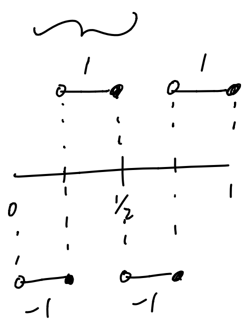
Lecture 7: Maximal inequality & the law of the iterated log

Throughout this lecture $(\Omega, \mathcal{F}, P)$ will denote the uniform prob measure on $\Omega = (0,1]$ 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 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

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

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. $E(Y_k) = 0$.

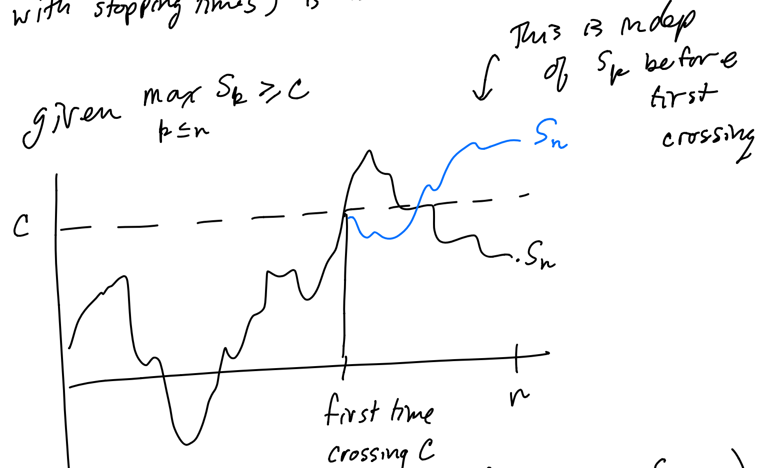
Also generalizes to martingale sequences $S_1, S_2, \ldots$

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 $-Y_k$.

The heuristic argument (done rigorously with stopping times) is as follows

given $\max_{k \leq n} S_k \geq c$



$$\frac{1}{2} = P\left(S_n \geq c \,\middle|\, \max_{k \leq n} S_k \geq c\right) = \frac{P\left(S_n \geq c, \max_{k \leq n} S_k \geq c\right)}{P\left(\max_{k \leq n} S_k \geq c\right)}$$

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

Proof of Thm 1:

trick 1: partition via $S_n < a$ vrs $S_n \geq a$

$$P(\max_{k \le n} S_k \geq a) = P(\max_{k\le n} S_k \geq a, S_n < a) + P(\underbrace{\max_{k \le n} S_k \geq a}_{\text{This set}}, \underbrace{S_n \geq a}_{\text{This set}})$$

This set $\supset$ This set

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

So we just need to show

(*) $$P(\max_{k\le n} S_k \geq a, S_n < a) \le P(S_n \geq a).$$

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

$$\{\max_{k\le n} S_k \geq a\} = \bigcup_{k=1}^{n} \underbrace{\{S_1 < a, \dots, S_{k-1} < a, S_k \geq a\}}_{=: F_k \text{ are disjoint.}}$$

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

$$= \sum_{k=1}^{n-1} P(F_k \cap \{S_n < a\}), \quad F_n \cap \{S_n < a\} = \emptyset$$

$$= \sum_{k=1}^{n-1} P(F_k \cap \{S_n - S_k < 0\})$$

since $S_k \geq a$ on F_k

$\therefore S_n < a \le S_k$

$\therefore S_n - S_k < 0$

$$= \sum_{k=1}^{n-1} P(\underbrace{F_k}_{\in\sigma\langle H_1,\dots,H_k\rangle}) P(\underbrace{S_n - S_k < 0}_{\in\sigma\langle H_{k+1},\dots,H_n\rangle})$$

are indep

$$= \sum_{k=1}^{n-1} P(F_k) P(S_k - S_n < 0)$$

since probabilities for $S_1, S_2, \dots$ are the same for $-S_1, S_2, \dots$

$$= \sum_{k=1}^{n-1} P(F_k \cap \{S_k - S_n < 0\})$$

again by indep.

$$= \sum_{k=1}^{n-1} P(F_k \cap \{S_k < S_n\})$$

$$= \sum_{k=1}^{n-1} P(F_k \cap \{a < S_n\})$$

since $a \le S_k < S_n$ ($a \le S_k$ on F_k)

$$\le \sum_{k=1}^{n} P(F_k \cap \{a < S_n\})$$

$$= P\left(\left(\bigcup_{k=1}^{n} F_k\right) \cap \{a < S_n\}\right)$$

$$= P(\underbrace{\max_{k\le n} S_k \geq a}_{\text{this}}, \underbrace{a < S_n}_{\text{this}})$$

this $\supset$ this

$$= P(a < S_n) \le P(a \le S_n)$$

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

Proof of Kolmogorov's inequality:

Use same trick #2

$$\{\max_{k\le n} |S_k| \geq a\} = \bigcup_{k=1}^{n} \underbrace{\{|S_1| < a, \dots, |S_{k-1}| < a, |S_k| \geq a\}}_{=: F_k, \text{ disjoint}}$$

Note: if $A_i \in \sigma\langle H_1,\dots,H_k\rangle$ & $B_j \in \sigma\langle H_{k+1},\dots,H_n\rangle$ s.t.

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

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

Similarly for $E(X)$.

Moreover

$$E(XY) = \sum_{ij} a_i b_j \underbrace{\int_0^1 I_{A_i}(w) I_{B_j}(w)\, dw}_{P(A_i \cap B_j)}$$

$$= \sum_{ij} a_i b_j P(A_i) P(B_j)$$ by indep.

$\vdots$

$$= E(X)E(Y).$$

Now

$$\sum_{p=1}^{n} E(R_p^2) = E(S_n^2)$$ since R_p's are orthonormal.

$$= \sum_{p=1}^{n} E(I_{F_p} S_n^2)$$ since $\sum I_{F_p} \equiv 1$

expand $(S_n \pm S_p)^2 I_{F_p}$ & drop $E(S_n - S_p)^2 I_{F_p}$

$$\geq \sum_{p=1}^{n} E(I_{F_p} S_p^2) + 2\sum_{p=1}^{n} E(S_p I_{F_p}(S_n - S_p))$$

where

$$E(\underbrace{S_p I_{F_p}}_{\text{constant over } \sigma\langle H_1,\dots,H_p\rangle \text{ sets}} \underbrace{(S_n - S_p)}_{\text{constant over } \sigma\langle H_{p+1},\dots,H_n\rangle \text{ sets}}) = E(S_p I_{F_p}) \underbrace{E(S_n - S_p)}_{=0}$$

$\therefore$

$$\sum_{p=1}^{n} E(R_p^2) \geq \sum_{p=1}^{n} E(I_{F_p} S_p^2)$$ on F_p, $|S_p| \geq a$

$$\geq a^2 \sum_{p=1}^{n} E(I_{F_p})$$

$$= a^2 P(\max_{p \leq n} |S_p| \geq a)$$

QED

Application: Series with random signs

$\sum_n \frac{1}{n} = \infty$ } No cancelation at all

$\sum_n R_n \frac{1}{n} = ?$, $\sum_n R_n \frac{1}{\sqrt{n}} = ?$ } cancelation is given by random coin flips

$\sum_n (-1)^n \frac{1}{\sqrt{n}} < \infty$ } cancelation is finely tuned

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

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

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

$$= \frac{1}{a^2} \sum_{n=N}^{M} \frac{1}{n^2}$$

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

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

as $M \to \infty$

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

by continuity from below

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

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

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

$\therefore \quad \forall a > 0$

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

$$\parallel$$

$$P\left(\lim_{N\to\infty}\downarrow \left\{\sup_{N\le n,m} |\tilde{S}_n - \tilde{S}_m| \geq 2a\right\}\right)$$

$\therefore$

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

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

$\Rightarrow$ $\tilde{S}_n$ 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 $\sum_{k=1}^{\infty} X_k$ where the X_k's are indep, $E(X_k) = 0$. See duritt p. 79.

The law of the iterated logorithm

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

Modivation:

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

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

So the WLLN describes the ensemble

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

at a large n, where the SLLN describles the ensemble

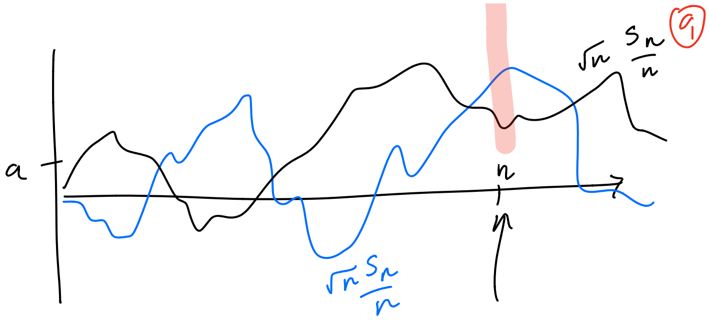
$$\left\{\frac{S_n(w)}{n} : n \in \mathbb{N}\right\}$$

for nearly all $w \in \Omega$.

However both results say $\frac{S_n(w)}{n}$ is "near" 0. To get an idea of rates consider scaling $a_n \frac{S_n(w)}{n}$ 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

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



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

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

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

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

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

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

(i) $P\left(\limsup\limits_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} = 1\right) = 1$

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

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

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

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

(& similarly for (ii)) which follows since

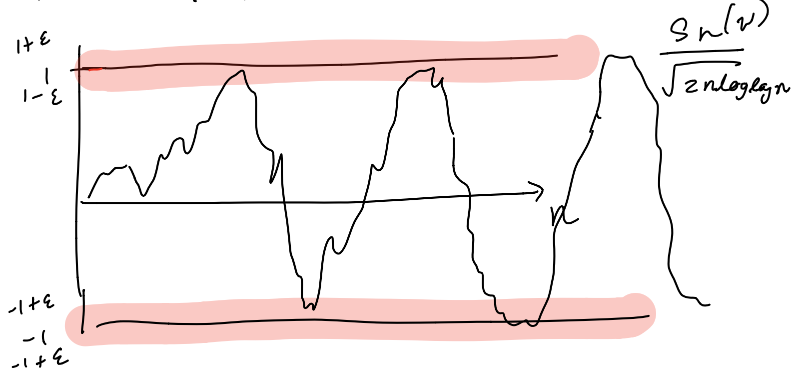
$$\left\{\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} = 1\right\}$$

implies $\limsup > 1-\varepsilon$

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

implies $\limsup < 1+\varepsilon$

This gives a nice picture: for almost all w



11

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 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_k.

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

$$I_k \cap A_k \subset \left\{ \frac{S_n}{\sqrt{2n\log\log n}} > 1+\varepsilon \right\}$$

12

where the SBCL can be applied to I_k & $P(A_k \text{ a.a. } k) = 1$

so $P(I_k \cap A_k \text{ i.o. } k) = 1$.

Instead of writing the proof on the black board lets move to the projector