

Lecture 7: Maximal inequality & ①
the law of the iterated log

Throughout this lecture

(Ω, \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

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

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

Theorem 1 (symmetric random walk Maximal Ineq)
if $a \geq 0$ then

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

Theorem 2 (Kolmogorov's Maximal inequality)

If $a \geq 0$ then

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

Remark 1 ②:

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

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:

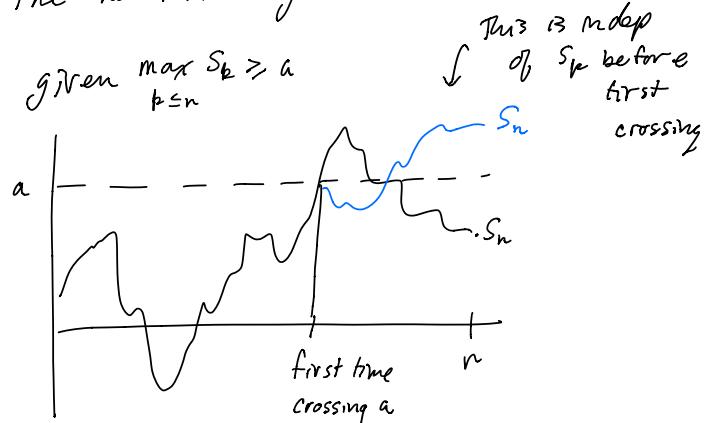
Thm 1 will generalize when S_n is a "symmetric random walk"

$$S_n := \sum_{k=1}^n Y_k \text{ where } Y_k \text{'s are indep}$$

& Y_k has the same distribution as $-Y_k$.

Also generalizes to Brownian motion.

The heuristic argument is as follows:



$$\frac{1}{2} = P(S_n \geq a \mid \max_{k \leq n} S_k \geq a) = \frac{P(S_n \geq a, \max_{k \leq n} S_k \geq a)}{P(\max_{k \leq n} S_k \geq a)}$$

$$\therefore P(\max_{k \leq n} S_k \geq a) = 2P(S_n \geq a, \max_{k \leq n} S_k \geq a) \leq 2P(S_n \geq a).$$

Proof of Thm 1: (3)

$P(\max_{k \leq n} S_k \geq a) = P(\max_{k \leq n} S_k \geq a, S_n < a)$

$\quad + P(\max_{k \leq n} S_k \geq a, S_n \geq a)$

$\quad \quad \quad \text{This set} \supset \text{This set}$

$= P(\max_{k \leq n} S_k \geq a, S_n < a)$

$\quad + P(S_n \geq a)$

$\quad \quad \quad \text{idea ... try to flip this to } a \leq S_n$

So we just need to show $P(\max_{k \leq n} S_k \geq a, S_n < a) \leq P(S_n \geq a)$. (*)

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

$\{\max_{k \leq n} S_k \geq a\} = \bigcup_{k=1}^n \left\{ S_1 < a, \dots, S_{k-1} < a, S_k \geq a \right\}$

$\quad \quad \quad \text{=: } F_k \text{ are disjoint.}$

$$\begin{aligned} & \therefore P(\max_{k \leq n} S_k \geq a, S_n < a) \\ &= \sum_{k=1}^{n-1} P(F_k \cap \{S_n < a\}), \\ & \quad F_n \cap \{S_n < a\} = \emptyset \\ &\leq \sum_{k=1}^{n-1} P(F_k \cap \{S_n - S_k < 0\}) \\ & \quad \text{since } S_k \geq a \text{ on } F_k \\ & \quad \therefore S_n < a \leq S_k \\ & \quad \therefore S_n - S_k < 0 \\ &= \sum_{k=1}^{n-1} P(F_k) P(S_n - S_k < 0) \\ & \quad \in \sigma(H_1, \dots, H_k) \quad \in \sigma(H_{k+1}, \dots, H_n) \\ & \quad \text{are indep.} \\ &= \sum_{k=1}^{n-1} P(F_k) P(S_k - S_n < 0) \\ & \quad \text{since probabilities for } S_1, S_2, \dots \text{ are the same for } -S_1, -S_2, \dots \\ &= \sum_{k=1}^{n-1} P(F_k \cap \{S_k - S_n < 0\}) \\ & \quad \text{again by indep.} \end{aligned}$$

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

$\therefore (*)$ holds as was to be shown. QED

Proof of Kolmogorov's inequality:

Use same trick #2

$$\{\max_{k \leq n} |S_k| \geq a\} = \bigcup_{k=1}^n \left\{ |S_1| < a, \dots, |S_{k-1}| < a, |S_k| \geq a \right\}$$

$\quad \quad \quad \text{=: } F_k \text{, disjoint}$

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

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

$$\begin{aligned}
 \sum_{k=1}^n E(R_k^2) &= E(S_n^2) \quad \text{since } R_k \text{'s are orthonormal.} \\
 &\geq \sum_{k=1}^n E(I_{F_k} S_k^2) \\
 &\quad \text{since } \sum_{k=1}^n I_{F_k} = I_{\{\max_{k \leq n} |S_k| \geq a\}} \\
 &\text{expand } (S_n + S_k)^2 I_{F_k} \rightsquigarrow \geq \sum_{k=1}^n E(I_{F_k} S_k^2) \\
 &\text{drop } E(S_n - S_k)^2 I_{F_k} \\
 &\quad + 2 \sum_{k=1}^n E(S_k I_{F_k} (S_n - S_k))
 \end{aligned}$$

where

$$\begin{aligned}
 E(S_k I_{F_k} (S_n - S_k)) &= E(S_k I_{F_k}) E(S_n - S_k) \\
 &\quad \text{constant over sets } \{H_{k+1}, \dots, H_n\} \\
 &= 0
 \end{aligned}$$

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

Application: Series with random signs

(6)

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

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

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

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

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

as $M \rightarrow \infty$

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

by continuity from below

$$\begin{aligned}
 &\therefore P\left(\sup_{n,m: N \leq n, m} |\tilde{S}_n - \tilde{S}_m| > 2a\right) \\
 &\quad \leftarrow \uparrow \tilde{S}_N \\
 &\leq P(2 \sup_{n:N \leq n} |\tilde{S}_n - \tilde{S}_N| > 2a) \\
 &\leq \frac{1}{a^2} \sum_{n=N+1}^{\infty} \frac{1}{n^2} \rightarrow 0 \text{ as } n \rightarrow \infty
 \end{aligned}$$

$\therefore \forall a > 0$

(7)

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

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

\therefore

$$P\left(\bigcup_{a \in \mathbb{Q}^+} \bigcap_{N=1}^{\infty} \left\{ \sup_{N \leq n, m} |\tilde{S}_n - \tilde{S}_m| > 2a \right\}\right) = 0$$

$$\therefore P\left(\bigcap_{a \in \mathbb{Q}^+} \bigcup_{N=1}^{\infty} \left\{ \sup_{N \leq n, m} |\tilde{S}_n - \tilde{S}_m| \leq 2a \right\}\right) = 1$$

Notice: we didn't need a summable decay rate in the upper bound

$P(X_N > a) \leq b(a, N)$ to conclude $P(X_N \rightarrow 0) = 1$

where $X_N := \sup_{N \leq n, m} |\tilde{S}_n - \tilde{S}_m|$ since

$X_N \downarrow$ monotonically.

$\Rightarrow \tilde{S}_n$ is a Cauchy sequence
 $\Rightarrow \lim_n \tilde{S}_n(w)$ exists & is finite

QED

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.

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

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

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

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

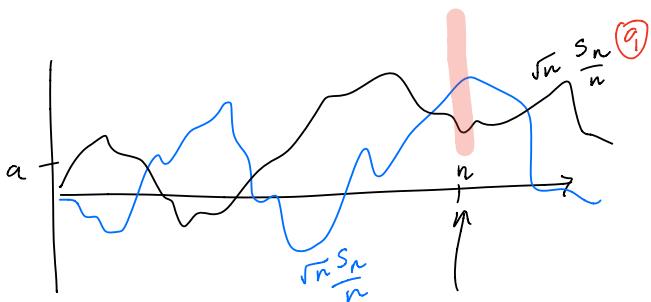
for a fixed $w \in \mathbb{J}$.

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.

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

CLT:

$$\lim_{n \rightarrow \infty} P\left(\sqrt{n} \frac{S_n}{n} \geq a\right) = \frac{1}{\sqrt{2\pi}} \int_a^{\infty} e^{-t^2/2} dt$$



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

Unfortunately these rates are not right for $\left\{ \sqrt{n} \frac{S_n(w)}{n} : n \in \mathbb{N} \right\}$ since the sup of \mathbb{P} is ∞ 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 $w \in \mathcal{D}$ (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: $\forall \varepsilon > 0$

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

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

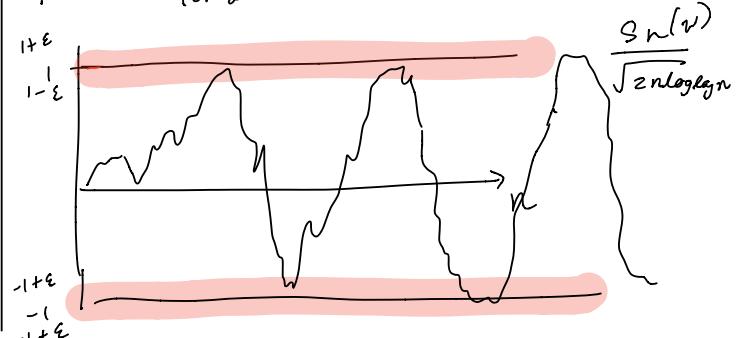
(& similarly for (ii)) which follows since

$$\left\{ \limsup_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \log \log n}} = 1 \right\} \quad \begin{matrix} \text{implies} \\ \limsup > 1 - \varepsilon \end{matrix}$$

$$= \bigcap_{\varepsilon \in (0,1) \cap \mathbb{Q}} \left\{ \frac{S_n}{\sqrt{2n \log \log n}} > 1 - \varepsilon \text{ i.o.n} \right\}$$

$$\text{and } \frac{S_n}{\sqrt{2n \log \log n}} < 1 + \varepsilon \text{ a.a.n} \quad \begin{matrix} \uparrow \\ \text{implies} \\ \limsup < 1 + \varepsilon \end{matrix}$$

This gives a nice picture: for almost all w



The LIL is a very detailed/fine analysis ... but it's a bit tedious. (11)

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

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

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

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

independent. The trick is to again look at subsequences &

find $I_k \cap A_k \subset \left\{ \frac{s_{n_k}}{\sqrt{2n_k \log \log n_k}} > 1+\varepsilon \right\}$

where the SBCL can be applied to I_k & $P(A_k \text{ a.a.p}) = 1$
so $P(I_k \cap A_k \text{ i.o.p}) = 1$. (12)

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