THE HONG KONG
POLYTECHNIC UNIVERSITY
香港理工大學

AMA 2233

Data Analytics and Visualization

Yancheng Yuan

Department of Data Science and Artificial Intelligence

Recap of last lecture

Connection to probability

Riemann integral: Let $f: R^d \to R$ be an integrable function, we have

$$\iiint_V f(x_1, \cdots, x_d) dx_1 \cdots dx_d = vol(V) E[f(X)]$$

where X is a uniformly distributed random variable on V.

MC method

Monte Carlo Method for estimating integral

$$\iiint_V f(x_1, \cdots, x_d) dx_1 \cdots dx_d$$

Step 1: Set N.

Step 2: Random sample N points uniformly on V.

Step 3: Estimate $vol(V)$.

Step 4: Compute $vol(V) * \frac{f(X_1) + \cdots + f(X_N)}{N}$.

Convergence rate

Let $f: R^d \to R$ be an integrable function and $V \in R^d$ is the unit cube, we have

$$\boldsymbol{E\left[\left(\int_V f(x)dx - \frac{1}{N}\sum_{i=1}^{N} f(X_i)\right)^2\right] \leq O(\frac{1}{N}).}$$

Remark: The convergence rate is independent of d !!!

Application in financial forecasting

Geometric Brownian Motion (GBM) model: The GBM model assumes that the future is independent of the past. If this model is applied in the context of asset pricing, it assumes that the past price information has already been incorporated into the current price and the price movement in the future is conditionally independent of any price movements in the past.

$$S_t = S_{t-1} \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)dt - \sigma\sqrt{dt}Z_t\right)$$

where $Z_t \sim N(0,1)$, μ is the drift rate which determines the growth in the expected value of the random process, σ is the volatility of the drift rate.

Application in financial forecasting

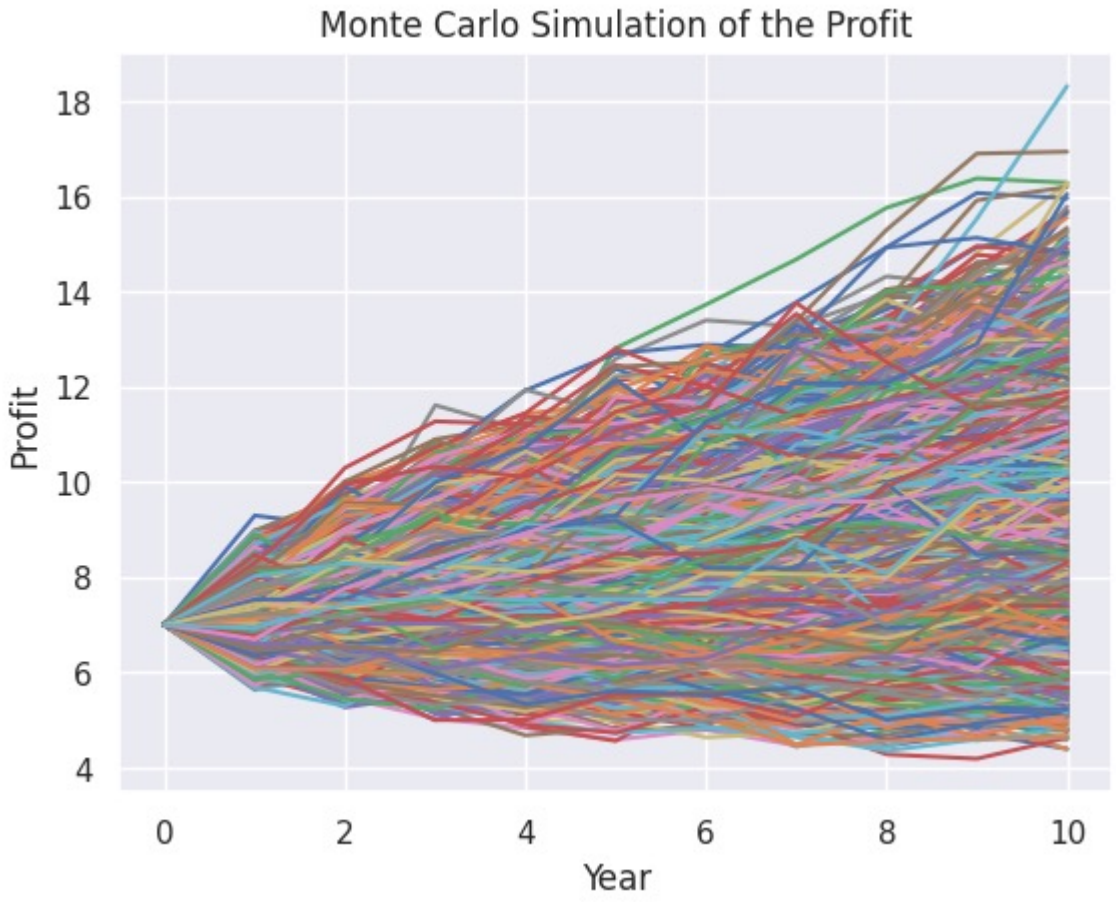
Model settings: **Profit = Revenue — Fixed Cost — Variable Cost**.

Assume the fixed cost is 2 and the revenue follows a GBM model with

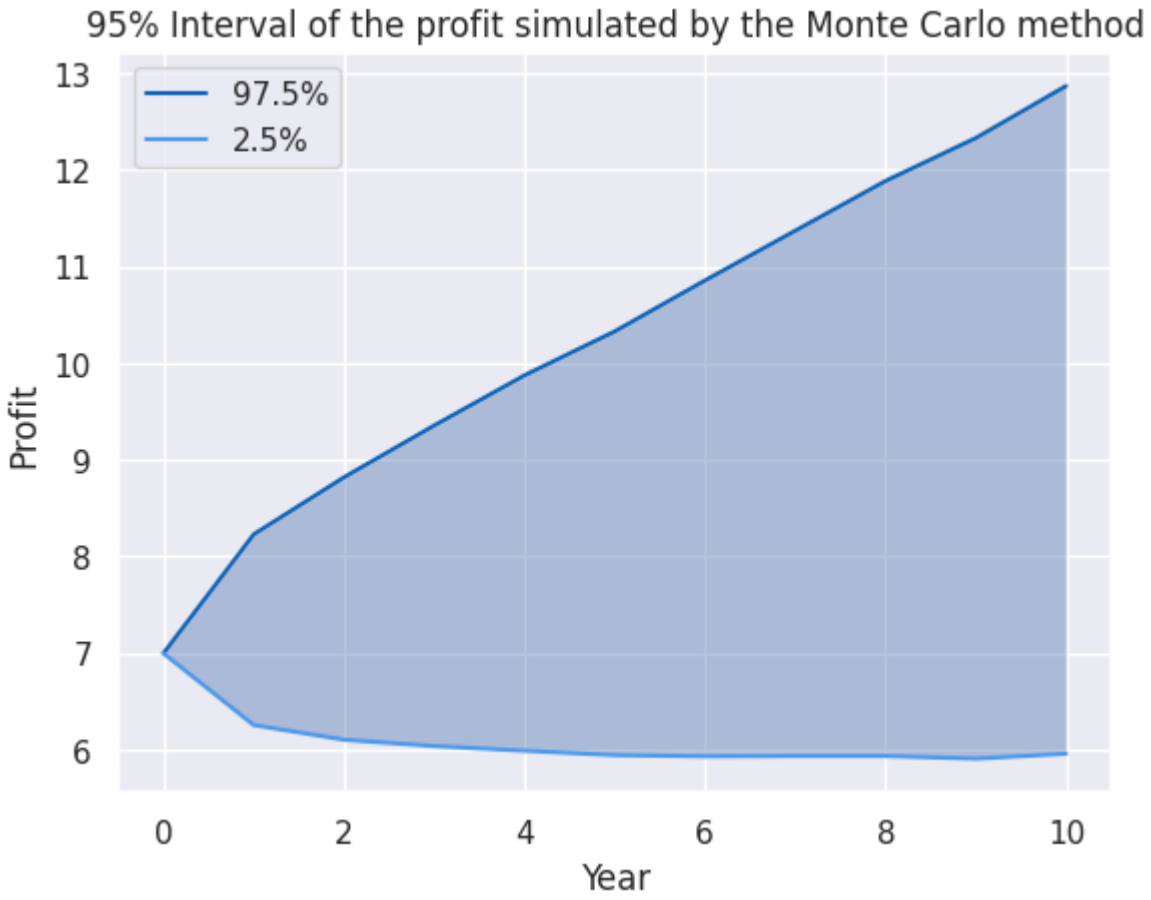
$$S_0 = 10\ \mu = 0.02, \sigma = 0.05,$$

and the variable cost is 10% of the revenue.

Application in financial forecasting



Application in financial forecasting



Agenda

1. Simple Random Walk: An Introduction
2. The Reflection Principle and its Applications
3. Visualization and Applications

Simple Random Walk

Simple Random Walk

Simple Random Walk: A walker stands at the position S_0 at time t = 0. Each time step, the walker will randomly choose to walk up (with probability p) or walk down for one unit. Mathematically, assume X_t be independent identical Bernoulli random variable with mass function

$$P(X_t = 1) = p, \quad P(X_t = -1) = 1 - p,$$

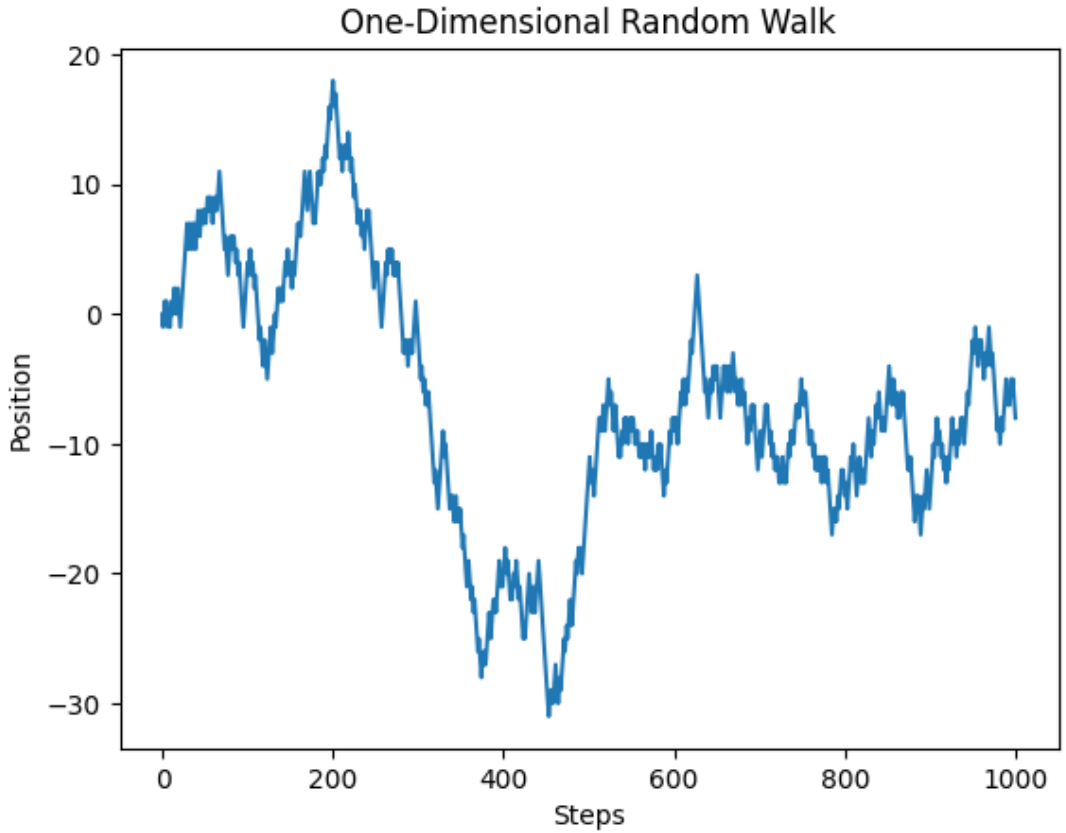
then

$$S_n = S_0 + \sum_{t=1}^{n} X_t$$

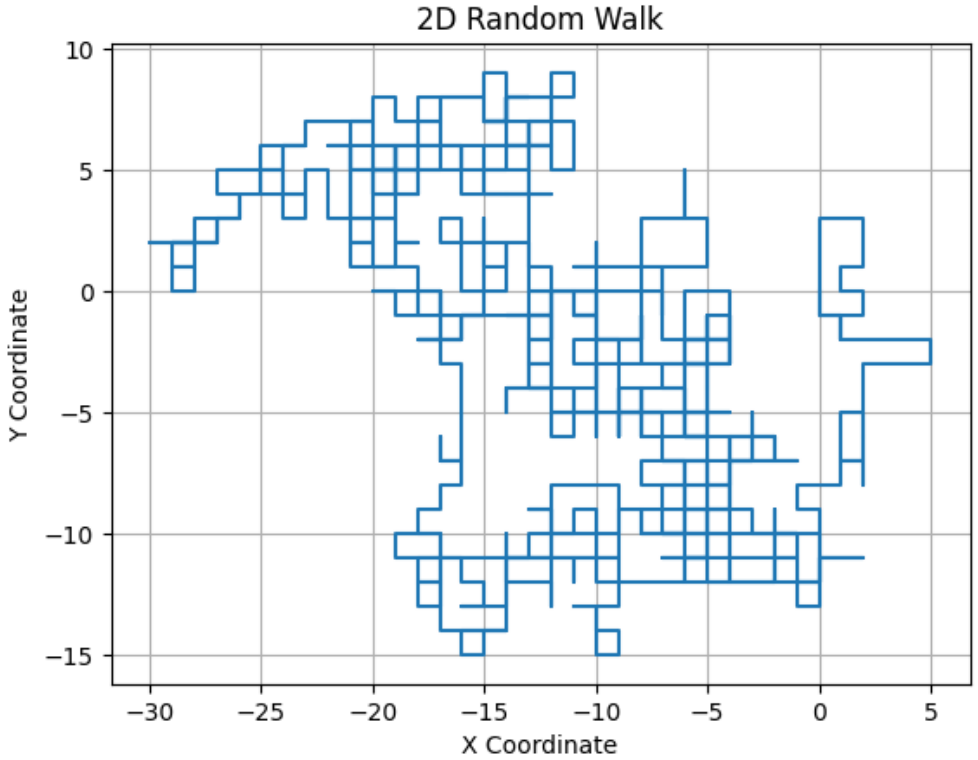
is the position of the walker located after n steps of random walks starting from the initial position S_0.

Remark: We can regard S_n as a time series.

Visualization of 1D Random Walk



Visualization of 2D Random Walk



An application

Gambler's ruin problem: A man is saving up to buy a new watch at a cost of N dollars. He starts with k dollars ($0 < k < N$), and tries to win the reminder by the following gamble in the casino. He tosses a fair coin repeatedly; if it comes up heads, then he gains one dollar; otherwise, he loses one dollar. He plays this game repeatedly until one of two events occurs: either he runs out of money and is bankrupted or he wins enough to buy the watch.

Question: What is the probability that he is ultimately bankrupted?

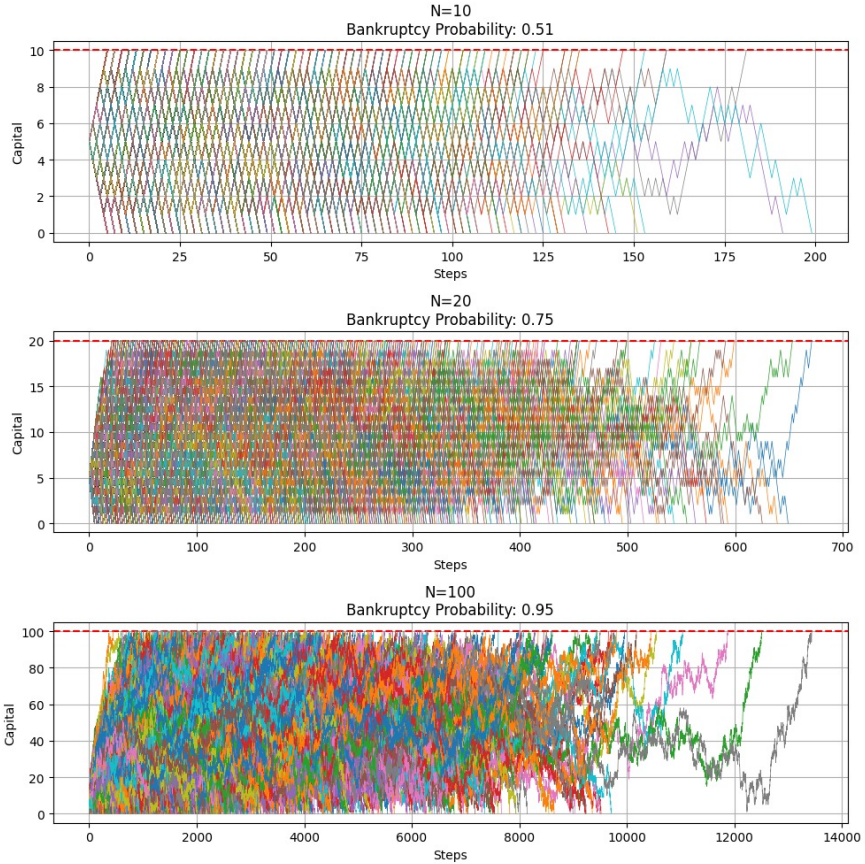
An application

Gambler's ruin Theorem: To buy a new watch at price N dollars, the probability that he is ultimately bankrupted starting with k dollars is

$$1 - \frac{k}{N}$$

Implication: Without a good strategy, "to use a small matter to achieve great things" is almost impossible.

Visualization for k = 5



The Reflection Principle and its Applications

Random walk Setting

Simple Random Walk: A walker stands at the position S_0 at time t = 0. Each time step, the walker will randomly choose to walk up (with probability p) or walk down for one unit. Mathematically, assume X_t be independent identical Bernoulli random variable with mass function

$$P(X_t = 1) = p, \quad P(X_t = -1) = 1 - p,$$

then

$$S_n = S_0 + \sum_{t=1}^{n} X_t$$

is the position of the walker located after n steps of random walks starting from the initial position S_0.

Random walk Setting

Simple Random Walk: A walker stands at the position S_0 at time t = 0. Each time step, the walker will randomly choose to walk up (with probability p) or walk down for one unit. Mathematically, assume X_t be independent identical Bernoulli random variable with mass function

$$P(X_t = 1) = p, \quad P(X_t = -1) = 1 - p,$$

then

$$S_n = S_0 + \sum_{t=1}^{n} X_t$$

is the position of the walker located after n steps of random walks starting from the initial position S_0.

Question 1: What is the probability that $S_0 = a$ and $S_n = b$ $(\mathrm{a} - \mathrm{n} \leq b \leq a + n)$?

Question 2: What is the probability that $S_0 = a > 0$ and $S_n = b > 0$ and the random walk intersects with the x-axis somewhere?

Motivating Examples I

Portfolio management with Maximum Drawdown Constraint: If you are a portfolio manager with initial funding M dollars from the investors. The investors put a constraint that the maximum drawdown during the investment period should not exceed 10% of the funding M. If the profit of your portfolio is just a random walk (increase absolute 1% or drop absolute 1% daily with equal probability), What is the probability is your portfolio ends with at least 10% profit after 364 trading days?

Motivating Examples II

Ballot Problem: Suppose that in a ballot, candidate A scores α votes and candidate B scores β votes where $\alpha > \beta$. What is the probability that, during the ballot, A was always ahead of B?

The simple random walk is spatially homogeneous

Theorem: The simple random walk is spatially homogeneous, that is

$$P(S_n = j \mid S_0 = a) = P(S_n = j + b \mid S_0 = a + b).$$

The simple random walk is temporally homogeneous



Theorem: The simple random walk is temporally homogeneous, that is

$$P(S_n = j \mid S_0 = a) = P(S_{n+m} = j \mid S_m = a).$$

The simple random walk has Markov property



Theorem: The simple random walk has Markov property, that is

$$P(S_{n+m} = j \mid S_0, S_1, \dots, S_m) = P(S_{n+m} = j \mid S_m).$$

Answer to Question 1

Theorem: For a simple random walk,

$$P(S_n = b \mid S_0 = a) = \binom{n}{\frac{1}{2}(n+b-a)} p^{\frac{1}{2}(n+b-a)}(1-p)^{\frac{1}{2}(n+a-b)}.$$

Count the paths

Motivation: To answer Question 2, we need to count the number of random walk paths $N_n(a,b)$ that $S_0 = a$ and $S_n = b$. Also, we need to count the number of random paths among them that revisits the x-axis.

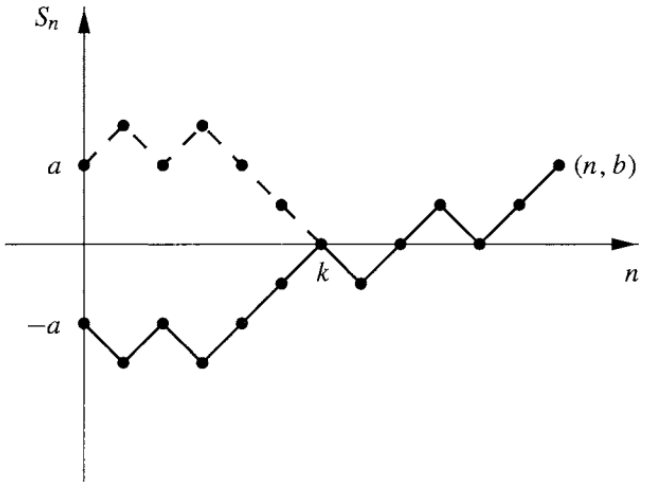
Theorem: For a simple random walk,

$$N_n(a,b) = \binom{n}{\frac{1}{2}(n+b-a)}.$$

The Reflection Principle

Theorem (The Reflection Principle): For a simple random walk and some $a, b > 0$, the number of random paths satisfying $S_0 = a$ and $S_n = b$ and intersecting with the x-axis somewhere is given by

$$N_n^0(a, b) = N_n(-a, b).$$



Ballot Theorem

Ballot Theorem: For a simple random walk and some $b > 0$, the number of random paths satisfying $S_0 = 0$ and $S_n = b$ and does not revisit the x-axis is

$$\frac{\boldsymbol{b}}{\boldsymbol{n}} * \boldsymbol{N_n(0, b)}.$$

Visualization and Applications

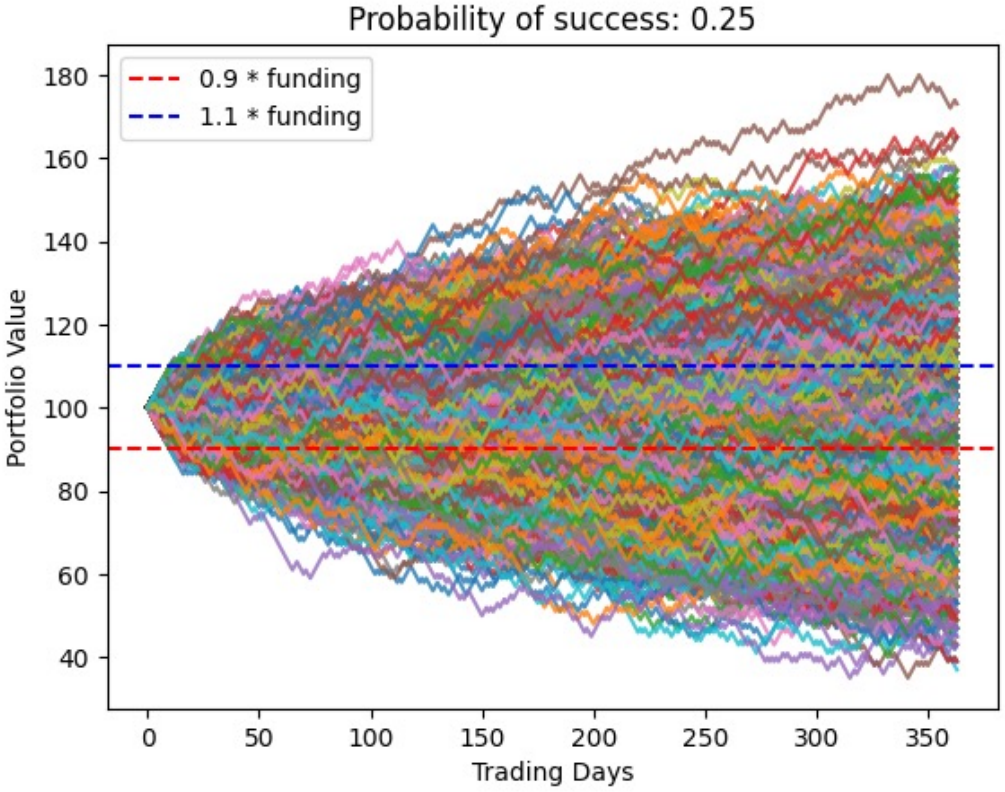
Recap the portfolio management problem

Portfolio management with Maximum Drawdown Constraint: If you are a portfolio manager with initial funding M dollars from the investors. The investors put a constraint that the maximum drawdown during the investment period should not exceed 10% of the funding M. If the profit of your portfolio is just a random walk (increase absolute 1% or drop absolute 1% daily with equal probability), What is the probability is your portfolio ends with at least 10% profit after 364 trading days?

Question: What is the probability value in theory?

Answer: **0.25439**

What is the value by simulation?




Thank You!