Part I

Section 5.1 Random walk in $\mathbb{Z}$

- Later this week random walk in $\mathbb{Z}^d$, $d=1,2,3$ and the dependence of RW on d qualitatively

Motivation

- Models diffusion
- Its continuum version is Brownian motion
- Diffusion appears in mixing of fluids, heat equation...

Definition

Let $0<p<1$ and $q=1-p$. Let $(X_k)_{k\geq 1}$ be independent and identically distributed (i.i.d) random variables such that $\mathbb{P}[X_1=+1]=p$, $\mathbb{P}[X_1=-1]=q$. Random walk with parameter p sent from $x_0\in\mathbb{Z}$ is defined as $S=(S_t)_{t\in\mathbb{N}}$

$$S_t = x_0 + \sum_{k=1}^{t} X_k$$

If $p=q=\frac{1}{2}$, we call it symmetric or simple random walk.

Remarks
- t is called time
- $S=(S_t)_{t\in\mathbb{N}}$ is a stochastic process, a collection of random variables indexed by time
- For S time is discrete $t\in\mathbb{N}=\{0,1,2,\dots\}$ and the state space is discrete $S_t\in\mathbb{Z}$.

- The natural space for defining the law of $S = (S_t)_{t\in\mathbb{N}}$ is $\mathbb{Z}^{\mathbb{N}}$

- For fixed t, the distribution of S_t is a kind of binomial distribution. If $x_0, x_1, \dots, x_t$ is a path with $n_+ = \#\{j : x_j - x_{j-1} = +1\}$, $n_- = \#\{j : x_j - x_{j-1} = -1\}$, $n_+ + n_- = t$, then

$$\mathbb{P}[S_1 = x_1, S_2 = x_2, \dots, S_t = x_t] = p^{n_+} q^{n_-}$$

by independence of $X_1, \dots, X_t$. If $S_t = x_0 + k$, then

$$\begin{cases} n_+ + n_- = t \\ n_+ - n_- = k \end{cases} \Rightarrow \begin{cases} n_+ = \frac{t+k}{2} \\ n_- = \frac{t-k}{2} \end{cases}$$

Number of walks of step t with $S_t = x_0 + k$ is then $\binom{t}{n_+} = \binom{t}{(t+k)/2}$

$$\Rightarrow \mathbb{P}[S_t = x_0 + k] = \binom{t}{(t+k)/2} p^{\frac{t+k}{2}} q^{\frac{t-k}{2}}$$

when $t+k$ is even and 0 otherwise.

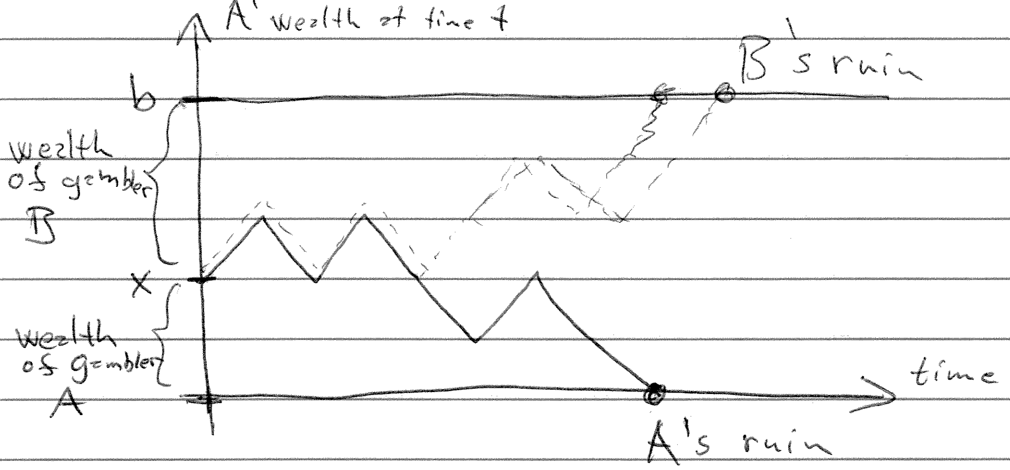
- By law of large numbers

$$\lim_{n\to\infty} S_n = +\infty \quad \text{a.s. when } p > q$$

$$\lim_{n\to\infty} S_n = -\infty \quad \text{a.s. when } p < q$$

<u>Definition</u> Discrete time <u>stochastic process</u> on a state space X is a family of random variables $S = (S_t)_{t \in \mathbb{N}}$ indexed by time $t \in \mathbb{N}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $S_t : \Omega \to X$.

5.1.1 First step analysis and gambler's ruin



Let $a < b$, $a, b \in \mathbb{Z}$. Define

$$\tau_{\{a,b\}} = \min\{t \geq 0 : S_t \in \{a,b\}\}$$

If the set is $\emptyset$, then set $\tau_{\{a,b\}} = \infty$.

It is easy to see that $\tau_{\{a,b\}} < \infty$ a.s. (exercise)

<u>Theorem</u> If $a < b$, $x \in [\![a,b]\!]$, $h(x) = \mathbb{P}[S_{\tau_{\{a,b\}}} = b \mid S_0 = x]$,
~~then~~ for $0 < p < 1$

$$h(x) = \begin{cases} \frac{x-a}{b-a}, & \text{when } p = q = \frac{1}{2} \\ \frac{1 - (\frac{q}{p})^{x-a}}{1 - (\frac{q}{p})^{b-a}}, & \text{otherwise} \end{cases}$$

<u>Proof</u> Clearly $h(a) = 0$, $h(b) = 1$. Let $a < x < b$ now.

By <u>taking one step</u> $S_0 \rightsquigarrow S_1$ we find

$$h(x) = \sum_{y = x \pm 1} \mathbb{P}[\{S_t \in \{a,b\}\} \cap \{S_1 = y\} \mid S_0 = x]$$

$$= \sum_{y = x \pm 1} \mathbb{P}[\{S_t \in \{a,b\}\} \mid S_1 = y] \, \mathbb{P}[S_1 = y \mid S_0 = x]$$

$$= p\, h(x+1) + q\, h(x-1)$$

This is a difference equation which could be solved with trial solution $h(x) = z^x$. We follow another approach

$$\Rightarrow \quad h(x+1) - h(x) = \frac{q}{p}\Big(h(x) - h(x-1)\Big)$$

$$\Rightarrow \quad h(x+1) - h(x) = C\left(\frac{q}{p}\right)^{x-a}, \quad C = h(a+1) - h(a) = h(a+1)$$

$$\Rightarrow \quad h(x) = C \sum_{k=0}^{x-a-1} \left(\frac{q}{p}\right)^k \qquad (*)$$

When $p = q = \frac{1}{2}$, by $(*)$ and $h(b) = 1$

$$h(x) = C \cdot (x-a) = \frac{x-a}{b-a}$$

When $p \neq q$, by $(*)$ and $h(b) = 1$

$$h(x) = C\,\frac{1-\left(\frac{q}{p}\right)^{x-a}}{1-\left(\frac{q}{p}\right)} = \frac{1-\left(\frac{q}{p}\right)^{x-a}}{1-\left(\frac{q}{p}\right)^{b-a}}$$

$\square$

Section 5.2

Stirling's approximation and RW

Stirling, de Moivre (≈ 1700)

Lemma (Stirling's approximation)

As $n \to \infty$ $\quad n! \simeq n^n e^{-n} \sqrt{2\pi n}$

In fact $\quad n! = n^n e^{-n} \sqrt{2\pi n}\,(1 + O(\tfrac{1}{n}))$

The notation $a_n \simeq b_n \iff \frac{a_n}{b_n} \to 1$ as $n \to \infty$.

Proof Exercise.

Suppose for simplicity that $\boxed{p = q = \frac{1}{2}}$.

Theorem Let $x \in \mathbb{R}$ and $k = k_n \in \mathbb{Z}$ s.t. $2k \simeq x\sqrt{2n}$.

Then $P[S_{2n} = x_0 + 2k \mid S_0 = x_0] \simeq \frac{1}{\sqrt{\pi n}} \exp(-\frac{x^2}{2})$

Proof Let $k \in [\![-n, n]\!]$ s.t. $2k \simeq x\sqrt{2n} \Rightarrow \boxed{\frac{k^2}{n} \to \frac{x^2}{2}}$

$$\binom{2n}{n+k} = \frac{(2n)!}{(n-k)!\,(n+k)!}$$

$$\overset{\text{Stirling}}{\simeq} \frac{(2n)^{2n} e^{-2n} \sqrt{2\pi(2n)}}{(n-k)^{n-k} (n+k)^{n+k} e^{-2n}\, 2\pi\sqrt{(n-k)(n+k)}}$$

$$= 2^{2n} \underbrace{\sqrt{\frac{2n}{2\pi(n^2-k^2)}}}_{\simeq (\pi n)^{-1/2}} \underbrace{\left(1 - \frac{k^2}{n^2}\right)^{-n}}_{\simeq e^{+\frac{x^2}{2}}} \underbrace{\left(1 - \frac{k}{n}\right)^{k}}_{=\left(1-\frac{1}{k}\frac{k^2}{n}\right)^{k} \simeq e^{-\frac{x^2}{2}}} \underbrace{\left(1 + \frac{k}{n}\right)^{-k}}_{=\left(1+\frac{k}{n}\right)^{\dots} \simeq e^{-\frac{x^2}{2}}}$$

$$\simeq 2^{2n} \frac{1}{\sqrt{\pi n}} e^{-\frac{x^2}{2}} \qquad \square$$

These estimates are uniform for $x \in [a, b]$

$\Rightarrow$ a version of CLT $\quad \lim_{n\to\infty} \mathbb{P}\left[a \le \frac{S_n}{\sqrt{n}} \le b\right] = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$

Large deviations

In general, if $E[X_1] = \mu$, then $a > \mu$

$$-\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}[S_n \geq an] = \rho(a)$$

is the large deviations rate function.

Theorem (Cramér's entropy)
If $x \in (-1,1)$ and $k_n = n\frac{1+x}{2} + \mathcal{O}(1)$, $n \to \infty$
then

$$\log \binom{n}{k_n} = n\left(\log 2 - I(x)\right) + \mathcal{O}(\log n)$$

where $I(x) = \frac{1+x}{2}\log(1+x) + \frac{1-x}{2}\log(1-x)$.

Proof Let $k = k_n$ be as above. By Stirling's approximation

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$$= \frac{n^n e^{-n}\sqrt{2\pi n}\,(1+\mathcal{O}(\frac{1}{n}))}{k^k e^{-k}(n-k)^{n-k} e^{-n+k}\, 2\pi\sqrt{k(n-k)}}$$

$$= 2^n \underbrace{\sqrt{\frac{n}{2\pi\, k(n-k)}}}_{J_1}\ \underbrace{\left(2\frac{k}{n}\right)^{-k}}_{J_2}\ \underbrace{\left(2\frac{n-k}{n}\right)^{-n+k}}_{J_3}\ \underbrace{\left(1+\mathcal{O}(\tfrac{1}{n})\right)}_{J_4}$$

$$\log J_1 = \mathcal{O}(\log n)$$
$$\log J_2 = -n\left(\frac{1+x}{2} + \mathcal{O}(\tfrac{1}{n})\right)\log\left(1+x + \mathcal{O}(\tfrac{1}{n})\right)$$
$$\log J_3 = -n\left(\frac{1-x}{2} + \mathcal{O}(\tfrac{1}{n})\right)\log\left(1-x + \mathcal{O}(\tfrac{1}{n})\right)$$
$$\log J_4 = \mathcal{O}(\tfrac{1}{n})$$

$$\Rightarrow \log\binom{n}{k} = n\left(\log 2 - I(x)\right) + \mathcal{O}(\log n) \quad \square$$

If $S=(S_t)_{t\in\mathbb{N}}$ is the symmetric random walk, then we can show the following result.

Theorem (Large deviation for SRW)
If $0<a<1$, then $\lim_{n\to\infty} \frac{1}{2n}\log \mathbb{P}[S_n \geq an] = -I(a)$.

Proof Let

$$k_n^* = \min\{k\in\mathbb{Z} : k \geq an\}$$

$\Rightarrow$ $2k_n^*$ smallest even number $\geq 2an$

$k_n^* = an + O(1) \Rightarrow n+k_n^* = 2n\frac{1+a}{2} + O(1)$.
By the previous theorem

$$\log\binom{2n}{n+k_n^*} = 2n\left(\log 2 - I(x)\right) + O(\log n)$$

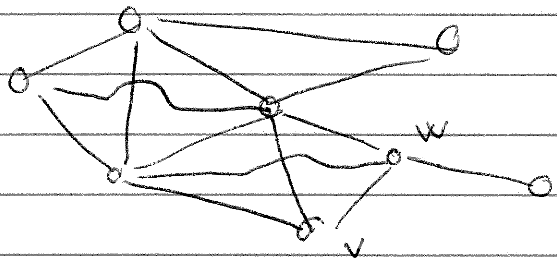
Let $p_{n,k} = \mathbb{P}[S_{2n} = x_0 + 2k] = \binom{2n}{n+k}2$
then $p_{n,k} \leq p_{n,k_n^*}$ $\forall\, k \geq k_n^*$

$$\Rightarrow p_{n,k_n^*} \leq \mathbb{P}(S_{2n} \geq x_0 + 2an)$$
$$\leq (n-k_n^*+1)\,p_{n,k_n^*} \leq n\,p_{n,k_n^*}$$

$$\Rightarrow \lim_{n\to\infty} \frac{1}{2n}\log\mathbb{P}(S_{2n}\geq x_0+2an)$$
$$= -\log 2 + \lim_{n\to\infty}\frac{1}{2n}\log\binom{2n}{n+k_n^*} = -I(a)$$
$\square$

Section 9.5 Random walk on graph G

A simple random walk on G is random nearest neighbor walk



$G = (V, E)$

$S_t \in V$

$$\mathbb{P}[S_{t+1} = w \mid S_t = v, \ldots, S_0 = v_0] = \frac{1}{\deg(v)} \quad \text{if } v \sim w$$