Lecture 25

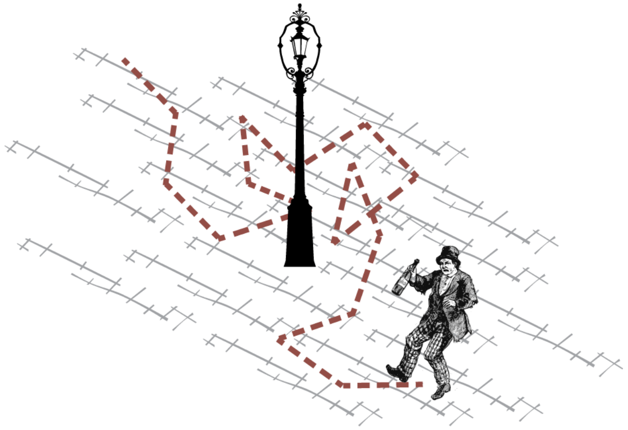
Module VI

03.05.2019

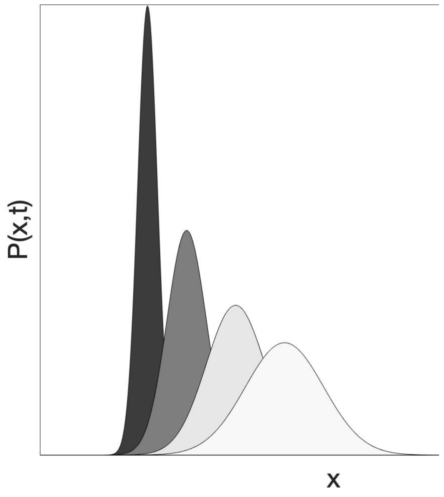
Random walk, Central limit theorem

Overview

- **What is a random walk**



- **Universality of the Gaussian distribution: Central limit theorem or the «law of large numbers»**

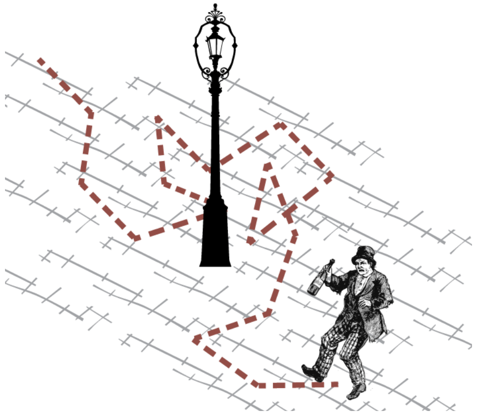


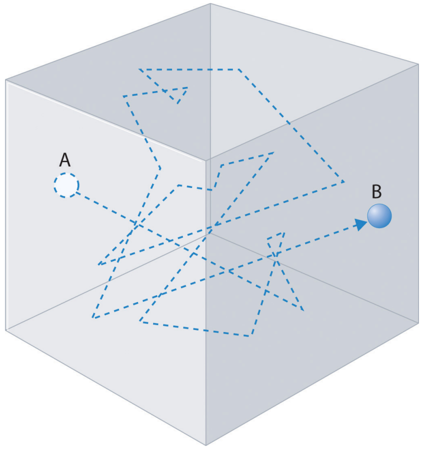
Random walk

- *Random* collision between particles in an ideal gas means that on lenghscales larger than the mean free path, the trajectory of a particle is a *random* meandering also called thermal motion

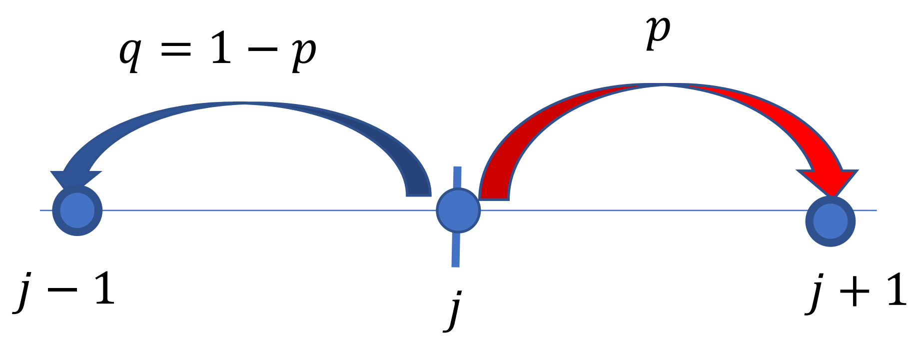
- Microscopic stochastic model for particle dynamics due to thermal motion in an ideal gas

- Each particle has a given probability to be within a small interval around a position r at time t, $P(r,t)dr$ where $P(r,t)$ is the Gaussian probability distribution function





1D Random walk (RW)



- Random motion on a line
  - Discrete time steps $N = 0,1,2 \cdots$ in units of $\Delta t = 1$
  - Discrete space: lattice index $j = 0, \pm 1, \pm 2$\cdots with increments $\Delta x = 1$

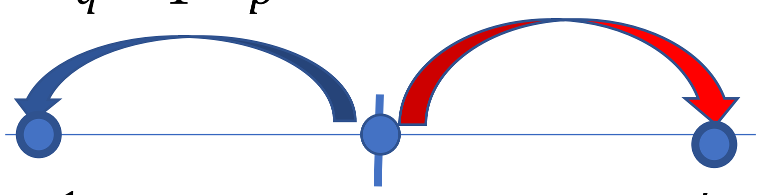
- At each timestep, the walker has probability p to the right $j \to j+1$ and probability $q = 1 - p$ to the left $j \to j-1$

What is the mean displacement $\langle S \rangle (N)$ of the RW after N steps?

What is the mean square displacement $\langle (\Delta S)^2 \rangle (N)$ of the RW after N steps?

What is the probability distribution for a displacement S after N steps, $P_N(S)$?

Bernoulli distribution



After N steps, the RW has made R steps to the right and L steps to the right, so

$$R + L = N, \qquad S = R - L$$

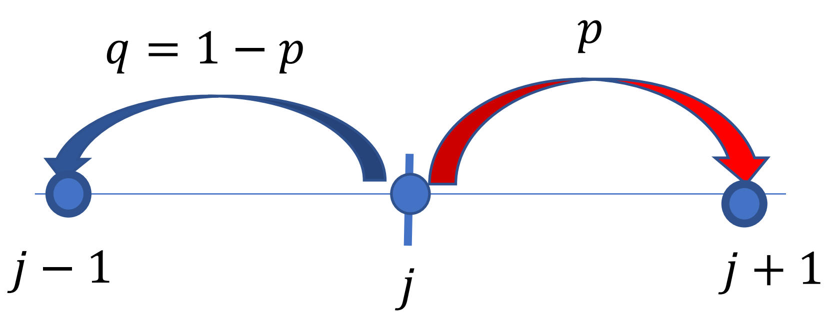
- Probability for R independent steps to the right out of N steps is given by the probability for a configuration with R out N steps, $p^R(1-p)^{N-R}$, times the number of possible configurations $\frac{N!}{R!(N-R)!}$

$$F_N(R) = \frac{N!}{R!(N-R)!} p^R q^{N-R}$$ **(Bernoulli distribution)**

Probability $F_N(R)$ is normalized: The probability for N steps (irrespective of right or left direction) is 1

$$\sum_{R=1}^{N} F_N(R) = \sum_{R=1}^{N} \frac{N!}{R!\,(N-R)!} p^R q^{N-R} = (p+q)^N = 1$$

Average displacement is zero



- Average number of steps to the right $\langle R \rangle = Np$

$$\langle R \rangle = \sum_{R=1}^{N} R F_N(R) = \sum_{R=1}^{N} \frac{N!}{R!\,(N-R)!} R p^R q^{N-R}$$

$$\left(use\ that\ Rp^R = p \frac{d}{dp} p^R\right)$$

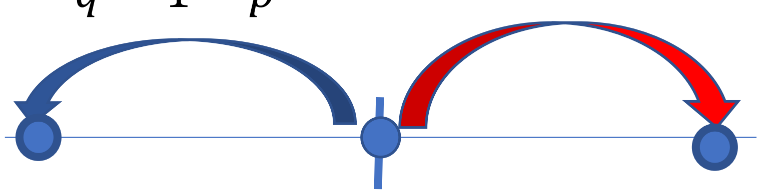
$$\langle R \rangle = p \frac{d}{dp} \sum_{R=1}^{N} \frac{N!}{R!\,(N-R)!} p^R q^{N-R} = p \frac{d}{dp} (p+q)^N = p\,N\,(p+q)^{N-1} = Np$$

- Average displacement from the origin is $\langle S \rangle = 0$. Not an efficient way to walk.

$$\langle S \rangle = \langle R \rangle - (N - \langle R \rangle) = 2\langle R \rangle - N$$

$$\langle S \rangle = N(2p-1) = N(p-q), \qquad for\ \ p = q = \frac{1}{2} \rightarrow \langle S \rangle = 0$$

Mean square displacement



Probability for R steps to the right out of N steps $F_N(R) = \frac{N!}{R!(N-R)!} p^R q^{N-R}$

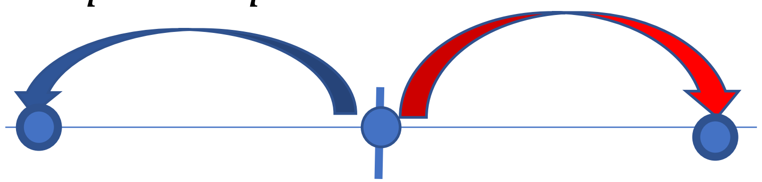
$$\langle R^2 \rangle = \sum_{R=1}^{N} R^2 P_N(R) = \sum_{R=1}^{N} \frac{N!}{R!\,(N-R)!} R^2 p^R q^{N-R}$$

$$\left(use\ that\ R^2 p^R = \left(p \frac{d}{dp} \right)^2 p^R \right)$$

$$\langle R^2 \rangle = \left(p \frac{d}{dp} \right)^2 \sum_{R=1}^{N} \frac{N!}{R!\,(N-R)!} p^R q^{N-R} = \left(p \frac{d}{dp} \right)^2 (p+q)^N = Np \frac{d}{dp} [(p+q)^{N-1}]$$

$$\langle R^2 \rangle = Np + \ N(N-1)p^2$$

Mean square displacement



$$\langle R^2 \rangle = Np + N(N-1)p^2, \qquad \langle R \rangle = Np$$

$$\langle S^2 \rangle = \langle (2R - N)^2 \rangle = 4\langle R^2 \rangle - 4N\langle R \rangle + N^2 = N^2(p-q)^2 + 4Npq$$

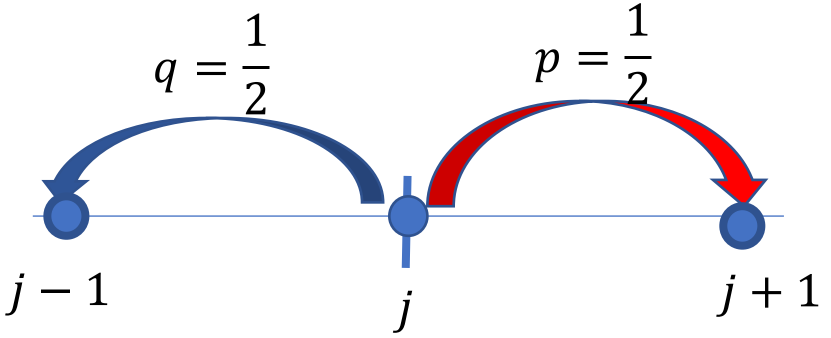
Mean square dislacement: $\langle \Delta S^2 \rangle = \langle S^2 \rangle - \langle S \rangle^2$

$$\langle \Delta S^2 \rangle = 4Npq,$$

For an unbias walk $p = q = \frac{1}{2} \rightarrow \langle \Delta S^2 \rangle = N$, the mean square displacement increases linearly with the total number of steps.

Bottom line: the walker does not go anywhere on average since $\langle S \rangle = 0$, but the area of its meandering around the origin increases proportial to the number of steps, $\langle \Delta S^2 \rangle = N$.

Displacement probability $P_N(S)$



Bernoulli distrubution:

Probability for R steps to the right out of N steps for $p = q = \frac{1}{2}$

$$F_N(R) = \frac{N!}{R!\,(N-R)!}\frac{1}{2^N}$$

Using Stirling approx: $n! = \sqrt{2\pi n}\, n^n e^{-n}$

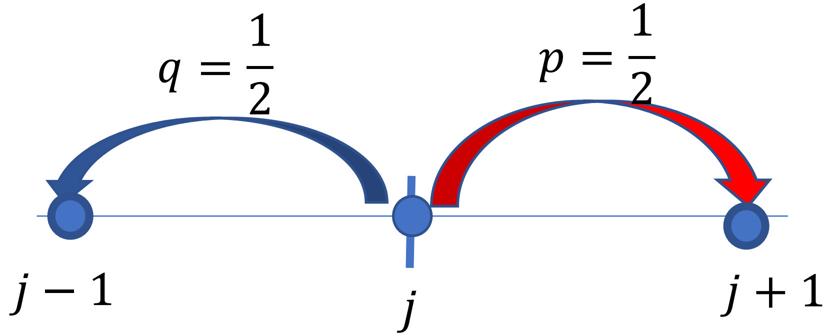
$$F_N(R) = \sqrt{\frac{N}{2\pi R(N-R)}}\, e^{N\log N - R\ln(R) - (N-R)\ln(N-R) - N\log 2} = \sqrt{\frac{N}{2\pi R(N-R)}}\, e^{-R\log\left(\frac{2R}{N}\right) - (N-R)\ln\left(\frac{2(N-R)}{N}\right)}$$

$2R = N + S$

Change of variables $P_N(S)dS = F_N(R)dR$

$$P_N(S) = \frac{1}{2}F_N\left(\frac{N+S}{2}\right) = \frac{1}{2}\sqrt{\frac{2N}{\pi(N^2-S^2)}}\, e^{-\frac{N+S}{2}\log\left(1+\frac{S}{N}\right) - \frac{N-S}{2}\log\left(1-\frac{S}{N}\right)}$$

Gaussian distribution $P_N(S)$



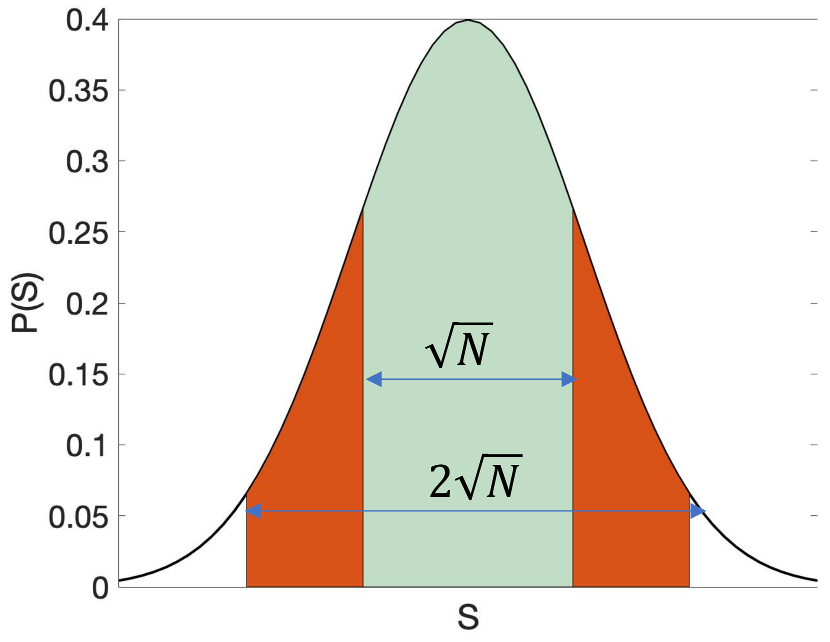
Probability for a net displacement S

$$P_N(S) = \sqrt{\frac{N}{2\pi(N^2 - S^2)}}\, e^{-\frac{N+S}{2}\log\left(1+\frac{S}{N}\right) - \frac{N-S}{2}\log\left(1-\frac{S}{N}\right)}$$

$$\log\left(1+\frac{S}{N}\right) \approx \frac{S}{N} - \frac{1}{2}\left(\frac{S}{N}\right)^2, \qquad \log\left(1-\frac{S}{N}\right) \approx -\frac{S}{N} - \frac{1}{2}\left(\frac{S}{N}\right)^2,$$

$$P_N(S) = \sqrt{\frac{1}{2\pi N}}\, e^{-\frac{S}{2}\left(1+\frac{S}{N}\right)\left(1-\frac{1}{2}\frac{S}{N}\right) + \frac{S}{2}\left(1-\frac{S}{N}\right)\left(1+\frac{1}{2}\frac{S}{N}\right)} = \sqrt{\frac{1}{2\pi N}}\, e^{-\frac{S}{2}\left(1+\frac{S}{2N}\right) + \frac{S}{2}\left(1-\frac{S}{2N}\right)}$$

$$P_N(S) = \sqrt{\frac{1}{2\pi N}}\, e^{-\frac{S^2}{2N}}$$

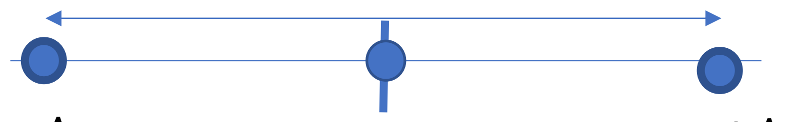


Mean $\langle S \rangle = 0$

Standard deviation $\langle (\Delta S)^2 \rangle = N$

$$P_N(S) = \sqrt{\frac{1}{2\pi\langle \Delta S^2 \rangle}}\, e^{-\frac{(S - \langle S \rangle)^2}{2\langle \Delta S^2 \rangle}}$$

Continuous space-time RW



RW displacememt is the position $X = S\Delta x$ *the continuous line*

Number of steps *determine the time* $t = N\Delta t$

Probability distribution to find the RW at a give position x at time t:

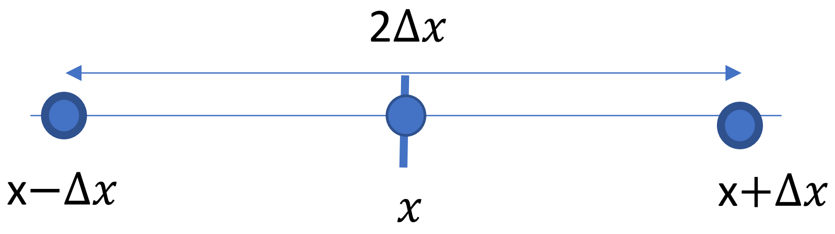
$$P(x,t)dx = P_N(S)dS \quad \rightarrow \quad P(x,t) = P_{\frac{t}{\Delta t}}(X/\Delta x)\frac{1}{\Delta x}$$

$$P(X,t) = \sqrt{\frac{\Delta x^2}{2\pi\langle \Delta X^2\rangle}}\, e^{-\frac{(X-\langle X\rangle)^2}{2\langle \Delta X^2\rangle}} \frac{1}{\Delta x}, \qquad \int_{-\infty}^{+\infty} dX \;\; P(X,t) = 1, at\ any\ t$$

Mean displacement $\langle X\rangle = \Delta x\langle S\rangle$

Standard deviation $\langle(\Delta X)^2\rangle = \Delta x^2\langle(\Delta S)^2\rangle$

Gaussian distribution



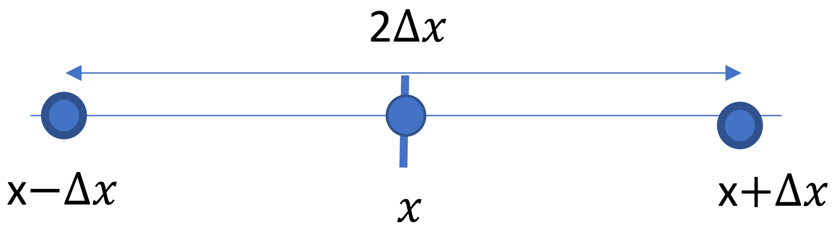
Probability distribution to be within an internal of width $2\Delta x$ around X and t

$$P(X,t) = \sqrt{\frac{1}{2\pi\langle \Delta X^2 \rangle}}\, e^{-\frac{(X-\langle X \rangle)^2}{2\langle \Delta X^2 \rangle}}$$

Mean displacement $\langle X \rangle = vt,\ where\ v = (p-q)\frac{\Delta x}{\Delta t}$ **drift velocity**

Standard deviation $\langle (\Delta X)^2 \rangle = 2Dt,\ where\ D = 2pq\frac{\Delta x^2}{\Delta t}$ **diffusion coefficient**

Gaussian distribution



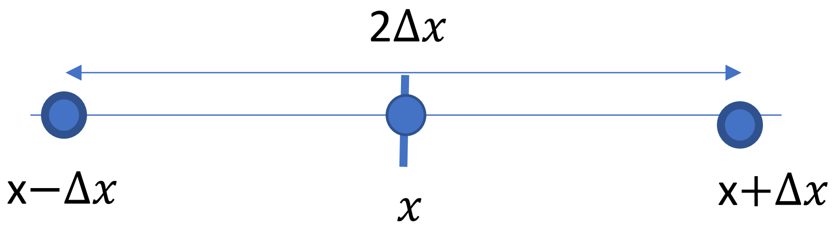
Probability distribution for RW with drift velocity v and diffusion coefficient D

$$P(X,t) = \sqrt{\frac{1}{4\pi Dt}}\, e^{-\frac{(X-vt)^2}{4Dt}}, \qquad \int_{-\infty}^{+\infty} dX\, P(X,t) = 0$$

Mean displacement $\langle X \rangle = \int_{-\infty}^{+\infty} dX\, X P(X,t) = \; vt$

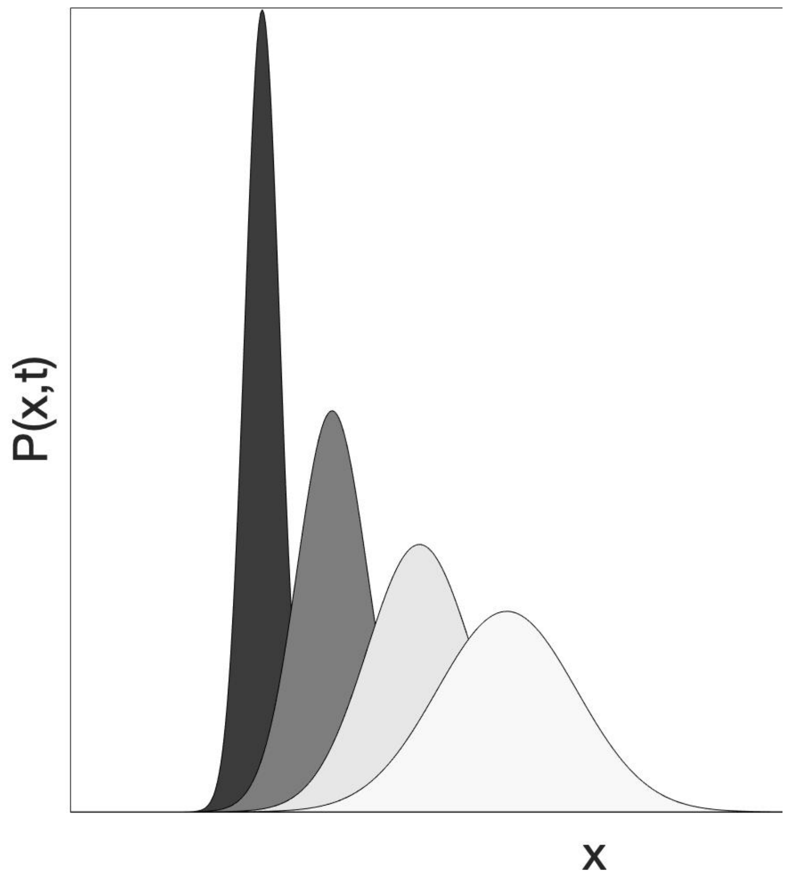
Standard deviation $\langle (\Delta X)^2 \rangle = \int_{-\infty}^{+\infty} dX\, (X - vt)^2 P(X,t) = \; 2Dt$

Continuous time-space RW

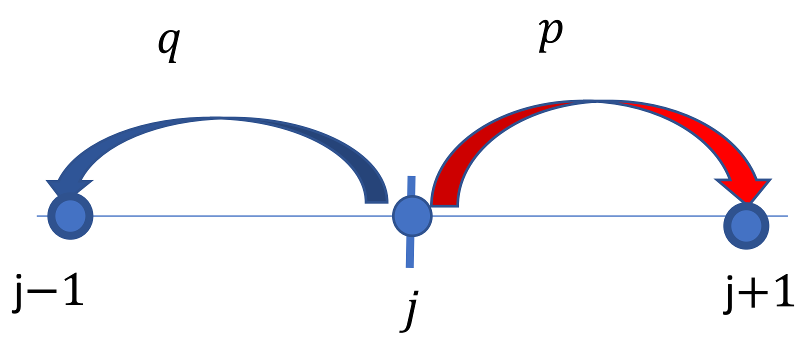


Probability distribution for RW with constant drift velocity v and diffusion coefficient D

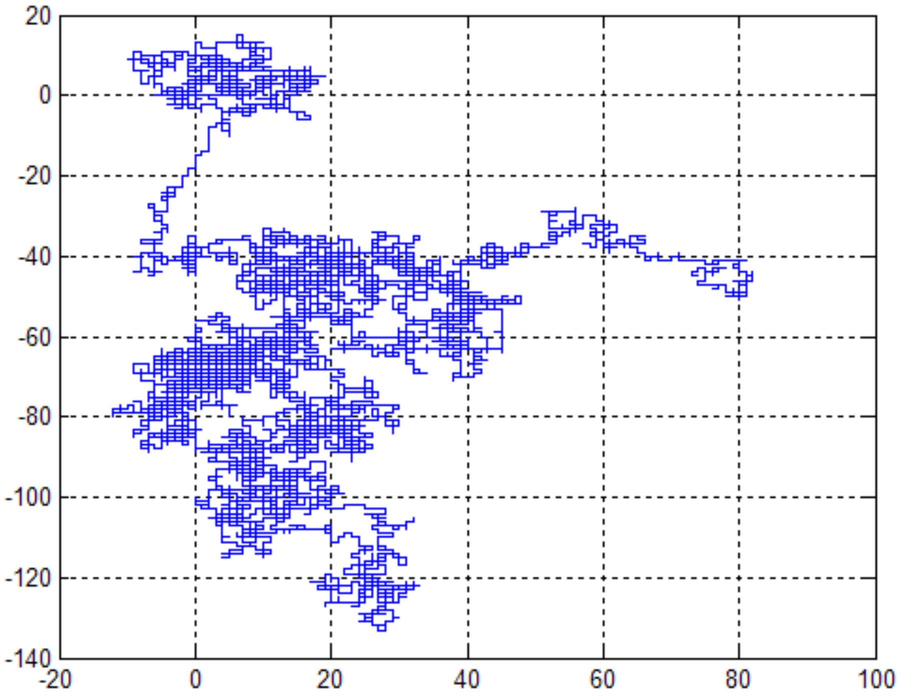
$$P(X,t) = \sqrt{\frac{1}{4\pi Dt}}\, e^{-\frac{(X-vt)^2}{4Dt}}$$



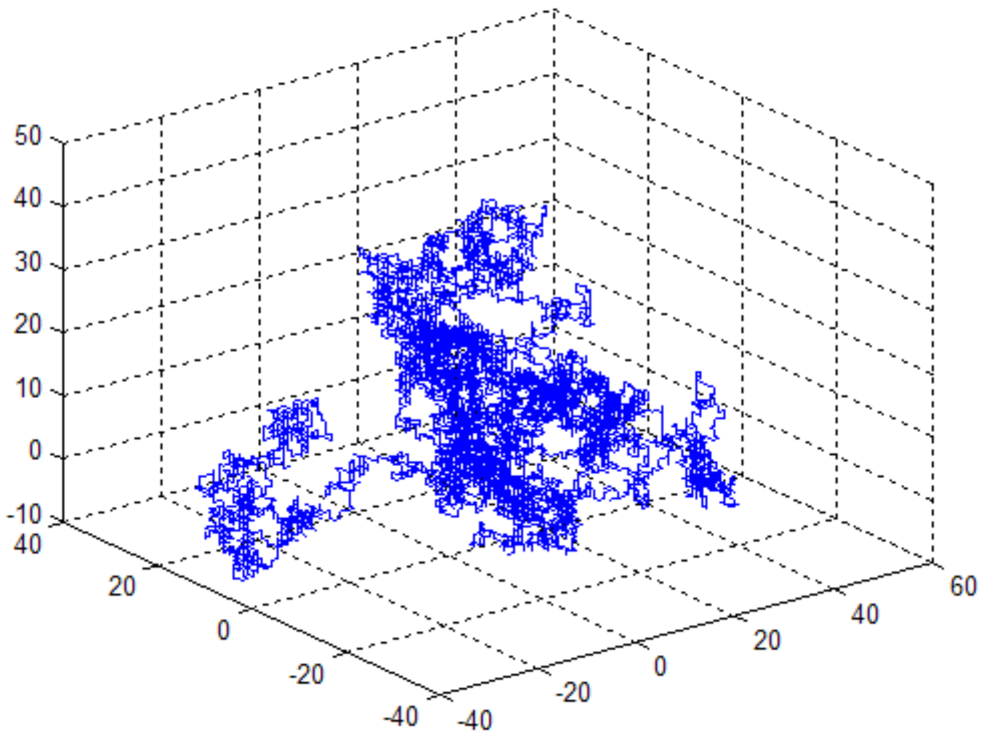
RW simulations



2D



3D



Central limit theorem: Limit distribution of sums

Suppose we have a set of N *independent, identically distributed (i.i.d.) variables* x_i drawn from the same parent distribution $p(x_i)$ *with*

1. *Zero mean* $\int dx\, x\, p(x) = \langle x \rangle = 0$
2. *Finite variance* $\int dx\, x^2\, p(x) = \sigma^2 < \infty$

The sum of N variables $X = \sum_{i=1}^{N} x_i$ is also a stochastic quantity which, in the limit of $N \gg 1$, is distributed according to the Gaussian distribution ***independent of the parent distributions*** $\boldsymbol{p(x_i)}$

Application to the RW

x_i are the independent random increments drawn from the same uniform distribution $p(x_i)$ (prob for an increment)

The sum of N increments is the net displacement

$$X(N) = \sum_{i=1}^{N} x_i$$

Central limit theorem gives us the limit distribution of displacement X regardless of the distribution of individual increment

$X(N) \rightarrow_{N \gg 1}$ *is Gaussian distribution*

Central limit theorem:
Proof using the characteristic function

Probability density $P_N(X)$ for the sum of random variables

$$X = \sum_{i=1}^{N} x_i \qquad (1)$$

depends on the product of the probability density for i.i.d. random variable, i.e. $\prod_{i=1}^{N} p(x_i)$ with the constraint that their sum is given by Eq. (1)

$$P_N(X) = \int dx_1 \cdots dx_N \, p(x_1) \cdots p(x_N) \delta\left(X - \sum_{i=1}^{N} x_i\right)$$

Method of chacteristic function

Fourier transform of $P_N(X)$ defines the characteristic function

$$\hat{P}(k) = \frac{1}{2\pi}\int dXe^{-ikX}P_N(X)$$

$$\hat{P}(k) = \frac{1}{2\pi}\int dXe^{-ikX}\int dx_1 \cdots dx_N\, p(x_1)\cdots p(x_N)\delta\left(X - \sum_{i=1}^{N} x_i\right)$$

$$\hat{P}(k) = \frac{1}{2\pi}\int dx_1 \cdots dx_N\, p(x_1)\cdots p(x_N)\left(\int dXe^{-ikX}\delta\left(X - \sum_{i=1}^{N} x_i\right)\right)$$

$$\hat{P}(k) = \frac{1}{2\pi}\int dx_1 \cdots dx_N\, p(x_1)\cdots p(x_N)e^{-ik\sum x_i}$$

$$\hat{P}(k) = \frac{1}{2\pi}\left(\int dx_1 p(x_1)e^{-ikx_1}\right)\cdots\left(\int dx_N p(x_N)e^{-ikx_N}\right)$$

Chacteristic function

$$2\pi\hat{P}(k) = \left(\int dxp(x)e^{-ikx}\right)^N = \left(2\pi\hat{p}(k)\right)^N$$

We Taylor expand e^{-ikx} since the wavenumber k scales as $1/N$ as being the reciprocal of $X(\mathrm{N})$. Hence,

$$\hat{p}(k) = \frac{1}{2\pi}\sum_n \frac{(-ik)^n}{n!}\int dxp(x)x^n$$

Characteristic function of the parent distribution can be written as a power series of its moments

$$\hat{p}(k) = \frac{1}{2\pi}\sum_n \frac{(-ik)^n}{n!}\langle x^n\rangle$$

Asymptotic behavior in the limit of large N

$$\hat{p}(k) = \frac{1}{2\pi}\sum_n \frac{(-ik)^n}{n!}\langle x^n\rangle \approx \frac{1}{2\pi}\left(1 - \frac{k^2}{2}\sigma^2\right), \text{ since } k \sim \frac{1}{N} \text{ is small}$$

k *is the wavenumver associated with* X *from* (e^{ikX}); $X = \sum_{i=1}^{N} x_i \sim N \rightarrow k \sim \frac{1}{N}$

$q = Nk$

$$2\pi\hat{P}\left(\frac{q}{N}\right) = \left(2\pi\hat{p}\left(\frac{q}{N}\right)\right)^N \approx \left(1 - \frac{(q\sigma)^2}{2N^2}\right)^N \rightarrow_{N\gg1} e^{-\frac{q^2\sigma^2}{2N}}$$

Central limit theorem:

$$2\pi\hat{P}(k) \to_{N\gg 1} e^{-\frac{k^2 N\sigma^2}{2}}$$

$$P(X) = \int dk e^{ikX}\hat{P}(k) \approx \frac{1}{2\pi}\int dk e^{ikX} e^{-\frac{k^2 N\sigma^2}{2}}$$

$$P(X) = \frac{1}{2\pi}\int dk e^{ikX-k^2\frac{N\sigma^2}{2}}$$

Complete the square $e^{ikX-k^2\frac{N\sigma^2}{2}} = e^{-\left(k^2\frac{N\sigma^2}{2}-2\left(k\sqrt{\frac{N\sigma^2}{2}}\right)\left(iX\sqrt{\frac{1}{2N\sigma^2}}\right)-X^2\frac{1}{2N\sigma^2}\right)} e^{-\frac{X^2}{2N\sigma^2}}$

$= e^{-\left(k\sqrt{\frac{N\sigma^2}{2}}-iX\sqrt{\frac{1}{2N\sigma^2}}\right)^2} e^{-\frac{X^2}{2N\sigma^2}}$

$$P(X) = \frac{1}{2\pi}e^{-\frac{X^2}{2N\sigma^2}}\int dk e^{-\left(k\sqrt{\frac{N\sigma^2}{2}}-iX\sqrt{\frac{1}{2N\sigma^2}}\right)^2} = \frac{1}{2\pi}\sqrt{\frac{2\pi}{N\sigma^2}}e^{-\frac{X^2}{2N\sigma^2}}$$

Central limit theorem: law of large number

$$P(X) = \sqrt{\frac{1}{2\pi N\sigma^2}}\, e^{-\frac{X^2}{2N\sigma^2}}$$

With

$$\int dX\, X\, P(X) = 0$$

$$\langle X^2 \rangle = \int dX\, X^2\, P(X)$$

$$\langle X^2 \rangle = N\sigma^2$$

- *Gaussian distribution captures universal fluctuations about a mean*
- *As N increases the sample mean approachs the gaussian mean*
- *Any probability distribution can be approxiated near its mean by a Gaussian distribution*