

# Lecture 1: Maximal inequality & the law of the iterated log

①

②

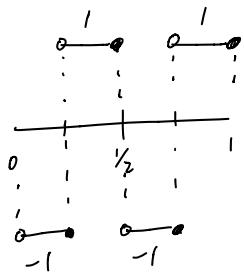
Throughout this lecture

$(\Omega, \mathcal{F}, P)$  will denote the

uniform prob measure on  $\Omega = \{0, 1\}^{\mathbb{N}}$   
generated by the binary digit  
coinflips from lecture 1.

Recall

$$S_n(w) := \sum_{k=1}^n R_k(w)$$



We will show two "maximal inequalities" which are useful for studying random series, martingales etc...  
the law of the iterated log etc...

Thm 1.

If  $a \geq 0$  then

$$P\left(\max_{k \leq n} S_k \geq a\right) \leq 2P(S_n \geq a).$$

Thm 2 (Kolmogorov's inequality).

If  $a \geq 0$  then

$$P\left(\max_{k \leq n} |S_k| \geq a\right) \leq \frac{1}{a^2} \sum_{k=1}^n E(R_k^2).$$

Remark 1:

After the proof notice Thm 1 & 2 both hold when  $R_k$  is replaced with  $\frac{R_k}{k}$  &  $S_n$  is replaced with  $S'_n := \sum_{k=1}^n \frac{R_k}{k}$

Remark 2:

Thm 2 will generalize to the case  $S_n := \sum_{k=1}^n Y_k$  for indep  $Y_k$  s.t.  $E(Y_k) = 0$ .

Also generalizes to martingale sequences  $S_1, S_2, \dots$

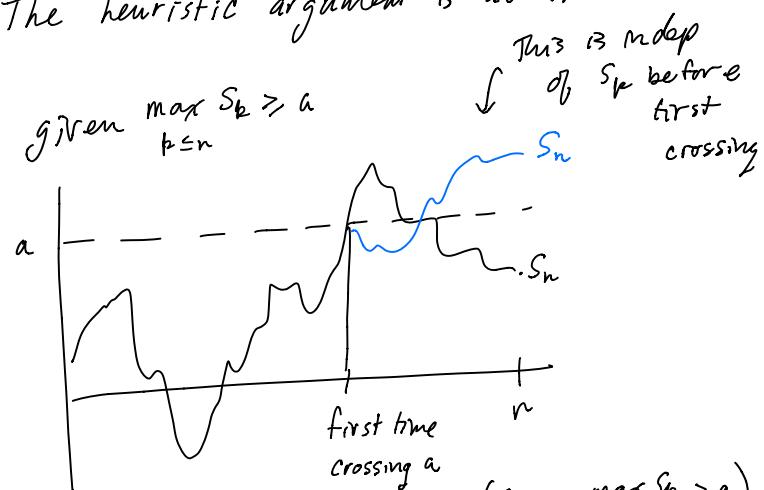
Remark 3:

Thm 1 will generalize when  $S_n$  is a "symmetric random walk"

$S_n := \sum_{k=1}^n Y_k$  where  $Y_k$ 's are indep &  $Y_k$  has the same distribution as  $-Y_k$ .

Also generalizes to Brownian motion.

The heuristic argument is as follows:



$$\begin{aligned} \frac{1}{2} &= P(S_n \geq a \mid \max_{k \leq n} S_k \geq a) = \frac{P(S_n \geq a, \max_{k \leq n} S_k \geq a)}{P(\max_{k \leq n} S_k \geq a)} \\ &\therefore P(\max_{k \leq n} S_k \geq a) = 2P(S_n \geq a, \max_{k \leq n} S_k \geq a) \\ &\leq 2P(S_n \geq a). \end{aligned}$$

Proof of Thm 1: ✓ trick #1  
partition via  $s_n > a$  (3)

$$\begin{aligned} P(\max_{k \leq n} s_k \geq a) &= P(\max_{k \leq n} s_k \geq a, s_n < a) \\ &\quad + P(\underbrace{\max_{k \leq n} s_k \geq a}_{\text{This set}}, \underbrace{s_n \geq a}_{\text{This set}}) \\ &= P(\max_{k \leq n} s_k \geq a, s_n < a) \\ &\quad + P(s_n \geq a) \end{aligned}$$

So we just need to show  $\downarrow$

$$P(\max_{k \leq n} s_k \geq a, s_n < a) \leq P(s_n \geq a). \quad (*)$$

Trick #2, partition  $\{\max_{k \leq n} s_k \geq a\}$  by  
the first time  $s_k \geq a$ :

$$\{\max_{k \leq n} s_k \geq a\} = \bigcup_{k=1}^n \left\{ s_1 < a, \dots, s_{k-1} < a, s_k \geq a \right\} \quad := F_k \text{ are disjoint.}$$

$$\begin{aligned} \therefore P(\max_{k \leq n} s_k \geq a, s_n < a) &= \sum_{k=1}^{n-1} P(F_k \cap \{s_n < a\}), \\ &\quad F_n \cap \{s_n < a\} = \emptyset \\ &= \sum_{k=1}^{n-1} P(F_k \cap \{s_n - s_k < 0\}) \\ &\quad \text{since } s_k \geq a \text{ on } F_k \\ &\quad \therefore s_n < a \leq s_k \\ &\quad \therefore s_n - s_k < 0 \\ &= \sum_{k=1}^{n-1} P(F_k) P(s_n - s_k < 0) \\ &\quad \in \sigma(H_1, \dots, H_n) \quad \in \sigma(H_{k+1}, \dots, H_n) \\ &\quad \text{are indep} \\ &= \sum_{k=1}^{n-1} P(F_k) P(s_k - s_n < 0) \\ &\quad \text{since probabilities} \\ &\quad \text{for } s_1, s_2, \dots \text{ are} \\ &\quad \text{the same for} \\ &\quad -s_1, -s_2, \dots \\ &= \sum_{k=1}^{n-1} P(F_k \cap \{s_k - s_n < 0\}) \quad \text{again by indep.} \end{aligned}$$

$$\begin{aligned} &= \sum_{k=1}^{n-1} P(F_k \cap \{s_k < s_n\}) \quad (4) \\ &= \sum_{k=1}^{n-1} P(F_k \cap \{a < s_n\}) \\ &\quad \text{since } a \leq s_k < s_n \\ &\quad \uparrow \text{on } F_k \\ &\leq \sum_{k=1}^n P(F_k \cap \{a < s_n\}) \\ &= P\left(\left(\bigcup_{k=1}^n F_k\right) \cap \{a < s_n\}\right) \\ &= P\left(\max_{k \leq n} s_k \geq a, a < s_n\right) \\ &\quad \text{this } \rightarrow \text{this} \\ &= P(a < s_n) \leq P(a \leq s_n) \end{aligned}$$

$\therefore (*)$  holds as was to be shown.  $\square \text{ED}$

Proof of Kolmogorov's inequality:

Use same trick #2

$$\{\max_{k \leq n} |s_k| \geq a\} = \bigcup_{k=1}^n \left\{ |s_1| < a, \dots, |s_{k-1}| < a, |s_k| \geq a \right\} \quad := F_k, \text{ disjoint}$$

Note: if  $A_i \in \sigma(H_1, \dots, H_n)$  &

$B_j \in \sigma(H_{k+1}, \dots, H_n)$  s.t.

$$\begin{aligned} Y(w) &:= \sum a_i I_{A_i}(w) \\ X(w) &:= \sum b_j I_{B_j}(w) \end{aligned} \quad \left. \begin{array}{l} \text{Finite sums} \\ \text{with} \\ a_i, b_j \in \mathbb{R} \end{array} \right\}$$

Then  $E(Y) := \int_0^1 Y(w) dw$  is well defined  
as a Riemann integral.

Similarly for  $E(X)$ .

Moreover

$$\begin{aligned}
 E(XY) &= \sum_{ij} a_i b_j \int_0^1 I_{A_i}(w) I_{B_j}(w) dw \\
 &\quad \underbrace{\qquad\qquad}_{P(A_i \cap B_j)} \\
 &= \sum_{ij} a_i b_j P(A_i) P(B_j) \\
 &\quad \vdots \qquad \text{by Indep.} \\
 &= E(X)E(Y).
 \end{aligned}
 \tag{5}$$

Now

$$\sum_{k=1}^n E(R_k^2) = E(S_n^2) \quad \text{since } R_k \text{'s are orthonormal.}$$

$$\geq \sum_{k=1}^n E(I_{F_k} S_n^2)$$

since  $\sum_{k=1}^n I_{F_k} = I_{\{\max_{k \leq n} |S_k| \geq a\}}$

$$\begin{aligned}
 &\text{expand } (S_n + S_k)^2 I_{F_k} \rightsquigarrow \geq \sum_{k=1}^n E(I_{F_k} S_k^2) \\
 &\text{& drop } E(S_n - S_k)^2 I_{F_k} \\
 &E(S_n - S_k)^2 I_{F_k} + 2 \sum_{k=1}^n E(S_k I_{F_k} (S_n - S_k))
 \end{aligned}$$

where

$$\begin{aligned}
 E(S_k I_{F_k} (S_n - S_k)) &= E(S_k I_{F_k}) E(S_n - S_k) \\
 &\quad \text{constant over sets } \{H_1, \dots, H_n\} \\
 &= 0
 \end{aligned}$$

$$\begin{aligned}
 \sum_{k=1}^n E(R_k^2) &\geq \sum_{k=1}^n E(I_{F_k} S_k^2) \\
 &\geq a^2 \sum_{k=1}^n E(I_{F_k}) \\
 &= a^2 P(\max_{k \leq n} |S_k| \geq a) \quad \text{QED}
 \end{aligned}$$

(5)

Application: Series with random signs

(6)

$$\begin{array}{lcl}
 \sum_n \frac{1}{n} = \infty & \left\{ \begin{array}{l} \text{no cancellation at all} \end{array} \right. \\
 \sum_n R_n \frac{1}{n} = ? & \left\{ \begin{array}{l} \text{cancellation B} \\ \text{given by random coin flips} \end{array} \right. \\
 \sum_n R_n \frac{1}{\sqrt{n}} = ? & \left\{ \begin{array}{l} \text{cancellation is finely tuned} \end{array} \right. \\
 \sum_n (-1)^n \frac{1}{\sqrt{n}} < \infty & \left\{ \begin{array}{l} \text{cancellation is finely tuned} \end{array} \right.
 \end{array}$$

For now we will only analyze the second. The third is answered by the Rademacher-Paley-Zygmund Thm.

$$\text{Let } \tilde{S}_N = \sum_{n=1}^N R_n \cdot \frac{1}{n}.$$

Notice that the proof of Kolmogorov's inequality goes through for

$$\begin{aligned}
 P\left(\max_{N \leq n \leq M} |\tilde{S}_n - \tilde{S}_N| \geq a\right) &\leq \frac{1}{a^2} \sum_{n=N}^M E(R_n^2 \frac{1}{n^2}) \\
 &= \frac{1}{a^2} \sum_{n=N}^M \frac{1}{n^2}
 \end{aligned}$$

Note  $\max_{N \leq n \leq M} |\tilde{S}_n - \tilde{S}_N| \uparrow$  as  $M \rightarrow \infty$ .

$$\therefore \left\{ \max_{N \leq n \leq M} |\tilde{S}_n - \tilde{S}_N| \geq a \right\} \uparrow \left\{ \sup_{N \leq n} |\tilde{S}_n - \tilde{S}_N| \geq a \right\}$$

as  $M \rightarrow \infty$

$$\therefore P\left(\sup_{N \leq n} |\tilde{S}_n - \tilde{S}_N| \geq a\right) \leq \frac{1}{a^2} \sum_{n=N}^{\infty} \frac{1}{n^2}$$

by continuity from below

$$\therefore P\left(\sup_{N \leq n, m} |\tilde{S}_n - \tilde{S}_m| \geq 2a\right) \xrightarrow{\substack{\uparrow \\ \tilde{S}_N}} \tilde{S}_N$$

$$\leq P\left(2 \sup_{N \leq n} |\tilde{S}_n - \tilde{S}_N| \geq 2a\right)$$

$$\leq \frac{1}{a^2} \sum_{n=N}^{\infty} \frac{1}{n^2} \rightarrow 0 \text{ as } N \rightarrow \infty$$

$\forall \alpha > 0$

$$\lim_{n \rightarrow \infty} P\left(\sup_{N \leq n, m} |\tilde{S}_n - \tilde{S}_m| \geq 2\alpha\right) = 0$$

$$P\left(\liminf_{n \rightarrow \infty} \left\{ \sup_{N \leq n, m} |\tilde{S}_n - \tilde{S}_m| \geq 2\alpha\right\}\right)$$

⑦

The law of the iterated logarithm

⑧

Reminder:  $P$  is the uniform measure on  $(0, 1]$  &  $S_n(w) := \sum_{k=1}^n R_k(w)$ .

Motivation:

WLLN:  $\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) \rightarrow 0$  as  $n \rightarrow \infty$

SLLN:  $P\left(\lim_{n \rightarrow \infty} \frac{S_n}{n} = 0\right) = 1$ .

So the WLLN describes the ensemble

$$\left\{ \frac{S_n(w)}{n} : w \in \mathbb{J} \right\}$$

for a fixed  $w \in \mathbb{J}$ .

The SLLN describes the ensemble

$$\left\{ \frac{S_n(w)}{n} : n \in \mathbb{N} \right\}$$

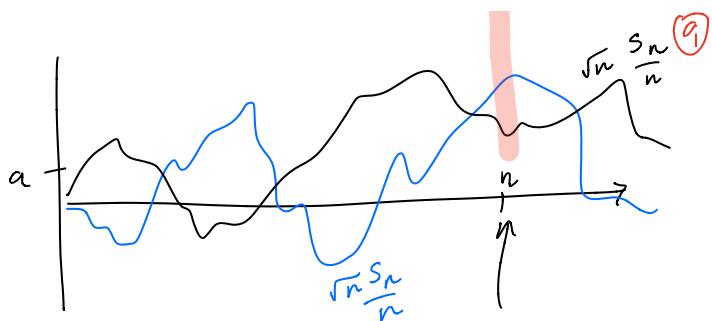
for a fixed  $w \in \mathbb{J}$ .

However both results say  $\frac{S_n(w)}{n}$  is "near" 0. To get an idea of rates consider scaling an  $\frac{S_n(w)}{n}$  for some  $a_n \uparrow \infty$  to lift  $\frac{S_n(w)}{n}$  away from 0.

Remark: This was a lot of work for one series... However the proofs are exactly similar for general random sums  $\sum_{k=1}^{\infty} Y_k$  where the  $Y_k$ 's are indep,  $E(Y_k) = 0$ . See durrett p. 79.

Later we will show the CLT (central limit thm) which says

$$\text{CLT: } \lim_{n \rightarrow \infty} P\left(\sqrt{n} \frac{S_n}{n} \geq a\right) = \int_a^{\infty} e^{-t^2/2} dt$$



The CLT describes the probability a random from  $\{\sqrt{n} \frac{S_n(w)}{n}; n \in \mathbb{N}\}$  is greater than a. For a fixed but large  $n$ .

Unfortunately these rates are not right for  $\{\sqrt{n} \frac{S_n(w)}{n}; n \in \mathbb{N}\}$  since the sup of  $\sigma$  is  $\infty$  for nearly all  $w$ .

The LIL (law of the iterated log) gives the right "almost sure rates".

$$\sup \left\{ \frac{\sqrt{n}}{\sqrt{\log \log n}} \frac{S_n(w)}{n}; n \in \mathbb{N} \right\} = \sqrt{2}$$

$$\inf \left\{ \frac{\sqrt{n}}{\sqrt{\log \log n}} \frac{S_n(w)}{n}; n \in \mathbb{N} \right\} = -\sqrt{2}$$

for nearly all  $w \in \Omega$  (i.e. with prob 1).

Intuitively the  $\sqrt{\log \log n}$  term controls the rare gaussian excursions  $\sqrt{n} \frac{S_n}{n}$  takes under the CLT.

Another way to write this is (10)

$$(i) P \left( \limsup_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \log \log n}} = 1 \right) = 1$$

$$(ii) P \left( \liminf_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \log \log n}} = -1 \right) = 1$$

Notice that (i) is equivalently written:  $\forall \varepsilon > 0$

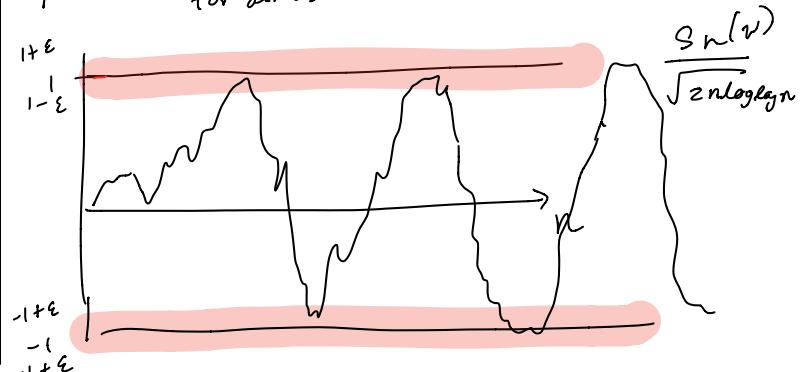
$$P \left( \frac{S_n}{\sqrt{2n \log \log n}} > 1 - \varepsilon \text{ i.o.n} \right) = 1$$

$$\text{&} P \left( \frac{S_n}{\sqrt{2n \log \log n}} < 1 + \varepsilon \text{ a.a.n} \right) = 1$$

(& similarly for (ii)) which follows since

$$\left\{ \limsup_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \log \log n}} = 1 \right\} = \bigcap_{\varepsilon \in (0, 1) \cap \mathbb{Q}^+} \left\{ \frac{S_n}{\sqrt{2n \log \log n}} > 1 - \varepsilon \text{ i.o.n} \text{ and } \frac{S_n}{\sqrt{2n \log \log n}} < 1 + \varepsilon \text{ a.a.n} \right\}$$

This gives a nice picture: for almost all  $w$  implies  $\limsup < 1 + \varepsilon$



The LIL is a very detailed/fine analysis ... but it's a bit tedious.

I like covering it, however, since it gives a very advanced usage of both the First & second Borel-Cantelli lemmas.

The FBCL is used to show

$$P\left(\frac{s_n}{\sqrt{2n \log \log n}} \geq 1+\varepsilon \text{ i.o.n}\right) = 0$$

but it can't be applied directly

$$\text{since } \sum_{n=1}^{\infty} P\left(\frac{s_n}{\sqrt{2n \log \log n}} \geq 1+\varepsilon\right) = \infty.$$

The trick is to study sub-sequences

$n_p$ .

The SBCL is used to show

$$P\left(\frac{s_n}{\sqrt{2n \log \log n}} \geq 1+\varepsilon \text{ i.o.n}\right) = 1$$

but, again, can't be applied directly since the events

$\left\{ \frac{s_n}{\sqrt{2n \log \log n}} \geq 1+\varepsilon \right\}$  are not

independent. The trick is to again look at subsequences &

find  $I_{k \cap A_k} \subset \left\{ \frac{s_n}{\sqrt{2n \log \log n}} \geq 1+\varepsilon \right\}$

where the SBCL can be applied to  $I_k$  &  $P(A_k \text{ a.a.p.}) = 1$   
so  $P(I_k \cap A_k \text{ i.o.p.}) = 1$ .

Instead of writing the proof on the black board let's move to the projector