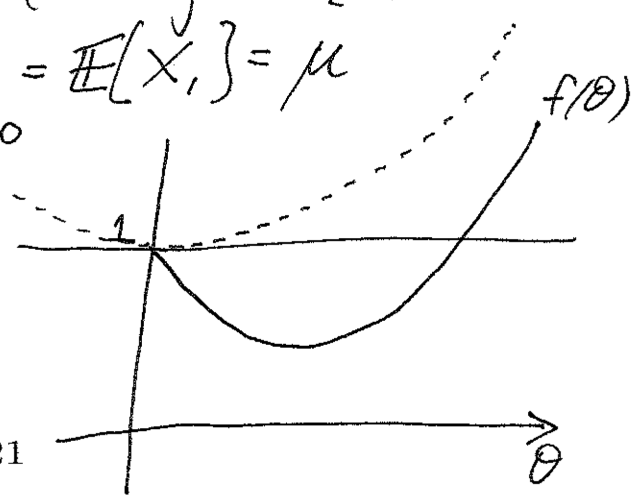
Martingales to Analyze Random Walks

- The general random walk, $\{S_n, n \geq 0\}$, is defined by $S_0 = 0$ and $S_n = \sum_{i=1}^{n} X_i$ for $n > 0$, where $X_1, X_2, \ldots$ are iid.
- A random walk can be considered as a generalization of a renewal process, where we drop the requirement that $X_i \geq 0$.
- The most obvious martingale is $S_n - n\mu$ where $\mu = \mathbb{E}[X_1]$. Here, μ is called the **drift**.
- Another useful martingale is $\exp\{\theta S_n\}$ where θ solves $\mathbb{E}[e^{\theta X_1}] = 1$. This equation has one solution at $\theta = 0$, and it usually has exactly one other solution, with $\theta > 0$, if $\mathbb{E}[X_1] < 0$. Why?

Let $f(\theta) = \mathbb{E}[e^{\theta X_1}]$. Then $f(\theta)$ is strictly convex, since $\frac{d^2 f}{d\theta^2} = \mathbb{E}[X_1^2 e^{\theta X_1}] > 0$ (as long as $\mathbb{P}[X_1 \neq 0] > 0$)

Also, $f(0) = 1$, and $\left.\frac{df}{d\theta}\right|_{\theta=0} = \mathbb{E}[X_1] = \mu$

If $\mu < 0$, the picture looks like this:



What if $\mu = 0$?

Example: Let $N = \min\{n : S_n \geq A \text{ or } S_n \leq -B\}$. Use martingale arguments to find (approximately) $\mathbb{P}[S_n \geq A]$ and $\mathbb{E}[N]$.

- Note: this models a general situation where we accumulate rewards, and at some point we quit and declare failure (if $S_n \leq -B$), or quite having achieved our goal (if $S_n \geq A$). An example is sequential analysis of clinical trails.

Solution Let $Z_n = \exp\{\theta S_n\}$ with $\mathbb{E}[e^{\theta X_1}] = 1$.

So, $\{Z_n\}$ is a Martingale and N is a stopping time.

Check: $\mathbb{E}[|Z_{n+1} - Z_n| \,|\, Z_1, \ldots, Z_n] \leq \mathbb{E}[Z_n + Z_{n+1} | Z_1, \ldots, Z_n]$ since $\{Z_n\}$ is positive.

$= 2Z_n$

So, for $n < N$, $\mathbb{E}[|Z_{n+1} - Z_n| \,|\, Z_1, \ldots, Z_n] \leq 2e^{\theta A}$

Also, $\mathbb{E}[N] < \infty$ (check, e.g. by bounding N with a negative binomial random variable).

So, the Martingale stopping theorem applies, and

$$\mathbb{E}[Z_N] = \mathbb{E}[Z_0] = 1$$

Let $P_A = \mathbb{P}[S_N \geq A]$ and $P_B = \mathbb{P}[S_N \leq -B]$

So, $\mathbb{E}[Z_N] = P_A\, \mathbb{E}[e^{\theta S_N} | S_N \geq A] + P_B\, \mathbb{E}[e^{\theta S_N} | S_N \leq -B]$ (*)

Ignoring overshoot & undershoot,

$$\mathbb{E}[e^{\theta S_N} \mid S_N \geq A] \approx e^{\theta A}$$

Solution continued

$$\mathbb{E}[e^{\theta S_N} \mid S_N \leq -B] \approx e^{-\theta B}$$

In (*), this gives

$$P_A e^{\theta A} + P_B e^{-\theta B} \approx 1$$

$$P_A \approx \frac{1-e^{-\theta B}}{e^{\theta A} - e^{-\theta B}}$$

Now, employ a similar martingale argument for $M_n = S_n - n\mu$. Then, by the stopping theorem,

$$\mathbb{E}[M_N] = \mathbb{E}[S_N] - \mu\mathbb{E}[N] = 0$$

(check the conditions!)

Ignoring overshoot & undershoot,

$$\mu\mathbb{E}[N] \approx AP_A - BP_B$$

$$\mathbb{E}[N] \approx \frac{1}{\mu}[AP_A - BP_B]$$

If $\mu = 0$, the exponential martingale doesn't exist. We can use a quadratic martingale, correcting S_n^2 by its conditional expectation.

7. Random Walks

The Duality Principle for Random Walks

- For $S_n = \sum_{i=1}^{n} X_i$ with $X_1, X_2, \ldots$ iid, we note that $(X_1, \ldots, X_n)$ has the same joint distribution as $(X_n, \ldots, X_1)$
- This obvious property has surprising consequences!

includes equality, not a strictly new low.

Example 1. Show that

$\mathbb{P}$[random walk doesn't exceed 0 by time n]

$= \mathbb{P}$[random walk hits a new low at time n].

$$\mathbb{P}[S_1 \le 0,\ S_2 \le 0,\ \ldots,\ S_n \le 0]$$

$$= \mathbb{P}[X_1 \le 0,\ X_1 + X_2 \le 0,\ \ldots,\ X_1 + X_2 + \cdots + X_n \le 0]$$

now apply duality

$$= \mathbb{P}[X_n \le 0,\ X_n + X_{n-1} \le 0, \ldots,\ X_n + X_{n-1} + \cdots + X_1 \le 0]$$

$$= \mathbb{P}[S_n \le S_{n-1},\ S_n \le S_{n-2}, \ldots,\ S_n \le 0]$$

$$= \mathbb{P}[S_n = \min_{0 \le k \le n} S_k]$$

- Now, notice that the times at which a random walk hits a new low are arrival times for a renewal process (possibly a defective renewal process, with positive probability of infinite arrival times). Why?

At each record low, the time until the next record low has the same distribution as

$$N = \left\{ \inf_{n>0} S_n \le 0 \right\} = \inf\{n : n>0 \text{ and } S_n \le 0\}$$

If $\mathbb{E}[X_1] > 0$, then the strong law of large numbers says $\lim_{n\to\infty} \frac{S_n}{n} = \mathbb{E}[X_1] > 0$ w.p. 1 and so there are only finitely many n with $S_n \le 0$, so there are only finitely many record lows (w.p. 1). This is possible only for a defective renewal process, so

$$\mathbb{P}[N=\infty] > 0.$$

<u>Example 2.</u> Use Example 1 to show that, for a random walk with **positive drift** (i.e., $\mathbb{E}[X_1] > 0$) $N = \min\{n : S_n > 0\}$ has $\mathbb{E}[N] < \infty$.

$$\mathbb{E}[N] = \sum_{n=0}^{\infty} \mathbb{P}[N > n]$$

$$= \sum_{n=0}^{\infty} \mathbb{P}[\text{Random walk doesn't exceed } 0 \text{ by time } n]$$

(now apply example 1)

$$= \sum_{n=0}^{\infty} \mathbb{P}[\text{RW hits a new low at time } n]$$

$$= \sum_{n=0}^{\infty} \mathbb{P}[n \text{ is a renewal time for the renewal process counting record lows}]$$

So, $\mathbb{E}[N] = \mathbb{E}[\#\text{ of renewal times}]$ $\quad (*)$

It may be surprising that $(*)$ relates the expected time until a positive value is observed to the ~~total~~ expected <u>number</u> of record low values.

Since $\mathbb{E}[X_1] > 0$, we argued previously that the renewal process is defective, so the # of renewals is geometrically distributed and $\mathbb{E}[\#\text{ of renewal times}] < \infty$.

Example 3. Let S_n be a random walk on the integers (i.e., X_1 takes integer values). Show that

$$\mathbb{P}\big[S_n = k,\ \text{no return to zero before time } n\big]$$
$$= \mathbb{P}\big[\text{random walk first hits } k \text{ at time } n\big].$$

$\mathbb{P}[S_n = k,$ no return to 0 before time $n]$

$= \mathbb{P}[X_1 + X_2 + \dots + X_n = k,\ X_1 \neq 0,\ X_1 + X_2 \neq 0, \dots, X_1 + \dots + X_{n-1} \neq 0]$

(now apply duality)

$= \mathbb{P}[X_n + X_{n-1} + \dots + X_1 = k,\ X_n \neq 0,\ X_n + X_{n-1} \neq 0, \dots, X_n + \dots + X_2 \neq 0]$

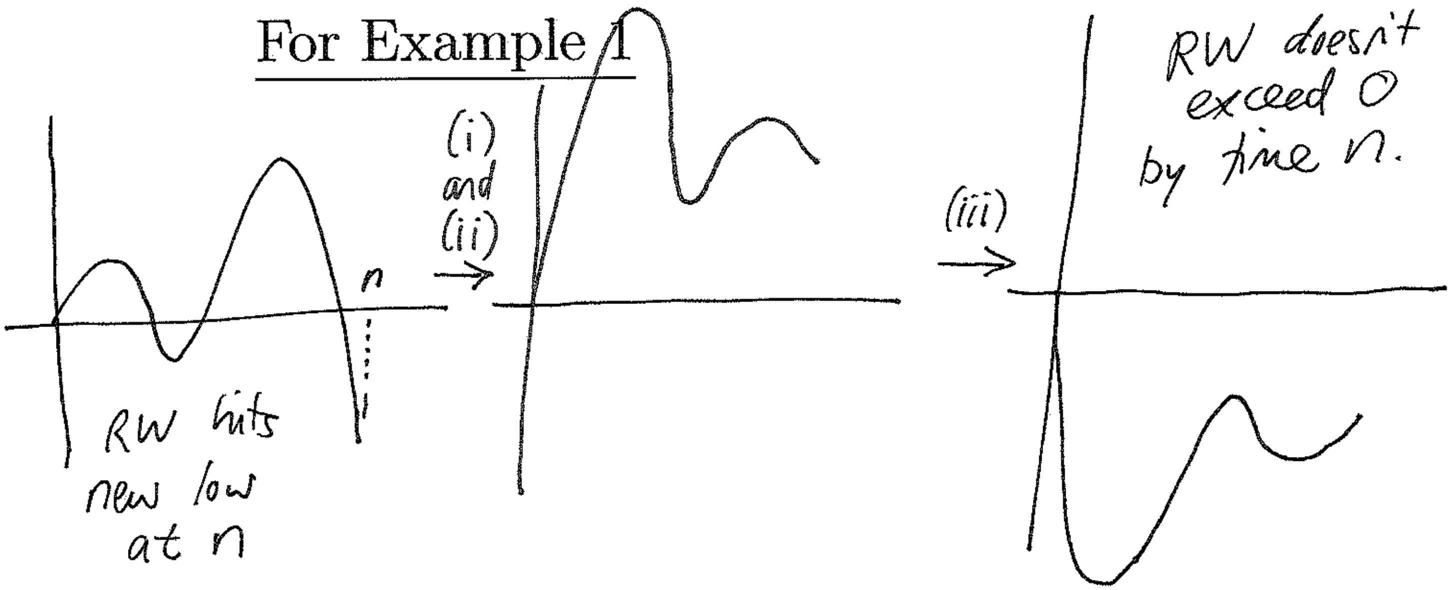
$= \mathbb{P}[\ S_n = k,\ S_n \neq S_{n-1},\ S_n \neq S_{n-2}, \dots,\ S_n \neq S_1\ ]$

$= \mathbb{P}[\ S_n = k,\ S_{n-1} \neq k,\ S_{n-2} \neq k, \dots,\ S_1 \neq k]$

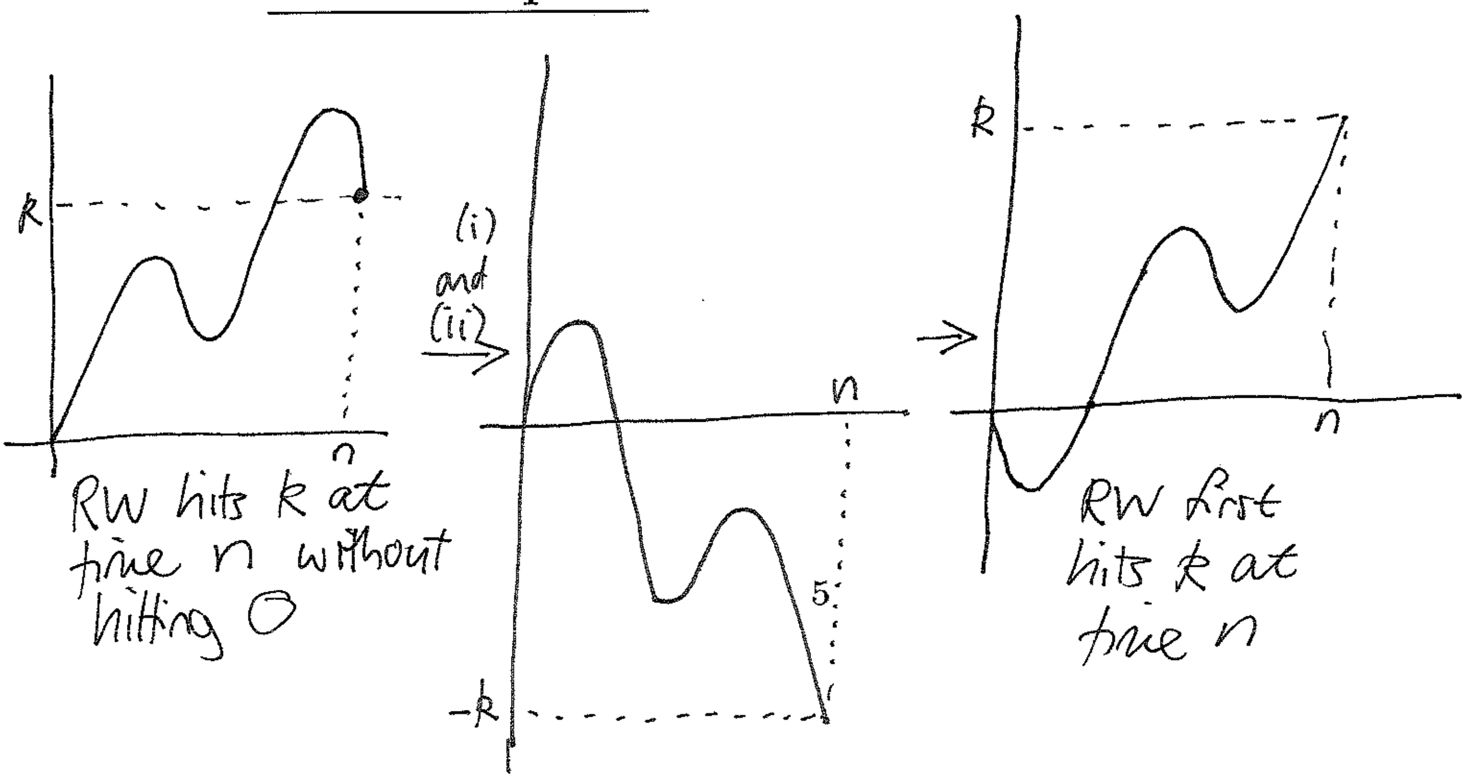
$= \mathbb{P}[$RW first hits k at time $n]$

- Duality is distantly related to time reversal. Sample paths of the dual process can be obtained by:
  (i) Look backwards in time, starting at time n.
  (ii) Shift the trajectory so its initial value is 0.
  (iii) Reflect the trajectory about the x-axis.

For Example 1



For Example 3



<u>Example 4</u>: Use Example 3 to show that

$\mathbb{E}\big[\#$ of visits to k before returning to $0\big]$

$= \mathbb{P}\big[$random walk ever hits $k\big]$

$(= 1$ for a recurrent random walk).

$\mathbb{E}[\#$ of visits to k before returning to $0]$

$= \sum_n \mathbb{P}[S_n = k, S_1 \neq 0, S_2 \neq 0, \ldots, S_{n-1} \neq 0]$

now apply duality via Example 3

$= \sum_n \mathbb{P}[$ 1^{st} arrival at k occurs at time $n]$

$= \mathbb{P}[$Random walk ever hits $k]$.