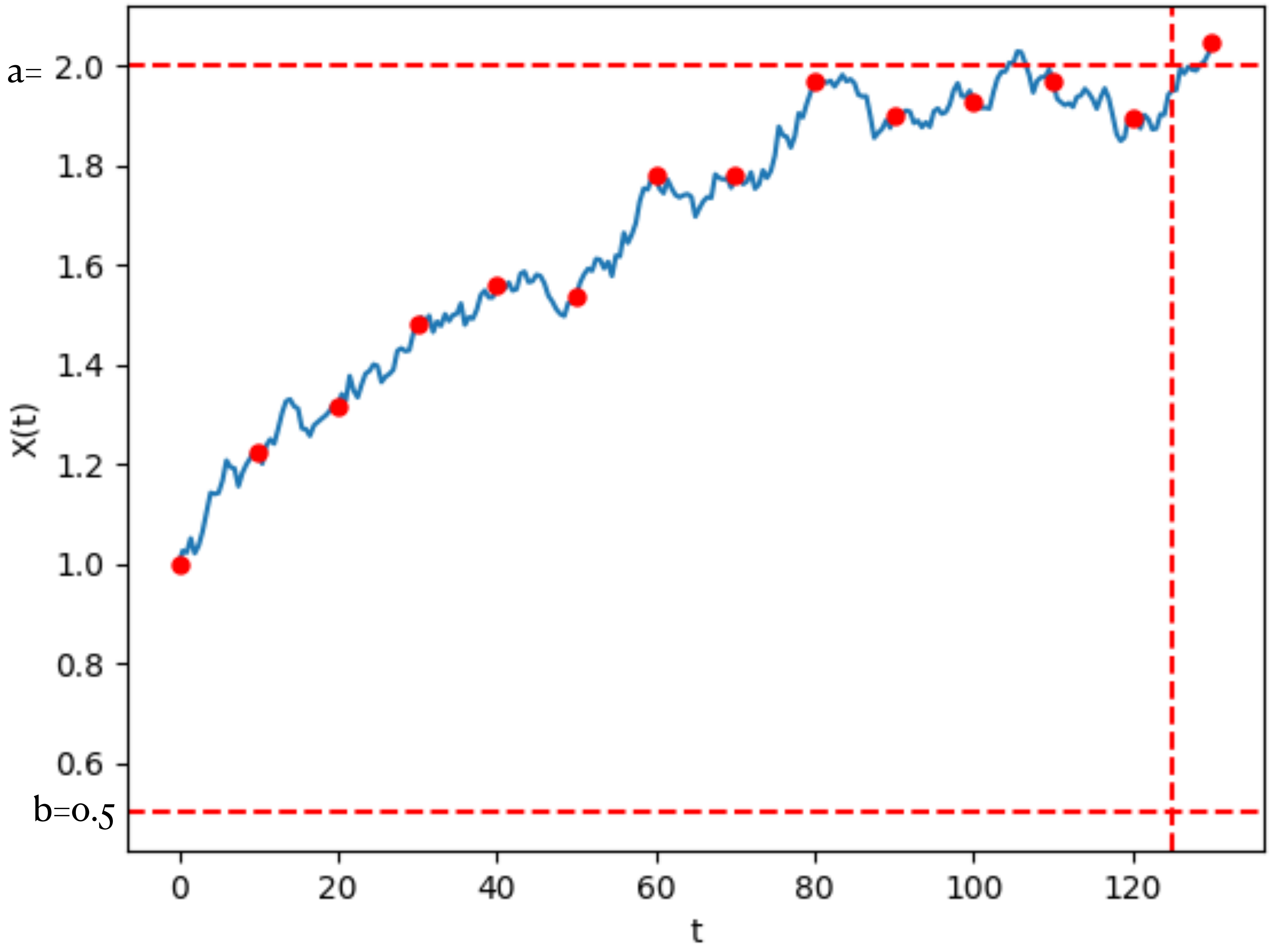
$$a_M = \frac{1}{M}\sum_{s=1}^{M} T^s_{\text{exit}}$$

$$b_M^2 = \frac{1}{M-1}\sum_{s=1}^{M} (T^s_{\text{exit}} - a_M)^2$$

Sources of Error:

- Sampling error (approximate expectations with MC)
- Discretisation error (explicit Euler)
- Only observe process on a grid (not continuously)

Monte Carlo for Mean Exit Time — Example

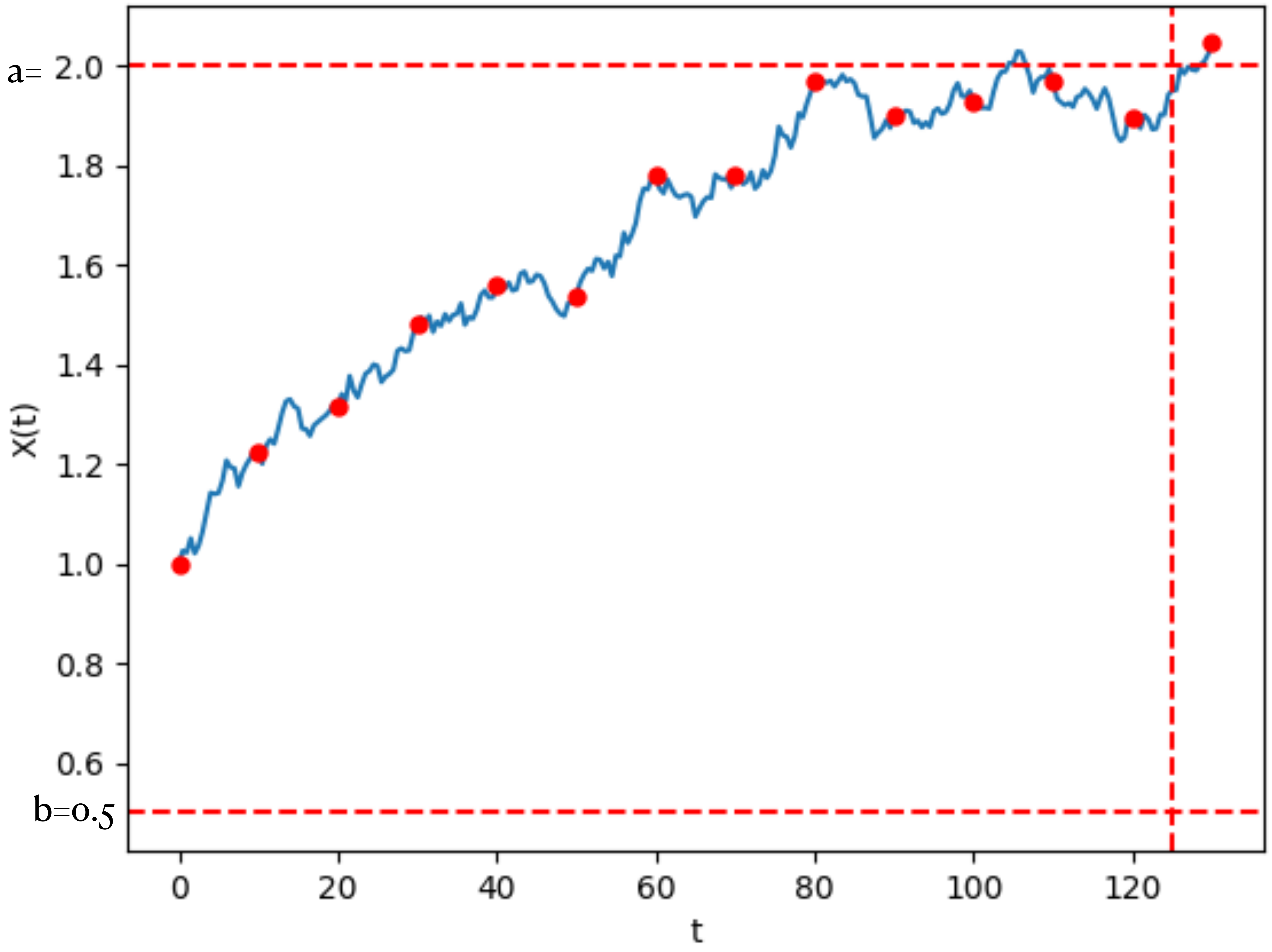


Red points are observations

Blue is the “true” path

Monte Carlo for Mean Exit Time — Example

$dX(t) = \mu X(t)dt + \sigma X(t)dW(t),$ with $X(0) = 1$ $a = 0.5, b = 2, \mu = 0.1, \sigma = 0.2$



Red points are observations

Blue is the "true" path

Systems of SDEs

Systems of SDEs arise when we model dynamics with more than one state e.g.

- Option pricing with the BSM for more than one asset
- Option pricing for a single asset but with stochastic volatility or interest rates
- Portfolio optimisation problems
- All the algorithms/ideas extend to systems with some change in notation

$$\mathbf{X}(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_d(t) \end{bmatrix} \quad \mathbf{f}(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_d(t) \end{bmatrix} \quad \mathbf{W}(t) = \begin{bmatrix} W_1(t) \\ W_2(t) \\ \vdots \\ W_m(t) \end{bmatrix} \quad \mathbf{g}(t) = \begin{bmatrix} g_{1,1}(t) & g_{1,2}(t) & \dots & g_{1,m}(t) \\ g_{2,1}(t) & g_{2,2}(t) & \dots & g_{2,m}(t) \\ \vdots & \vdots & \ddots & \vdots \\ g_{d,1}(t) & g_{d,2}(t) & \dots & g_{d,m}(t) \end{bmatrix}$$

$$d\mathbf{X}(t) = \mathbf{f}(\mathbf{X}(t))dt + \mathbf{g}(\mathbf{X}(t))d\mathbf{W}(t)$$

Example: Stochastic Volatility

Stochastic Volatility model with the Ornstein-Uhlenbeck (OU) process:

$$dS(t) = rS(t)dt + \sigma(t)S(t)dW_x(t)$$ Asset Price

$$d\sigma(t) = \kappa(\bar{\sigma} - \sigma(t))dt + \gamma dW_\sigma(t)$$ Stochastic volatility

with all parameters positive and correlation $\rho dt = dW_x(t)dW_\sigma(t)$

Example: Stochastic Volatility

Stochastic Volatility model with the Ornstein-Uhlenbeck (OU) process:

$$dS(t) = rS(t)dt + \sigma(t)S(t)dW_x(t)$$ Asset Price

$$d\sigma(t) = \kappa(\bar{\sigma} - \sigma(t))dt + \gamma dW_\sigma(t)$$ Stochastic volatility

with all parameters positive and correlation $\rho dt = dW_x(t)dW_\sigma(t)$

The OU process can be negative, a more appropriate model is the Heston SV model

$$dS(t) = rS(t)dt + \sqrt{v(t)}S(t)dW_x(t)$$

$$dv(t) = \kappa(\bar{v} - \sigma(t))dt + \gamma\sqrt{v(t)}dW_v(t)$$

with all parameters positive and correlation $\rho dt = dW_x(t)dW_v(t)$